Formalization of the Lebesgue Measure in MathComp-Analysis

Reynald Affeldt and Cyril Cohen

National Institute of Advanced Industrial Science and Technology (AIST), Tokyo, Japan

Université Côte d'Azur, Inria, Sophia Antipolis, France

July 2, 2021

Motivation

Why formalize the Lebesgue measure?

1. Develop integration and probability theories on top of MathComp
2. More generally: development of reusable machinery for MathComp-Analysis [1, 2]
   - extension of MathComp for classical analysis (topology, real and complex analysis, etc.)

Problem statement

Construction of the Lebesgue measure which is:

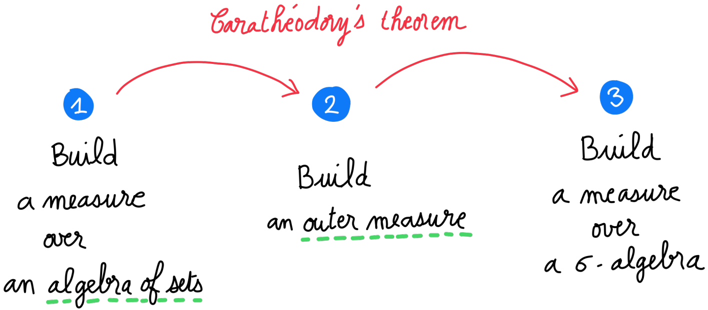
- a function μ
  - domain: sets that form a σ-*algebra*
  - codomain: extended real numbers (can be $+\infty$)
- which is a *measure*
  - in particular: it is σ-*additive*, i.e. $\mu(\bigcup_i F_i) = \sum_i \mu(F_i)$ when F_i are pairwise-disjoint

This is a long construction (our main reference [7] is 14 pages long and still glosses over many details)

This is non-trivial (the result can be admitted at the undergraduate-level in French universities)

Construction Approach

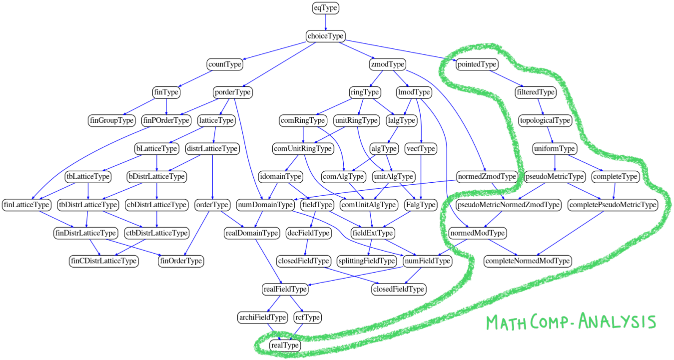
Standard, textbook approach, from the ground up:



Outline

Starting Point: MathComp and MathComp-Analysis

MathComp-Analysis adds several mathematical structures to MathComp:



Support for Extended Real Numbers

- $\infty - \infty$ is undefined in the mathematical practice
  - We define it as $-\infty$ so that the extended real numbers form a commutative monoid
  - **Important benefit:** we can use the bigop library of MathComp!
- Extended real numbers form a topological/uniform/pseudometric space
  - Tedious instantiation work using the following definition of ball:

```
Definition ereal_ball (x y : \bar R) (e : R) :=
  `|𝒞 x - 𝒞 y| < e.
```

  - **Reward:** we can reuse existing lemmas to develop the theory of sequences, e.g.:

```
Lemma ereal_limD (R : realType) (f g : (\bar R)^nat) :
  cvg f → cvg g → adde_def (lim f) (lim g) →
  lim (f \+ g) = lim f + lim g.
```

New Mathematical Structures

- *σ-algebra* (note the countable union *(**)*):

```
HB.factory Record isMeasurable T := {
 measurable : set (set T) ;
 measurable0 : measurable set0 ;
 measurableC : ∀ A, measurable A → measurable (~` A) ;
 measurable_bigcup : ∀ A : (set T)^nat,
  (∀ i, measurable (A i)) → measurable (\bigcup_i (A i)) (**) }.
```

- Algebra of sets (finite union only *(***)*):

```
HB.factory Record isAlgebraOfSets T := {
 measurable : set (set T) ;
 measurable0 : measurable set0 ;
 measurableC : ∀ A, measurable A → measurable (~` A) ;
 measurableU : ∀ A B, measurable A → measurable B →
  measurable (A `|` B) (***) }.
```

- These *factories* are implemented with Hierarchy-Builder [4]
  - *σ-algebras* actually extend algebras of sets, which extend *ring of sets*, which extend *semirings of sets*

Measures

- Measure:

```
Record axioms (mu : set T → \bar R) := Axioms {
  _ : mu set0 = 0 ;
  _ : ∀ x, measurable x → 0 ≤ mu x ;
  _ : semi_sigma_additive mu (**) }.
```

(**) $\stackrel{\text{def}}{=} \mu(\cup_n F_n) = \sum_i^\infty \mu(F_i)$ for any sequence F, s.t.

  - $\forall i, \texttt{measurable}(F_i)$
  - F_i pairwise-disjoint
  - $\texttt{measurable}(\cup_n F_n)$

- Outer measure:

```
Record axioms (mu : set T → \bar R) := Axioms {
  _ : mu set0 = 0 ;
  _ : ∀ x, 0 ≤ mu x ;
  _ : {homo mu : A B / A ⊆ B ↣ A ≤ B} ;
  _ : sigma_subadditive mu (***)}.
```

(***) $\stackrel{\text{def}}{=} \mu(\cup_n F_n) \leq \sum_i^\infty \mu(F_i)$ (no equality required when sets are pairwise-disjoint)

Outline

Carathéodory's Theorem (1/2)

- Goal: build a σ-*algebra* and a measure over it given an outer measure
- The resulting σ-*algebra* is composed of *Carathéodory measurable* sets, i.e., sets A s.t. $\forall X, \mu(X) = \mu(X \cap A) + \mu(X \cap \bar{A})$
- The resulting measure is over these Carathéodory measurable sets
- Coq proofs are a direct formalization of paper proofs
  - This is thanks to the newly developed support for sequences of extended real numbers!

Carathéodory's Theorem (2/2)

- Goal: define an outer measure μ^* given a measure μ over an *algebra of sets*

$$\mu^*(X) \stackrel{\text{def}}{=} \inf \left\{ \sum_i^{\infty} \mu(F_i) | (\forall i, \texttt{measurable}(F_i)) \wedge X \subseteq \bigcup_i F_i \right\}$$

- Again, Coq proofs are a direct formalization of paper proofs modulo a new development:
  - In the course of proving σ*-subadditivity*, we run into the following subgoal:

$$\mu^*(\cup_i F_i) \leq \sum_i^{\infty} \left(\mu^*(F_i) + \frac{\varepsilon}{2^i} \right)$$

  - The proof goes on by showing $\mu^*(\cup_i F_i) \leq \sum_{i,j} \mu(G_i j) \leq \sum_i \sum_j \mu(G_i\, j)$ for some well-chosen G, which requires *sums over general sets*

Sums (of non-negative terms) over General Sets

- Paper definition:

$$\sum_{i \in A} a_i \stackrel{\text{def}}{=} \sup \left\{ \sum_{i \in F} a_i \mid F \text{ non-empty finite subset of } A \right\}$$

- Coq definition:

```
Definition csum R (T : choiceType) (A : set T) (a : T → \bar R) :=
  if A == set0 then 0 else
  ereal_sup [set \sum_(i <- F) a i |
            F in [set F : {fset T} | [set i | i ∈ F] ⊆ A]].
```

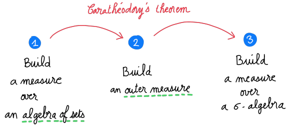
- Sample lemma:

$$(\forall n, 0 \leq a_n) \to (\forall k, J_k \neq \emptyset) \to J_k \text{ pairwise-disjoint} \to$$

$$\sum_{i \in \bigcup_{k \in K} J_k} a_i = \sum_{k \in K} \left(\sum_{j \in J_k} a_j \right)$$

Towards the Lebesgue Measure

- More formalized properties to complete Carathéodory's theorem:
  - The built measure coincides with the original measure
  - The built σ-*algebra* contains the smallest σ-*algebra* that contains the starting algebra of sets
  - The extension is unique provided the original measure is σ-finite

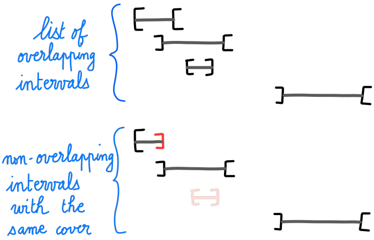


- Reminder:
  - Missing piece: a σ-finite measure over the algebra of sets generated by intervals...

Outline

The Algebra of Sets of Simple Sets

- Let us consider *simple sets*: sets that can be covered by a finite union of intervals
- Simple sets form an algebra of sets
    - Difficulty: stability by complement,
      better proved using non-overlapping intervals:



Decomposition of Intervals

Turn intervals into non-overlapping intervals with the same cover:

1. Ordering of intervals:

```
Definition lt_itv i j :=
  (i.1 < j.1)%O || ((i.1 == j.1) ∧ (i.2 < j.2)%O).
```

2. Recursive procedure to chop overlaps:

```
Lemma decompose_two i j t : decompose [:: i, j & t] =
  if (i ≤ j)%O then decompose (j :: t) else
  if (j ≤ i)%O then decompose (i :: t) else
  itv_diff i j :: decompose (j :: t).
```

3. Complete procedure:

```
Definition Decompose s :=
  decompose (sort le_itv [seq x ← s | neitv x]).
```

Length of Simple Sets

- Length of an interval:

```
Definition hlength (A : set R) : \bar R :=
 let i := Rhull A in i.2 - i.1.
```

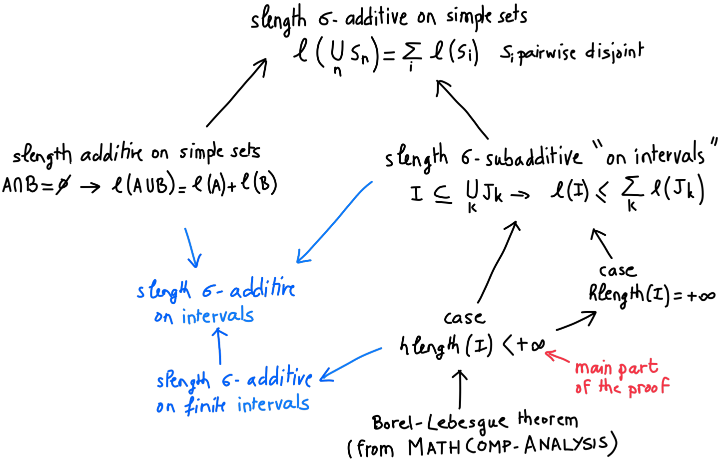
- Length of a simple set:

```
Definition slength (X : set R) : \bar R :=
 let s := xget [::] [set s | X = [sset of s] ] in
\sum_(i <- Decompose s) hlength (set_of_itv i).
```

- What is left to do?
  - prove that slength is σ-finite and that it is a measure
  - The difficulty is σ-*additivity*...

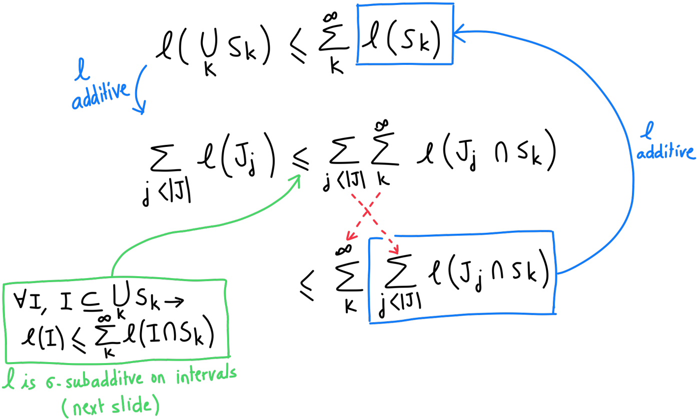
The σ-*additivity* of `slength`

Proof Overview



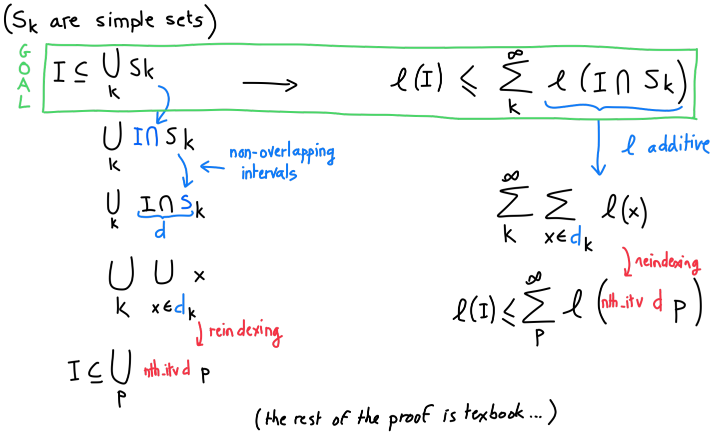
slength is σ-*additive*

Sample Technicality



slength is σ-*subadditive*

Sample Technicality



Outline

Related Work

About the Lebesgue measure:

- Lebesgue measure in Isabelle/HOL [6] defined using the gauge integral
- Lean defines the Lebesgue measure from the Lebesgue outer measure, not as an extension from an algebra of sets
- Lebesgue measure in Mizar as early as 1996 but needed a reconstruction in 2020 [5]

About Lebesgue integration in Coq:

- Recent work by Boldo et al. [3]: not directly compatible with our work (addition of extended real numbers is not associative because $\infty - \infty = 0$), Lebesgue measure is future work, σ-*algebra*'s defined as generated σ-*algebra*'s

Conclusions

We have a construction of the Lebesgue measure in MathComp-Analysis

- from the ground up (for documentation see, say, [7, 8, 9])
- principled construction
  - New mathematical structures (using Hierarchy-Builder)
  - Support for extended real numbers takes advantage of `bigop.v` and existing lemmas for topological structures
  - Sums of non-negative terms over general sets (using `finmap.v`)
  - Concrete instance of algebra of sets (using `interval.v`)

Current/future work: integration and probability theories

R. Affeldt, C. Cohen, M. Kerjean, A. Mahboubi, D. Rouhling, and K. Sakaguchi.
Competing inheritance paths in dependent type theory: a case study in functional analysis.
In *IJCAR 2020*, volume 12167(2) of *LNAI*, pages 3–20. Springer, Jul 2020.

R. Affeldt, C. Cohen, and D. Rouhling.
Formalization techniques for asymptotic reasoning in classical analysis.
Journal of Formalized Reasoning, 11(1):43–76, 2018.

S. Boldo, F. Clément, F. Faissole, V. Martin, and M. Mayero.
A Coq formalization of Lebesgue Integration of nonnegative functions.
Technical Report RR-9401, Inria, 2021.

C. Cohen, K. Sakaguchi, and E. Tassi.
Hierarchy Builder: Algebraic hierarchies made easy in Coq with Elpi (system description).
In *FSCD 2020*, volume 167 of *LIPIcs*, pages 34:1–34:21, 2020.

N. Endou.
Reconstruction of the one-dimensional Lebesgue measure.
Technical report, National Institute of Technology, Gifu College, 2020.
Formalized Mathematics 28(1):93–104.

J. Hölzl and A. Heller.
Three chapters of measure theory in Isabelle/HOL.
In *ITP 2011*, volume 6898 of *LNCS*, pages 135–151. Springer, 2011.

D. Li.
Construction de la mesure de Lebesgue.
Available online, Université d'Artois, Faculté des Sciences Jean Perrin, 2008.

M. Papadimitrakis.
Notes on measure theory.
Available online, Department of Mathematics, University of Crete, 2004.

A. Zeriahi.
Mesure et intégration (notes de cours de l3).
Available online, Oct 2011.