Formalizing a proof of Carleson's theorem

Floris van Doorn

University of Bonn

10 Oct 2024

j.w.w. Lars Becker, Leo Diedering, Asgar Jamneshan, Rajula Srivastava, Jeremy Tan, Christoph Thiele and others

Overview

Overview:

- Some highlights about Lean;
- Carleson's theorem;
- Discussion of the formalization.

History of Lean 4

- **2018** Lean 4 development started while the community was writing `Mathlib` in Lean 3.
- **2021** the first official release of Lean 4.
- **2022-2023** Porting `Mathlib` from Lean 3 to Lean 4 (∼8 months, >1 million lines of code).
- **now** New Lean release every month (currently version is Lean 4.12).
  Lean 3 is fully deprecated and unused.
  There is a lot of activity and Mathlib continues to grow.
  (>1.5 million lines of code, ∼20 commits/day)

Features of Lean 4

- Lean 4 is a fully-fledged programming language.
- Elaboration, printing and tactics are all implemented in Lean itself.
- `Mathlib` uses a lot of custom tactics, notation, pretty printers and linters.
- Lean 4 is very fast
- Features: hygienic macro expansion, high-performance, flexible language server protocol, widgets, code actions, indexing using discrimination trees, incremental compilation, efficient code generator, . . .

Exciting projects in Lean

- ∞-Cosmos project (∞-category theory) (led by Emily Riehl)
- Groupoid model of HoTT (Hazratpour, Awodey, and others)
- Formalization of the polynomial Freiman–Rusza conjecture (led by Terrence Tao)
- Prime Number Theorem+ project (led by Kontorovich and Tao)
- Google Deepmind used Lean to perform well at the international mathematics olympiad.
- Fermat's Last Theorem project (led by Kevin Buzzard)
- Equational theories (led by Terrence Tao)
- Carleson project (led by FvD)

Background: Fourier series

Carleson's theorem is an important theorem in Fourier analysis with a notoriously difficult proof.

It is a theorem about the Fourier transform or Fourier series.

Background: Fourier series

Carleson's theorem is an important theorem in Fourier analysis with a notoriously difficult proof.

It is a theorem about the Fourier transform or Fourier series.

Definition

Let $f : [0, 2\pi] \to \mathbb{C}$ be an integrable function. Then its Fourier coefficients for $n \in \mathbb{Z}$ are defined as

$$\hat{f}_n := \frac{1}{2\pi} \int_{[0,2\pi]} f(x) e^{-inx}\, dx.$$

Its (partial) Fourier series is

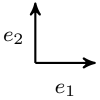
$$S_N f(x) := \sum_{n=-N}^{N} \hat{f}_n e^{inx}.$$

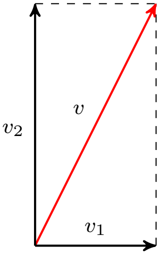
For nice enough functions $\lim_{N\to\infty} S_N f(x) = f(x)$ (e.g. when $f \in C^1$).

Fourier series is splitting a vector into components

More abstractly, the Fourier series is the decomposition of $f \in L^2[0, 2\pi]$ (square integrable functions on $[0, 2\pi]$) into components.

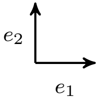
The basis functions we use are the functions $x \mapsto e^{inx}$ which form an orthonormal basis.





Fourier series is splitting a vector into components

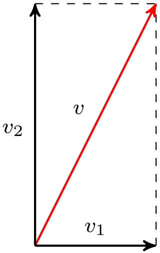
More abstractly, the Fourier series is the decomposition of $f \in L^2[0, 2\pi]$ (square integrable functions on $[0, 2\pi]$) into components.



The basis functions we use are the functions $x \mapsto e^{inx}$ which form an orthonormal basis.

These basis functions are also eigenfunctions of the differentiation operator. In particular we have $\widehat{f'}_n = in\hat{f}_n$.

This means that a differential equation turns into a polynomial equation in the Fourier coefficients.



Background: Fourier transform

We want to also do this for functions with domain $\mathbb{R}$ (i.e. nonperiodic functions).

In this case the situation is a bit more subtle, since the basis elements $x \mapsto e^{inx}$ are *not* square integrable on $\mathbb{R}$.

Background: Fourier transform

We want to also do this for functions with domain $\mathbb{R}$ (i.e. nonperiodic functions).

In this case the situation is a bit more subtle, since the basis elements $x \mapsto e^{inx}$ are *not* square integrable on $\mathbb{R}$.

Definition

Let $f : \mathbb{R} \to \mathbb{C}$ be an integrable function. Then its Fourier transform $\mathcal{F}f : \mathbb{R} \to \mathbb{C}$ is defined as

$$\mathcal{F}f(\xi) := \int_{\mathbb{R}} f(x) e^{-2\pi i \xi x}\, dx.$$

The inverse Fourier transform $\mathcal{F}^{-1}$ is

$$\mathcal{F}^{-1}g(x) := \int_{\mathbb{R}} g(\xi) e^{2\pi i x \xi}\, d\xi.$$

Properties of the Fourier transform

- If f is C^1 and both f and f' are integrable, then:

$$\mathcal{F}(f')(\xi) = 2\pi i\xi\mathcal{F}f(\xi).$$

- If f is C^1 and integrable, then

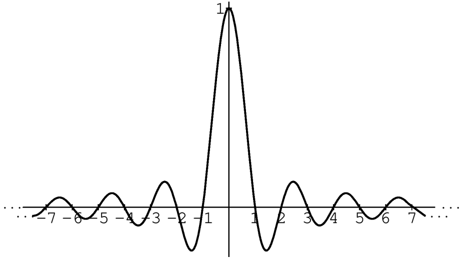
$$\mathcal{F}^{-1}\mathcal{F}f(x) = f(x).$$

Example: Fourier transform of a box

Example: Let $f := \chi_{[-\frac{1}{2},\frac{1}{2}]}$ be a box function. It has fourier transform

$$\mathcal{F}f(\xi) = \frac{\sin(\pi\xi)}{\pi\xi}.$$

Note: $\mathcal{F}f$ is not integrable on $\mathbb{R}$.



Improper integrals

We can define the (inverse) Fourier transform for a wider class of functions by defining it as an improper integral:

$$\mathcal{F}f(\xi) := \lim_{R\to\infty} \int_{[-R,R]} f(x)e^{-2\pi i\xi x}\,dx.$$

Improper integrals

We can define the (inverse) Fourier transform for a wider class of functions by defining it as an improper integral:

$$\mathcal{F}f(\xi) := \lim_{R\to\infty} \int_{[-R,R]} f(x)e^{-2\pi i\xi x}\,dx.$$

Important: Whether this limit converges depends on the topology you use for this limit:

- Pointwise convergence
- L^p-convergence: $\|f\|_{L^p}^p := \int |f(x)|^p dx$.

Fourier inversion with weaker conditions

In what generality does the Fourier inversion theorem hold?

- If $f \in L^2$ then $\mathcal{F}f$ is well-defined using the L^2-norm, and $\mathcal{F}f \in L^2$. In this case, we have $\mathcal{F}^{-1}\mathcal{F}f = f$ w.r.t. to the L^2-norm.

Carleson's theorem

Theorem (Lennart Carleson, 1966)

If $f \in L^2$. Then for almost every x we have $\mathcal{F}^{-1}\mathcal{F}f(x) = f(x)$.

Carleson's theorem

Theorem (Lennart Carleson, 1966)

If $f \in L^2$. Then for almost every x we have $\mathcal{F}^{-1}\mathcal{F}f(x) = f(x)$.

We also have the following generalization:

Theorem (Richard Hunt, 1968)

If $f \in L^p$ for some $1 < p \leq 2$. Then for almost every x we have $\mathcal{F}^{-1}\mathcal{F}f(x) = f(x)$.

These theorems have very hard proofs.

Counterexamples and remarks

- We cannot remove the "almost every" from the statement: even for continuous L^2 functions the limit might diverge for some x.
- There are L^1 functions where the limit defining $\mathcal{F}^{-1}\mathcal{F}f(x)$ diverges for all points x.
- There is a direct analogue of the Carleson–Hunt theorem to Fourier series of periodic functions. For L^1-periodic functions it also fails.
- If f is a function in multiple variables, versions of Carleson's theorem also hold. One has to be very careful about the shape of the integration domain that tends to infinity. If the shape is spherical, then this is still an open problem.

Carleson's operator

The Carleson operator is a sublinear operator that is roughly defined as

$$Tf(x) := \sup_{n \in \mathbb{Z}} \left| \int_{\mathbb{R}} f(y) \frac{1}{x-y} e^{iny} dy \right|$$

for $f : \mathbb{R} \to \mathbb{C}$.

Carleson's operator

The Carleson operator is a sublinear operator that is roughly defined as

$$Tf(x) := \sup_{n\in\mathbb{Z}} \left| \int_{\mathbb{R}} f(y) \frac{1}{x-y} e^{iny} dy \right|$$

for $f : \mathbb{R} \to \mathbb{C}$.

This operator is bounded from L^2 to itself, i.e. $\|Tf\|_{L^2} \leq C\|f\|_{L^2}$ for some constant C.

From this, we can show that Carleson's theorem holds.

Generalizing Carleson's operator

For $f : \mathbb{R} \to \mathbb{C}$

$$Tf(x) := \sup_{n \in \mathbb{Z}} \left| \int_{\mathbb{R}} f(y) \frac{1}{x - y} e^{iny} dy \right| \text{ (roughly).}$$

Generalizing Carleson's operator

For $f : \mathbb{R} \to \mathbb{C}$

$$Tf(x) := \sup_{n \in \mathbb{Z}} \sup_{0 < R_1 < R_2} \left| \int_{R_1 < |x-y| < R_2} f(y) \frac{1}{x-y} e^{iny} dy \right| .$$

Generalizing Carleson's operator

For $f : \mathbb{R} \to \mathbb{C}$

$$Tf(x) := \sup_{n \in \mathbb{Z}} \sup_{0 < R_1 < R_2} \left| \int_{R_1 < |x-y| < R_2} f(y) K(x,y) e^{iny} dy \right| .$$

Here $K : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ is a Calderón-Zygmund kernel.

Generalizing Carleson's operator

For $f : \mathbb{R} \to \mathbb{C}$

$$Tf(x) := \sup_{\theta \in \Theta} \sup_{0 < R_1 < R_2} \left| \int_{R_1 < |x-y| < R_2} f(y) K(x,y) e^{i\theta(y)} dy \right| .$$

Here $K : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ is a Calderón-Zygmund kernel.
Where Θ is a collection of compatible functions $\mathbb{R} \to \mathbb{R}$.

Generalizing Carleson's operator

For $f : X \to \mathbb{C}$

$$Tf(x) := \sup_{\theta \in \Theta} \sup_{0 < R_1 < R_2} \left| \int_{R_1 < d(x,y) < R_2} f(y) K(x,y) e^{i\theta(y)} d\mu(y) \right| .$$

Here $K : X \times X \to \mathbb{C}$ is a Calderón-Zygmund kernel.
Where Θ is a collection of compatible functions $X \to \mathbb{R}$.
X is a doubling metric measure space: a metric space with a Borel measure μ satisfying for some $a \geq 1$

$$\mu(B(x, 2r)) \leq 2^a \mu(B(x, r)) .$$

Generalizing Carleson's operator

For $f : X \to \mathbb{C}$

$$Tf(x) := \sup_{\theta \in \Theta} \sup_{0 < R_1 < R_2} \left| \int_{R_1 < d(x,y) < R_2} f(y) K(x,y) e^{i\theta(y)} d\mu(y) \right|.$$

Here $K : X \times X \to \mathbb{C}$ is a Calderón-Zygmund kernel.
Where Θ is a collection of compatible functions $X \to \mathbb{R}$.
X is a doubling metric measure space: a metric space with a Borel measure μ satisfying for some $a \geq 1$

$$\mu(B(x, 2r)) \leq 2^a \mu(B(x, r)).$$

In this generality, the generalized Carleson operator is not guaranteed to be bounded, it depends on the boundedness of another simpler operator.

Formalization

- I started the project to formalize this result last year with the harmonic analysis group in Bonn.
- There was a paper proof of the result, but with the usual omitted steps that are obvious for experts in the field, but that are hard to reproduce for non-experts.
- The harmonic analysis group then wrote a blueprint for the formalization with a detailed proof.
- This allows non-experts to take a single lemma and formalize it.
- The formalization started in June, and we're half done.
- Website: `https://florisvandoorn.com/carleson/`

Blueprint

The original proof was about 30 pages, but that became 120 pages when writing the proofs out in detail, plus 30 pages to prove classical Carleson's theorem as a corollary.

It has 11 chapters:

- Section 1: statement of the generalized metric Carleson's theorem;
- Section 2: statement of 6 propositions used in the proof;
- Section 3: proof of metric Carleson from the propositions;
- Sections 4-9: each section proves one of the 6 propositions;
- Sections 10-11: proof of the classical Carleson theorem.

Carleson's theorem in Lean

```
theorem classical_carleson {c : ℝ} (hc : c > 0)
    {f : ℝ → ℂ} (h1 : Continuous f) (h2 : f.Periodic c) :
    ∀ᵐ x : ℝ, Tendsto (partialFourierSum f · x) atTop (𝓝 (f x))
```

Carleson's theorem in Lean

```
theorem metric_carleson [CompatibleFunctions ℝ X (2 ^ a)]
    [IsCancellative X (2 ^ a)] [IsOneSidedKernel a K]
    (ha : 4 ≤ a) (hq : q ∈ Ioc 1 2) (hqq' : q.IsConjExponent q')
    (hF : MeasurableSet F) (hG : MeasurableSet G)
    (hT : HasBoundedStrongType (ANCZOperator K) 2 2 volume volume (C_Ts a))
    (f : X → ℂ) (hf : ∀ x, ‖f x‖ ≤ F.indicator 1 x) :
    ∫⁻ x in G, CarlesonOperator K f x ≤
    ENNReal.ofReal (C1_2 a q) * (volume G) ^ q'⁻¹ * (volume F) ^ q⁻¹ := by
  sorry
```

Collaboration with Lean

In June I posted this project in Lean's chat, asking volunteers to help.

A large number of volunteers helped, none of whom has a background in Fourier analysis.

Special thanks to María Inés de Frutos Fernández, Leo Diedering, Pietro Monticone, Jim Portegies, Michael Rothgang, James Sundstrom and especially Jeremy Tan.

Typically I stated the lemmas in Lean, and then contributors formalized the proof, following the blueprint.

Collaboration with Lean

1. ✅ (**Edward van de Meent**) Three simple lemmas about comparing volumes of balls w.r.t. a doubling measure: carleson#volume_ball_le_same, carleson#volume_ball_le_of_dist_le and carleson#volume_ball_le_of_subset. These are not explicitly in the blueprint, but will be needed for Lemma 4.1.1 and Lemma 4.1.2.
2. ✅ (**Jeremy Tan**) Lemma 2.1.1 is a combinatorial lemma. I expect it is easiest to prove carleson#Θ.mk_le_of_le_dist first and then conclude carleson#Θ.card_le_of_le_dist and carleson#Θ.finite_of_le_dist from it.
3. ▶️ (**Ruben van de Velde**) Lemma 2.1.2: two Lean lemmas about doing some approximations in a metric space.
4. ▶️ (**James Sundstrom**) Lemma 2.1.3: three Lean lemmas computing bounds on a binary function. It might be useful to also fill some other sorry's in the file `Psi.lean`.
5. ✅ (**Jeremy Tan**) A simple combinatorial lemma about balls covering other balls: carleson#CoveredByBalls.trans (used in Lemma 2.1.1)
6. 🆕 Lemma 4.0.3: this requires manipulating some integrals, and also require stating/proving that a bunch more things are integrable.
7. 🆕 Proposition 2.0.1: the proof is located above and below Lemma 4.0.3. It is not long, but requires quite some manipulation with integrals.
8. ✅ (**Jeremy Tan**) Show that `𝓓 X` is almost a docs#SuccOrder: Define `𝓓.succ` and prove the 4 lemmas below it.
9. ▶️ (**Bhavik Mehta**) Define `𝓩` by Zorn's lemma and prove the properties about it (up to and including the finiteness and inhabited instances). The finiteness comes from a variant of `Θ.finite_of_le_dist`.
10. ✅ (**Bhavik Mehta**) Prove Lemma 4.2.1 (simple exercise in a metric space)
11. ✅ (**Jeremy Tan**) Prove Lemma 4.2.2 (short, but maybe a bit tricky to use the definitions `Ω₁` and `Ω₁_aux`)
12. ✅ (**Jeremy Tan**) Prove Lemma 4.2.3
13. ✅ (**Jeremy Tan**) Prove Lemma 4.0.2 from the lemmas in Section 4.2 (quite long, but splits naturally into 5 parts).
14. ✅ (**Edward van de Meent**) Prove lemma 4.1.1 and 4.1.2.
15. ✅ (**Jim Portegies**) Proof basic properties about the distribution function carleson#MeasureTheory.distribution up to carleson#ContinuousLinearMap.distribution_le . This requires some simple measure theory. The proofs are in *Folland, Real Analysis. section 6.3.*
16. ✅ (**Mauricio Collares**) (roughly already in Mathlib) Proof a more general layer-cake principle carleson#MeasureTheory.lintegral_norm_pow_eq_measure_lt which generalizes docs#MeasureTheory.lintegral_comp_eq_lintegral_meas_lt_mul . Also please prove the three corollaries below.

Questions

- The work is separated by Harmonic analysists and Lean formalizers. Will this cause friction?
  - So far this is going really well
  - Sometimes it is tricky to formalize a result verbatim.
  - When reformulating results, it can be tricky as a non-expert to know if the result still holds.
- Can we ensure that this material is incorporated into Mathlib?
  - Some proofs are only done for a special case
  - For various preliminary results, we intentionally did them more generally than needed for this proof.
- How do we ensure that contributors get academic credit?

Why formalization?

- Verify correctness: not just of the final result, but the proof of each individual lemma.
- Easier peer-reviews
- Enable large-scale collaborations
- Lean makes it easy to *refactor* proofs.

In conclusion

- Lean is a language with a lot of exciting developments.
- Large formalization projects can be efficiently divided into small parts, with the help of a detailed blueprint.
- It is feasible to formalize hard theorems in harmonic analysis.

Thank you for listening

http://florisvandoorn.com/carleson/