Integrals Within Integrals: A Formalization of the Gagliardo-Nirenberg-Sobolev Inequality

Floris van Doorn
Mathematical Institute, University of Bonn, Germany

Heather Macbeth
Department of Mathematics, Fordham University, New York, NY, USA

Abstract

We introduce an abstraction which allows arguments involving iterated integrals to be formalized conveniently in type-theory-based proof assistants. We call this abstraction the *marginal construction*, since it is connected to the marginal distribution in probability theory. The marginal construction gracefully handles permutations to the order of integration (Tonelli's theorem in several variables), as well as arguments involving an induction over dimension.

We implement the marginal construction and several applications in the language Lean. The most difficult of these applications, the Gagliardo-Nirenberg-Sobolev inequality, is a foundational result in the theory of elliptic partial differential equations and has not previously been formalized.

2012 ACM Subject Classification Mathematics of computing → Continuous mathematics; Theory of computation → Logic and verification

Keywords and phrases Sobolev inequality, measure theory, Lean, formalized mathematics

Digital Object Identifier 10.4230/LIPIcs.ITP.2024.37

Supplementary Material
Software (Formalization code): `https://github.com/leanprover-community/mathlib4`
archived at `swh:1:dir:650d45022cc1853d3a91744387e43ff17631c638`

Acknowledgements We thank Patrick Massot for useful comments on a draft of this article.

1 Introduction

There are two major challenges which appear in the formalization of iterated integration. First, according to Tonelli's theorem, the order of integration does not matter on well-behaved integrands. A formalism for iterated integration should make this convenient to state and apply. Secondly, iterated integration often turns up in the wild in the context of analytic arguments involving induction on dimension. Experience suggests such arguments are intrinsically hard to formalize. A good formalism for iterated integration should provide auxiliary constructions which enable users to mimic such induction arguments.

In this article we introduce a framework for iterated integration in the `Mathlib` library of the interactive proof assistant Lean. We test this framework in several applications, most notably in a proof of the Gagliardo-Nirenberg-Sobolev inequality, a foundational result from the theory of elliptic partial differential equations. The proof of the inequality is a tricky argument whose details are often elided in the literature. It involves both the reordering of iterated integrals and (something akin to) induction on dimension.

The structure of the article is as follows. As a foundation for this project we construct the finitary product measure in Lean (Section 3). This is the most general context for which iterated integration can be considered. This setting includes $\mathbb{R}^n$, whose standard measure is built as the product of n copies of the Lebesgue measure on $\mathbb{R}$. This work builds on earlier work of van Doorn [19] defining the binary product measure and a pre-existing measure theory library developed over the previous several years [3].

15th International Conference on Interactive Theorem Proving (ITP 2024).
Editors: Yves Bertot, Temur Kutsia, and Michael Norrish; Article No. 37; pp. 37:1–37:18
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

We then develop (Section 4) a notion of iterated integration suitable for type theory, and implement this in Lean. We refer to our notion of iterated integration as the *marginal construction* because it is inspired by the marginal distribution in probability theory. Our construction permits expression of iterated integrals in a form which is closer to the on-paper notion, bringing the formalized math into better correspondence with the written form. Our framework, as mentioned above, allows for arguments involving induction on dimension to be expressed intuitively. Furthermore, our framework handles Tonelli's theorem silently, reducing it to a statement on equalities of sets.

We give several demonstrations (Section 5) of our iterated integration framework. First, as an example we compute the volume of a ball in $\mathbb{R}^n$ (Subsection 5.1), as well as a matrix change of coordinates argument previously formalized in Lean by Gouëzel [9] as a key step to prove the change of variables formula for integration. In these examples, we will show how our framework can handle these arguments elegantly.

Next we provide a more elaborate example (Subsection 5.3), which we name the "grid-lines lemma." This is an argument requiring both induction over dimension and Tonelli's theorem, and it would be extremely cumbersome to express without an explicit notion of iterated integration.

The grid-lines lemma is an abstraction of the key argument in the Gagliardo-Nirenberg-Sobolev inequality [17, 8, 7]. The final component of our project (Section 6) is the deduction of this inequality from our grid-lines lemma. The Gagliardo-Nirenberg-Sobolev inequality has not been previously formalized.

In Section 7 we give a sketch of the importance of the Gagliardo-Nirenberg-Sobolev inequality to the theory of elliptic partial differential equations and suggest future formalization work in this direction. Related work is discussed in Section 8.

2 Preliminaries

2.1 Lean and `Mathlib`

Lean [4] is a theorem proving language; its logical foundation is dependent type theory. `Mathlib` [3] is its standard mathematical library, currently totalling 1.4 million lines of code. The design of `Mathlib` prioritizes convenience and mathematical generality; a tradeoff is that no effort is made to work constructively. The development of `Mathlib` is a distributed project, with some 300 contributors over the seven years of its existence.

Recently a new version, Lean 4 [15], was introduced and the `Mathlib` library was ported from Lean 3 to Lean 4. We refer to the two versions of the library as `Mathlib3` and `Mathlib4` when there is a possibility of confusion.

The finitary product measure construction described in this article (Section 3) was written in 2021 in Lean 3 and contributed to the `Mathlib3`, and was ported to Lean 4 as part of the broader `Mathlib` porting effort.

The rest of the work described in the article was written in Lean 4. The marginal construction (Section 4) and its prerequisites were contributed to `Mathlib4` in 2023, and the Sobolev inequality was contributed in 2024. We will use clickable links to link our paper to specific results in the library (to the version of `Mathlib` of June 29, 2024). The application described in Section 5.1 is not part of Mathlib, but on a branch, since Xavier Roblot had already contributed the computation of the volume of a ball to `Mathlib4`, using different techniques.

We estimate that the whole project comprises some 3,500 lines of code.

2.2 Basic Measure Theory

In this section we briefly describe the most important parts of the measure theory library in `Mathlib` prior to our work. We refer to [19] for a fuller description.

In `Mathlib` we have the notion of *measurable space*, which is just a type equipped with a chosen σ-algebra of sets which we call the measurable sets. On a measurable space we can consider a measure μ which sends any measurable set A to a number in $[0, \infty]$ which is a monotone and countably additive. Here $[0, \infty]$ is the type of nonnegative real numbers extended by a single element ∞, and denoted `ℝ≥0∞` in Lean. In Lean we allow μ to be applied to any set A (even nonmeasurable ones), in which case it is defined as the infimum of the measures of measurable sets containing A. This makes μ an outer measure on all sets (i.e. monotone and countably subadditive function that sends $\varnothing$ to 0).

`Mathlib` contains two notions of integration. The *Lebesgue integral* is for functions $X \to [0, \infty]$ where X is a measurable space equipped with a measure μ. The *Bochner integral* is for integrable functions $X \to E$ where E is a Banach space. We will denote both of these integrals using any of the following notations:

$$\int_X f \,\mathrm{d}\mu = \int_X f(x) \,\mathrm{d}\mu(x) = \int_X f = \int_X f(x) \,\mathrm{d}x.$$

In this paper we will be almost exclusively working with the Lebesgue integral.

Given two measurable spaces X and Y, and two σ-finite measures μ and ν, we can construct a measure $\mu \times \nu$ on the measurable space $X \times Y$, which satisfies

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$$

for all sets A and B (we do not need to require that A and B are measurable).

Tonelli's theorem is an important theorems that states how to compute Lebesgue integrals with respect to the product measure.

▶ **Theorem 1** (Tonelli's theorem). *Let $f : X \times Y \to [0, \infty]$ be a measurable function. Then*

$$\int_{X \times Y} f \,\mathrm{d}(\mu \times \nu) = \int_X \int_Y f(x, y) \,\mathrm{d}\nu(y)\mathrm{d}\mu(x) = \int_Y \int_X f(x, y) \,\mathrm{d}\mu(x)\mathrm{d}\nu(y)$$

and all the functions in the integrals above are measurable.

There is also an analogous theorem for the Bochner integral in Mathlib, called Fubini's theorem, but we will not be using that in this paper.

2.3 Notation

In this paper, we will use set-theoretic notation for type-theoretic concepts, writing $i \in \iota$ or $A \subseteq \iota$ even when ι is a type.

Let ι be a type. If $\mathbf{x}$ is a vector of type $\prod_{i \in \iota} X_i$, and t is of type X_i, then $X_i(\mathbf{x}, t)$ denotes the vector whose i-th coordinate is t and whose j-th coordinate is $\mathbf{x}_j$ for $j \neq i$. In Lean this vector is denoted `Function.update x i t`.

Similarly, if $A \subseteq \iota$ and $y : \prod_{i \in \iota} X_i$, we use the same notation $X_A(\mathbf{x}, y)$ for the operation that updates $\mathbf{x}$ on A:

$$X_A(x, y)_i := \begin{cases} y_i & \text{if } i \in A \\ x_i & \text{otherwise.} \end{cases}$$

3 Finite Product Measures

There are a few ways to define the product measure on finite product spaces. Conceptually this can be done by iterating the binary product measure construction. However, some care is required, since the spaces $X \times (Y \times Z)$ and $(X \times Y) \times Z$ are not the same space, they are merely equivalent spaces.

Given a finite family of measurable spaces $(X_i)_{i\in\iota}$ with a σ-finite measure μ_i on X_i, we want to define the product measure on $\prod_{i\in\iota} X_i$. We could define the measure by choosing an arbitrary enumeration of ι as $\iota = \{i_1, \dots, i_k\}$, and then by using the measurable equivalence

$$\left(\prod_{i\in\iota} X_i\right) \simeq X_{i_1} \times X_{i_2} \times \cdots \times X_{i_k}, \tag{1}$$

to transport the iterated binary product measure from the right-hand side to the left-hand side. We decide not to do this in order to avoid arbitrary choices in the definition. Instead, we don't care too much how we define the measure, as long as it satisfies the property that if $A_i \subseteq X_i$ for all $i \in \iota$, then

$$(\Pi_i\mu_i)(\Pi_i A_i) = \prod_i \mu_i(A_i). \tag{2}$$

We will define the measure as the maximal measure satisfying (2). To do this, we first define the projection $\pi_i(A)$ of a subset $A \subseteq \prod_{i\in\iota} X_i$ as the image of A under the evaluation function $\pi_i : \left(\prod_{i\in\iota} X_i\right) \to X_i$. Then we define an auxiliary function n which sends a set $A \subseteq \prod_{i\in\iota} X_i$ to

$$n(A) := \prod_{i\in\iota} \mu_i(\pi_i(A)) \in [0, \infty].$$

Note that n will be equal to the measure of the smallest box containing A. Now there is a unique maximal outer measure m such that $m(A) \le n(A)$ for all sets A.

Next, we want to turn this outer measure m into a measure on the product space. We say that a subset $A \subseteq \prod_{i\in\iota} X_i$ is *Carathéodory-measurable w.r.t. m* if for all $B \subseteq \prod_{i\in\iota} X_i$ the following equality holds:

$$m(B) = m(B \cap A) + m(B \setminus A).$$

We then show that all measurable subsets of $\prod_{i\in\iota} X_i$ are actually Carathéodory-measurable w.r.t. m, and this shows that m allows us to get a measure $\Pi_i\mu_i$ on $\prod_{i\in\iota} X_i$, such that for each measurable set A we have $(\Pi_i\mu_i)(A) = m(A)$.

Finally, we need to show that this measure satisfies (2). We first do this in the case that each A_i is measurable. In this case, it is easy to show that

$$(\Pi_i\mu_i)(\Pi_i A_i) = m(\Pi_i A_i) \le n(\Pi_i A_i) = \prod_i \mu_i(A_i).$$

We can show the reverse inequality by giving a specific instance of a measure that is bounded by n and satisfies (2). To do this, we use the idea at the start of this section, by enumerating $\iota = \{i_1, \dots, i_k\}$ and using equivalence (1) to construct some specific measure ν. It is not too hard to show that $\nu \le n$ and that $\nu(\Pi_i A_i) = \prod_i \mu_i(A_i)$. Since $\Pi_i\mu_i$ is the maximal measure bounded by n, we have $\nu \le \Pi_i\mu_i$, hence reverse inequality follows.

There are some interesting observations in implementing ν, since it is naively defined by recursion on the cardinality of ι. However, it is not so easy to perform recursion on finite types, especially if you define something that depends on the ordering on the type. Furthermore, it is convenient if we don't transport along too many equivalences, during our construction.

Our solution was to define a new way to define product types as an auxiliary construction.

```
def TProd {ι : Type*} (α : ι → Type*) (l : List ι) : Type* :=
  l.foldr (fun i β ↦ α i × β) PUnit
```

So e.g. `TProd α [i, j, k]` is by definition `α i × α j × α k × PUnit`, where `PUnit` is the (universe polymorphic) unit type. This definition is convenient, since `TProd α (i::l)` is by definition the same as `α i × TProd α l`. This makes it very easy to define the product measure on `TProd α l` by induction on `l`, given measures on each `α i`. We prove that if `l` contains each element of ι exactly once, then `TProd α l` is equivalent to the usual product $\Pi_i \alpha_i$. Finally, we construct ν by transporting the measure on `TProd α l` along this equivalence. Finally, this definition makes it very easy to show that $\nu(\Pi_i A_i) = \prod_i \mu_i(A_i)$.

In `Mathlib` we generally try to avoid these auxiliary constructions, because it's yet another way to talk about the same mathematical object. The thing to be worried about is that we would want all the properties for product of types stated both for the usual Π-type and for `TProd`, leading to a large duplication of lemmas. However, in this case we explicitly mark `TProd` as an implementation detail, and avoid its usage unless you specifically want its precise definitional behavior. For our construction, this definitional behavior made the construction particularly easy, since we didn't have to transport anything along an equivalence in the recursion argument, only once at the end.

This definition of finitary product measures was completed in 2021 and has since been used in various analysis formalizations in Lean. It is used to define the Lebesgue measure on $\mathbb{R}^n$ (or, more precisely, $\iota \to \mathbb{R}$). There is another definition of this measure using the Haar measure, but we show that these give rise to the same measure. And this definition makes it a lot easier to show some simple properties about this measure, such as (2).

It is also used by Kudryashov [13] as the setting for the Bochner-integrability version of his divergence theorem, and by Gouëzel [9] in order to use $\mathbb{R}^n$ as a setting for certain measure-theoretic results which are subsequently transferred to a general finite-dimensional normed space by a choice of basis (see Subsection 5.2).

4 The Marginal Construction

4.1 Approaches to Formalization

We note that the fundamental issues in formalizing iterated integration are the same for systems based on simple type theory and dependent type theory. In dependent type theory it is possible directly to express a dependent finitary product $\prod_{j:\iota} A_j$ of measure spaces, but the issues we address appear already in the setting of a function type $\iota \to A$, which can be expressed in simple type theory.

When working with products of finitely many spaces, one also wants to use the Tonelli and Fubini theorems. For example, if $f : \mathbb{R}^{n+m} \to \mathbb{R}$, then one might want to write

$$\int_{\mathbb{R}^{n+m}} f(z)\,\mathrm{d}z = \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} f(x, y)\,\mathrm{d}y\mathrm{d}x.$$

People familiar with formalization will note that this is not simply an application of Fubini's theorem, since $\mathbb{R}^{n+m}$ is not the same entity as $\mathbb{R}^n \times \mathbb{R}^m$. One option is to show that there is a "canonical" measure-preserving equivalence $\mathbb{R}^{n+m} \simeq \mathbb{R}^n \times \mathbb{R}^m$. However, this is still a bit inconvenient to work with, since one has to work explicitly with this equivalence. Below we will develop a framework that does not require working with any of these measure-preserving equivalences (of course, to prove that the framework is correct, we will use such measure-preserving equivalences).

Another expression that one might want to deal with is for $f : \mathbb{R}^n \to \mathbb{R}$ an iterated integral of the form

$$\int \cdots \int f(x_1, \ldots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_k$$

where $k \leq n$. Here only some of the arguments of f are integration variables, and the remaining expression is still a function of the remaining variables. Manipulating iterated integrals like this is a key part of the proof of the grid-lines lemma discussed in Section 5.3.

The solution we implement was suggested in the concluding section of [19], which we will do in the next section.

4.2 Definition and Properties

We encapsulate this notion in the following definition. In this definition we will generalize $\mathbb{R}^n$ to an arbitrary product space $\prod_{i \in \iota} X_i$.

▶ **Definition 2.** *Let ι be a indexing set (not necessarily finite), $A \subseteq \iota$ a finite subset and E be a Banach space. For $i \in \iota$ suppose we are given a measure space (X_i, μ_i) and let $f : (\prod_{i \in \iota} X_i) \to [0, \infty]$ be a function. Then the* marginal of f w.r.t. A

$$\int \cdots \int_{i \in A} f \, \mathrm{d}\mu_i$$

is by definition another function $(\prod_{i \in \iota} X_i) \to [0, \infty]$ that is defined as (the notation is explained in Subsection 2.3).

$$x \mapsto \int_{\prod_{i \in A} X_i} f(X_A(x, y)) \, \mathrm{d}\Pi_{i \in A} \mu_i(y).$$

Note that $\int \cdots \int_{i \in A} f \, \mathrm{d}\mu_i$ is a function that does not depend on the arguments in A. We could also viewed this as a function on $\prod_{i \in \iota \setminus A} X_i$ instead of $\prod_{i \in \iota} X_i$. However, it is much more convenient to view it as a function on the whole product space $\prod_{i \in \iota} X_i$, since the alternative makes the statements of the lemmas below much more complicated.

We call this operation the marginal of f because of our intuition from probability theory. If all the μ_i are probability measures and f is a random variable, then $\int \cdots \int_{i \in A} f \, \mathrm{d}\mu_i$ is the marginal variable on $\prod_{i \in \iota \setminus A} X_i$.

Note that we do not assume that ι is finite: this construction works in an infinite product, as long as we only have finitely many integration variables.

▶ **Lemma 3.** *The following basic properties hold for any function $f : (\prod_{i \in \iota} X_i) \to [0, \infty]$.*

1. $\int \cdots \int_{i \in \varnothing} f \, \mathrm{d}\mu_i = f$.
2. *If $x, x' \in \prod_{i \in \iota} X_i$ and $x_i = x'_i$ for all $i \in \iota \setminus A$ then $\int \cdots \int_{i \in A} f \, \mathrm{d}\mu_i$ will have the same value on x and x'.*
3. $\int \cdots \int_{i \in A} f \, \mathrm{d}\mu_i$ *is monotone in f.*
4. *If ι is finite then $\int \cdots \int_{i \in \iota} f \, \mathrm{d}\mu_i$ is the constant function with value $\int f \, \mathrm{d}\Pi_i \mu_i$.*
5. *If f is measurable, then so is $\int \cdots \int_{i \in A} f \, \mathrm{d}\mu_i$.*

Proof. Parts 2, 3 and 4 follow immediately from the definition.[1]

For Part 1 note that the marginal of f w.r.t. $\varnothing$ is an integral over an empty product space. Since the empty product of measures is the Dirac measure on the unique point in the space, this equality follows easily.

For Part 5, to show that $\int\cdots\int_{i\in A} f\,\mathrm{d}\mu_i$ is measurable, by Tonelli's theorem it suffices to show measurability for $(x,y)\mapsto f(X_A(x,y))$, which is an easy exercise. ◀

Using the definition of marginal, we get a very nice formulation of Tonelli's theorem for finitary products.

▶ **Lemma 4.** *If f is measurable and A and B are disjoint finite subsets of ι, then*

$$\int\cdots\int_{i\in A\cup B} f\,\mathrm{d}\mu_i = \int\cdots\int_{i\in A}\int\cdots\int_{j\in B} f\,\mathrm{d}\mu_j\,\mathrm{d}\mu_i$$

Proof. Since A and B are disjoint, we have a measurable equivalence

$$e : \left(\prod_{i\in A} X_i\right)\times\left(\prod_{i\in B} X_i\right)\simeq\left(\prod_{i\in A\cup B} X_i\right).$$

Note that e maps the measure $(\Pi_{i\in A}\mu_i)\times(\Pi_{i\in B}\mu_i)$ to the measure $\Pi_{i\in A\cup B}\mu_i$. Therefore we compute

$$\begin{aligned}\int\cdots\int_{i\in A\cup B} f\,\mathrm{d}\mu_i &= \int_{\prod_{i\in A\cup B} X_i} f(X_{A\cup B}(x,y))\,\mathrm{d}\Pi_{i\in A\cup B}\mu_i(y)\\ &= \int_{(\prod_{i\in A} X_i)\times(\prod_{i\in B} X_i)} f(X_{A\cup B}(x,e(y)))\,\mathrm{d}(\Pi_{i\in A}\mu_i)\times(\Pi_{i\in B}\mu_i)(y)\\ &= \int_{\prod_{i\in A} X_i}\int_{\prod_{i\in B} X_i} f(X_{A\cup B}(x,e(y,z)))\,\mathrm{d}\Pi_{i\in B}\mu_i(z)\,\mathrm{d}\Pi_{i\in A}\mu_i(y)\\ &= \int_{\prod_{i\in A} X_i}\int_{\prod_{i\in B} X_i} f(X_B(X_A(x,y),z))\,\mathrm{d}\Pi_{i\in B}\mu_i(z)\,\mathrm{d}\Pi_{i\in A}\mu_i(y)\\ &= \int\cdots\int_{i\in A}\int\cdots\int_{j\in B} f\,\mathrm{d}\mu_j\,\mathrm{d}\mu_i.\end{aligned}$$

where the second step uses the properties of e and the third step uses Tonelli's theorem. ◀

▶ **Lemma 5.** *For $i_0\in\iota$,*

$$\int\cdots\int_{i\in\{i_0\}} f\,\mathrm{d}\mu_i = \int_{X_{i_0}} f(X_{i_0}(x,y))\mathrm{d}\mu_{i_0}(y)$$

Proof. We have a measurable equivalence $\left(\prod_{i\in\{i_0\}} X_i\right)\simeq X_{i_0}$ that maps the measure $\Pi_{i\in\{i_0\}}\mu_i$ to μ_{i_0}. Therefore,

$$\begin{aligned}\int\cdots\int_{i\in\{i_0\}} f\,\mathrm{d}\mu_i\,(x) &= \int_{\prod_{i\in\{i_0\}} X_i} f(X_{\{i_0\}}(x,y))\,\mathrm{d}\Pi_{i\in\{i_0\}}\mu_i(y)\\ &= \int_{X_{i_0}} f(X_{i_0}(x,y))\,\mathrm{d}\mu_{i_0}(y).\end{aligned}$$ ◀

[1] In Part 4 there is a slight complication in the formalization, because the type ι is not the same type as the universal subtype of ι (in the definition of marginal, A is used as a subtype of ι). This is not a problem: the proof is still only a few lines long in the formalization.

5 Applications of the Marginal Construction

5.1 Volume of an n-ball

Our first application is a loose port of Manuel Eberl's Isabelle formalization[2] (2017) of the formula for the volume of a ball in Euclidean n-space.

Let ι be a type of finite cardinality n. In this section $\mathbf{x}$ will denote a point in $\mathbb{R}^\iota$ and (x_j) the individual co-ordinates of such a point. Fix a real number $R \geq 0$. We will study the Euclidean ball in $\mathbb{R}^\iota$,

$$\{\mathbf{x} : \|\mathbf{x}\| \leq R\} = \left\{\mathbf{x} : \sum_{j:\iota} {x_j}^2 \leq R^2\right\}.$$

Define a constant $B_n := \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}$, where Γ denotes the gamma function (available in `Mathlib` due to work of David Loeffler[3]). In this section we prove:

▶ **Proposition 6.** volume $\{\mathbf{x} : \|\mathbf{x}\| \leq R\} = B_n R^n$.

We introduce the notation

$$I_k(t) := \begin{cases} 0, & t < 0 \\ t^{k/2}, & 0 \leq t, \end{cases}$$

for $k : \mathbb{N}$ and $t : \mathbb{R}$, and

$$A_s(\mathbf{x}) := B_{|s|} I_{|s|}\left(R^2 - \sum_{j \in s^c} {x_j}^2\right),$$

for a set s in ι and a vector $\mathbf{x} : \mathbb{R}^\iota$.

Observe that (denoting by χ_U the characteristic function of a set U)

$$\int \cdots \int_{\varnothing^c} A_\varnothing = \int_{\mathbf{x}:\mathbb{R}^\iota} B_{|\varnothing|}\chi_{\{\mathbf{x}:0 \leq R^2 - \|\mathbf{x}\|^2\}} = \text{volume}\,\{\mathbf{x} : \|\mathbf{x}\| \leq R\},$$

$$\int \cdots \int_{univ^c} A_{univ} = B_n R^n.$$

(A priori the left-hand sides are functions on $\mathbb{R}^\iota$. The statements are to be understood as saying that these functions are constant and equal to the expressions on the right-hand sides.) So Proposition 6 follows by induction from the following fact:

▶ **Proposition 7.** *For all sets s in ι and all $i \notin s$,*

$$\int \cdots \int_{s^c} A_s = \int \cdots \int_{(\{i\} \cup s)^c} A_{\{i\} \cup s}.$$

Proof. Given $i : \iota$, $\mathbf{x} : \mathbb{R}^\iota$ and $t : \mathbb{R}$,

A computation in single-variable calculus establishes that for all natural numbers k and all real numbers c,

$$\int B_k I_k\left(c - t^2\right)\, dt = B_{k+1} I_{k+1}(c).$$

[2] https://isabelle-dev.sketis.net/rISABELLEc60e3d615b8
[3] https://github.com/leanprover-community/mathlib/pull/12917

Therefore for any $\mathbf{x} : \mathbb{R}^\iota$,

$$\int A_s\left(X_i(\mathbf{x},t)\right)\ dt = \int B_{|s|}I_{|s|}\left(\left[R^2 - \sum_{j\in(\{i\}\cup s)^c} {x_j}^2\right] - t^2\right)\ dt$$
$$= B_{|\{i\}\cup s|}I_{|\{i\}\cup s|}\left(R^2 - \sum_{j\in(\{i\}\cup s)^c} {x_j}^2\right) = A_{\{i\}\cup s}\left(\mathbf{x}\right).$$

Integrating this fact over the variables in $(\{i\} \cup s)^c$,

$$\int\cdots\int_{s^c} A_s = \int\cdots\int_{(\{i\}\cup s)^c}\left(\mathbf{x} \mapsto \int A_s(X_i(\mathbf{x},t))\ dt\right) = \int\cdots\int_{(\{i\}\cup s)^c} A_{\{i\}\cup s}. \quad ◀$$

5.2 Transvections Preserve the Lebesgue Measure

A *transvection* is a matrix of the form $1 + A$, where 1 is the identity matrix and A is a matrix that has one, off-diagonal, non-zero entry. `Mathlib` contains the result that the linear transformation of $\mathbb{R}^n$ induced by a transvection preserves the Lebesgue measure.

This is one step in the proof of the corresponding result for a general matrix M (where a factor $|\det(M)|$ occurs). This result, the infinitesimal version of the change of variables formula, was contributed to `Mathlib` by Gouëzel [9, Section 5], and had previously been formalized in other systems. For example, Harrison [10, Section 7], working in HOL Light, calls it out as “quite hard work to formalize.” In both cases this result is proved along the way to a (non-infinitesimal) version of the change of variables formula.

The existing `Mathlib` proof was pretty long (34 lines) and required reasoning about explicit equivalences on the indexing set. Using the marginal construction, we gave a proof in 15 lines with the same mathematical argument. The main mathematical argument lies in proving equation (3) below. This remains roughly the same in both versions of the formalization, but the marginal construction allowed us to remove a lot of work for the remaining part of the argument.

▶ **Proposition 8.** *If ι is a finite indexing set, and T is a transvection on R^ι, then T preserves the Lebesgue measure.*

Proof. We have to show that $T_*\lambda = \lambda$ where λ is the Lebesgue measure. Since boxes form a basis of the σ-algebra on $\mathbb{R}^n$, it is sufficient to show that the measures agree on a box $A = \Pi_i A_i$, so we have to show that $\lambda(T^{-1}(A)) = \lambda(A)$. Suppose that $T = 1 + M$ where M has entry $c \neq 0$ in position (i,j) for $i \neq j$. For a given $\mathbf{x} \in \mathbb{R}^n$ we will first show that the following equality holds:

$$\lambda(\{y \mid X_i(\mathbf{x},y) \in T^{-1}(A)\}) = \lambda(\{y \mid X_i(\mathbf{x},y) \in A\}). \tag{3}$$

The intuition of this equality is that we're fixing all but one of the coordinates, and looking at the length A when varying only coordinate i. We calculate:

$$\lambda(\{y \mid X_i(\mathbf{x},y) \in T^{-1}(A)\}) = \lambda(\{y \mid T(X_i(\mathbf{x},y)) \in A\})$$
$$= \lambda(\{y \mid X_i(\mathbf{x},y) + c\mathbf{x}_j\mathbf{e_i} \in A\}) = \lambda(\{y \mid X_i(\mathbf{x},y + c\mathbf{x}_j) \in A\}) = \lambda(\{y \mid X_i(\mathbf{x},y) \in A\}),$$

where $\mathbf{e_i}$ is the i-th standard vector and where in the last inequality we use the translation-invariance of λ, showing (3).

To finish the proof, notice that we want to prove that

$$\int \cdots \int_{\{1,\ldots,n\}} \chi_A = \int \cdots \int_{\{1,\ldots,n\}} \chi_{T^{-1}(A)},$$

where χ_X is the characteristic function of X. Equation (3) can be rewritten as

$$\int \cdots \int_{\{i\}} \chi_A = \int \cdots \int_{\{i\}} \chi_{T^{-1}(A)},$$

and the claim follows from Proposition 4, in the following form:

$$\int \cdots \int_{\{1,\ldots,n\}} f = \int \cdots \int_{\{1,\ldots,n\}\setminus\{i\}} \int \cdots \int_{\{i\}} f.$$ ◀

5.3 Grid-lines Lemma

In this section we present our most intricate application of the marginal construction: the key argument in the Gagliardo-Nirenberg-Sobolev inequality (see Section 6), which for clarity we have abstracted as a separate proposition and baptized the *grid-lines lemma*.

This key argument is quite an illuminating test case for the difference between informal and formal mathematical practice. So before discussing our approach, we describe the presentations available in the mathematical literature. The argument involves a succession of uses of Hölder's inequality with respect to different variables of integration. In the literature, the argument is either presented in a particular low dimension and left for the reader to extrapolate, or described as an implicit induction with the actual structure of the induction being left unstated.

- **Nirenberg, 1959 [17]**: "We shall prove (2.4)' here for $n = 3$ .... For general n the inequality is proved in the same way."
- **Gilbarg-Trudinger, 1977 [8]**: "The inequality (7.27) is now integrated successively over each variable x_i, $i = 1, \ldots, n$, the generalized Hölder inequality (7.11) for $m = p_1 = \cdots = p_m = n - 1$ then being applied after each integration. Accordingly we obtain ... "
- **Evans, 1998 [7]**: "We continue by integrating with respect to $x_3, \ldots x_n$, eventually to find ... "
- **Tsui, 2008 [18]**: "To illustrate the main ideas, we discuss the case when $n = 3$ .... For the general case, we start with .... Repeating this process, we get ... "
- **Liu, 2023 [14]**: "[T]he inequality (1) for $p = 1$ is proved by integrating ... with respect to x_1 and applying the extended Hölder inequality, then repeating this procedure with respect to $x_2, x_3, \ldots x_n$ successively .... This tedious procedure is not very transparent, and is not easy to follow."

To formalize this argument, we need an explicit statement in general dimension. Given the appeals to the extrapolation in the presentations quoted above, it is perhaps not surprising that we did not find this explicit statement in the literature!

We need an ι-indexed family of sigma-finite measure spaces $(A_i)_{i:\iota}$, where ι is a type of finite cardinality n. The reader may wish to imagine for concreteness that each factor A_i is $\mathbb{R}$, so that the product type $\prod_{i:\iota} A_i$ is the function type $\iota \to \mathbb{R}$ (or $\mathbb{R}^\iota$ for short). The essential points of the argument remain unchanged in this special case, which is in fact the case needed for the Gagliardo-Nirenberg-Sobolev inequality.

We furthermore need a nonnegative real parameter p, which at different times will have different upper bounds (specified explicitly).

The statement of the lemma in general dimension is as follows.

▶ **Proposition 9** (Grid-lines lemma). *Suppose that $(n-1)p \leq 1$. If $f : \prod_{j:\iota} A_j \to [0,\infty]$ is a measurable function, then (the notation is explained in Subsection 2.3)*

$$\int_{\mathbf{x}:\prod_{j:\iota} A_j} f(\mathbf{x})^{1-(n-1)p} \prod_i \left(\int_{t:A_i} f(X_i(\mathbf{x},t))\right)^p \leq \left(\int_{\mathbf{x}:\prod_{j:\iota} A_j} f(\mathbf{x})\right)^{1+p}.$$

```
theorem lintegral_mul_prod_lintegral_pow_le (hp : (#ι - 1 : ℝ) * p ≤ 1)
    {f : (∀ i : ι, A i) → ℝ≥0∞} (hf : Measurable f) :
    ∫⁻ x, f x ^ (1 - (#ι - 1 : ℝ) * p)
    * ∏ i, (∫⁻ xᵢ, f (update x i xᵢ) ∂μ i) ^ p ∂.pi μ
    ≤ (∫⁻ x, f x ∂.pi μ) ^ (1 + p)
```

Our name for this lemma comes from the integrand on the left-hand side. Note that at $\mathbf{x} : \prod_{j:\iota} A_j$ the integrand is a weighted product of $f(\mathbf{x})$ and expressions of the form

$$\int_{t:A_i} f(X_i(\mathbf{x},t));$$

this expression is the integral of f in the single co-ordinate i, and in $\mathbb{R}^\iota$ such an expression represents the integral of f over the "grid line" through $\mathbf{x}$ obtained by varying the i-th co-ordinate while fixing the others. See Figure 1.

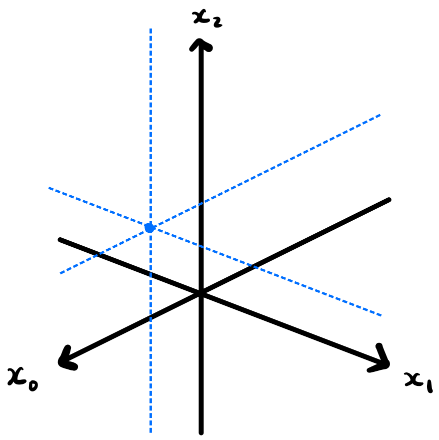

Figure 1 The left-hand integrand of the grid-lines lemma at a fixed point (blue point) is a weighted product of the value there with integrals over the lines through it parallel to the axes (blue dotted lines).

We introduce the notation

$$I_j f := \int \cdots \int_{\{j\}} f$$

for the marginal integral over the singleton set $\{j\}$ in ι, and

$$T_{p,s}(f) := \int \cdots \int_s f^{1-(|s|-1)p} \prod_{j \in s} (I_j f)^p.$$

Observe that

$$T_{p,\varnothing}\left(\int_{\prod_{j:\iota} A_j} f\right) = \left(\int_{\prod_{j:\iota} A_j} f\right)^{1+p}$$

$$T_{p,univ}(f) = \int_{\mathbf{x}:\prod_{j:\iota} A_j} f(\mathbf{x})^{1-(n-1)p} \prod_{j:\iota}\left(\int_{t:A_i} f(X_i(\mathbf{x},t))\right)^p.$$

(A priori the left-hand sides are functions on $\prod_{j:\iota} A_j$. The statements are to be understood as saying that these functions are constant and equal to the expressions on the right-hand sides.) So Proposition 9 follows by induction from the following fact:

▶ **Proposition 10.** *For all sets s in ι and all $i \notin s$, and for all p such that $|s|p \leq 1$,*

$$T_{p,\{i\}\cup s}(f) \leq T_{p,s}\left(I_i f\right).$$

The proof is a tricky computation that relies on Hölder's inequality at its heart. Note that on the left-hand-side we have an $|s|+1$-times iterated integral with $|s|+2$ factors inside the integral. If x_i denotes the i-th variable, we want to move the integral over x_i inside, and apply Hölder's inequality to the $|s|+1$ factors that depend on x_i (whose powers sum exactly to 1).

Proof. We have that

$$[1-|s|p] + |s|p = 1,$$

so for any $\mathbf{x} : \prod_{j:\iota} A_j$, by Hölder's inequality,

$$\begin{aligned}
&\int_{t:A_i} f(X_i(\mathbf{x},t))^{1-|s|p} \prod_{j\in s} I_j f(X_i(\mathbf{x},t))^p \\
&\leq \left(\int_{t:A_i} f(X_i(\mathbf{x},t))\right)^{1-|s|p} \prod_{j\in s}\left(\int_{t:A_i} I_j f(X_i(\mathbf{x},t))\right)^p.
\end{aligned}$$

Therefore for any $\mathbf{x} : \prod_{j:\iota} A_j$,

$$\begin{aligned}
&\int_{t:A_i} f(X_i(\mathbf{x},t))^{1-|s|p} \prod_{j\in\{i\}\cup s} I_j f(X_i(\mathbf{x},t))^p \\
&= \int_{t:A_i} I_i f(\mathbf{x})^p \left(f(X_i(\mathbf{x},t))^{1-|s|p} \prod_{j\in s} I_j f(X_i(\mathbf{x},t))^p\right) \\
&= I_i f(\mathbf{x})^p \int_{t:A_i} f(X_i(\mathbf{x},t))^{1-|s|p} \prod_{j\in s} I_j f(X_i(\mathbf{x},t))^p \\
&\leq I_i f(\mathbf{x})^p \left(\int_{t:A_i} f(X_i(\mathbf{x},t))\right)^{1-|s|p} \prod_{j\in s}\left(\int_{t:A_i} I_j f(X_i(\mathbf{x},t))\right)^p \\
&= I_i f(\mathbf{x})^p \; I_i f(\mathbf{x})^{1-|s|p} \prod_{j\in s} I_i I_j f(\mathbf{x})^p \\
&= I_i f(\mathbf{x})^{1-(|s|-1)p} \prod_{j\in s} I_j I_i f(\mathbf{x})^p.
\end{aligned}$$

Integrating this over the variables in s,

$$\begin{aligned}T_{p,\{i\}\cup s}(f) &= \int\cdots\int_{\{i\}\cup s} f^{1-|s|p} \prod_{j\in\{i\}\cup s} (I_j f)^p \\ &= \int\cdots\int_s \left(\mathbf{x} \mapsto \int_{t:A_i} f(X_i(\mathbf{x},t))^{1-|s|p} \prod_{j\in\{i\}\cup s} I_j f(X_i(\mathbf{x},t))^p\right) \\ &\le \int\cdots\int_s (I_i f)^{1-(|s|-1)p} \prod_{j\in s} (I_j I_i f)^p \\ &= T_{p,s}(I_i f)\,.\end{aligned}$$

◀

6 Gagliardo-Nirenberg-Sobolev Inequality

The version of the inequality we prove is due independently to Nirenberg [17, Lecture II] and Gagliardo; a variant result with different exponents was proved earlier by Sobolev, and can be deduced from the Gagliardo-Nirenberg version (although we do not formalize this deduction).

The L^p norm of a function f w.r.t. a measure μ is defined to be

$$\|f\|_{L^p} := \left(\int |f|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}} \in [0,\infty].$$

▶ **Theorem 11** (Gagliardo-Nirenberg-Sobolev inequality)**.** *Let E be a real normed space of finite dimension $n \ge 2$ with Haar measure μ. Let $1 \le p < n$ be a real number with Sobolev conjugate $p^* = \frac{np}{n-p}$. Then there exists a nonnegative real number C such that for all compactly supported C^1 functions $u : E \to \mathbb{R}$,*

$$\|u\|_{L^{p^*}} \le C\|Du\|_{L^p}. \tag{4}$$

The Lean version, which is displayed below, features a zoo of type classes. The predicate `ContDiff ℝ 1 u` states that u is C^1, `snorm u p′ μ` is the $L^{p'}$ norm of u (w.r.t. μ) and `fderiv ℝ u` is the total derivative of u. The conclusion features a constant `SNormLESNormFDerivOfEqConst F μ p : ℝ≥0`. We don't care about the precise value of this constant, but it is important that it only depends on F, μ and p (and E, the space on which μ is a measure).

Note also that in the formalization we generalized the codomain of u to be any finite-dimensional normed vector space.

```
theorem lintegral_pow_le_pow_lintegral_fderiv [NormedAddCommGroup E]
    [NormedSpace ℝ E] [MeasurableSpace E] [BorelSpace E]
    [FiniteDimensional ℝ E] (μ : Measure E) [IsAddHaarMeasure μ]
    [NormedAddCommGroup F] [NormedSpace ℝ F] [FiniteDimensional ℝ F]
    {u : E → F} (hu : ContDiff ℝ 1 u) (h2u : HasCompactSupport u)
    {p p′ : ℝ≥0} (hp : 1 ≤ p) (h2p : 0 < finrank ℝ E)
    (hp′ : (p′ : ℝ)⁻¹ = p⁻¹ - (finrank ℝ E : ℝ)⁻¹) :
    snorm u p′ μ ≤
      SNormLESNormFDerivOfEqConst F μ p * snorm (fderiv ℝ u) p μ
```

The main difficulty of the proof is for $p = 1$, and we will first prove that in the case that $E = \mathbb{R}^\iota$, where ι is a type of finite cardinality n. In that case, we can prove the following result.

▶ **Proposition 12.** *Let ι be a finite type of cardinality $n \geq 2$. For all compactly supported C^1 functions $u : \mathbb{R}^\iota \to \mathbb{R}$,*

$$\int |u|^{n/(n-1)} \leq \left(\int \|Du\|\right)^{n/(n-1)}.$$

Proof. The key observation here is that, by a half-infinite version of the Fundamental Theorem of Calculus, a compactly supported function is bounded pointwise by the integral of the norm of its gradient along any co-ordinate line. To be precise, for a given $\mathbf{x} : \mathbb{R}^\iota$ and $i : \iota$,

$$\begin{aligned} |u(\mathbf{x})| &= \left| \int_{-\infty}^{\mathbf{x}(i)} (u \circ X_i(\mathbf{x}, \cdot))'(t)\ dt \right| \\ &\leq \int_{-\infty}^{\mathbf{x}(i)} \left|(u \circ X_i(\mathbf{x}, \cdot))'(t)\right| dt \\ &\leq \int_{-\infty}^{\infty} \left\|Du|_{X_i(\mathbf{x},t)}\right\| dt. \end{aligned}$$

Here we use $Du|_{\mathsf{F}}$ to denote evaluation of Du at point F.

We obtain the desired bound by taking the product over all $i : \iota$ of these inequalities, for each $\mathbf{x} : \mathbb{R}^\iota$:

$$\begin{aligned} \int |u(\mathbf{x})|^{n/(n-1)}\, d\mathbf{x} &= \int \prod_i |u(\mathbf{x})|^{1/(n-1)}\, d\mathbf{x} \\ &\leq \int \prod_i \left(\int \left\|Du|_{X_i(\mathbf{x},t)}\right\| dt\right)^{1/(n-1)} d\mathbf{x}; \end{aligned}$$

the last line has exactly the form of the left-hand side of the grid-lines lemma (Proposition 9), with $f(\mathbf{x}) = \|Du|_{\mathbf{x}}\|$, and so, by that Proposition, is bounded above by

$$\left(\int \|Du|_{\mathbf{x}}\|\, d\mathbf{x}\right)^{n/(n-1)}.$$ ◀

Proof of Theorem 11. For $p = 1$, we can raise Proposition 12 to the power $\frac{n-1}{n}$ to obtain (4) for $E = \mathbb{R}^\iota$. Then we want to transfer this statement to functions u with as domain an arbitrary finite-dimensional vector space. This argument is not hard: we choose a basis on E and then use a continuous linear equivalence $e : \mathbb{R}^n \simeq E$, where n is the dimension of E. Then the measures μ and the pushforward of the Lebesgue measure $e_*(\lambda)$ are both Haar measures, so they must be the same up to some constant factor. So let us assume that $\mu = c \cdot e_*(\lambda)$. Then let

$$C = \frac{\|e\|^p}{c^{p-1}},$$

where $\|e\|$ is the operator norm of e. We then apply 12 to the function $u \circ e : \mathbb{R}^n \to \mathbb{R}$. A straightforward calculation involving the chain rule then shows (4) for $p = 1$ with the aforementioned value of C.

For $p > 1$, Define $\gamma := \frac{p(n-1)}{n-p}$. A simple calculation shows that $\frac{\gamma n}{n-1} = p^* = \frac{(\gamma-1)p}{p-1}$. We now apply the version for $p = 1$ to the function $v := |u|^\gamma$.

$$\|v\|_{L^{\frac{n}{n-1}}} \leq C\|Dv\|_{L^1} \leq C\gamma \int |u|^{\gamma-1}\|Du\| \leq C\gamma \left(\int |u|^{p^*}\right)^{\frac{p-1}{p}} \left(\int \|Du\|^p\right)^{\frac{1}{p}}$$

where in the second inequality we use the chain rule and in the third inequality we use Hölder's inequality. Hence by using that $\frac{n-1}{n} - \frac{p-1}{p} = \frac{1}{p^*}$ we compute

$$\|u\|_{L^{p^*}} = \|v\|_{L^{\frac{n}{n-1}}}^{\frac{1}{p^*}} \leq C\gamma \left(\int \|Du\|^p \right)^{\frac{1}{p}} = C\gamma \|Du\|_{L^p}.$$

This finishes the proof. ◀

In the formalization, there is one additional step in the proof, since we generalize the codomain of u. In the proof we use the fact that $x \mapsto |x|^\gamma$ is differentiable with derivative bounded by $\gamma|x|^{\gamma-1}$. This is still true in Hilbert spaces, but not generally in normed spaces, since the norm there need not be differentiable at all. To solve this, we first prove it for arbitrary Hilbert spaces, and then use the fact that for every finite-dimensional normed vector space there is a continuous linear equivalence to a Hilbert space (namely $\mathbb{R}^n$). We can then transfer the result along this equivalence.

Note that in the formalization we transferred the inequality twice along continuous linear equivalences. However, because of the nature of the statement, this transfer is not at all easy: it involves steps like the chain rule, the uniqueness of Haar measures and estimates using the operator norm of a linear map. It would be interesting to see to what extend automated transfer tactics (such as [20]) would be able to transfer a result like this.

The Gagliardo-Nirenberg-Sobolev inequality (Theorem 11) holds uniformly for all functions on a normed space: the supports of the functions considered must be compact, but they can be arbitrarily large. For fixed p and n, the Sobolev conjugate $p^* = \frac{np}{n-p}$ is the unique exponent for which such an inequality could be true; this is easily seen by a scaling argument.

On the other hand, if one restricts consideration to functions supported within a fixed bounded region s, there is more flexibility in the choice of exponent. A variant of Theorem 11 then holds for any q such that $1 \leq q \leq p^*$. This follows immediately from the monotonicity of L^p-membership, which is a consequence of Hölder's inequality:

$$\left(\int_s |f|^q \right)^{1/q} \leq \mathrm{Vol}(s)^{1/q - 1/p^*} \left(\int_s |f|^{p^*} \right)^{1/p^*}.$$

The most important special case is the case when q is p itself, which is valid since $p \leq p\frac{n}{n-p} = p^*$.

▶ **Theorem 13.** *Let s be a bounded measurable set in $\mathbb{R}^\iota$. Let $1 \leq p < |\iota|$. There exists a constant C, such that for all C^1 functions $u : \mathbb{R}^\iota \to \mathbb{R}$ with support in s,*

$$\|u\|_{L^p} \leq C \|Du\|_{L^p}.$$

7 Future Prospects: Sketch of some PDE Theory

Sobolev spaces are a longstanding goal for formalization [1, 2]. They are a standard setting for the solution of elliptic partial differential equations.

We outline a little of this theory to motivate our interest in the Gagliardo-Nirenberg-Sobolev inequality. Sobolev spaces are certain Banach spaces of functions, let us say for "nice" domain $\Omega \subseteq \mathbb{R}^n$ and codomain $\mathbb{R}$. The simplest example, the Sobolev space $H_0^1(\Omega, \mathbb{R})$, is (to give a nonstandard description) the subspace of the Hilbert space $L^2(\Omega, \mathbb{R}^n)$ consisting of functions U which are equal to Du, in the sense of weak (distributional) derivatives, for some element u in $L^2(\Omega, \mathbb{R})$, and which are L^2-approximated by C^1 compactly-supported functions in Ω. The Gagliardo-Nirenberg-Sobolev inequality (in the variant Theorem 13) implies that this subspace is a *closed* subspace of $L^2(\Omega, \mathbb{R}^n)$, thus Banach, and the (necessarily linear) operation sending U to a suitable u is a *bounded* linear map, which we notate P_0.

It follows by the Fréchet-Riesz representation theorem (see formalizations [16, 1, 6]) that for any function f in $L^2(\Omega, \mathbb{R})$, there exists a unique element U of the Sobolev space H_0^1 such that for all V in H_0^1,

$$\int_\Omega \langle U, V\rangle = \int_\Omega f P_0(V). \tag{5}$$

If $U = Du$ for a smooth (not just L^2) function u, this condition implies that for all smooth compactly-supported v,

$$\int \langle Du, Dv\rangle = \int fv.$$

By integrating by parts, this implies that u solves the *Poisson equation* $-\Delta u = f$. Motivated by this, we define a solution (5) to be a *weak solution* to this Poisson equation, even when U is not smooth, and thus we have proved existence and uniqueness of weak solutions to this Poisson equation.

In fact, the constant coefficients and high degree of symmetry in Poisson's equation make it rather special: it can be solved by a variety of methods and in many cases its solutions can be represented by semi-explicit formulas. See [5] for a formalization in this spirit of some theory of the heat equation, another PDE with constant coefficients and a high degree of symmetry. The notable point of the method described above is that it does not really exploit these constant coefficients or symmetries, so it continues to work for a large class of other *elliptic second-order linear* partial differential equations.

The argument above is representative of the subject as a whole. Most PDEs do not admit explicit solutions. Instead, researchers prove nonconstructive existence, uniqueness and regularity theorems for solutions of such PDEs (and, in the best-case scenario, also prove results about the convergence properties of numerical methods for approximating these solutions). Inequalities such as the Gagliardo-Nirenberg-Sobolev inequality play a crucial role, in establishing the functional-analysis preconditions for the nonconstructive existence theorems which are invoked.

8 Related Work

We refer to [2, Section 1] for a survey of the available formalizations of the binary Tonelli and/or Fubini theorems.

The first work implementing integration on finitary product types in formal theorem provers was carried out by Harrison [11, Section 5], whose work in HOL Light covers the specific case of $\mathbb{R}^n$ as part of a broader theory covering calculus on finite-dimensional vector spaces.

Hölzl and Heller [12] implemented integration on a general finitary product type as part of a full development of measure theory in the language Isabelle. Their framework for measure spaces is flexible: σ-algebras and measures are naturally defined on a subspace of a type. Similarly, when taking product measures, there is a subset of the indexing set that is considered when taking the product measure. This approach is similar to the marginal construction described in Section 4, since both allow for proofs by induction over the dimension. Such an induction is used to calculate the volume of the Euclidean ball in general dimension, formalized by Manuel Eberl. We re-formalize this in our own framework in Section 5.1.

The approaches are not the same; the marginal construction is more expressive than taking integrals with the notion of measure in Isabelle/HOL. Given a finite family of measurable spaces $(X_i)_{i\in\iota}$ and a measure μ_i on each X_i, the framework in Isabelle/HOL allows one to define the integral

$$\int_{\prod_{i\in A} X_i} f(X_A(\overline{x}, y))\, \mathrm{d}\Pi_{i\in A}\mu_i(y)$$

(for sets A in ι, and for $\overline{x}$ some fixed default element of $\prod_{i\in\iota} X_i$). In contrast, our marginal construction $\overline{x}$ is a variable, which allows us to define the function

$$x \mapsto \int_{\prod_{i\in A} X_i} f(X_A(x, y))\, \mathrm{d}\Pi_{i\in A}\mu_i(y).$$

In this comparison we have translated the concepts from Isabelle/HOL to our framework and notation, but because of differences in foundations, this translation is not exact.

The computation of the volume of a ball is simple enough that this can be conveniently done in Isabelle/HOL. For more complicated cases like the grid-lines lemma, a more expressive notion is needed. It would be interesting to see if the marginal construction can be conveniently adapted to Isabelle/HOL, by defining a variant of the product measure that depends on a point in the product space, used to take the "default values".

References

1. Sylvie Boldo, François Clément, Florian Faissole, Vincent Martin, and Micaela Mayero. A Coq formal proof of the Lax–Milgram theorem. In *6th ACM SIGPLAN Conference on Certified Programs and Proofs*, Paris, France, January 2017. doi:10.1145/3018610.3018625.
2. Sylvie Boldo, François Clément, Vincent Martin, Micaela Mayero, and Houda Mouhcine. A Coq formalization of Lebesgue induction principle and Tonelli's theorem. In *25th International Symposium on Formal Methods (FM 2023)*, volume 14000 of *Lecture Notes in Computer Science*, pages 39–55, Lübeck, Germany, March 2023. doi:10.1007/978-3-031-27481-7_4.
3. The mathlib Community. The Lean mathematical library. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs*, CPP 2020, pages 367–381, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3372885.3373824.
4. Leonardo Mendonça de Moura, Soonho Kong, Jeremy Avigad, Floris van Doorn, and Jakob von Raumer. The Lean Theorem Prover (System Description). In *Automated Deduction - CADE-25 - 25th International Conference on Automated Deduction, Berlin, Germany, August 1-7, 2015, Proceedings*, pages 378–388, Cham, 2015. Springer International Publishing. doi:10.1007/978-3-319-21401-6_26.
5. Elif Deniz, Adnan Rashid, Osman Hasan, and Sofiène Tahar. On the formalization of the heat conduction problem in HOL. In *Intelligent Computer Mathematics: 15th International Conference, CICM 2022, Tbilisi, Georgia, September 19-23, 2022, Proceedings*, pages 21–37, Berlin, Heidelberg, 2022. Springer-Verlag. doi:10.1007/978-3-031-16681-5_2.
6. Frédéric Dupuis, Robert Y. Lewis, and Heather Macbeth. Formalized functional analysis with semilinear maps. *Journal of Automated Reasoning*, 237:10:1–10:19, 2022. doi:10.4230/LIPIcs.ITP.2022.10.
7. Lawrence C. Evans. *Partial differential equations*, volume 19 of *Grad. Stud. Math.* Providence, RI: American Mathematical Society (AMS), 2nd ed. edition, 2010.
8. David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*, volume 224 of *Grundlehren Math. Wiss.* Springer, Cham, 1977. doi:10.1007/978-3-642-61798-0.

9 Sébastien Gouëzel. A formalization of the change of variables formula for integrals in mathlib. In Kevin Buzzard and Temur Kutsia, editors, *Intelligent Computer Mathematics - 15th International Conference, CICM 2022, Tbilisi, Georgia, September 19-23, 2022, Proceedings*, volume 13467 of *Lecture Notes in Computer Science*, pages 3–18. Springer, 2022. doi:10.1007/978-3-031-16681-5_1.

10 John Harrison. Formalizing basic complex analysis. In R. Matuszewski and A. Zalewska, editors, *From Insight to Proof: Festschrift in Honour of Andrzej Trybulec*, volume 10(23) of *Studies in Logic, Grammar and Rhetoric*, pages 151–165. University of Białystok, 2007. URL: http://mizar.org/trybulec65/.

11 John Harrison. The HOL Light theory of Euclidean space. *Journal of Automated Reasoning*, 50(2):173–190, February 2013. doi:10.1007/s10817-012-9250-9.

12 Johannes Hölzl and Armin Heller. Three chapters of measure theory in Isabelle/HOL. In *Interactive Theorem Proving*, pages 135–151, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg. doi:10.1007/978-3-642-22863-6_12.

13 Yury Kudryashov. Formalizing the divergence theorem and the Cauchy integral formula in lean. In June Andronick and Leonardo de Moura, editors, *13th International Conference on Interactive Theorem Proving (ITP 2022)*, volume 237 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 23:1–23:19, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ITP.2022.23.

14 Shibo Liu. Gagliardo-Nirenberg-Sobolev inequality: An induction proof. *The American Mathematical Monthly*, 130(9):859–861, 2023. doi:10.1080/00029890.2023.2240683.

15 Leonardo de Moura and Sebastian Ullrich. The Lean 4 theorem prover and programming language. In *Automated Deduction – CADE 28: 28th International Conference on Automated Deduction, Virtual Event, July 12–15, 2021, Proceedings*, pages 625–635, Berlin, Heidelberg, 2021. Springer-Verlag. doi:10.1007/978-3-030-79876-5_37.

16 Keiko Narita, Noboru Endou, and Yasunari Shidama. The orthogonal projection and the Riesz representation theorem. *Formaliz. Math.*, 23(3):243–252, 2015. doi:10.1515/forma-2015-0020.

17 L. Nirenberg. On elliptic partial differential equations. *Annali della Scuola Normale Superiore di Pisa - Scienze Fisiche e Matematiche*, Ser. 3, 13(2):115–162, 1959. URL: http://www.numdam.org/item/ASNSP_1959_3_13_2_115_0/.

18 Mao-Pei Tsui. Partial differential equations I. Lecture notes for Math 6540, University of Toledo, 2008. URL: http://math.utoledo.edu/~mtsui/8540f08/hw/Sobolev-Inequality.pdf.

19 Floris van Doorn. Formalized Haar measure. In Liron Cohen and Cezary Kaliszyk, editors, *12th International Conference on Interactive Theorem Proving (ITP 2021)*, volume 193 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 18:1–18:17, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ITP.2021.18.

20 Théo Zimmermann and Hugo Herbelin. Automatic and Transparent Transfer of Theorems along Isomorphisms in the Coq Proof Assistant. In *Conference on Intelligent Computer Mathematics*, Washington, D.C., United States, 2015. URL: https://hal.science/hal-01152588.