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

Floris van Doorn[1], Heather Macbeth[2]

[1]University of Bonn, [2]Fordham University

12 September 2024

Overview

Question: How to generalize Tonelli's/Fubini's theorem to finitely many variables?

$$\int_{X\times Y} f(z)\,dz = \int_X \int_Y f(x,y)\,dy\,dx$$

Overview

Question: How to generalize Tonelli's/Fubini's theorem to finitely many variables?

$$\int_{X\times Y} f(z)\,dz = \int_X \int_Y f(x,y)\,dy\,dx$$

Overview:

- Products of measures
- The marginal construction
- The Gagliardo-Nirenberg-Sobolev (GNS) inequality

Background: Measure Theory

Definition

A σ-algebra Σ on X is a collection of subsets of X that contains the empty set and is closed under complements and countable unions.

Definition

If Σ is a σ-algebra on X, then a measure on Σ is a function $\mu : \Sigma \to [0, \infty]$ such that $\mu(\varnothing) = 0$ and μ is *countably additive*: For pairwise disjoint sets $\{A_i\}_i$ we have

$$\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \sum_{i\in\mathbb{N}} \mu(A_i).$$

Given a measure on X we can define the (lower) Lebesgue integral

$$\int f\, d\mu = \int f(x)\, \mu(dx)$$

for functions $f : X \to [0, \infty]$.

Product measures

If μ is a measure on X and ν a measure on Y then we can define the product measure $\mu \times \nu$ on $X \times Y$. It can be defined as

$$(\mu \times \nu)(C) = \int_X \nu\{y \mid (x, y) \in C\} \, \mu(dx).$$

For general measures there are multiple product measures, but if μ and ν are σ-finite, then there is a unique product measure satisfying

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

Tonelli's theorem

Theorem (Tonelli's theorem)

Let $f : X \times Y \to [0, \infty]$ *be a measurable function.*
Then

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

and all the functions in the integrals above are measurable.

Iterated products

Suppose we have finitely many measure spaces $(X_i, \mu_i)_{i \in I}$ and we want to define a measure on $\prod_i X_i$. We could choose an ordering $I = \{i_1, \ldots, i_n\}$ and define the measure as (roughly) $\mu_{i_1} \times \cdots \times \mu_{i_n}$, but that means we have a non-canonical choice in the definition.

Iterated products

Suppose we have finitely many measure spaces $(X_i, \mu_i)_{i \in I}$ and we want to define a measure on $\prod_i X_i$. We could choose an ordering $I = \{i_1, \dots, i_n\}$ and define the measure as (roughly) $\mu_{i_1} \times \cdots \times \mu_{i_n}$, but that means we have a non-canonical choice in the definition.

We take the maximal measure μ such that for all $A_i \subseteq X_i$ we have

$$\mu(\Pi_i A_i) \leq \prod_i \mu_i(A_i).$$

We then show that equality holds by using the above non-canonical measure.

Iterated products

How would we state Tonelli's theorem for iterated products? We want something like

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

Iterated products

How would we state Tonelli's theorem for iterated products? We want something like

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

This is not the same as the statement we had for products.

Iterated products

How would we state Tonelli's theorem for iterated products? We want something like

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

This is not the same as the statement we had for products.

The equivalence $e : \mathbb{R}^{n+m} \simeq \mathbb{R}^n \times \mathbb{R}^m$ preserves the Lebesgue measure, so

$$\begin{aligned}\int_{\mathbb{R}^{n+m}} f(z)\,dz &= \int_{\mathbb{R}^n \times \mathbb{R}^m} f(e^{-1}(z))\,dz \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} f(x,y)\,dydx.\end{aligned}$$

Iterated products

How would we state Tonelli's theorem for iterated products? We want something like

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

This is not the same as the statement we had for products.

The equivalence $e : \mathbb{R}^{n+m} \simeq \mathbb{R}^n \times \mathbb{R}^m$ preserves the Lebesgue measure, so

$$\begin{aligned}\int_{\mathbb{R}^{n+m}} f(z)\,dz &= \int_{\mathbb{R}^n \times \mathbb{R}^m} f(e^{-1}(z))\,dz \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} f(x,y)\,dydx.\end{aligned}$$

Note that there are many equivalences e, which pull out different variables.

Iterated products

We should be able to pull out any k of the components and integrate only over those variables.

Iterated products

We should be able to pull out any k of the components and integrate only over those variables.

So if $f : \mathbb{R}^n \to [0, \infty]$ we want to be able to write an integral like

$$\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(\mathbf{x})\, dx_{a_1} dx_{a_2} \cdots dx_{a_k}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ and $A = \{a_1, \ldots, a_k\}$ is a subset of the variables.

Iterated products

We should be able to pull out any k of the components and integrate only over those variables.

So if $f : \mathbb{R}^n \to [0, \infty]$ we want to be able to write an integral like

$$\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(\mathbf{x})\, dx_{a_1} dx_{a_2} \cdots dx_{a_k}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ and $A = \{a_1, \ldots, a_k\}$ is a subset of the variables.

Note that the integral does not depend on the ordering of the a_i.
This is a function in the remaining variables.

Marginal construction: Definition

We denote by

$$\int \cdots \int_A f$$

the function $\mathbb{R}^n \to [0, \infty]$

$$x \mapsto \int_{\mathbb{R}^k} f(\mathbf{x}[\mathbf{y}/A])\, d\mathbf{y}.$$

where the vector $\mathbf{x}[\mathbf{y}/A]$ is defined as

$$\mathbf{x}[\mathbf{y}/A]_i := \begin{cases} y_i & \text{if } i \in A \\ x_i & \text{otherwise.} \end{cases}$$

Marginal construction: Remarks

- Note that
$$\int \cdots \int_A f : \mathbb{R}^k \to [0, \infty],$$
but it only depends on the variables x_i for $i \notin A$.

Marginal construction: Remarks

- Note that
$$\int \cdots \int_A f : \mathbb{R}^k \to [0, \infty],$$
but it only depends on the variables x_i for $i \notin A$.
- This generalizes nicely to any product of measure spaces, by replacing $\mathbb{R}^n$ by $\prod_{i \in I} X_i$ and $\mathbb{R}^k$ by $\prod_{i \in A} X_i$ and taking the integral w.r.t. a (finite) product measure (of σ-finite measures).

$$\int \cdots \int_{i \in A} f \, d\mu_i : \Big(\prod_{i \in I} X_i \Big) \to [0, \infty]$$

Marginal construction: Remarks

- Note that
$$\int \cdots \int_A f : \mathbb{R}^k \to [0, \infty],$$
but it only depends on the variables x_i for $i \notin A$.
- This generalizes nicely to any product of measure spaces, by replacing $\mathbb{R}^n$ by $\prod_{i \in I} X_i$ and $\mathbb{R}^k$ by $\prod_{i \in A} X_i$ and taking the integral w.r.t. a (finite) product measure (of σ-finite measures).
$$\int \cdots \int_{i \in A} f \, d\mu_i : \Big(\prod_{i \in I} X_i \Big) \to [0, \infty]$$
- We call this the marginal construction in reference to probability theory. If all the measures are probability measures and f is a random variable, then $\int \cdots \int_{i \in A} f \, d\mu_i$ is the marginal variable on $\prod_{i \in I \setminus A} X_i$.

Marginal construction: Remarks

- Note that
$$\int \cdots \int_A f : \mathbb{R}^k \to [0, \infty],$$
but it only depends on the variables x_i for $i \notin A$.
- This generalizes nicely to any product of measure spaces, by replacing $\mathbb{R}^n$ by $\prod_{i \in I} X_i$ and $\mathbb{R}^k$ by $\prod_{i \in A} X_i$ and taking the integral w.r.t. a (finite) product measure (of σ-finite measures).
$$\int \cdots \int_{i \in A} f \, d\mu_i : \Big(\prod_{i \in I} X_i \Big) \to [0, \infty]$$
- We call this the marginal construction in reference to probability theory. If all the measures are probability measures and f is a random variable, then $\int \cdots \int_{i \in A} f \, d\mu_i$ is the marginal variable on $\prod_{i \in I \smallsetminus A} X_i$.
- The definition is ugly.

Marginal construction: Remarks

- Note that
$$\int \cdots \int_A f : \mathbb{R}^k \to [0, \infty],$$
but it only depends on the variables x_i for $i \notin A$.
- This generalizes nicely to any product of measure spaces, by replacing $\mathbb{R}^n$ by $\prod_{i \in I} X_i$ and $\mathbb{R}^k$ by $\prod_{i \in A} X_i$ and taking the integral w.r.t. a (finite) product measure (of σ-finite measures).
$$\int \cdots \int_{i \in A} f \, d\mu_i : \Big(\prod_{i \in I} X_i \Big) \to [0, \infty]$$
- We call this the marginal construction in reference to probability theory. If all the measures are probability measures and f is a random variable, then $\int \cdots \int_{i \in A} f \, d\mu_i$ is the marginal variable on $\prod_{i \in I \setminus A} X_i$.
- The definition is ugly.
- It has very nice properties.

Marginal construction: Properties

Proposition

Let $f : \mathbb{R}^n \to [0, \infty]$*.*

1. *If* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* $x_i = y_i$ *for all* $i \notin A$ *then* $\int \cdots \int_A f$ *will have the same value on* $\mathbf{x}$ *and* $\mathbf{y}$*.*

Marginal construction: Properties

Proposition

Let $f : \mathbb{R}^n \to [0, \infty]$.

1. *If* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* $x_i = y_i$ *for all* $i \notin A$ *then* $\int \cdots \int_A f$ *will have the same value on* $\mathbf{x}$ *and* $\mathbf{y}$.
2. $\int \cdots \int_{\varnothing} f = f$

Marginal construction: Properties

Proposition

Let $f : \mathbb{R}^n \to [0, \infty]$.

1. *If* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* $x_i = y_i$ *for all* $i \notin A$ *then* $\int \cdots \int_A f$ *will have the same value on* $\mathbf{x}$ *and* $\mathbf{y}$.
2. $\int \cdots \int_\varnothing f = f$
3. *If* A *contains all variables, then* $\int \cdots \int_A f$ *is a constant functions with value* $\int f(\mathbf{x})\, d\mathbf{x}$.

Marginal construction: Properties

Proposition

Let $f : \mathbb{R}^n \to [0, \infty]$.

1. *If* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* $x_i = y_i$ *for all* $i \notin A$ *then* $\int\cdots\int_A f$ *will have the same value on* $\mathbf{x}$ *and* $\mathbf{y}$.
2. $\int\cdots\int_\varnothing f = f$
3. *If* A *contains all variables, then* $\int\cdots\int_A f$ *is a constant functions with value* $\int f(\mathbf{x})\,d\mathbf{x}$.
4. *If* $i_0 \in I$ *then*

$$\int\cdots\int_{\{i_0\}} f = \int_{\mathbb{R}} f(\mathbf{x}[y/\{i_0\}])dy$$

Marginal construction: Properties

Proposition

Let $f : \mathbb{R}^n \to [0, \infty]$.

1. *If* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* $x_i = y_i$ *for all* $i \notin A$ *then* $\int \cdots \int_A f$ *will have the same value on* $\mathbf{x}$ *and* $\mathbf{y}$.
2. $\int \cdots \int_\varnothing f = f$
3. *If* A *contains all variables, then* $\int \cdots \int_A f$ *is a constant functions with value* $\int f(\mathbf{x})\, d\mathbf{x}$.
4. *If* $i_0 \in I$ *then*
$$\int \cdots \int_{\{i_0\}} f = \int_{\mathbb{R}} f(\mathbf{x}[y/\{i_0\}]) dy$$
5. *If* f *is measurable and* A *and* B *are disjoint subsets of variables, then*
$$\int \cdots \int_{A \cup B} f = \int \cdots \int_A \int \cdots \int_B f$$

Marginal Construction: applications

- The GNS-inequality

Marginal Construction: applications

- The GNS-inequality
- Compute the volume of the unit ball in $\mathbb{R}^n$ (this was also done independently in Lean by Xavier Roblot), and is also proven in HOL Light and Isabelle/HOL.

Marginal Construction: applications

- The GNS-inequality
- Compute the volume of the unit ball in $\mathbb{R}^n$ (this was also done independently in Lean by Xavier Roblot), and is also proven in HOL Light and Isabelle/HOL.
- We can shorten the proof of a lemma in `Mathlib` that proves that transvections preserve the Lebesgue measure.

Marginal Construction: applications

- The GNS-inequality
- Compute the volume of the unit ball in $\mathbb{R}^n$ (this was also done independently in Lean by Xavier Roblot), and is also proven in HOL Light and Isabelle/HOL.
- We can shorten the proof of a lemma in `Mathlib` that proves that transvections preserve the Lebesgue measure.
- We translated the lemma from HOL Light to Lean. It states that to compute a $\mu(s)$ for a measurable $s \subseteq \mathbb{R}^{n+1}$ we can take the measure of the set for a fixed i-th coordinate x and then integrate over x. This was 35 lines in Lean, 300-600 in HOL Light (the HOL Light version assumes that s is bounded).

L^p-norm

The L^p-norm of a function f is

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

Gagliardo-Nirenberg-Sobolev inequality

The Gagliardo-Nirenberg-Sobolev inequality shows that for sufficiently nice functions f the norm $\|f\|_{L^p}$ is bounded by the L^q-norm of the derivative $\|Du\|_{L^q}$.

Gagliardo-Nirenberg-Sobolev inequality

The Gagliardo-Nirenberg-Sobolev inequality shows that for sufficiently nice functions f the norm $\|f\|_{L^p}$ is bounded by the L^q-norm of the derivative $\|Du\|_{L^q}$.

Theorem (Gagliardo-Nirenberg-Sobolev inequality)

Let E be a real normed space of finite dimension $n \geq 2$. Let $1 \leq p < n$ be a real number and $p^* = \frac{np}{n-p}$ its Sobolev conjugate. Then there exists a nonnegative real number C such that for all compactly supported C^1 functions $u : E \to \mathbb{R}$, we have

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

Gagliardo-Nirenberg-Sobolev inequality

The Gagliardo-Nirenberg-Sobolev inequality shows that for sufficiently nice functions f the norm $\|f\|_{L^p}$ is bounded by the L^q-norm of the derivative $\|Du\|_{L^q}$.

Theorem (Gagliardo-Nirenberg-Sobolev inequality)

Let E be a real normed space of finite dimension $n \geq 2$. Let $1 \leq p < n$ be a real number and $p^* = \frac{np}{n-p}$ its Sobolev conjugate. Then there exists a nonnegative real number C such that for all compactly supported C^1 functions $u : E \to \mathbb{R}$, we have

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

The GNS-inequality has many applications, e.g.

- define Sobolev spaces
- find (weak) solutions to elliptic second-order linear partial differential equations
  - e.g. the Poisson equation $\Delta u = f$.

Proof of the GNS-inequality

We prove the GNS-inequality first for $p = 1$ and then use that in the general case.

For $p = 1$ one estimates

$$|u(x)| = \left| \int_{-\infty}^{x_i} \frac{\partial u}{\partial x_i}(x) dx_i \right| \leq \int_{\mathbb{R}} |Du(x)| dx_i.$$

Then one inductively integrates over the variables x_1, x_2, x_3, ... and applies Hölder's inequality multiple times.

Proof of the GNS-inequality

We prove the GNS-inequality first for $p = 1$ and then use that in the general case.

For $p = 1$ one estimates

$$|u(x)|^{\frac{n}{n-1}} = \left| \int_{-\infty}^{x_i} \frac{\partial u}{\partial x_i}(x) dx_i \right|^{\frac{n}{n-1}} \leq \left(\int_{\mathbb{R}} |Du(x)| dx_i \right)^{\frac{n}{n-1}}.$$

Then one inductively integrates over the variables x_1, x_2, x_3, ... and applies Hölder's inequality multiple times.

Proof of the GNS-inequality

We prove the GNS-inequality first for $p = 1$ and then use that in the general case.

For $p = 1$ one estimates

$$|u(x)|^{\frac{n}{n-1}} = \left|\int_{-\infty}^{x_i} \frac{\partial u}{\partial x_i}(x)dx_i\right|^{\frac{n}{n-1}} \leq \left(\int_{\mathbb{R}} |Du(x)|dx_i\right)^{\frac{n}{n-1}}.$$

Then one inductively integrates over the variables x_1, x_2, x_3, ... and applies Hölder's inequality multiple times.

The induction hypothesis involves expressions like

$$\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} |Du(x)| dx_i dx_1 dx_2 \cdots dx_k$$

Proof of the GNS-inequality

- **Nirenberg, 1959**: "We shall prove (2.4)' here for $n = 3$ .... For general n the inequality is proved in the same way."
- **Gilbarg-Trudinger, 1977**: "The inequality (7.27) is now integrated successively over each variable x_i, $i = 1, \dots, n$, 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**: "We continue by integrating with respect to $x_3, \dots x_n$, eventually to find ..."
- **Tsui, 2008**: "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**: "[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 $x_2, x_3, \dots x_n$ successively .... This tedious procedure is not very transparent, and is not easy to follow."

```
calc ∫⁻ x : α i, Measure.pi (μ ∘' succAbove i) (insertNth i x ⁻
    = ∫⁻ x : α i, (∫⋯∫⁻_.univ, indicator (insertNth i x ⁻¹' s)
      simp_rw [← lintegral_indicator_one (measurable_insertNth
        lintegral_eq_lmarginal_univ y]
  _ = ∫⁻ x : α i, (∫⋯∫⁻_.univ, indicator (insertNth i x ⁻¹' s)
        (z ∘' i.succAbove) ∂μ i := by
      rw [← insertNth_dcomp_succAbove i x y]
  _ = ∫⁻ x : α i, (∫⋯∫⁻_{i}ᶜ,
        indicator (insertNth i x ⁻¹' s) 1 ∘ (· ∘' succAbove i) ∂
      simp_rw [← λ x ↦ lmarginal_image succAbove_right_injecti
        (f := indicator (insertNth i x ⁻¹' s) (1 : ((j : Fin n)
        (measurable_one.indicator (measurable_insertNth _ hs))
  _ = ∫⁻ x : α i, (∫⋯∫⁻_{i}ᶜ,
        indicator (insertNth i x ∘ (· ∘' succAbove i) ⁻¹' s) 1 ∂
      rfl
  _ = ∫⁻ x : α i, (∫⋯∫⁻_{i}ᶜ,
        indicator ((Function.update · i x) ⁻¹' s) 1 ∂μ) z ∂μ i
```

Side conditions are painful

```
let n' := NNReal.conjExponent n
have h0n : 2 ≤ n := Nat.succ_le_of_lt <| Nat.one_lt_cast.mp <|
have hn : NNReal.IsConjExponent n n' := .conjExponent (by norm
have h1n : 1 ≤ (n : ℝ≥0) := hn.one_le
have h2n : (0 : ℝ) < n - 1 := by simp_rw [sub_pos]; exact hn.c
have hnp : (0 : ℝ) < n - p := by simp_rw [sub_pos]; exact h2p
rcases hp.eq_or_lt with rfl|hp
let q := Real.conjExponent p
have hq : Real.IsConjExponent p q := .conjExponent hp
have h0p : p ≠ 0 := zero_lt_one.trans hp |>.ne'
have h1p : (p : ℝ) ≠ 1 := hq.one_lt.ne'
have h3p : (p : ℝ) - 1 ≠ 0 := sub_ne_zero_of_ne h1p
have h0p' : p' ≠ 0 := by
  suffices 0 < (p' : ℝ) from (show 0 < p' from this) |>.ne'
  rw [← inv_pos, hp', sub_pos]
  exact inv_lt_inv_of_lt hq.pos h2p
have h2q : 1 / n' - 1 / q = 1 / p' := by
```

Final thoughts

- The marginal construction is a very convenient tool to deal with iterated integrals.
- It allows one to conveniently apply Tonelli's theorem
- We have formalized this in Lean, and this work is now part of `Mathlib`

- As future work, we want to properly formalize Sobolev spaces and their applications to PDEs.

Final thoughts

- The marginal construction is a very convenient tool to deal with iterated integrals.
- It allows one to conveniently apply Tonelli's theorem
- We have formalized this in Lean, and this work is now part of `Mathlib`

- As future work, we want to properly formalize Sobolev spaces and their applications to PDEs.

Thank You

Application: Sobolev spaces

Let $u : \Omega \to \mathbb{R}$ for some nice open $\Omega \subseteq \mathbb{R}^n$.
If u is differentiable with derivative $v = Du : \Omega \to \mathbb{R}^n$, then for all compactly supported smooth functions $\varphi : \Omega \to \mathbb{R}$

$$\int_\Omega v\varphi = -\int_\Omega uD\varphi.$$

This equation also makes sense if u is not differentiable, and v is a weak derivative of u if it holds for all such φ. v is also denoted Du.

The Sobolev space $H_0^1(\Omega, \mathbb{R})$ consists of all functions that have an L^2 weak derivative and are a L^2 limit of smooth functions with compact support.

Application: Sobolev spaces

The Sobolev inequality implies that for $u \in H_0^1(\Omega)$

$$\|u\|_{L^2} \le \|u\|_{L^{\frac{2n}{n-2}}} \le C\|Du\|_{L^2}.$$

The Sobolev space is a Hilbert space with

$$\|u\|_{H_0^1} := \sqrt{\|u\|_{L^2}^2 + \|Du\|_{L^2}^2} \le C'\|Du\|_{L^2}.$$

This means that $(u, v) \mapsto \langle Du, Dv\rangle_{L^2}$ forms an inner product on H_0^1 that gives the same topology.

For $f \in L^2(\Omega)$ we have that the operator $v \mapsto \langle f, v\rangle_{L^2} : H_0^1(\Omega) \to \mathbb{R}$ is a bounded linear functional on $H_0^1(\Omega)$. By the Riesz representation theorem there is a unique element $u \in H_0^1(\Omega)$ such that for all $v \in H_0^1(\Omega)$ we have

$$\langle Du, Dv\rangle_{L^2} = \langle f, v\rangle_{L^2}.$$