Three Chapters of Measure Theory in Isabelle/HOL

Johannes Hölzl* and Armin Heller

Technische Universität München
Institut für Informatik, Boltzmannstr. 3, 85748 Garching bei München, Germany
`http://home.in.tum.de/~hoelzl/`

Abstract. Currently published HOL formalizations of measure theory concentrate on the Lebesgue integral and they are restricted to real-valued measures. We lift this restriction by introducing the extended real numbers. We define the Borel σ-algebra for an arbitrary type forming a topological space. Then, we introduce measure spaces with extended real numbers as measure values. After defining the Lebesgue integral and verifying its linearity and monotone convergence property, we prove the Radon-Nikodým theorem (which shows the maturity of our framework). Moreover, we formalize product measures and prove Fubini's theorem. We define the Lebesgue measure using the gauge integral available in Isabelle's multivariate analysis. Finally, we relate both integrals and equate the integral on Euclidean spaces with iterated integrals. This work covers most of the first three chapters of Bauer's measure theory textbook.

1 Introduction

Measure theory plays an important role in modeling the physical world, and in particular is the foundation of probability theory. Current HOL formalizations of measure theory mostly concentrate on the Lebesgue integral [2, 10, 11]. We extend this by a number of fundamental concepts:

Lebesgue measure To use the Lebesgue integral for functions on a real domain we need to introduce a measure on $\mathbb{R}$. The Lebesgue measure $\boldsymbol{\lambda}$ assigns the length $b - a$ to every interval $[a, b]$, and is closed under countable union and difference. The Lebesgue integral on $\boldsymbol{\lambda}$ is an extension of the Riemann integral.

Product measure Defines a measure on the product of two or more measure spaces. We can also represent Euclidean spaces as products of the Lebesgue measure on $\mathbb{R}$. This is also necessary to prove Fubini's theorem, i.e., the commutativity of integrals.

---

* Supported by the DFG Graduiertenkolleg 1480 (PUMA).

[0] Published at ITP 2011.
The original publication is available at `www.springerlink.com`

Extended real numbers The introduction of Lebesgue measure requires infinite measure values, hence we introduce the extended real numbers and use them as measure values.

Radon-Nikodým derivative Given two measures ν and μ, we can represent ν with density f with respect to μ, under certain assumptions.

$$\nu\ A = \int_A f\ d\mu$$

This density f is called the Radon-Nikodým derivative. The existence of such a density is used in information theory to define *mutual information* and in probability theory to define *conditional expectation.*

Restricted forms of these concepts where already formalized in HOL theorem provers [2,5,6,8,10,11]. By formalizing these concepts in a more generic way, it is now possible to combine the results of these works. With the Lebesgue measure, the Radon-Nikodým theorem and the product measure we formalize most of the first three chapters ($\sim$ 70 pages) of Bauer's textbook about measure theory [1]. We only show the theorem statements, but the full proofs are publicly available in the current development version of Isabelle.[1]

2 Preliminaries

We use the following concepts and notations: We write the power set as $\mathcal{P}(A) = \{B \mid B \subseteq A\}$, the universe for type α as $\mathcal{U} :: \alpha\ set$, $\perp :: \alpha$ is an arbitrary element of type α, the image of A under f is $f[A] = \{f\ x | x \in A\}$, the range of f is $rng\ f = f[\mathcal{U}]$, the preimage of B under f is $f^{-1}[B] = \{x |\ f\ x \in B\}$, the cartesian product is $A \times B = \{(a,b) \mid a \in A, b \in B\}$, and the dependent product is $\Pi x{\in}A.\ B\ x = \{f |\ \forall x \in A.\ f\ x \in B\ x\}$ and $A \rightarrow B = \Pi x{\in}A. B$. The indicator function $\chi_A\ x = 1$ if $x \in A$ otherwise $\chi_A\ x = 0$. With $f \uparrow x$ we state that f converges monotonically from below to x, this is defined for functions with range $\overline{\mathbb{R}}$ and sets. Hilbert choice is $SOME\ x.\ P\ x$, i.e. $(\exists x.Px) \longrightarrow P(SOME\ x.P\ x)$ holds. For the product spaces we also use the extensional dependent product

$$\Pi_E x \in A.\ B\ x = \{f |\ (\forall x \in A.\ f\ x \in B\ x) \wedge (\forall x \notin A.\ f\ x = \perp)\}$$

We need to enforce having exactly one value outside of the domain, otherwise there is more than one function with the same values on A. We use $t :: \alpha$ to annotate types and $\alpha \Rightarrow \beta$ for function types. These should not be confused with set membership $x \in A$ or the dependent function space $A \rightarrow B$, which are predicates and not type annotations. We use $t :: (\alpha :: \text{type\_class})$ to annotate type classes.

The **locale**-command introduces a new locale [4]. We use it to define the concept of algebras, σ-algebras, measure spaces etc.

locale $l = bs +$ **fixes** $x :: \alpha$ **assumes** $P_1\ x$ **and** $\ldots$ **and** $P_n\ x$

[1] `http://isabelle.in.tum.de/repos/isabelle/file/tip/src/HOL/Probability`

This introduces the locale l with a variable x and the assumptions $P_1, \ldots, P_n$. It inherits the context like variables and assumptions, but also theorems, abbreviations, definitions, setup for the proof methods and more from bs. We get the theorems about a specific instantiation $l\ x$ by proving $P_1\ x \wedge \cdots \wedge P_n\ x$. When we prove a theorem in a locale we have access to the theorems of bs, i.e. a lemma in *algebra* is immediately available in the *sigma_algebra* locale.

3 Extended Reals

The Lebesgue measure $\boldsymbol{\lambda}$ takes infinite values, as there is no real number we can reasonably assign to $\boldsymbol{\lambda}(\mathbb{R})$. So we need a type containing the real numbers and a distinct value for infinity. We introduce the type $\overline{\mathbb{R}}$ as the reals extended with a positive and a negative infinite element.

Definition 1 (Extended reals $\overline{\mathbb{R}}$).

$$\textit{datatype}\ \overline{\mathbb{R}} = \infty \mid (\mathbb{R})_\infty \mid -\infty \qquad \textit{real} :: \overline{\mathbb{R}} \Rightarrow \mathbb{R}$$
$$\textit{real}\ (r_\infty) = r \qquad \textit{real}\ \infty = 0 \qquad \textit{real}\ (-\infty) = 0$$

The conversion function *real* restricts the extended reals to the real numbers and maps $\pm\infty$ to 0. For the sake of readability we hide this conversion function.

Definition 2 (Order and arithmetic operations on $\overline{\mathbb{R}}$).

$$r_\infty \leqslant p_\infty \longleftrightarrow r \leqslant p \qquad x \leqslant \infty \qquad -\infty \leqslant x$$
$$-(r_\infty) = (-r)_\infty \qquad -(-\infty) = \infty$$
$$r_\infty + p_\infty = (r+p)_\infty \qquad \infty + x = \infty \qquad x + \infty = \infty$$
$$r_\infty \cdot p_\infty = (r \cdot p)_\infty \qquad x \cdot \pm\infty = \pm\infty \cdot x = \begin{cases} 0 & \textit{if}\ \ x = 0 \\ \operatorname{sgn} x \cdot \pm\infty & \textit{otherwise} \end{cases}$$

For measure theory it is suitable to define $\infty \cdot 0 = 0$. Using *min* and *max* as join and meet, we get that $\overline{\mathbb{R}}$ is a complete lattice where *bot* is $-\infty$ and *top* is ∞.

Our next step is to define the topological structure on $\overline{\mathbb{R}}$. This is an extension of the topological structure on real numbers. However we need to take care of what happens when $\pm\infty$ is in the set.

Definition 3. $\textit{open}\ A \longleftrightarrow \textit{open}\ \{r \mid r_\infty \in A\}\ \wedge$
$(\infty \in A \longrightarrow \exists x. \forall y > x.\ y_\infty \in A) \wedge (-\infty \in A \longrightarrow \exists x. \forall y < x.\ y_\infty \in A)$

From this definition the continuity of $\cdot_\infty$ follows directly. The definition of limits of sequences in Isabelle/HOL is based on topological spaces. This allows us to reuse these definitions and also some of the proofs such as uniqueness of limits. We also verify that the limits and infinite sums on real numbers are the same as the limits and sums on extended reals:

Lemma 1. $(\lambda n. (f\ n)_\infty) \xrightarrow[n\to\infty]{} r_\infty \quad \longleftrightarrow \quad (\lambda n.\ f\ n) \xrightarrow[n\to\infty]{} r$

Corollary 1. *If f is summable, then* $\sum_n (f\ n)_\infty = (\sum_n f\ n)_\infty$.

Hurd [7] formalizes similar positive extended reals and also defines a complete lattice on them. Our $\overline{\mathbb{R}}$ includes negative numbers and we not only show that it forms a complete lattice but also that it forms a topological space. The complete lattice is used for monotone convergence and the topological space is used to define a Borel σ-algebra on $\overline{\mathbb{R}}$.

4 Measure Theory

We largely follow Bauer's textbook [1] for our formalization of measure theory. An exception is the definition of the Lebesgue integral which is taken from Schilling [12].

4.1 The σ-algebra

We use records to represent (σ-)algebras and measure spaces. We define measure spaces as extensions to algebras, hence we can use measure spaces as algebras.

$$\textbf{record}\ \alpha\ algebra = space :: \alpha\ set \quad sets :: \alpha\ set\ set$$

To represent the algebra $M = (\Omega, \mathcal{A})$ we write $M = (\!|space = \Omega, sets = \mathcal{A}|\!)$. We use this type to introduce the concept of (σ-)algebras. The set Ω is typically but not necessarily the universe $\mathcal{U}$. For probability theory in particular, it is often $[0, 1]$ instead of $\mathbb{R}$. The sets in $\mathcal{A}$ are the measurable sets.

Definition 4 (σ-algebra).

locale $algebra$ =
  fixes $M :: \alpha\ algebra$
  assumes $sets\ M \subseteq \mathcal{P}(space\ M)$
  and $\emptyset \in sets\ M$
  and $\forall a \in sets\ M.\ space\ M - a \in sets\ M$
  and $\forall a, b \in sets\ M.\ a \cup b \in sets\ M$

locale $sigma\_algebra$ = $algebra$ +
  assumes $\forall F :: nat \Rightarrow \alpha\ set.\ rng\ F \subseteq sets\ M \longrightarrow (\bigcup_i F\ i) \in sets\ M$

The easiest way to define a σ-algebra (other than the power set) is to give a generator $\mathcal{G}$ and use the smallest σ-algebra containing $\mathcal{G}$ (called its σ-*closure*).

Definition 5 (σ-closure). *$sigma\_sets\ \mathcal{G}\ \Omega$ denotes the smallest superset of $\mathcal{G}$ containing $\emptyset$ and is closed under Ω-complement and countable union.*

inductive $sigma\_sets$ **for** $\mathcal{G}$ **and** Ω **where**
$a \in \mathcal{G} \longrightarrow a \in sigma\_sets\ \Omega\ \mathcal{G}$
$\emptyset \in sigma\_sets\ \Omega\ \mathcal{G}$
$a \in sigma\_sets\ \Omega\ \mathcal{G} \longrightarrow \Omega - a \in sigma\_sets\ \Omega\ \mathcal{G}$
$rng\ (F :: nat \Rightarrow \alpha\ set) \subseteq sigma\_sets\ \Omega\ \mathcal{G} \longrightarrow (\bigcup_i F\ i) \in sigma\_sets\ \Omega\ \mathcal{G}$

$sigma\ M = (\!|space = space\ M, sets = sigma\_sets\ (space\ M)\ (sets\ M)|\!)$

We define the σ-closure inductively to get a nice induction rule. Then we show that it actually is the smallest σ-algebra containing $\mathcal{G}$.

Lemma 2. *The sigma operator generates a σ-algebra.*
$\mathit{sets}\ M \subseteq \mathcal{P}(\mathit{space}\ M) \longrightarrow \mathit{sigma\_algebra}\ (\mathit{sigma}\ M)$

Lemma 3. *If $\mathcal{G} \subseteq \mathcal{P}(\Omega)$ then*
$\mathit{sigma\_sets}\ \Omega\ \mathcal{G} = \bigcap \{B \supseteq \mathcal{G} |\ \mathit{sigma\_algebra}\ (\!|\mathit{space} = \Omega, \mathit{sets} = B|\!)\}$

Measurable functions. When preimages of measurable sets in M_2 under f are measurable sets in M_1 we say f is M_1-M_2-*measurable*. We use the function-type to represent them, but restrict it to the functions from *space* M_1 to *space* M_2. We also need to intersect the preimage under f with *space* M_1.

Definition 6 (Measurable). $\mathit{measurable}\ M_1\ M_2 =$
$\{f \in \mathit{space}\ M_1 \to \mathit{space}\ M_2 |\ \forall A \in \mathit{sets}\ M_2.\ f^{-1}[A] \cap \mathit{space}\ M_1 \in \mathit{sets}\ M_1\}$

When M_2 is generated by a σ-closure it is enough to show that it is measurable on the generator:

Lemma 4. *If $\mathit{sets}\ G \subseteq \mathcal{P}(\mathit{space}\ G)$ and $f \in \mathit{measurable}\ M_1\ G$ then*
$f \in \mathit{measurable}\ M_1\ (\mathit{sigma}\ G)$.

Borel σ-algebra. The σ-algebra generated by the open sets of a topological space is called a Borel σ-algebra. In Isabelle/HOL topological spaces form a type class defining the open sets. Instances are Euclidean spaces (hence $\mathbb{R}$) and $\overline{\mathbb{R}}$.

Definition 7 (Borel sets).
$\mathit{borel} = \mathit{sigma}\ (\!|\mathit{space} = \mathcal{U} :: (\alpha :: \mathit{topological\_space})\ \mathit{set}, \mathit{sets} = \{S |\ \mathit{open}\ S\}|\!)$

As a first step we show that the Borel sets on real numbers are not only generated by all the open sets, but also by all the intervals $]-\infty, a[$. Then we show the Borel measurability of arithmetic operations, min, max, etc. To show the measurability of these operations on $\overline{\mathbb{R}}$ we first show that $\cdot_\infty$ and *real* are Borel-Borel-measurable (which follows from their continuity). The operations on $\overline{\mathbb{R}}$ are compositions of $\cdot_\infty$ and *real* with operations on real numbers. We use "M-measurable" as abbreviation for "M-Borel-measurable."

Dynkin systems. We use Dynkin systems to prove the uniqueness of measures. Compared with σ-algebras, they are only closed under countable unions *if the sets are disjoint.*

Definition 8. $\mathit{disjoint}\ F \longleftrightarrow (\forall i, j.\ i \neq j \longrightarrow F\ i \cap F\ j = \emptyset)$

Definition 9 (Dynkin system).

$$\begin{aligned}
&\textbf{locale } \textit{dynkin_system} \ = \\
&\qquad \textbf{fixes } D :: \alpha \ \textit{algebra} \\
&\quad \textbf{assumes } \textit{sets } D \subseteq \mathcal{P}(\textit{space } D) \\
&\qquad \textbf{and } \varnothing \in \textit{sets } D \\
&\qquad \textbf{and } \forall a \in \textit{sets } D.\ \textit{space } D - a \in \textit{sets } D \\
&\qquad \textbf{and } \forall F.\ \textit{disjoint } F \wedge \textit{rng } F \subseteq \textit{sets } D \longrightarrow (\textstyle\bigcup_i F\ i) \in \textit{sets } D
\end{aligned}$$

Definition 10 (Closed under intersection).
$\cap$-*stable* $G \longleftrightarrow (\forall A, B \in \textit{sets } G.\ A \cap B \in \textit{sets } G)$

Dynkin systems are now used to prove Dynkin's lemma, which helps to generalize statements about all sets of a $\cap$-*stable* set to the σ-closure of that set. We use Dynkin's lemma to prove the uniqueness of measures.

Theorem 1 (Dynkin's lemma). *For any Dynkin system D and $\cap$-stable system G, if sets $G \subseteq$ sets $D \subseteq$ sets (sigma G), then sigma $G = D$.*

4.2 Measure spaces

A measure space is a σ-algebra extended with a measure which maps measurable sets to nonnegative, possibly infinite *measure* values. We introduce a new type *measure_space* which extends the *algebra* record. We represent measure values with $\overline{\mathbb{R}}$, and abbreviate $\Omega = \textit{space } M$, $\mathcal{A} = \textit{sets } M$, and $\mu = \textit{measure } M$.

Definition 11 (Measure space).

$$\begin{aligned}
&\textbf{record } \textit{measure_space} = \alpha \ \textit{algebra} \ + \ \textit{measure} :: (\alpha \ \textit{set}) \Rightarrow \overline{\mathbb{R}} \\
&\textbf{locale } \textit{measure_space} = \textit{sigma_algebra } M \textbf{ for } M :: \alpha \ \textit{measure_space} + \\
&\quad \textbf{assumes } \mu\ \varnothing = 0 \\
&\qquad \textbf{and } \forall A \in \mathcal{A}.\ 0 \leqslant \mu\ A \\
&\qquad \textbf{and } \forall F.\ \textit{disjoint } F \wedge \textit{rng } F \subseteq \mathcal{A} \longrightarrow \mu\ \left(\textstyle\bigcup_i F\ i\right) = \textstyle\sum_i \mu\ (F\ i)
\end{aligned}$$

In the remaining sections we fix the measure space M. We prove the additivity, monotonicity, and continuity from above and below for measures. For proving the existence of a measure we provide Caratheodory's theorem, which was ported from Hurd [6] and Coble [2].

Theorem 2 (Caratheodory). *Assume G is an algebra, and let be f a function such that f is nonnegative on $\mathcal{A}$, $f\ \varnothing = 0$, and is f countably additive, i.e.,*

$$\forall F.\ \textit{disjoint } F \wedge \textit{rng } F \subseteq \textit{sets } G \longrightarrow f\ \left(\textstyle\bigcup_i F\ i\right) = \textstyle\sum_i f\ (F\ i)$$

then there exists a ν s.t. $\forall A \in$ sets $G.\ \nu\ A = f\ A$ and sigma $G(\!|\textit{measure} := \nu|\!)$ is a measure space.

For our purposes to formalize product measures and to equate the products of the Lebesgue measure, we prove the uniqueness of measures.

Theorem 3 (Uniqueness of measures). *Assume*

- μ *and* ν *are two measures on sigma* G
- G *is* $\cap$*-stable*
- C *is a* σ*-finite cover of* G*:* $rng\ C \subseteq sets\ G$, $C \uparrow space\ G$, *and* $\forall i.\ \mu(C\ i) < \infty$
- μ *and* ν *are equal on* G*:* $\forall X \in sets\ G.\ \mu\ X = \nu\ X$

then μ *and* ν *are equal on sigma* G*.*

An important class of measure spaces are σ-finite measure spaces. It requires a sequence of finitely measurable sets which cover the entire space. The product measure and the Radon-Nikodým theorem assume a σ-finite measure.

Definition 12 (σ-finite measure space).

$$\textbf{locale}\ sigma\_finite\_measure = measure\_space\ +$$
$$\textbf{assumes}\ \exists F.\ rng\ F \subseteq \mathcal{A} \wedge F \uparrow \Omega\ \wedge \forall i.\ \mu\ (F\ i) < \infty$$

Almost everywhere. Often predicates on measure spaces do not hold for all elements in a measure space, but the elements where they do not hold form a subset of a null set. In textbooks this is often written without an explicitly quantified variable but rather with an appended "a.e." (standing for "almost everywhere"). We use a syntax with an explicit binder.

Definition 13 (Almost everywhere).
$(AE\,x.\ P\ x) \longleftrightarrow \exists N \in \mathcal{A}.\ \{x \in \Omega |\ \neg\ P\ x\} \subseteq N \wedge \mu\ N = 0$

The definition of almost everywhere in [6] and [10] assumes that $\{x \in \Omega |\ \neg\ P\ x\}$ is a null set, i.e. it is also measurable.

Theorem 4 (AE modus ponens).
$(AE\,x.\ P\ x) \longrightarrow (AE\,x.\ P\ x \longrightarrow Q\ x) \longrightarrow (AE\,x.\ Q\ x)$

Our relaxed definition requires no measurability of Q in the *modus ponens* rule of almost everywhere.

Theorem 5. $(\forall x \in \Omega.\ P\ x) \longrightarrow (AE\,x.\ P\ x)$

Let us take a look at the statement $(AE\,x.\ f\ x < g\ x) \longrightarrow (AE\,x.\ f\ x \leqslant g\ x)$. This can be directly solved by AE modus ponens and theorem 5. The measurability of f and g is not required.

4.3 Lebesgue Integral

The definition of the Lebesgue integral requires the concept of simple functions. A simple function is a Borel-measurable step function (i.e. its range is a finite set), or equivalently a step function where the preimage of every singleton set containing an element of the range is measurable. The second formulation has the advantage that the definition does not require the notion of Borel σ-algebras and is thus more general, as it allows arbitrary ranges. The predicate *simple_function* is defined as follows:

Definition 14 (Simple function).

$$\textit{simple\_function } f \longleftrightarrow \textit{finite } f[\Omega] \wedge \forall x \in f[\Omega].\ f^{-1}[\{x\}] \cap \Omega \in \mathcal{A}$$

While we use this definition only for functions $f :: \alpha \Rightarrow \overline{\mathbb{R}}$, this is a nice characterisation for finite random variables in probability theory. When the range of f is $\overline{\mathbb{R}}$ it is also representable as sum:

Lemma 5.

$$\forall x \in \Omega.\ f\ x = \sum_{y \in f[\Omega]} y \cdot \chi_{f^{-1}[\{y\}] \cap \Omega}\ x$$

This already suggests the definition of the integral $\int^S$ of a simple function f with respect to the measure space M:

Definition 15 (Simple integral). *Let f be a simple function.*

$$\int^S f\ dM = \sum_{y \in f[\Omega]} y \cdot \mu(f^{-1}[\{y\}] \cap \Omega)$$

To state the definition of the integral of functions $f :: \alpha \Rightarrow \overline{\mathbb{R}}$, simple functions have to be used as approximations from below of f. Then the integral is defined as the supremum of all the simple integrals of the approximations.

Definition 16 (Positive integral).

$$\int^+ f\ dM = \sup \left\{ \int^S g\ dM \ \middle|\ g \leqslant f^+ \wedge \textit{simple\_function } g \right\}$$

The function f^+ is the nonnegative part of f, i.e. f^+ is zero when f is negative, otherwise it is equal to f. Hence the positive integral is equal when the integrating functions are almost everywhere equal. Finally integration can be defined for functions $f :: \alpha \Rightarrow \mathbb{R}$ as usual.

Definition 17 (Lebesgue Integrability and Integral).

$$\begin{aligned} \textit{integrable } M\ f \longleftrightarrow\ & f \in \textit{measurable } M \textit{ borel} \wedge \\ & \left(\int^+ x.\ (f\ x)_\infty\ dM\right) < \infty \wedge \left(\int^+ x.\ (-f\ x)_\infty\ dM\right) < \infty \\ \int f\ dM \quad = \quad & \left(\int^+ x.\ (f\ x)_\infty\ dM\right) - \left(\int^+ x.\ (-f\ x)_\infty\ dM\right) \end{aligned}$$

(Note that explicit type conversions from $\overline{\mathbb{R}}$ to $\mathbb{R}$ have been omitted in this definition for the sake of readability.)

Remark: Textbooks usually write $\int f d\mu(x)$, where we instead specify the entire measure space M and optionally bind the variable x directly after the integral symbol, $\int x.\, f\, x\, dM$. If no variable is needed we write $\int f\, dM$, and a restricted integral is abbreviated as $\int x \in A. f\ x\ dM = \int x. f\ x \cdot \chi_A\ x\ dM$.

Many proofs of properties about the integral follow the scheme of the definitions and first establish the desired property for $\int^S$, then for $\int^+$, and eventually for $\int$. The monotonicity of the integral is proven this way, for example.

Lemma 6 (Monotonicity). *If f and g are measurable functions, then*

$$(AE\,x.\ f\ x \leqslant g\ x) \longrightarrow \int f\,dM \leqslant \int g\,dM$$

Another way of constructing proofs about Borel-measurable functions $u :: \alpha \Rightarrow \overline{\mathbb{R}}$ is: first, prove the desired property about finite simple functions, then, prove that the property is preserved under the pointwise monotone limit of functions. For this to work, we need a lemma stating that every Borel-measurable function $u :: \alpha \Rightarrow \overline{\mathbb{R}}$ can be seen as the limit of a monotone sequence of finite simple functions.

Lemma 7. *Let u be a nonnegative and measurable function.*

$$\exists f.\ (\forall i.\ \mathit{simple\_function}\ (f\ i) \wedge (\forall x \in \Omega.\ 0 \leqslant\ f\ i\ x\ \neq \infty)) \wedge f \uparrow u$$

To use this with the Lebesgue integral, there is a compatibility theorem, called the monotone convergence theorem, which allows switching the supremum operator and the positive integral.

Lemma 8 (Monotone convergence theorem). *Let $f :: \mathbb{N} \Rightarrow \alpha \Rightarrow \overline{\mathbb{R}}$ be a sequence of nonnegative Borel-measurable functions, such that $\forall i.\forall x \in \Omega.\ f\ i\ x \leqslant f\ (i+1)\ x$. Then it holds that:*

$$\sup\ i.\ \int^{+} f\ i\ dM = \int^{+} (\sup\ i.\ f\ i)\ dM$$

The Monotone convergence theorem is used in the proof of Fubini's theorem. Another useful convergence theorem is the dominated convergence theorem. It can be used when the monotonicity of the function sequence does not hold.

Lemma 9 (Dominated convergence theorem). *Let $u :: \mathbb{N} \Rightarrow \alpha \Rightarrow \mathbb{R}$ be a sequence of integrable functions, $w :: \alpha \Rightarrow \mathbb{R}$ be an integrable function, and $v :: \alpha \Rightarrow \mathbb{R}$ be a function. If $(\forall i.\ |u\ i\ x|\ \leqslant\ w\ x)$ and $(\lambda i.\ u\ i\ x) \longrightarrow_{\infty} v\ x$ for all $x \in \Omega$ then integrable $M\ v$ and $\left(\lambda i.\ \int u\ i\ dM\right) \longrightarrow_{\infty} \int v\ dM$.*

To transfer results about integrals from one measure space to another one, the following transformation lemma can be used.

Lemma 10. *If T is M-M'-measurable and measure $M'\ A$ equals $\mu\ (T^{-1}[A] \cap \Omega)$ for all $A \in$ sets M' and f is M'-integrable, then $f \circ T$ is M'-integrable and*

$$\int f\ dM' = \int f \circ T\ dM$$

4.4 Radon-Nikodým derivative

The Radon-Nikodým theorem states that for each measure ν that is absolutely continuous on M there exists an a.e.-unique density function to represent ν on M. This is needed to define *conditional expectation* in probability theory and *mutual information* in information theory. In this section we assume that M is σ-finite.

Definition 18 (Radon-Nikodým derivative).

$$
\begin{aligned}
&\mathit{RN\_deriv}\ M\ \nu = \mathit{SOME}\ f \in \mathit{measurable}\ M\ \mathit{borel}.\\
&\quad (\forall x \in \Omega.\ 0 \leqslant f\ x) \wedge \left(\forall X \in \mathcal{A}.\ \nu\ X = \left(\textstyle\int^{+} x \in X.\ f\ x\ dM\right)\right)
\end{aligned}
$$

To work with this definition we need to prove the existence of such a function.

Theorem 6 (Radon-Nikodým). *If ν is a measure on M and ν is absolutely continuous w.r.t. M, i.e., $\forall A \in \mathcal{A}.\ \mu\ A = 0 \longrightarrow \nu\ A = 0$ then*

$$
\begin{aligned}
&\mathit{RN\_deriv}\ M\ \nu \in \mathit{measurable}\ M\ \mathit{borel}\\
&\wedge\ \forall X \in \mathcal{A}.\ \nu\ X = \left(\textstyle\int^{+} x \in X.\ \mathit{RN\_deriv}\ M\ \nu\ x\ dM\right)
\end{aligned}
$$

The next theorem shows that two functions are a.e.-equal when they are equal on all measurable sets, hence follows the uniqueness of *RN_deriv*.

Theorem 7. *If f and g are nonnegative and M-measurable and* $\forall A \in \mathcal{A}.\ \left(\int^{+} x \in A.\ f\ x\ dM\right) = \left(\int^{+} x \in A.\ g\ x\ dM\right)$ *then* $(AE\, x.\ f\ x = g\ x)$

4.5 Product Measure and Fubini's theorem

We first introduce the binary product of measure spaces, and later extend this to arbitrary, finite products of measure spaces.

Binary product measure. The definition of a measure on a binary product σ-algebra is straightforward. All we need to do is compose the integrals of both measure spaces. With Fubini's theorems we later show that the result is independent of the order of integration.

Definition 19.

$$
\begin{aligned}
\mathit{bin\_algebra}_G &:: \alpha\ \mathit{measure\_space} \Rightarrow \beta\ \mathit{measure\_space}\\
&\quad \Rightarrow (\alpha \times \beta)\ \mathit{measure\_space}\\
\mathit{bin\_algebra}_G\ M_1\ M_2 &= (\!|\mathit{space} = \mathit{space}\ M_1 \times \mathit{space}\ M_2,\\
&\quad \mathit{sets} = \{A \times B | A \in \mathit{sets}\ M_1, B \in \mathit{sets}\ M_2\},\\
&\quad \mathit{measure} = \textstyle\int^{+} x.\ \left(\int^{+} y.\ \chi_{A \times B}\ (x, y)\ dM_2\right) dM_1|\!)\\
M_1 \bigotimes\nolimits_m M_2 &= \mathit{sigma}(\mathit{bin\_algebra}_G\ M_1\ M_2)
\end{aligned}
$$

In this section we assume that M_1 and M_2 are σ-finite measure spaces. We verify the definition of the binary product measure by applying the measure to an element $A \times B$ from the generating set of $M_1 \bigotimes_m M_2$.

Lemma 11. *If $A \in \mathit{sets}\ M_1$ and $B \in \mathit{sets}\ M_2$ then* $\mathit{measure}\ (M_1 \bigotimes_m M_2)\ (A \times B) = \mathit{measure}\ M_1\ A \cdot \mathit{measure}\ M_2\ B$.

Lemma 12. *Show the measurability of the cut* $\{y|(x,y) \in Q\}$ *for all* $Q \in$ *sets* $(M_1 \bigotimes_m M_2)$ *and all* x.

$$\{y|(x,y) \in Q\} \in \textit{sets } M_2 \tag{1}$$
$$(\lambda x.\ \textit{measure } M_2\ \{y|(x,y) \in Q\}) \in \textit{measurable } M_1\ \textit{borel} \tag{2}$$
$$\textit{measure } (M_1 \bigotimes_m M_2)\ Q = \textstyle\int^{+} x.\ \textit{measure } M_2\ \{y|(x,y) \in Q\}\ dM_1 \tag{3}$$

Theorem 8. *sigma_finite_measure* $(M_1 \bigotimes_m M_2)$

Fubini's theorem. From the product measure we get directly to the fact that integrals on σ-finite measure spaces are commutative.

Lemma 13. *If* f *is* $M_1 \bigotimes_m M_2$*-measurable then*

$$\left(\lambda x. \textstyle\int^{+} y.\ f\ (x,y)\ dM_2\right) \in \textit{measurable } M_1\ \textit{borel} \quad \textit{and}$$

$$\textstyle\int^{+} x.\ \left(\int^{+} y.\ f\ (x,y)\ dM_1\right) dM_2 = \int^{+}\ f\ d\,(M_1 \bigotimes_m M_2)\ .$$

With theorem 3 we show that the pair swap function $(\lambda(x,y).(y,x))$ is measure preserving between $M_1 \bigotimes_m M_2$ and $M_2 \bigotimes_m M_1$. This allows us to get symmetric variants of (1), (2), and (3) without reproducing a symmetric proof.

Corollary 2 (Fubini's theorem on $\overline{\mathbb{R}}$). *If* f *is* $M_1 \bigotimes_m M_2$*-measurable then*

$$\textstyle\int^{+} x.\ \left(\int^{+} y.\ f\ (x,y)\ dM_2\right)\ dM_1 = \int^{+} y.\ \left(\int^{+} x.\ f\ (x,y)\ dM_1\right)\ dM_2$$

Lemma 13 can be extended to integrability on real numbers.

Lemma 14. *If* f *is* $M_1 \bigotimes_m M_2$*-integrable then*

$$M_1\text{-}\textit{AE}\, x.\ \textit{integrable } M_2\ (\lambda y.\ f\ (x,y)) \quad \textit{and}$$

$$\textstyle\int x.\ \left(\int y.\ f(x,y)\ dM_2\right)\ dM_1 = \int f\ d(M_1 \bigotimes_m M_2)\ .$$

Finally, we prove Fubini's theorem by this lemma and its symmetric variant.

Corollary 3 (Fubini's theorem). *If* f *is* $M_1 \bigotimes_m M_2$*-integrable then*

$$\textstyle\int x. \left(\int y.\ f\ (x,y)\ dM_2\right)\ dM_1 = \int y. \left(\int x.\ f\ (x,y)\ dM_1\right)\ dM_2\ .$$

Product measures. Product spaces are modeled as function space, i.e. the space of dependent products. In this section we assume $M\ i$ is a σ-finite measure space for all i. Product spaces can also be defined on arbitrary index sets I, however this holds only on probability spaces. We assume a finite index set I.

Definition 20.

$$\begin{aligned}
&\mathit{prod_algebra}_G \quad :: \iota\ \mathit{set} \Rightarrow (\iota \Rightarrow \alpha\ \mathit{algebra}) \Rightarrow (\iota \Rightarrow \alpha)\ \mathit{algebra}\\
&\mathit{prod_algebra}_G\ I\ M = (\!|\mathit{space} = (\Pi_E i \in I.\ \mathit{space}\ (M\ i)),\\
&\qquad\qquad \mathit{sets} = \left\{(\Pi_E i \in I.\ E\ i) \,\middle|\, E.\ \forall i \in I.\ E\ i \in \mathit{sets}\ (M\ i)\right\}|\!)\\
&\Pi_m\ i \in I.\ M\ i \quad = \mathit{sigma}\ (\mathit{prod_algebra}_G\ I\ M)(\!|\mathit{measure} := \mathit{SOME}\ \nu.\\
&\quad \mathit{sigma_finite_measure}\ (\mathit{sigma}\ (\mathit{prod_algebra}_G\ I\ M)(\!|\mathit{measure} := \nu|\!)) \wedge\\
&\quad \forall E.\ (\forall i \in I.\ E\ i \in \mathit{sets}\ (M\ i))\\
&\qquad \longrightarrow \nu\,(\Pi_E i \in I.\ E\ i) = \textstyle\prod_{i \in I} \mathit{measure}\ (M\ i)\ (E\ i)|\!)
\end{aligned}$$

We abbreviate $P_I = (\Pi_m\ i \in I.\ M\ i)$ and $\pi_I = \mathit{measure}\ P_I$. The definition of P_I takes $\mathit{sigma}\ (\mathit{prod\_algebra}_G\ I\ M)$ and extends it with *some* measure ν which forms a σ-finite measure space and which is uniquely defined on $\mathit{prod\_algebra}_G\ I\ M$, i.e., the generating set. These properties only holds for P_I when such a measure function exists, we prove the existence by induction over the finite set I.

Theorem 9. *If I is a finite set then sigma_finite_measure P_I and*
$\forall E.\ (\forall i.\ E\ i \in \mathit{sets}\ (M\ i)) \longrightarrow \pi_I\ (\Pi_E i \in I.\ E\ i) = \prod_{i \in I} \mathit{measure}\ (M\ i)\ (E\ i)$

We use $\mathit{merge}\ I\ J = (\lambda(x, y)\ i.\ \mathit{if}\ i \in I\ \mathit{then}\ x\ i\ \mathit{else}\ \mathit{if}\ i \in J\ \mathit{then}\ y\ i\ \mathit{else}\ \bot)$ as measure preserving function from $P_I \bigotimes_m P_J$ to $P_{I \cup J}$.

Lemma 15. *If I and J are two disjoint finite sets and $A \in \mathit{sets}\ P_{I \cup J}$ then*

$$\pi_{I \cup J}\ A = \mathit{measure}\ (P_I \textstyle\bigotimes_m P_J)\ ((\mathit{merge}\ I\ J)^{-1}[A] \cup \mathit{space}\ (P_I \bigotimes_m P_J))$$

A finite index set I' is either represented as $I' = I \cup J$, wih I and J finite, or $I' = \{i\}$. We give rules how to handle integrals in both cases, this allows us to iterate the Lebesgue integral on nonnegative functions in an inductive proof.

Lemma 16. *If I and J are disjoint finite sets and f is $P_{I \cup J}$-measurable then*

$$\int^+ f\ dP_{I \cup J} = \int^+ x.\ \left(\int^+ y.\ f\ (\mathit{merge}\ I\ J\ (x, y))\ dP_J\right)\ dP_I$$

Lemma 17. *If f is $M\ i$-measurable then*

$$\int^+ x.\ f\ (x\ i)\ dP_{\{i\}} = \int^+ f\ d(M\ i)$$

We extend these two lemmas to Lebesgue integrable functions. This helps us to prove the distributivity of multiplication and integration by induction on I.

Corollary 4. *If $I \neq \emptyset$ is finite and $f\ i$ is $M\ i$-integrable for all $i \in I$ then*

$$\int x.\, (\textstyle\prod_{i \in I} f\ i\ (x\ i))\, dP_I = \prod_{i \in I} \left(\int (f\ i)\ d(M\ i)\right)\ .$$

4.6 Lebesgue Measure

We have now formalized the concepts of measure spaces, Lebesgue integration and product spaces. An important measure space is the one on $\mathbb{R}$, where each interval $[a, b]$ has as measure value the length of the interval, $b - a$. The Borel σ-algebra is generated by these intervals. The corresponding measure is called the Lebesgue-Borel measure, its completion is the Lebesgue measure.

Instead of following the usual construction of the Lebesgue measure as the σ-extension of an interval measure we use the gauge integral[2] available in the multivariate analysis in Isabelle/HOL.[3] The gauge integral is an extension of the Riemann and also of the Lebesgue integral on Euclidean vector spaces. In Isabelle/HOL the predicate *integrable_on* A f states that the function f is gauge integrable on the set A, in which case the gauge integral of f on the set A has the real value *integral* A f. The gauge measure of a set A is the gauge integral of the constant 1 function on A.

Since the gauge measure is restricted to finitely measurable sets, it cannot be used directly as Lebesgue measure. However we can measure the indicator function χ_A on the intervals $[-n, n]$ for all natural numbers n. When χ_A is measurable on all intervals, we define it as Lebesgue measurable and the Lebesgue measure is the supremum of the gauge measures for all intervals $[-n, n]$. To define the Lebesgue measure on multidimensional Euclidean spaces we use hypercubes $\{x|\forall i.\, |x_i| \leqslant n\}$. The σ-algebra of the Lebesgue measure on a Euclidean space α consists of all $A{::}\alpha$ *set* which are gauge measurable on all intervals.

Definition 21 (Lebesgue measure).

$$\mathit{lebesgue}_\alpha = (\!| \mathit{space} = \mathcal{U}, \\ \mathit{sets} = \{A|\forall n.\ \mathit{integrable\_on}\ \{x|\forall\ i.\, |x_i| \leqslant n\}\ (\chi_A)\} \\ \mathit{measure} = \sup\ n.\ \mathit{integral}\ \{x|\forall\ i.\, |x_i| \leqslant n\}\ (\chi_A) |\!)$$

Theorem 10. *The Lebesgue measure forms a σ-finite measure space.*
sigma_finite_measure $\mathit{lebesgue}_\alpha$

From the definition of the Lebesgue measure it is easy to see that all Lebesgue measurable simple functions whose integral is finite are also gauge integrable. With the monotone convergence of the gauge integral we show that all nonnegative Lebesgue measurable functions with a finite integral are gauge integrable. And finally we show that all Lebesgue integrable functions are gauge integrable.

Theorem 11. *If f is Lebesgue integrable then integrable_on $\mathcal{U}$ f and* $\mathit{integral}\,\mathcal{U} f = \int f\ d(\mathit{lebesgue}_\alpha)$.

We know that lebesgue_α is a σ-algebra, and since all intervals $[a, b]$ are Lebesgue measurable all Borel sets are Lebesgue measurable.

Lemma 18. $A \in \mathit{sets\ borel} \longrightarrow A \in \mathit{sets\ lebesgue}_\alpha$

[2] The gauge integral is also called the Henstock-Kurzweil integral.
[3] The multivariate analysis in Isabelle/HOL is ported from a later version of [5].

We introduce the Lebesgue-Borel measure by changing the measurable sets from the Lebesgue sets to the Borel sets.

Definition 22 (Lebesgue-Borel measure). $\boldsymbol{\lambda}_\alpha = \mathit{lebesgue}_\alpha (\!| \mathit{sets} := \mathit{sets\ borel} |\!)$

Theorem 12. *sigma_finite_measure* $\boldsymbol{\lambda}_\alpha$

With theorem 3 we know that $\boldsymbol{\lambda}_\alpha$ is equal to other measures introduced on the Borel sets and based on the interval length. The Lebesgue-Borel measure is defined as a sub-σ-algebra of the Lebesgue measure, hence Lebesgue-Borel integrability induces gauge integrability.

Theorem 13. *If f is $\boldsymbol{\lambda}_\alpha$-integrable then integrable_on $\mathcal{U} f$ and integral $\mathcal{U} f = \int f d\boldsymbol{\lambda}_\alpha$.*

Euclidean vector spaces and product measures. We relate the Euclidean space α with the n-dimensional Lebesgue measure $\boldsymbol{\lambda}^n = (\Pi_m\ i \in \{1, \dots, n\}.\ \boldsymbol{\lambda}_\mathbb{R})$. The function $p2e :: (\mathbb{N} \Rightarrow \mathbb{R}) \Rightarrow \alpha$ maps functions to vectors with $(p2e\ f)_i = f\ i$. Theorem 3 helps us to show that it is measure preserving between $\boldsymbol{\lambda}^{\mathcal{D}(\alpha)}$ and α.[4]

Lemma 19. *Any $\boldsymbol{\lambda}_\alpha$-measurable set A satisfies*

$$\mathit{measure}\ \boldsymbol{\lambda}_\alpha\ A = \mathit{measure}\ \boldsymbol{\lambda}^{\mathcal{D}(\alpha)}\ (p2e^{-1}[A] \cap \mathit{space}\ \boldsymbol{\lambda}^{\mathcal{D}(\alpha)}).$$

From this follows the equivalence of integrals.

Theorem 14. *If f is $\boldsymbol{\lambda}_\alpha$-measurable then*

$$\begin{aligned} \int^+ f\ d\boldsymbol{\lambda}_\alpha &= \int^+ x.\ f\ (p2e\ x)\ d\boldsymbol{\lambda}^{\mathcal{D}(\alpha)} \\ \mathit{integrable}\ \boldsymbol{\lambda}_\alpha\ f &\longleftrightarrow \mathit{integrable}\ \boldsymbol{\lambda}^{\mathcal{D}(\alpha)}\ (f \circ p2e) \\ \int f\ d\boldsymbol{\lambda}_\alpha &= \int x.\ f\ (p2e\ x)\ d\boldsymbol{\lambda}^{\mathcal{D}(\alpha)} \end{aligned}$$

The Euclidean vector space formalizations in Isabelle/HOL include the dimensionality in the vector type. Here it is not possible to use induction over the dimensionality of the Euclidean space. With theorems 13 and 14 we equate the gauge integral to the Lebesgue integral over $\boldsymbol{\lambda}^n$, we then use induction over n.

5 Discussion

Most measure theory textbooks assume that product spaces are built by iterating binary products and that the Euclidean space is equivalent to the product of the Lebesgue measure. In our case these are three different types, which we need to relate. Using theorem 3 we show the equivalence of measure spaces of theses types. This not only helps us to transfer between different types but also to avoid repeated proofs. For example Fubini's theorem needs the symmetric variant of

[4] $\mathcal{D}(\alpha)$ is the dimension of the euclidean space α.

some theorems. Instead of repeating these proofs as the text books suggest we show that the measure is equal under the pair swap function.

We also diverge from text books by directly constructing the binary product measure and the Lebesgue measure. Usually text books show the existence of a measure and then choose one meeting the specification. This is difficult in theorem provers as the definition is not usable until the existence is proven. Otherwise, we prefer to stay close to the standard formalizations of measure theory concepts. Sometimes this requires more work if we only want to prove one specific lemma, but it is easier to find textbook proofs usable for formalization.

Locales as mechanism for theory interpretation are convenient when proving the Radon-Nikodým theorem and product measures. We instantiate measure spaces restricted to sets obtained in the proof. By interpretation inside the proof we have full access to the automation and lemmas provided by the locale.

Type classes simplify the introduction of $\overline{\mathbb{R}}$ as it allows us to reuse syntax and some theorems about lattices, arithmetic operations, topological spaces, limits, and infinite series. We use the topological space type class to define the Borel σ-algebra. This allows us to state theorems about Borel sets for $\mathbb{R}$ and $\overline{\mathbb{R}}$.

6 Related Work

Our work started as an Isabelle/HOL port of the HOL4 formalization done by Coble [2]. We later reworked most of it to use the extended reals as measure values and open sets as generator for the Borel σ-algebra. We also changed the definition of the Lebesgue integral to the one found in Schilling's textbook [12]. We define the integral of f as the supremum of all simple functions bounded by f. Coble used the limit of the simple functions converging to f.

	Hurd	Richter	Coble	Mhamdi	Lester	PVS	Mizar	HOL-Light	Isabelle
$\overline{\mathbb{R}}$					✓	✓	✓		✓
Borel (open)				✓	✓	✓			✓
Integral		✓	✓	✓	✓	✓	✓	✓	✓
$\boldsymbol{\lambda}$	$[0,1]$					✓	✓	✓	✓
Products						✓		$\mathbb{R}^{\mathcal{D}(\alpha+\beta)}$	✓
Dynkin							✓		✓

Table 1. Overview of the current formalizations of measure theory.

Table 1 gives an overview of the current formalizations of measure theory we are aware of. The columns list first the work of Hurd [6], Richter [11], Coble [2], Mhamdi et al. [10], and Lester [8]. The second part of the columns list theorem provers or libraries formalizing measure theory. Beginning with the PVS-NASA library,[5] the Mizar Mathematical Library (MML), the multivari-

[5] `http://shemesh.larc.nasa.gov/fm/ftp/larc/PVS-library/pvslib.html`

ate analysis found in HOL Light and finally the work presented in this paper. Mhamdi et al. represents the current state of HOL4, hence HOL4 is not listed. The rows correspond to different measure theoretic concepts and features.

Hurd [6] formalizes a measure space on infinite boolean streams isomorphic to the Lebesgue measure on the unit interval $[0, 1]$. His positive extended reals [7] are unrelated to this. Richter [11] formalizes the Lebesgue integral in Isabelle/HOL and uses it together with Hurd's bitstream measure. Richter introduces the Borel σ-algebra, but only on right-bounded intervals in $\mathbb{R}$.

Coble [2] uses product spaces and the Radon-Nikodým derivative on finite sets to define mutual information for his formalization of information theory. He ports Richter's formalization of the Lebesgue integral to HOL4 and generalizes the definition of σ-algebras to be defined on arbitrary spaces $\Omega \neq \mathcal{U}$. While his formalizations of the Lebesgue integral is on arbitrary measure spaces, he states the Radon-Nikodým theorem and the product measure for finite sets only.

The work by Mhamdi et. al [10] extends Coble's [2]. Their definitions of the Lebesgue integral and Borel σ-algebra are comparable to the ones in this paper. However, they do not formalize measure values as extended real numbers, but only as plain reals. They define a more restricted version of the almost everywhere predicate, and do not give rules for the interaction with logical connectives. They prove Markov's inequality and the weak law of large numbers.

There is also the PVS formalization of topology by Lester [8]. He gives a short overview of the formalized measure theory, which includes measures using extended real numbers, a definition of almost everywhere, Borel σ-algebras on topological spaces, and the Lebesgue integral. In recent developments the PVS-NASA library contains binary product spaces and the proof that the Lebesgue integral extends the Riemann integral. In PVS, abstract reasoning is performed using parametrized theories, similar to our usage of locales.

Endou et. al. [3] proves monotone convergence of the Lebesgue integral in the MML. It also contains measure spaces with extended real numbers, and the Lebesgue measure. Merkl [9] formalized Dynkin systems and Dynkin's lemma in MML, however without a concrete application.

In HOL Light an extended version of Harrison's work [5] introduces gauge integration on finitely-dimensional Euclidean spaces which is similar to the product space of Lebesgue measures. This is then used to define a subset of the Lebesgue measure, missing infinite measure values. The definition of Euclidean spaces $\mathbb{R}^{\mathcal{D}(\alpha)}$ and $\mathbb{R}^{\mathcal{D}(\alpha)}$ allows to create the product $\mathbb{R}^{\mathcal{D}(\alpha+\beta)}$. His theories are now available in Isabelle/HOL and we use them to introduce the Lebesgue measure and to show that Lebesgue integrability implies gauge integrability and that in this case both integrals are equal.

7 Conclusion

The formalizations presented in this paper form the foundations of measure theory. Looking at the table of contents of Bauer's textbook [1] we formalized almost all of the first three chapters. What is missing are the function spaces

$\mathcal{L}^p$, stochastic convergence, and the convolution of finite Borel measures. Isabelle supported us with its type classes allowing to reuse definitions and theorems for limits and arithmetic operations on $\overline{\mathbb{R}}$. We used Isabelle's locales to introduce the concepts of the different set systems and spaces used in measure theory.

With product spaces and the Radon-Nikodým derivative it is possible to combine the concepts introduced by Hurd [6] and Coble [2]. We can now verify information theoretic properties of probabilistic programs.

This paper is the first to derive the Radon-Nikodým theorem and the multi-dimensional version of Fubini's theorem. Our next step concerns the development of probability theory. We already formalized conditional expectation, Kullback-Leibler divergence, mutual information, and infinite products measure using the measure theory presented in this paper. The details are available at the URL given on page 2. The future goals concern the formalization of infinite sequences of independent random variables and the central limit theorem as well as Markov chains and Markov decision processes.

Acknowledgments We would like to thank Lawrence C. Paulson for porting Coble's measure space formalization from HOL4 to Isabelle/HOL and to Robert Himmelmann for formalizing the first iteration of the Lebesgue measure.

References

1. Bauer, H.: Measure and Integration theory. de Gruyter (2001)
2. Coble, A.R.: Anonymity, Information, and Machine-Assisted Proof. Ph.D. thesis, King's College, University of Cambridge (2009)
3. Endou, N., Narita, K., Shidama, Y.: The Lebesgue monotone convergence theorem. Formal Mathematics 16(2), 167–175 (2008)
4. Haftmann, F., Wenzel, M.: Local theory specifications in Isabelle/Isar. In: TYPES 2008, LNCS, vol. 5497, pp. 153–168 (2009)
5. Harrison, J.: A HOL theory of Euclidean space. In: Proceedings of TPHOLs 2005. LNCS, vol. 3603, pp. 114–129 (2005)
6. Hurd, J.: Formal Verification of Probabilistic Algorithms. Ph.D. thesis, University of Cambridge (2002)
7. Hurd, J., McIver, A., Morgan, C.: Probabilistic guarded commands mechanized in HOL. Theoretical Computer Science 346(1), 96–112 (Nov 2005)
8. Lester, D.R.: Topology in PVS: continuous mathematics with applications. In: Proceedings of AFM'07. pp. 11–20 (2007)
9. Merkl, F.: Dynkin's lemma in measure theory. FM 9(3), 591–595 (2001)
10. Mhamdi, T., Hasan, O., Tahar, S.: On the formalization of the Lebesgue integration theory in HOL. In: Proceedings of ITP 2010. LNCS, vol. 6172, pp. 387–402 (2010)
11. Richer, S.: Formalizing integration theory with an application to probabilistic algorithms. In: Proceedings of TPHOLs 2004. LNCS, vol. 3223, pp. 271–286 (2004)
12. Schilling, R.L.: Measures, Integrals and Martingales. Cambridge Univ. Press (2005)