



SCHOOL OF COMPUTATION,
INFORMATION AND TECHNOLOGY —
INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Bachelor's Thesis in Informatics

On the Formalization of Martingales

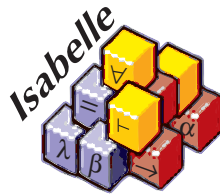
Ata Keskin



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 `filtered_sigma_finite_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

Acknowledgements	iii
Abstract	iv
1 Introduction	1
2 Background and Related Work	3
2.1 Existing Formalizations	3
2.1.1 Lean Mathematical Library	3
2.1.2 Archive of Formal Proofs	5
2.2 Mathematical Foundations and Reference Material	6
3 Conditional Expectation in Banach Spaces	8
3.1 Preliminaries	9
3.1.1 Averaging Theorem	10
3.1.2 Diameter Lemma	12
3.1.3 Induction Schemes for Integrable Simple Functions	14
3.1.4 Bochner Integration on Linearly Ordered Banach Spaces	15
3.2 Constructing the Conditional Expectation	18
3.2.1 Uniqueness	18
3.2.2 Existence	19
3.2.3 Properties of the Conditional Expectation	23
3.3 Conditional Expectation on Linearly Ordered Banach Spaces	26
4 Stochastic Processes	27
4.1 Filtered Measure Spaces	27
4.2 Adapted Processes	27
4.3 Progressively Measurable Processes	27
4.4 Predictable Processes	27
4.5 Discrete Time Processes	27
5 Martingales	28
5.1 Definitions	28

5.2	Necessary and Sufficient Conditions	28
5.3	Discrete-Time Martingales	28
6	Discussion	29
6.1	Formalization Approach	29
6.2	Comparison with Existing Formalizations	29
6.3	Challenges and Limitations	31
6.4	Future Research	31
7	Conclusion	32
	Abbreviations	33
	List of Figures	34
	List of Tables	35
	Bibliography	36

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 integration theory 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]. Also stated on the official documentation on mathlib, 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` $f \mathcal{F} \mu$:

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 S  bastien Gou  zel, 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. Similarly, we define adapted and progressively measurable process. One of the major advantages of our formalization is the use of locales and sublocale relations. Concretely, we will show the relationship

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

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 $f : \iota \rightarrow \Omega \rightarrow E$. 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 $f : 't \rightarrow 'a \text{ measure} \rightarrow 'b$,

where the type $'t$ is an instance of the class `linorder_topology` and the type $'b$ is an instance of the class `banach`. Furthermore, we fix an element t_0 of type $'t$ which represents the smallest element for which the function should be integrable. With this approach we can restrict ourselves to the index set $\{i \mid t_0 \leq i\}$. The fact that we only need to consider this index set, which is bounded from below, will prove to be crucial in certain steps in our development. It will also provide us with the ability to use the reals as an index set. With this specification, our approach mirrors the `mathlib` formalization, since σ -algebras are encoded as measures 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.

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

First, we consider simple functions $s : \Omega \rightarrow E$. 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 $i \in \{1, \dots, n\}$. In this case, we define the Bochner integral simply as the sum

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

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 $f : \Omega \rightarrow E$ 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 $f_n : \Omega \rightarrow E$ such that

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

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 $\int f_n \, d\mu$ form a Cauchy sequence. By completeness, this sequence converges to some element $\lim_{n \rightarrow \infty} \int f_n \, d\mu \in E$. This limit is called the Bochner integral of f with respect to

the measure μ

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

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 $f : \Omega \rightarrow E$ is strongly measurable if and only if it is Σ -measurable.

Consequently, we don't need to concern ourselves with defining 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 $f : \Omega \rightarrow E$ is a Bochner integrable function. In this case, the conditional expectation can be thought of as a linear operator $\mathbb{E}(\cdot | \mathcal{F}) : L^1(E) \rightarrow L^1(E)$ 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 $L^2(E)$ and $L^0(E)$, which is the space of strongly measurable functions with codomain E . Unsurprisingly, the definition of conditional expectation in the last case is a bit more complicated, since it has to take into account the case where f is not integrable.

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.

Without loss of generality, 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. Given a random variable $X : \Omega \rightarrow S$ and an event $A \in \mathcal{F}$, 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:

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

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

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

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 $X : \Omega \rightarrow E$. Given a sub- σ -algebra $\mathcal{H} \subseteq \mathcal{F}$, we call an \mathcal{H} -measurable function $g : \Omega \rightarrow E$ 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 $A \in \mathcal{H}$

$$\int_A X \, d\mu = \int_A g \, d\mu$$

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 $\Omega \rightarrow E$, 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 $\{\emptyset, \Omega\}$. 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 $X : \Omega \rightarrow \mathbb{R}$, we can define a measure ν on (Ω, \mathcal{F}) via

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

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 $g : \Omega \rightarrow \mathbb{R}$ such that

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

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

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

In the second equality, we use the fact that g is \mathcal{H} -measurable. Radon-Nikodym also guarantees 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 $\varepsilon/3 < r < \varepsilon/2$. 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. \square

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 $B_r(v) \cap S = \emptyset$, 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

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

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

Similar to the notation in Isabelle, we will use $-S$ to denote the complement of S . $-S$ 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 $B_{r_i}(w_i) \cap S = \emptyset$ for $i \in \mathbb{N}$. It follows

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

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 \Omega. finite_measure N
    \ N = restrict_space M \Omega
    \ \Omega \in sets M
    \ emeasure N \Omega \neq \infty
    \ emeasure N \Omega \neq 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 ease of use we include additional assumptions such as $\text{emeasure } N \Omega \neq 0$ 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 $d : E \times E \rightarrow \mathbb{R}$. The diameter of a set A is defined as

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

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 $d : E \times E \rightarrow \mathbb{R}$ 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

$$\lim_{n \rightarrow \infty} \text{diam}(S_n) = 0$$

Proof.

\Rightarrow : 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 $(s_i)_{i \in \mathbb{N}}$ is Cauchy, there exists some $N \in \mathbb{N}$ such that $d(s_n, s_m) < 1$ for all $n, m \geq N$. The set $\{s_i \mid i \in \{0, \dots, N\}\}$ is bounded since it is finite. Thus there exists some $a \in \mathbb{R}$ such that $d(s_N, s_i) < a$ for all $i \in \{0, \dots, N\}$. 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 $\text{diam}(S_n) < \infty$ 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

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

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

\Leftarrow : Assume $\lim_{n \rightarrow \infty} \text{diam}(S_n) = 0$ and that S_0 is bounded.

Hence $\text{diam}(S_n) < \infty$ 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. \square

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 \in space M \implies bounded (range (\lambda i. s i x))"
        "\i. (s i) \in borel_measurable M"
shows   "(\lambda x. diameter {s i x \mid i. n \leq i}) \in borel_measurable M"
...

```

Lemma 3.1.10

```
lemma integrable_bound_diameter:
  assumes "integrable M f"
    "\i. (s i) \in borel_measurable M"
    "\x i. x \in space M \implies norm (s i x) \leq f x"
  shows "integrable M (\lambda x. diameter {s i x | i. n \leq 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 $s : \Omega \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$, 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 $s : \Omega \rightarrow E$. 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 simple_integrable_function_induct[case_names cong indicator add]:
  assumes "simple_function M f" "emeasure M {y \in space M. f y \neq 0} \neq \infty"
  assumes cong: "\f g. simple_function M f \implies emeasure M {y \in space M. f y \neq 0} \neq \infty
    \implies simple_function M g \implies emeasure M {y \in space M. g y \neq 0} \neq \infty
    \implies (\lambda x. x \in space M \implies f x = g x) \implies P f \implies P g"
  assumes indicator: "\A y. A \in sets M \implies emeasure M A < \infty
    \implies P (\lambda x. indicator A x \cdot_{\mathbb{R}} y)"
  assumes add: "\f g. simple_function M f \implies emeasure M {y \in space M. f y \neq 0} \neq \infty
    \implies simple_function M g \implies emeasure M {y \in space M. g y \neq 0} \neq \infty
    \implies (\lambda z. z \in space M \implies norm (f z + g z) = norm (f z) + norm (g z))
    \implies P f \implies P g \implies P (\lambda x. 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 $(A_i)_{i=1,\dots,n}$ 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 simple_integrable_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

In the context of real numbers, the following statement is easy to show.

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

In this subsection, we aim to provide similar results for functions $f, g : \Omega \rightarrow E$ 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 $\lim_{n \rightarrow \infty} s_n(x) = f(x)$ μ -a.e. and $\lim_{n \rightarrow \infty} \int s_n \, d\mu = \int f \, d\mu$. 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

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

and

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

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 $\int h \, d\mu \geq 0$. 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 $\int g \, d\mu \geq 0$. It follows directly

$$\int h \, d\mu = \int g \, d\mu \geq 0$$

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

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

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 $\lim_{n \rightarrow \infty} \int \max(0, s_n) \, d\mu = \int \max(0, f) \, d\mu$. Since $f = \max(0, f)$ μ -a.e., we have $\int f \, d\mu = \int \max(0, f) \, d\mu$ and the statement follows. \square

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 $(\forall i. x_i \geq 0) \implies \lim_{i \rightarrow \infty} x_i \geq 0$ 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 ≤ integralL 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 "integrable M f" and nonneg: "AE x in M. 0 ≤ f x"
  shows "integralL M f = 0 ⟷ (AE x in M. f x = 0)"
proof
  assume *: "integralL M f = 0"
  {
    fix A assume asm: "A ∈ sets M"
    have "0 ≤ integralL M (λx. indicator A x · f x)" using nonneg by ...
    moreover have "... ≤ integralL 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"
  "integralL M f = integralL 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 “ $\text{has_cond_exp } M F f g$ ” holds. Then “ $\text{has_cond_exp } M F f (\text{cond_exp } M F f)$ ” and*

$$\text{cond_exp } M F f = g \text{ } \mu\text{-a.e.}$$

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 $A \in \text{sets } F$

$$\int_A f \, d\mu = \int_A g \, d\mu$$

and

$$\int_A f \, d\mu = \int_A \text{cond_exp } M F f \, d\mu$$

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. \square

Hence, we have the following succinct characterization.

Lemma 3.2.6

```
lemma cond_exp_charact:
  assumes "\A \in sets F. \int_A f \, d\mu = \int_A g \, d\mu"
    "integrable M f"
    "integrable M g"
    "g \in 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 S bastien Gou zel 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*

$$\text{has_cond_exp } M F (\mathbf{1}_A \cdot y) ((\text{real_cond_exp } M F \mathbf{1}_A) \cdot y)$$

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

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

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

$$\text{has_cond_exp } M F (f + g) (f' + g')$$

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

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 \in space M. f y \neq 0} \neq \infty"
  shows "has_cond_exp M F f (cond_exp M F f)"
  using assms
proof (induction rule: simple_integrable_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 $(\text{cond_exp } M F s_n)_{n \in \mathbb{N}}$ 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 converges in L^1 . Luckily, this is enough to show that the operator $\text{cond_exp } M F$ preserves limits as a function $L^1(E) \rightarrow L^1(E)$. We need the following lemma for this purpose

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

$$\|\text{cond_exp } M F s\| \leq \text{cond_exp } M F (\lambda x. \|s\|)$$

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`. \square

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

Lemma 3.2.13. *Let $f : \Omega \rightarrow E$ be an integrable function. Let $(s_n)_{n \in \mathbb{N}}$ be a sequence of integrable simple functions, such that $\lim_{n \rightarrow \infty} s_n(x) = f(x)$ 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*

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

and

$$\text{has_cond_exp } M F f \left(\lim_{n \rightarrow \infty} \text{cond_exp } M F s_{r_n} \right)$$

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

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

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

$$\lim_{n \rightarrow \infty} \int \text{diam}(S_n(x)) \, d\mu = 0$$

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

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

by the monotonicity of the integral. Furthermore

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

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

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

Together with the dominated convergence theorem, this implies that

$\lim_{n \rightarrow \infty} (\text{cond_exp } M F s_{r_n})$ is integrable.

As the limit of F -measurable functions, $\lim_{n \rightarrow \infty} (\text{cond_exp } M F s_{r_n})$ is also F -measurable.

Finally, we have for $A \in \text{sets } F$

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

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

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  $\Rightarrow$  'a  $\Rightarrow$  'b"
  assumes " $\bigwedge n$ . integrable M (u n)"
  and " $(\lambda n. (\int \text{norm } (u n x) \, \partial M)) \longrightarrow 0$ "
  shows " $\exists r :: \text{nat} \Rightarrow \text{nat}.$  strict_mono r  $\wedge$  (AE x in M.  $(\lambda n. u (r n) x) \longrightarrow 0$ )"
```

In our case, we can't easily formulate the convergence of $(\text{cond_exp } M F s_n)_{n \in \mathbb{N}}$ 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  $\Rightarrow$  'a  $\Rightarrow$  'b"
  assumes " $\bigwedge n$ . integrable M (s n)"
    and " $\bigwedge e. e > 0 \implies \exists N. \forall i \geq N. \forall j \geq N. \text{LINT } x | M. \text{norm } (s i x - s j x) < e$ "
  obtains r where "strict_mono r" "AE x in M. Cauchy ( $\lambda i. s (r i) x$ )"
```

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: " $\bigwedge i$ . simple_function M (s i)"
    " $\bigwedge i$ . emeasure M {y  $\in$  space M. s i y  $\neq$  0}  $\neq$   $\infty$ "
    " $\bigwedge x. x \in \text{space } M \implies (\lambda i. s i x) \longrightarrow f x$ "
    " $\bigwedge x i. x \in \text{space } M \implies \text{norm } (s i x) \leq 2 \cdot \text{norm } (f x)$ "
    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 $\text{cond_exp } M F f = f$ 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  $\in$  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. $F \subseteq G \subseteq M$. Then, for any $f \in L^1(E)$, we have*

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

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

$$\begin{aligned} \int_A \text{cond_exp } M G f \, d\mu &= \int_A f \, d\mu \\ &= \int_A \text{cond_exp } M F f \, d\mu \end{aligned}$$

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

Contractivity

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

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

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*

$$\|\text{cond_exp } M F f\| \leq \text{cond_exp } M F (\lambda x. \|f(x)\|)$$

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

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

and

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

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

$$\lim_{n \rightarrow \infty} (\text{cond_exp } M F s_{r_n}) = \text{cond_exp } M F f \text{ } \mu\text{-a.e.}$$

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 $(s_{r'_{n'}})_{n \in \mathbb{N}}$ such that

$$\lim_{n \rightarrow \infty} (\text{cond_exp } M F (\lambda x. \|s_{r'_{n'}}(x)\|)) = \text{cond_exp } M F (\lambda x. \|f(x)\|) \quad \mu\text{-a.e.}$$

Furthermore, we have

$$\|\text{cond_exp } M F s_{r'_{n'}}\| \leq \text{cond_exp } M F (\lambda x. \|s_{r'_{n'}}(x)\|) \quad \mu\text{-a.e.}$$

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

Pulling Out What's Known

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

Lemma 3.2.19. *Let $f : \Omega \rightarrow \mathbb{R}$ be an F -measurable function. Let $g \in L^1(E)$ and $f \cdot g \in L^1(E)$. Then*

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

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 $z : \Omega \rightarrow \mathbb{R}$ F -measurable and $z \cdot g \in L^1(E)$:

$$\int z \cdot g \, d\mu = \int z \cdot \text{cond_exp } M F g \, d\mu$$

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

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

and

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

In this case one can easily check that

$$\int s_n \cdot g \, d\mu = \int s_n \cdot \text{cond_exp } M F g \, d\mu$$

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

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

$$|z \cdot \text{cond_exp } M F (\lambda x. \|g(x)\|)| = \text{cond_exp } M F (\lambda x. |z(x) \cdot \|g(x)\||)$$

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

$$\|s_n \cdot \text{cond_exp } M F g\| \leq 2 \cdot \text{cond_exp } M F (\lambda x. |z(x) \cdot \|g(x)\||)$$

Applying the dominated convergence theorem twice, we get

$$\lim_{n \rightarrow \infty} \int s_n \cdot g \, d\mu = \int z \cdot g \, d\mu$$

and

$$\lim_{n \rightarrow \infty} \int s_n \cdot \text{cond_exp } M F g \, d\mu = \int z \cdot \text{cond_exp } M F g \, d\mu$$

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

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

4 Stochastic Processes

4.1 Filtered Measure Spaces

4.2 Adapted Processes

4.3 Progressively Measurable Processes

4.4 Predictable Processes

4.5 Discrete Time Processes

5 Martingales

5.1 Definitions

5.2 Necessary and Sufficient Conditions

5.3 Discrete-Time Martingales

6 Discussion

6.1 Formalization Approach

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
<code>martingale</code>	<code>martingale (locale)</code>
<code>martingale.adapted</code>	<code>adapted_process.adapted</code>
<code>martingale.add</code>	<code>martingale.add</code>
<code>martingale.condexp_ae_eq</code>	<code>martingale.martingale_property</code>
<code>martingale.eq_zero_of_predictable</code>	<code>martingale.predictable_eq_zero</code>
<code>martingale.integrable</code>	<code>martingale.integrable</code>
<code>martingale.neg</code>	<code>martingale.uminus</code>
<code>martingale.set_integral_eq</code>	<code>martingale.set_integral_eq</code>
<code>martingale.smul</code>	<code>martingale.scaleR</code>
<code>martingale.strongly_measurable</code>	<code>stochastic_process.random_variable</code>
<code>martingale.sub</code>	<code>martingale.diff</code>
<code>martingale.submartingale</code>	via sublocale relation
<code>martingale.supermartingale</code>	via sublocale relation
<code>martingale_condexp</code>	<code>filtered_sigma_finite_measure.martingale_cond_exp</code>
<code>martingale_const</code>	<code>filtered_sigma_finite_measure.martingale_const</code>
<code>martingale_const_fun</code>	<code>filtered_sigma_finite_measure.martingale_const</code>
<code>martingale_iff</code>	<code>martingale_iff</code>
<code>martingale_nat</code>	<code>nat_sigma_finite_adapted_process.martingale_nat</code>
<code>martingale_of_condexp_sub_eq_zero_nat</code>	<code>nat_sigma_finite_adapted_process.martingale_of_cond_exp_diff_Suc_eq_zero</code>
<code>martingale_of_set_integral_eq_succ</code>	<code>nat_sigma_finite_adapted_process.martingale_of_set_integral_eq_Suc</code>
<code>martingale_zero</code>	<code>filtered_sigma_finite_measure.martingale_zero</code>

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.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.max
submartingale.zero_le_of_predictable	nat_submartingale.predictable_ge_bot
submartingale_nat	nat_sigma_finite_adapted_process.submartingale_nat
submartingale_of_condexp_sub_nonneg	sigma_finite_adapted_process.submartingale_of _cond_exp_diff_nonneg
submartingale_of_condexp_sub_nonneg_nat	nat_sigma_finite_adapted_process.submartingale_of _cond_exp_diff_Suc_nonneg
submartingale_of_set_integral_le	sigma_finite_adapted_process.submartingale_of _set_integral_le
submartingale_of_set_integral_le_succ	nat_sigma_finite_adapted_process.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.set_integral_ge

<code>supermartingale.smul_nonneg</code>	<code>supermartingale.scaleR_nonneg</code>
<code>supermartingale.smul_nonpos</code>	<code>supermartingale.scaleR_nonpos</code>
<code>supermartingale.strongly_measurable</code>	<code>stochastic_process.random_variable</code>
<code>supermartingale.sub_martingale</code>	<code>supermartingale.diff</code>
<code>supermartingale.sub_submartingale</code>	<code>supermartingale.diff</code>
<code>supermartingale_nat</code>	<code>nat_sigma_finite_adapted_process.supermartingale_nat</code>
<code>supermartingale_of_condexp_sub_nonneg_nat</code>	<code>nat_sigma_finite_adapted_process.supermartingale_of</code>
	<code>_cond_exp_diff_Suc_nonneg</code>
<code>supermartingale_of_set_integral_succ_le</code>	<code>nat_sigma_finite_adapted_process.supermartingale_of</code>
	<code>_set_integral_le_Suc</code>

Table 6.3: Lookup table for supermartingale lemmas and definitions

6.3 Challenges and Limitations

6.3.1

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

6.1	Lookup Table for Martingale Lemmas and Definitions	29
6.2	Lookup Table for Submartingale Lemmas and Definitions	30
6.3	Lookup Table for Supermartingale Lemmas and Definitions	31

Bibliography

- [DY22a] R. Degenne and K. Ying. *A Formalization of Doob's Martingale Convergence Theorems in mathlib*. 2022. arXiv: 2212.05578 [cs.LG].
- [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.
- [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/CB09780511813603.
- [NPW02] T. Nipkow, L. C. Paulson, and M. Wenzel. *Isabelle/HOL — A Proof Assistant for Higher-Order Logic*. Vol. 2283. LNCS. Springer, 2002.
- [Pis16] G. Pisier. *Martingales in Banach Spaces*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2016. DOI: 10.1017/CB09781316480588.
- [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.
- [YB89] R. J. Yaes and A. S. Bechhoefer. “The Efficient Market Hypothesis.” In: *Science* 244.4911 (1989), pp. 1424–1424. ISSN: 00368075, 10959203.