Almost-Everywhere Convergence and the Magic of Maximal Functions

Sam Thiel
Honors Thesis in Mathematics
April 30, 2026

Oberlin College

Central Question

Does a given sequence (of functions) converge?

Central Question

Does a given sequence (of functions) converge? If so:

1. What is that limit?

Central Question

Does a given sequence (of functions) converge? If so:

1. What is that limit?
2. In the case of functions, in what sense? (pointwise, uniform, almost-everywhere, etc.)

Central Question

Does a given sequence (of functions) converge? If so:

1. What is that limit?
2. In the case of functions, in what sense? (pointwise, uniform, almost-everywhere, etc.)

Our goal: given a sequence of functions, can we prove existence of, and identify, their limit in the sense of almost-everywhere convergence.

Central Question

Does a given sequence (of functions) converge? If so:

1. What is that limit?
2. In the case of functions, in what sense? (pointwise, uniform, almost-everywhere, etc.)

Our goal: given a sequence of functions, can we prove existence of, and identify, their limit in the sense of almost-everywhere convergence.

Maximal Functions are a technique for achieving this goal.

Almost-Everywhere Convergence

A sequence of functions f_n converges *pointwise* if there is a function f such that for each fixed x,

$$\lim_{n \to \infty} f_n(x) = f(x)$$

Almost-Everywhere Convergence

A sequence of functions f_n converges *pointwise* if there is a function f such that for each fixed x,

$$\lim_{n \to \infty} f_n(x) = f(x)$$

Definition
A sequence of functions f_n converges *almost-everywhere* if there exists a function f such that $f_n \to f$ pointwise EXCEPT possibly on a set of measure zero.

That is, f_n converges pointwise except on a "negligible" set.

Almost-Everywhere Convergence

A sequence of functions f_n converges *pointwise* if there is a function f such that for each fixed x,

$$\lim_{n \to \infty} f_n(x) = f(x)$$

Definition
A sequence of functions f_n converges *almost-everywhere* if there exists a function f such that $f_n \to f$ pointwise EXCEPT possibly on a set of measure zero.

That is, f_n converges pointwise except on a "negligible" set.
"negligible" $\sim$ "probability 0".

Measures and Integrals

A *measure* μ on Ω assigns to each "reasonable" subset $A \subseteq \Omega$ a nonnegative real number or ∞. We say that a set A has measure zero if $\mu(A) = 0$.

Measures and Integrals

A *measure* μ on Ω assigns to each "reasonable" subset $A \subseteq \Omega$ a nonnegative real number or ∞. We say that a set A has measure zero if $\mu(A) = 0$.

Both discrete probabilities and continuous probabilities with density are examples of measures (where $\mu(\Omega) = 1$).

We can define an integral with respect to any measure. If the measure μ has a density, $m(x)$, then $d\mu = m(x)dx$.

Measures and Integrals

A *measure* μ on Ω assigns to each "reasonable" subset $A \subseteq \Omega$ a nonnegative real number or ∞. We say that a set A has measure zero if $\mu(A) = 0$.

Both discrete probabilities and continuous probabilities with density are examples of measures (where $\mu(\Omega) = 1$).

We can define an integral with respect to any measure. If the measure μ has a density, $m(x)$, then $d\mu = m(x)dx$.

If $\mu(A) = 0$, then

$$\int_{\Omega} f(x) d\mu = \int_{\Omega \setminus A} f(x) d\mu \ .$$

We say that $f = g$ *almost-everywhere* if

$$\int f(x) d\mu = \int g(x) d\mu \ .$$

Example: The Lebesgue Measure

The *Lebesgue measure* (denoted λ) is a measure on $\mathbb{R}$ that expands what the type of sets that are measure zero for dx. Since $\lambda([a, b]) = |b - a|$, if f is Riemann integrable, $d\lambda = dx$.

Recall that if $f = g$, except at finitely many points, then

$$\int f(x)dx = \int g(x)dx$$

However,

Theorem
If $f = g$, except at a countably infinite number of points, then

$$\int f(x)d\lambda = \int g(x)d\lambda$$

For example, $\chi_{\mathbb{Q}} = 0$ almost-everywhere with respect to the Lebesgue measure.

Sequences and Operators

What sequences do we want to find the almost-everywhere (a.e.) limit of?

Sequences and Operators

What sequences do we want to find the almost-everywhere (a.e.) limit of?

We consider sequences of functions generated by a sequence of linear operators T_n. Our central question then becomes about the sequence $T_n f$, given an arbitrary f: does $\lim_{n\to\infty} T_n f$ exist a.e.?

What sequences do we want to find the almost-everywhere (a.e.) limit of?

We consider sequences of functions generated by a sequence of linear operators T_n. Our central question then becomes about the sequence $T_n f$, given an arbitrary f: does $\lim_{n\to\infty} T_n f$ exist a.e.?

This allows us to find limits for a large class of functions all at once. For example, for $f \in \mathcal{C}^1$, we might consider

$$T_n f := \frac{f(x+1/n) - f(x)}{1/n}$$

and ask if $T_n f \to f'$ almost everywhere.

L^p functions

One particularly interesting class of functions for a.e. convergence is the space of L^p functions.

L^p functions

One particularly interesting class of functions for a.e. convergence is the space of L^p functions.

Given a set Ω and a measure μ on Ω, we define for each $1 \leq p < \infty$

$$L^p(\Omega, \mu) := \left\{ f : \Omega \to \mathbb{C} \;\middle|\; \int_\Omega |f(x)|^p d\mu < \infty \right\}$$

with the additional condition that two functions in L^p are the same function if they are equal a.e.; L^p is a space of equivalence classes, much like $\mathbb{Q}$.

L^p-norm and Random Variables

Regardless of the underlying set and measure L^p is a normed vector space. The norm is given by

$$\|f\|_p := \sqrt[p]{\int_\Omega |f(x)|^p d\mu}$$

L^p-norm and Random Variables

Regardless of the underlying set and measure L^p is a normed vector space. The norm is given by

$$\|f\|_p := \sqrt[p]{\int_\Omega |f(x)|^p d\mu}$$

We say that a sequence of functions f_n *converges in L^p-norm* to a function $f \in L^p$ if $\lim_{n\to\infty} \|f_n - f\|_p = 0$

L^p-norm and Random Variables

Regardless of the underlying set and measure L^p is a normed vector space. The norm is given by

$$\|f\|_p := \sqrt[p]{\int_\Omega |f(x)|^p d\mu}$$

We say that a sequence of functions f_n *converges in L^p-norm* to a function $f \in L^p$ if $\lim_{n\to\infty} \|f_n - f\|_p = 0$

Continuous functions (of a variety of flavors) are dense in L^p; in the same way every real number can be approximated by a sequence of rationals, every L^p function can be approximated by a sequence of continuous functions (in the L^p-norm sense).

Relation to Probability

The L^p-norm is highly related to the moments of a random variable. In fact, since every probability density defines a measure, for a random variable $X : \Omega \to [0, \infty)$,

$$\mathbb{E}(X^p) = \|X\|_p^p$$

A random variable having a finite pth moment means that random variable is an L^p function.

Relation to Probability

The L^p-norm is highly related to the moments of a random variable. In fact, since every probability density defines a measure, for a random variable $X : \Omega \to [0, \infty)$,

$$\mathbb{E}(X^p) = \|X\|_p^p$$

A random variable having a finite pth moment means that random variable is an L^p function. L^p functions for all $1 \leq p < \infty$ also satisify the Chebyshev-Markov inequality:

$$\mu(|f| \geq C) \leq \left(\frac{\|f\|_p}{C}\right)^p$$

Where $\mu(|f| \geq C)$ denotes the measure of the set $\{x \mid |f(x)| \geq C\}$. Measures of these sets will be of particular importance to us.

The Full Setup

We have a sequence of functions $T_n f$, where T_n are a sequence of operators mapping from $L^p(\Omega, \mu)$ functions to functions in general. We want to know if

$$\lim_{n \to \infty} T_n f =: T_\infty f$$

exists almost-everywhere with respect to the measure μ. If it does, we want to know what that limit is.

Definition (Maximal Functions)

For each $f \in L^p(\Omega, \mu)$ and $x \in \Omega$, we define

$$T^*f(x) := \sup_n |T_n f(x)| \ .$$

We call T^*f a *maximal function*, and T^* is called a *maximal operator*.

Theorem (The Magic of Maximal Functions)
Let T_n be a sequence of linear operators and let T^* be the associated maximal operator. Suppose that:

(H1) *There exists a dense subset $C \subseteq L^p(\Omega, \mu)$ where for each $f \in C$, $T_n f$ has a measurable limit for a.e. $x \in \Omega$.*

(H2) *T^* satisfies a Maximal Inequality*

Then, we have:

(i) *For every $f \in L^p(\Omega, \mu)$, $T_n f(x)$ has a measurable limit, $T_\infty f(x)$, for a.e. $x \in \Omega$. T_∞ is an operator from $L^p(\Omega, \mu)$ to functions on Ω.* **(Existence)**

(ii) *The operator T_∞ is continuous: that is, if $f_n \to f$ in L^p-norm, then $T_\infty f_n(x) \to T_\infty f(x)$ in measure.* **(Identification)**

Maximal Inequalities

Definition (Maximal Inequalities)
Let T^* be a maximal operator. A *maximal inequality* is a bound of the form

$$\mu(T^* f > C) \leq \left(\frac{K \|f\|_p}{C} \right)^p$$

Maximal Inequalities

Definition (Maximal Inequalities)
Let T^* be a maximal operator. A *maximal inequality* is a bound of the form

$$\mu(T^* f > C) \leq \left(\frac{K \|f\|_p}{C} \right)^p$$

Note that if $K = 1$, then this inequality appears like Chebyshev-Markov for the function $T^* f$. So, if we know that $T^* f \in L^p$ for each $f \in L^p$, we have a maximal inequality.

Maximal Inequalities

Definition (Maximal Inequalities)
Let T^* be a maximal operator. A *maximal inequality* is a bound of the form

$$\mu(T^* f > C) \leq \left(\frac{K \|f\|_p}{C} \right)^p$$

Note that if $K = 1$, then this inequality appears like Chebyshev-Markov for the function $T^* f$. So, if we know that $T^* f \in L^p$ for each $f \in L^p$, we have a maximal inequality.

Unfortunately, this is not often the case. The space of functions that satisfy maximal inequalities is called *weak-L^p*, which is a normed vector space that contains the L^p functions (but is a much larger space overall).

Recap of the Main Framework

The Magic of Maximal Functions requires three things:

- Linear operators T_n
- (H1) $\lim_n T_n f$ exists for $f \in C$, $\overline{C} = L^p$
- (H2) T^* satisfies a Maximal Inequality

Once these three hypotheses are satisfied, the proof is done.

The rest of the talk is dedicated to applications of this proof technique.

Example: Carleson's Theorem

Question: is every periodic function a (perhaps, infinite) sum of sine and cosine functions? Musically, can every possible sound be approximated by single-frequency tones? (Additive Synthesis)

Example: Carleson's Theorem

Question: is every periodic function a (perhaps, infinite) sum of sine and cosine functions? Musically, can every possible sound be approximated by single-frequency tones? (Additive Synthesis)

Given a 2π-periodic function f, we define the *Fourier coefficients* of f by

$$\widehat{f_k} := \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-2\pi ikx}dx$$

for each $k \in \mathbb{Z}$. $e^{i\theta} := \cos(\theta) + i\sin(\theta)$.

Example: Carleson's Theorem

Question: is every periodic function a (perhaps, infinite) sum of sine and cosine functions? Musically, can every possible sound be approximated by single-frequency tones? (Additive Synthesis)

Given a 2π-periodic function f, we define the *Fourier coefficients* of f by

$$\widehat{f_k} := \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-2\pi ikx}\,dx$$

for each $k \in \mathbb{Z}$. $e^{i\theta} := \cos(\theta) + i\sin(\theta)$.

Fourier's idea:

$$f(x) \stackrel{?}{=} \sum_{-\infty}^{\infty} \widehat{f_k} e^{2\pi ikx}$$

Example: Carleson's Theorem

Theorem (Carleson–Hunt)

Let $f \in L^p([0, 2\pi], \lambda)$, $1 < p < \infty$. Then,

$$\sum_{-n}^{n} \widehat{f}_k e^{2\pi ikx} =: S_n[f] \to f \ , \ \textit{Lebesgue a.e.}$$

The original proof of this theorem was an 100+ page proof.

Example: Carleson's Theorem

Theorem (Carleson–Hunt)

Let $f \in L^p([0, 2\pi], \lambda)$, $1 < p < \infty$. *Then,*

$$\sum_{-n}^{n} \widehat{f}_k e^{2\pi ikx} =: S_n[f] \to f \text{ , } \textit{Lebesgue a.e.}$$

The original proof of this theorem was an 100+ page proof.

S_n are linear, $\mathcal{C}^2$ functions are dense in L^p.

Example: Carleson's Theorem

Theorem (Carleson–Hunt)

Let $f \in L^p([0, 2\pi], \lambda)$, $1 < p < \infty$. Then,

$$\sum_{-n}^{n} \widehat{f}_k e^{2\pi ikx} =: S_n[f] \to f \ , \ \textit{Lebesgue a.e.}$$

The original proof of this theorem was an 100+ page proof.

S_n are linear, $\mathcal{C}^2$ functions are dense in L^p.

Theorem (L. Carleson, R. Hunt, 1967)
We define the Carleson maximal operator by

$$M_{\mathrm{C}} f := \sup_{n \in \mathbb{N}} |S_n[f]| \ .$$

$M_{\mathrm{C}} f$ is in weak-L^p for every $f \in L^p$.

Example 2: Lebesgue Differentiation Theorem

Question: By the Fundamental Theorem of Calculus, continuous functions are approximately equal to their local averages. Can we extend this to other functions?

Theorem (Lebesgue Differentiation Theorem)

Let $f \in L^1(\mathbb{R}, \lambda)$. Then, for Lebesgue a.e. $x \in \mathbb{R}$,

$$\lim_{\epsilon \to 0^+} \frac{1}{2\epsilon} \int_{x-\epsilon}^{x+\epsilon} f(y) d\lambda(y) = f(x) \ .$$

Example 2: Lebesgue Differentiation Theorem

Question: By the Fundamental Theorem of Calculus, continuous functions are approximately equal to their local averages. Can we extend this to other functions?

Theorem (Lebesgue Differentiation Theorem)

Let $f \in L^1(\mathbb{R}, \lambda)$. Then, for Lebesgue a.e. $x \in \mathbb{R}$,

$$\lim_{\epsilon \to 0^+} \frac{1}{2\epsilon} \int_{x-\epsilon}^{x+\epsilon} f(y) d\lambda(y) = f(x) .$$

Define

$$T_\epsilon f := \frac{1}{2\epsilon} \int_{x-\epsilon}^{x+\epsilon} f(y) d\lambda(y)$$

T_ϵ are linear, continuous, compactly supported functions are dense in L^p.

Example 2: Lebesgue Differentiation Theorem

Definition (Hardy–Littlewood Maximal Function)
For each $f \in L^1(\mathbb{R}, \lambda)$ define

$$M_{\mathrm{HL}} f := \sup_{\epsilon > 0} \frac{1}{2\epsilon} \int_{x-\epsilon}^{x+\epsilon} |f(y)| d\lambda(y) \ .$$

Theorem (Hardy-Littlewood Maximal Inequality)

For each $C > 0$, $f \in L^1$,

$$\lambda(M_{\mathrm{HL}} f(x) > C) \leq \frac{3 \|f\|_1}{C} \ .$$

Example 3: The Strong Law of Large Numbers

Question: How can we approximate the probability of a given event (e.g. a coin flip)?

Example 3: The Strong Law of Large Numbers

Question: How can we approximate the probability of a given event (e.g. a coin flip)?

Theorem (The Strong Law of Large Numbers)

Let $(\Omega, \mathbb{P})$ *be a probability space, and let* f_n *be iidrv's with the additional condition that* $\mathbb{E}(f_n) < \infty$ *for each n. Then,*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} f_j(x) = \mathbb{E}(f_1)$$

with probability 1. That is, a.e. with respect to $\mathbb{P}$*.*

Example 3: The Strong Law of Large Numbers

Question: How can we approximate the probability of a given event (e.g. a coin flip)?

Theorem (The Strong Law of Large Numbers)

Let $(\Omega, \mathbb{P})$ *be a probability space, and let* f_n *be iidrv's with the additional condition that* $\mathbb{E}(f_n) < \infty$ *for each* n. *Then,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} f_j(x) = \mathbb{E}(f_1)$$

with probability 1. That is, a.e. with respect to $\mathbb{P}$.

A typical first proof of the Strong Law of Large Numbers—based on Chebyshev's inequality—requires that the random variables have fourth moments (that is, $f_n \in L^4$). Here we only require $f_n \in L^1$.

Example 3: The Strong Law of Large Numbers

Let $f = (f_1, f_2, \dots)$ and define

$$Av_n[f] := \frac{1}{n} \sum_{1}^{n} f_j$$

Example 3: The Strong Law of Large Numbers

Let $f = (f_1, f_2, \ldots)$ and define

$$Av_n[f] := \frac{1}{n} \sum_{1}^{n} f_j$$

Av_n are linear

Example 3: The Strong Law of Large Numbers

Let $f = (f_1, f_2, \dots)$ and define

$$Av_n[f] := \frac{1}{n}\sum_{1}^{n} f_j$$

Av_n are linear, $\lim_{n\to\infty} Av_n f = \mathbb{E}(f_1)$ for $f \in I \oplus I^{\perp}$ where I is the set of bounded, Bernoulli-shift-invariant functions.

Example 3: The Strong Law of Large Numbers

Let $f = (f_1, f_2, \dots)$ and define

$$Av_n[f] := \frac{1}{n} \sum_1^n f_j$$

Av_n are linear, $\lim_{n \to \infty} Av_n f = \mathbb{E}(f_1)$ for $f \in I \oplus I^\perp$ where I is the set of bounded, Bernoulli-shift-invariant functions.

Define

$$M_{\mathrm{E}} f := \sup_n |Av_n[f]|$$

Theorem

Let $(\Omega, \mathbb{P})$ be a probability space. For any $\alpha > 0$, we have that

$$\mathbb{P}(M_{\mathrm{E}} f > \alpha) \leq \frac{\mathbb{E}(f_1)}{\alpha}.$$

Thank You!