Towards formalizing variation estimates for commuting transformations

Floris van Doorn

j.w.w. Polona Durcik, Felix Pernegger, Jim Portegies, Joris Roos, Michael Rothgang, Lenka Slavíková, Christoph Thiele

15 May 2026

Overview

Goal: solve problems in Ergodic theory using harmonic analysis, and formalize these proofs.

Ongoing collaboration with Polona Durcik, Joris Roos, Lenka Slavíková and Christoph Thiele on the pure mathematical research.

Working with multiple people, including Felix Pernegger, Jim Portegies and Michael Rothgang on required preliminaries.

Disclaimer: I'm neither an expert in ergodic theory nor harmonic analysis.

Outline

1. Commuting transformations in ergodic theory
2. Calderón transference principle
3. Multilinear interpolation
4. Refactoring Mathlib

What is Ergodic theory?

Study of statistical properties of a dynamical system.

Think: cloud of particles moving in a bounded region and we are interested in some quantity f (e.g. position, speed).

Questions:

- What is the time average of f for a single particle?
- What is the average of f for all particles at a given time?

What is Ergodic theory?

Mathematically, we have:

- Let (X, Σ, μ) be a probability space;
- $T : X \to X$ be a measure-preserving transformation, i.e.

$$\mu(T^{-1}(A)) = \mu(A);$$

- $f : X \to \mathbb{R}$ a measurable map.

What is Ergodic theory?

Mathematically, we have:

- Let (X, Σ, μ) be a probability space;
- $T : X \to X$ be a measure-preserving transformation, i.e.

$$\mu(T^{-1}(A)) = \mu(A);$$

- $f : X \to \mathbb{R}$ a measurable map.

An important concept is the ergodic average of a point x:

$$M_k f(x) := \frac{1}{k} \sum_{i=0}^{k-1} f(T^i x).$$

Ergodic averages

Ergodic average: $M_k f(x) := \frac{1}{k} \sum_{i=0}^{k-1} f(T^i x)$.

Recall: $f \in L^p(X)$ when $\|f\|_{L^p(X)} < \infty$ where

$$\|f\|_{L^p(X)} = \left(\int_X |f(x)|^p \, d\mu(x) \right)^{\frac{1}{p}} .$$

Ergodic averages

Ergodic average: $M_k f(x) := \frac{1}{k}\sum_{i=0}^{k-1} f(T^i x)$.
Recall: $f \in L^p(X)$ when $\|f\|_{L^p(X)} < \infty$ where

$$\|f\|_{L^p(X)} = \left(\int_X |f(x)|^p \, d\mu(x)\right)^{\frac{1}{p}}.$$

Theorem (Mean Ergodic Theorem)

If $f \in L^p(X)$, then $M_k f$ converges w.r.t. the L^p-norm as $k \to \infty$ (for $p < \infty$).

Proven in 1931 by von Neumann for $p = 2$. Other versions were proven later in the 1930s.
The limit is (constantly) $\|f\|_{L^p(X)}$ is T is *ergodic*.

Ergodic averages

Ergodic average: $M_k f(x) := \frac{1}{k} \sum_{i=0}^{k-1} f(T^i x)$.
Recall: $f \in L^p(X)$ when $\|f\|_{L^p(X)} < \infty$ where

$$\|f\|_{L^p(X)} = \left(\int_X |f(x)|^p \, d\mu(x) \right)^{\frac{1}{p}} .$$

Theorem (Mean Ergodic Theorem)

If $f \in L^p(X)$, then $M_k f$ converges w.r.t. the L^p-norm as $k \to \infty$ (for $p < \infty$).

Proven in 1931 by von Neumann for $p = 2$. Other versions were proven later in the 1930s.
The limit is (constantly) $\|f\|_{L^p(X)}$ is T is *ergodic*.

Theorem (Pointwise Ergodic Theorem – Birkhoff, 1931)

If $f \in L^1(X)$ then $M_k f(x)$ converges for almost all x as $k \to \infty$.

Commuting transformations

Now consider $T_1, \ldots, T_n : X \to X$ to be measure preserving and (pairwise) commuting. We now consider the multiple ergodic average

$$M_k(f_1, \ldots, f_n)(x) := \frac{1}{k} \sum_{i=0}^{k-1} f_1(T_1^i x) \cdots f_n(T_n^i x).$$

These were used in a new proof of Szemerédi's Theorem by Furstenberg (1977).

Commuting transformations

Now consider $T_1, \ldots, T_n : X \to X$ to be measure preserving and (pairwise) commuting. We now consider the multiple ergodic average

$$M_k(f_1, \ldots, f_n)(x) := \frac{1}{k} \sum_{i=0}^{k-1} f_1(T_1^i x) \cdots f_n(T_n^i x).$$

These were used in a new proof of Szemerédi's Theorem by Furstenberg (1977).

Theorem (Tao, 2007)

If $n \geq 1$ and $f_i \in L^\infty(X)$, then $M_k(f_1, \ldots, f_n)$ converges w.r.t. the L^p-norm as $k \to \infty$.

Commuting transformations

Now consider $T_1, \ldots, T_n : X \to X$ to be measure preserving and (pairwise) commuting. We now consider the multiple ergodic average

$$M_k(f_1, \ldots, f_n)(x) := \frac{1}{k} \sum_{i=0}^{k-1} f_1(T_1^i x) \cdots f_n(T_n^i x).$$

These were used in a new proof of Szemerédi's Theorem by Furstenberg (1977).

Theorem (Tao, 2007)

If $n \geq 1$ and $f_i \in L^\infty(X)$, then $M_k(f_1, \ldots, f_n)$ converges w.r.t. the L^p-norm as $k \to \infty$.

Conjecture

If $n = 2$ and $f_1, f_2 \in L^\infty(X)$, then $M_k(f_1, f_2)(x)$ converges for almost all x as $k \to \infty$.

Variation estimates

A more quantitative estimate is given by norm variation bounds.

Theorem (Jones, Ostrovskii, Rosenblatt, 1996)

For $n = 1$ and $u_0 < u_1 < \ldots < u_m$ and $f \in L^2(X)$ and $C = 5^4$ we have

$$\sum_{j=0}^{m-1} \|M_{u_{i+1}}(f) - M_{u_i}(f)\|_{L^2(X)}^2 \le C\|f\|_{L^2(X)}^2.$$

Variation estimates

A more quantitative estimate is given by norm variation bounds.

Theorem (Jones, Ostrovskii, Rosenblatt, 1996)

For $n = 1$ and $u_0 < u_1 < \ldots < u_m$ and $f \in L^2(X)$ and $C = 5^4$ we have

$$\sum_{j=0}^{m-1} \|M_{u_{i+1}}(f) - M_{u_i}(f)\|^2_{L^2(X)} \leq C\|f\|^2_{L^2(X)}.$$

If s is a sequence taking values in a normed space B, its r-variation is

$$\|s\|_{V^r(B)} := \sup_{\substack{m \in \mathbb{N} \\ u_0 \leq u_1 \leq \cdots \leq u_m}} \left(\sum_{j=0}^{m-1} \|s(u_{i+1}) - s(u_i)\|^r_B \right)^{\frac{1}{r}}.$$

Note: if $\|s\|_{V^r(B)}$ is finite, then s is a Cauchy sequence.

Variation estimates for commuting transformations

We also have norm variation estimates for $n = 2$ and $n = 3$.

Theorem (Durcik, Kovač, Škreb, Thiele, 2016)

For $n = 2$ and $f_1, f_2 \in L^4(X)$ with norm 1 we have

$$\|M_\bullet(f_1, f_2)\|_{V^2(L^2(X))} \leq C.$$

Theorem (Durcik, Slavíková, Thiele, 2023)

For $n = 3$ and $r > 4$ and $f_1, f_2 \in L^8(X)$ and $f_3 \in L^4(X)$ with respective norms 1 we have

$$\|M_\bullet(f_1, f_2, f_3)\|_{V^2(L^2(X))} \leq C_r.$$

Using interpolation, we can replace the exponents $(8, 8, 4)$ by $(6, 6, 6)$.

Variation estimates for commuting transformations

We also have norm variation estimates for $n = 2$ and $n = 3$.

Theorem (Durcik, Kovač, Škreb, Thiele, 2016)

For $n = 2$ *and* $u_0 < u_1 < \ldots < u_m$ *and* $f_1, f_2 \in L^4(X)$ *with norm 1 we have*

$$\sum_{j=0}^{m-1} \|M_{u_{i+1}}(f_1, f_2) - M_{u_i}(f_1, f_2)\|_{L^2(X)}^2 \leq C.$$

Theorem (Durcik, Slavíková, Thiele, 2023)

For $n = 3$ *and* $r > 4$ *and* $u_0 < u_1 < \ldots < u_m$ *and* $f_1, f_2 \in L^8(X)$ *and* $f_3 \in L^4(X)$ *with respective norms 1 we have*

$$\sum_{j=0}^{m-1} \|M_{u_{i+1}}(f_1, f_2, f_3) - M_{u_i}(f_1, f_2, f_3)\|_{L^2(X)}^2 \leq C_r.$$

Using interpolation, we can replace the exponents $(8, 8, 4)$ by $(6, 6, 6)$.

Summary of estimates

n	norm convergence	norm variation bound	a.e. pointwise
1	1930s	1996	1931
2	2007	2016	open
3		2023	
4+		open	

Variation estimates for commuting transformations

The proofs are highly technical, and require technical steps ingredients.

- The Calderón transference principle;
- Many singular Brascamp–Lieb inequalities;
- Lots of tricky computations;
- Multilinear interpolation.

Outline

1 Commuting transformations in ergodic theory

2 Calderón transference principle

3 Multilinear interpolation

4 Refactoring Mathlib

Calderón transference principle

Goal: transform the statement to a statement about functions on $\mathbb{R}^n$, so that you can use harmonic analysis to analyze the problem. For $n = 2$ the statement is reduced to the following.

Proposition

Let $F_1, F_2 \in L^4(\mathbb{R}^2)$ *with norm 1, and define for* $t > 0$

$$A_t^{\varphi}(F_1, F_2)(x, y) := \int_{\mathbb{R}} F_1(x + s, y) F_2(x, y + s) t^{-1} \varphi(t^{-1} s) \, ds.$$

Then

$$\| A_{\bullet}^{\mathbb{1}_{[0,1)}}(F_1, F_2) \|_{V^2(L^2(\mathbb{R}^2))} \leq C$$

Proof idea

Theorem

For $f_1, f_2 \in L^4(X)$ with norm 1 we have

$$\|M_\bullet(f_1, f_2)\|_{V^2(L^2(X))} \leq C.$$

Proposition

For $F_1, F_2 \in L^4(\mathbb{R}^2)$ with norm 1 we have

$$\|A_\bullet^{\mathbb{1}[0,1)}(F_1, F_2)\|_{V^2(L^2(\mathbb{R}^2))} \leq C$$

Proof idea of Proposition ⇒ Theorem:
Let f_i be given, and apply the proposition to

$$F_{i,x_0,N}(x,y) := \begin{cases} f_i(T_1^k T_2^\ell x_0) & \text{where } k = \lfloor x \rfloor,\ \ell = \lfloor y \rfloor \text{ and } 0 \leq k, \ell \leq N \\ 0 & \text{otherwise.} \end{cases}$$

Formalization

The Calderón transference principle was formalized by Felix Pernegger.

Feature: to work with functions like

$$x \mapsto T_1^{k_1} T_2^{k_2} \cdots T_n^{k_n} x$$

we use the formalized version

```
def iterateFamily {ι X : Type*} [Fintype ι] :
  (ι → X → X) → (ι → ℕ) → X → X

theorem iterateFamily_add {T : ι → X → X}
  (hT : PairwiseCommute T) (k k' : ι → ℕ) :
  iterateFamily T (k + k') =
  iterateFamily T k ∘ iterateFamily T k'
```

Outline

1. Commuting transformations in ergodic theory
2. Calderón transference principle
3. Multilinear interpolation
4. Refactoring Mathlib

Weak-L^p norm

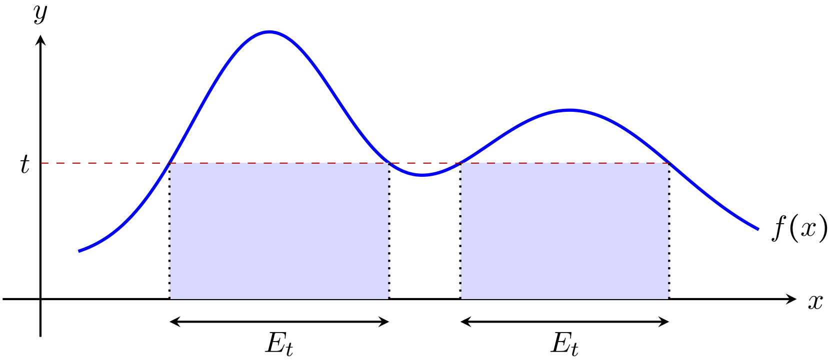
Real interpolation was formalized by Jim Portegies for the Carleson project. Michael Rothgang significantly cleaned and streamlined the proof.

To state it, we have to define the weak-L^p norm:

$$\|f\|_{L^{p,\infty}(X)} \coloneqq \sup_{t \geq 0} t \cdot \mu\{|f| > t\}^{\frac{1}{p}}.$$

Weak-L^p norm

$$\|f\|_{L^{p,\infty}(X)} := \sup_{t\geq 0} t \cdot \mu \underbrace{\{|f| > t\}}_{E_t}{}^{\frac{1}{p}}.$$



$$\|f\|_{L^{p,\infty}(X)} \leq \|f\|_{L^p(X)}.$$

Real interpolation

Basic fact: if $f \in L^p(X) \cap L^q(X)$, then $f \in L^r(X)$ for $r \in [p, q]$.

Real interpolation

Basic fact: if $f \in L^p(X) \cap L^q(X)$, then $f \in L^r(X)$ for $r \in [p, q]$.

Also, if $f \in L^{p,\infty}(X) \cap L^{q,\infty}(X)$, then $f \in L^r(X)$ for $r \in (p, q)$.

Real interpolation

Basic fact: if $f \in L^p(X) \cap L^q(X)$, then $f \in L^r(X)$ for $r \in [p, q]$.

Also, if $f \in L^{p,\infty}(X) \cap L^{q,\infty}(X)$, then $f \in L^r(X)$ for $r \in (p, q)$.

The interpolation theorem talks about operators on functions. Let $T : \mathbb{R}^X \to \mathbb{R}^Y$.

- T has strong type (p, q) if T is a bounded operator $L^p(X) \to L^q(Y)$.
- T has weak type (p, q) if T is a bounded operator $L^p(X) \to L^{q,\infty}(Y)$.

Real interpolation

Basic fact: if $f \in L^p(X) \cap L^q(X)$, then $f \in L^r(X)$ for $r \in [p, q]$.

Also, if $f \in L^{p,\infty}(X) \cap L^{q,\infty}(X)$, then $f \in L^r(X)$ for $r \in (p, q)$.

The interpolation theorem talks about operators on functions. Let $T : \mathbb{R}^X \to \mathbb{R}^Y$.

- T has strong type (p, q) if T is a bounded operator $L^p(X) \to L^q(Y)$.
- T has weak type (p, q) if T is a bounded operator $L^p(X) \to L^{q,\infty}(Y)$.

Theorem (Marcinkiewicz interpolation theorem)

If T is a sublinear operator with weak type (p, p) and weak type (q, q), then it has strong type (r, r) for $p < r < q$.

Formalization

The formalization uses extended norms, which take values in $[0, \infty]$, a.k.a. ENNReal or ℝ≥0∞.

```
class ENorm (E : Type*) where
  enorm : E → ℝ≥0∞

class ESeminormedAddMonoid (E : Type*) [TopologicalSpace E]
    extends ENorm E, AddMonoid E where
  continuous_enorm : Continuous enorm
  enorm_zero : ‖(0 : E)‖ = 0
  enorm_add_le : ∀ x y : E, ‖x + y‖ₑ ≤ ‖x‖ₑ + ‖y‖ₑ
```

The main reason to work with these notions is to work in either normed spaces or $[0, \infty]$ itself.

Jim noticed that you don't need sublinearity, just subadditivity.

```
def AESubadditiveOn [ENorm ε'] (T : (α → ε) → (α' → ε'))
    (P : (α → ε) → Prop) (A : ℝ≥0∞) (ν : Measure α') :
    Prop :=
  ∀ (f g : α → ε), P f → P g →
    ∀ᵐ x ∂ν, ‖T (f + g) x‖ₑ ≤ A * (‖T f x‖ₑ + ‖T g x‖ₑ)
```

Formalization

```
def HasStrongType (T : (α → ε) → (α' → ε')) (p p' : ℝ≥0∞)
    (μ : Measure α) (ν : Measure α') (c : ℝ≥0∞) : Prop :=
  ∀ f : α → ε, MemLp f p μ →
    AEStronglyMeasurable (T f) ν ∧
    eLpNorm (T f) p' ν ≤ c * eLpNorm f p μ
```

Notice that we're *not* working with (bundled) L^p-functions.

Real interpolation

```
theorem exists_hasStrongType_real_interpolation
  {T : (α → E₁) → (α' → E₂)} {p₀ p₁ q₀ q₁ p q : ℝ≥0∞}
  [TopologicalSpace E₁] [ENormedAddCommMonoid E₁]
  [MeasurableSpace E₁] [BorelSpace E₁]
  [PseudoMetrizableSpace E₁]
  [TopologicalSpace E₂] [ENormedAddCommMonoid E₂]
  (hp₀ : p₀ ∈ Ioc 0 q₀) (hp₁ : p₁ ∈ Ioc 0 q₁) (hq₀q₁ : q₀ ≠ q₁)
  {C₀ C₁ A : ℝ≥0} (hA : 1 ≤ A) (ht : t ∈ Ioo 0 1)
  (hC₀ : 0 < C₀) (hC₁ : 0 < C₁)
  (hp : p⁻¹ = (1 - t) / p₀ + t / p₁)
  (hq : q⁻¹ = (1 - t) / q₀ + t / q₁)
  (hmT : ∀ f, MemLp f p μ → AEStronglyMeasurable (T f) ν)
  (hT : AESubadditiveOn T
    (fun f ↦ MemLp f p₀ μ ∨ MemLp f p₁ μ) A ν)
  (h₀T : HasWeakType T p₀ q₀ μ ν C₀)
  (h₁T : HasWeakType T p₁ q₁ μ ν C₁) :
  HasStrongType T p q μ ν
    (C_realInterpolation p₀ p₁ q₀ q₁ q C₀ C₁ A t)
```

Generalizations

There is a generalization of the Marcinkiewicz interpolation theorem to Lorentz spaces $L^{p,q}$.

Lorentz spaces are formalized in the Carleson project by Leo Diedering. This can be used to prove more variants of Carleson's theorem.

Generalizations

There is a generalization of the Marcinkiewicz interpolation theorem to Lorentz spaces $L^{p,q}$.

Lorentz spaces are formalized in the Carleson project by Leo Diedering. This can be used to prove more variants of Carleson's theorem.

A generalization that will be useful for us is multilinear interpolation:

Theorem (Multilinear interpolation theorem)

If T is a sub-multilinear operator

$$T : L^p(X_1) \times \cdots \times L^p(X_n) \to L^{p,\infty}(Y)$$
$$T : L^q(X_1) \times \cdots \times L^q(X_n) \to L^{q,\infty}(Y)$$

Then also, for $p < r < q$

$$T : L^r(X_1) \times \cdots \times L^r(X_n) \to L^r(Y)$$

Outline

1 Commuting transformations in ergodic theory

2 Calderón transference principle

3 Multilinear interpolation

4 Refactoring Mathlib

r-variation

Mathlib has a notion of the variation norm.

```
noncomputable def eVariationOn {α E : Type*}
  [LinearOrder α] [PseudoEMetricSpace E]
  (f : α → E) (s : Set α) : ℝ≥0∞ :=
  ⨆ p : ℕ × { u : ℕ → α // Monotone u ∧ ∀ i, u i ∈ s },
  ∑ i ∈ Finset.range p.1,
    edist (f (p.2.1 (i + 1))) (f (p.2.1 i))
```

We will generalize this to the r-variation norm.

Weak extended metric spaces

We want to apply `eVariationOn` when E is $[0, \infty]$.

Problem: $[0, \infty]$ is not an `EMetricSpace`. In such spaces we have:

```
Tendsto u f (𝓝 a) ↔ ∀ ε > 0, ∀ᶠ x in f, edist (u x) a < ε
```

This is false in $[0, \infty]$.

Extended metric spaces

```
class PseudoEMetricSpace (α : Type u) : Type u
    extends EDist α where
  edist_self : ∀ x : α, edist x x = 0
  edist_comm : ∀ x y : α, edist x y = edist y x
  edist_triangle : ∀ x y z : α,
    edist x z ≤ edist x y + edist y z
  toUniformSpace : UniformSpace α :=
    uniformSpaceOfEDist edist edist_self edist_comm
   edist_triangle
  uniformity_edist : 𝓤 α =
    ⨅ ε > 0, 𝓟 { p : α × α | edist p.1 p.2 < ε } := by rfl
```

Weak extended metric spaces

```
class WeakPseudoEMetricSpace (α : Type u)
  [τ : TopologicalSpace α] : Type u extends EDist α  where
  edist_self : ∀ x : α, edist x x = 0
  edist_comm : ∀ x y : α, edist x y = edist y x
  edist_triangle : ∀ x y z : α,
    edist x z ≤ edist x y + edist y z
  /-- The topology on `α` is at most as fine as the
    topology generated by the `edist`. -/
  topology_le :
    (uniformSpaceOfEDist edist edist_self edist_comm
   edist_triangle).toTopologicalSpace ≤ τ
  /-- The ambient topology on `α` matches the `edist`
    topology on eballs. -/
  topology_eq_on_restrict : ∀ (x : α) (r : ℝ≥0∞),
    IsOpen ((Metric.eball x ⊤) ↓∩ (Metric.eball x r))
```

Weak extended metric spaces

This refactor was performed by Felix Pernegger. It was merged a few weeks ago.

Currently in an open PR:

```
weakPseudoEMetricSpace_of_isOpenEmbedding {α : Type*}
  [TopologicalSpace α] [TopologicalSpace (Option α)]
  [WeakPseudoEMetricSpace α]
  (h : IsOpenEmbedding (some (α := α))) :
  WeakPseudoEMetricSpace (Option α)
```

Other refactors

Other ongoing refactors or refactors under consideration:

- L^p-norm for $p = 0$ should be

$$\|f\|_{L^p(X,\mu)} = \mu(\text{support}(f)).$$

Other refactors

Other ongoing refactors or refactors under consideration:

- L^p-norm for $p = 0$ should be

$$\|f\|_{L^p(X,\mu)} = \mu(\mathsf{support}(f)).$$

- L^p-norm for $0 < p < 1$ should (maybe) be

$$\|f\|_{L^p(X,\mu)} = \int_X |f(x)|^p \, d\mu(x),$$

  i.e. without the $\frac{1}{p}$ exponent.

Other refactors

Other ongoing refactors or refactors under consideration:

- L^p-norm for $p = 0$ should be

$$\|f\|_{L^p(X,\mu)} = \mu(\text{support}(f)).$$

- L^p-norm for $0 < p < 1$ should (maybe) be

$$\|f\|_{L^p(X,\mu)} = \int_X |f(x)|^p \, d\mu(x),$$

  i.e. without the $\frac{1}{p}$ exponent.
- For nonmeasurable functions, maybe we should set

$$\|f\|_{L^p(X,\mu)} = \infty$$

  so that

$$f \in L^p(X,\mu) \Leftrightarrow \|f\|_{L^p(X,\mu)} < \infty.$$