Maximal Functions in Analysis

Robert Fefferman

June 20, 2005
The University of Chicago REU

Scribe: Philip Ascher

Abstract

This will be a self-contained introduction to the theory of maximal functions, which are some of the most important objects in modern harmonic analysis and partial differential equations. We shall consider various generalizations of the Fundamental Theorem of Calculus, and wind up with an elementary introduction to Calderon-Zygmund Theory. One of the basic ingredients of the study of maximal functions is a deep understanding of the geometry of collections of simple sets such as balls or rectangles in Euclidean spaces. We shall investigate this geometry in some detail.

Contents

1 The Fundamental Theorem of Calculus 2
2 Lebesgue Measure 2
3 The Generalized FTC 4
4 The Hardy-Littlewood Maximal Theorem 4
5 The Vitali Covering Lemma 5
6 Proof of The Hardy-Littlewood Theorem 5
7 The Generalized FTC Revisited 6
8 Proof of Hardy-Littlewood when $p > 1$ 8
9 Some Closing Questions 10
1 The Fundamental Theorem of Calculus

Theorem 1.1 (The Fundamental Theorem of Calculus). If \( f \) is continuous then \( (\int_a^x f \, dx)' = f \) and \( \int_a^b F' \, dx = F(b) - F(a) \).

Proof. \[
\lim_{h \to 0} \frac{\int_a^{x+h} f \, dx - \int_a^x f \, dx}{h} = \lim_{h \to 0} \frac{\int_x^{x+h} f \, dx}{h} = f
\]

Later, we’ll generalize the FTC to \( \mathbb{R}^n \) (\( n > 1 \)), and to the Lebesgue Integral.

Definition 1.2. The Dirichlet Function is defined as:

\[
D(x) = \begin{cases} 
1 & x \in \mathbb{Q} \\
0 & x \notin \mathbb{Q} 
\end{cases}
\]

Since it is measurable, the Dirichlet function has the property that

\[
\int_0^1 D \, dx = 1 \cdot \mu(\{D = 1\}) + 0 \cdot \mu(\{D = 0\})
\]

where \( \mu \) is Lebesgue Measure.

2 Lebesgue Measure

Let’s recall some basic facts about Lebesgue Measure on \( \mathbb{R} \). If we have

\[
\bigcup_{k=1}^{\infty} I_k \supseteq S \quad \text{then} \quad \mu(S) \leq \sum_{k=1}^{\infty} \text{length}(I_k)
\]

Definition 2.1. The Lebesgue Measure of a set \( S \) is defined as:

\[
\mu(S) = \inf_{\text{all covers}} \sum_{k=1}^{\infty} \text{length}(I_k)
\]

Remark 2.2.

\[
\mu(\mathbb{Q} \cap [0,1]) < \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon \to 0 \quad \forall \varepsilon > 0
\]
Remark 2.3.

\[ \mu \left( \mathbb{Q} \cap [0,1] \right) = 1 \]

Remark 2.4. For disjoint measurable sets

\[ \mu \left( \bigcup_{k=1}^{\infty} E_k \right) = \sum_{k=1}^{\infty} \mu(E_k) \]

Definition 2.5. A Lebesgue Simple Function \( S \) is a measurable function with finitely many values \( v_1, \ldots, v_N \).

Remark 2.6. For a simple function \( S \)

\[ \int_{0}^{1} S \, dx = \sum_{k=1}^{N} v_k \cdot \mu(\{S = v_k\}) \]

Definition 2.7 (Lebesgue Integral). Let \( f(x) \geq 0 \) be a measurable function on \([0,1]\), then

\[ \int_{0}^{1} f(x) \, dx \overset{\text{def}}{=} \sup_{0 \leq S \leq f} \int_{0}^{1} S \, dx \]

Definition 2.8 (\( L^p \) spaces). For \( 1 \leq p < \infty \) define:

\[ L^1([0,1]) \overset{\text{def}}{=} \left\{ f \text{ measurable} \mid \int_{0}^{1} |f| \, dx < \infty \right\} \]

\[ L^p([0,1]) \overset{\text{def}}{=} \left\{ f \text{ measurable} \mid \int_{0}^{1} |f|^p \, dx < \infty \right\} \]

Recall the Dirichlet Function

\[ D(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases} \]

We have that \( \left( \int_{0}^{x} D \, dx \right)' = D \) almost everywhere (that is, except on a set of measure 0). We can verify this: if \( x \notin \mathbb{Q} \) then \( \left( \int_{0}^{x} D(x) \, dx \right)' = 0 = D(x) = 0 \), so it’s false only when \( x \in \mathbb{Q} \) (which has measure 0).
3 The Generalized FTC

Theorem 3.1 (The Generalized FTC for the Lebesgue Integral). Let \( f \in L^1([0, 1]) \). Then

\[
\left( \int_0^x f \, dx \right)' = f \quad \text{almost everywhere}
\]

Definition 3.2. If \( f \in L^1([0, 1] \times [0, 1]) \) then

\[
\sup_{h>0} \frac{1}{2h} \int_{x-h}^{x+h} f \, dx = \sup_{x \in I} \frac{1}{|I|} \int_I f \, dx = Mf(x)
\]

Definition 3.3. If \( f \in L^1(\mathbb{R}^n) \) then the Maximal Function of \( f \) is

\[
M(f)(x) \overset{def}{=} \sup_{x \in B} \frac{1}{\mu(B)} \int_B |f| \, dx
\]

4 The Hardy-Littlewood Maximal Theorem

Theorem 4.1 (Hardy-Littlewood Maximal Theorem). If \( f \in L^p(\mathbb{R}^n) \)

\begin{enumerate}
  \item \( p = 1 \) \quad \( \mu\{M(f)(x) > \alpha\} \leq \frac{C}{\alpha} \|f\|_1 \) \quad (a weak-type on \( L^1 \))
  \item \( p > 1 \) \quad \( \|M(f)\|_p \leq C_{p,n} \|f\|_p \) \quad (bounded on \( L^p \))
\end{enumerate}

The first part of this theorem is related to Chebychev’s Inequality.

Theorem 4.2 (Chebychev’s Inequality). If \( Q \geq 0 \) then

\[
\mu(\{Q > \alpha\}) \leq \frac{1}{\alpha} \int_{\mathbb{R}^n} Q \, dx
\]

Proof. Starting out with a trivial inequality,

\[
\alpha \cdot \chi_{\{Q > \alpha\}}(x) \leq Q(x)
\]

\[
\alpha \int_{\mathbb{R}^n} \chi_{\{Q > \alpha\}} \, dx \leq \int_{\mathbb{R}^n} Q \, dx
\]

\[
\alpha \cdot \mu(\{Q > \alpha\}) \leq \int_{\mathbb{R}^n} Q \, dx
\]

\[
\mu(\{Q > \alpha\}) \leq \frac{1}{\alpha} \int_{\mathbb{R}^n} Q \, dx
\]

\[\square\]
5 The Vitali Covering Lemma

Question: can you get a subset of an arbitrary collection of finitely many balls in $\mathbb{R}^n$ that is disjoint and counts a fixed fraction of their total area?

Lemma 5.1 (The Vitali Covering Lemma). Let $B_1, B_2, \ldots, B_N$ be balls in $\mathbb{R}^n$. Then $\exists$ a sub-collection of balls, $\tilde{B}_1, \tilde{B}_2, \ldots, \tilde{B}_M$, such that:

- $\tilde{B}_k$ are pairwise disjoint
- $\mu\left(\bigcup_{k=1}^{M} \tilde{B}_k\right) \geq \frac{1}{3^n} \cdot \mu\left(\bigcup_{k=1}^{N} B_k\right)$

Proof. First, arrange the balls by decreasing radius. Choose $B_1$, so that $\tilde{B}_1 = B_1$. Choose $\tilde{B}_2 = B_2$ if and only if it is disjoint from $\tilde{B}_1 = B_1$. Continue this - so given a ball $B_k$, choose it for the sub-collection if and only if it is disjoint from all previously chosen balls.

Claim:

$\bigcup B_k \subseteq \bigcup (\tilde{B}_j)_3$

Where the notation $(\tilde{B}_j)_3$ means a concentric enlargement of the ball by a factor of 3.

Pf of Claim: Take one $B_k$ and ask: why is $B_k \subseteq \bigcup (\tilde{B}_j)_3$? This is easy to see by drawing a picture, and keeping in mind the way we chose the $\{\tilde{B}_j\}$ above. So:

$\mu\left(\bigcup B_k\right) \leq \mu\left(\bigcup (\tilde{B}_j)_3\right) \leq \sum \mu\left((\tilde{B}_j)_3\right)
= \sum 3^n \cdot \mu(\tilde{B}_j) = 3^n \cdot \mu\left(\bigcup \tilde{B}_j\right)$

since the $\{\tilde{B}_j\}$ were chosen to be disjoint. \hfill $\Box$

6 Proof of The Hardy-Littlewood Theorem

Theorem 6.1 (Restating Thm 4.1). If $f \in L^p(\mathbb{R}^n)$

- $p = 1 \quad \mu\{M(f)(x) > \alpha\} \leq \frac{C}{\alpha} ||f||_1$ (a weak-type on $L^1$)
- $p > 1 \quad ||M(f)||_p \leq C_{p,n} ||f||_p$ (bounded on $L^p$)
Proof. For \( p = \infty \), it’s easy to check that \( \|M(f)\|_\infty \leq 1 \cdot \|f\|_\infty \). Now, we’ll use the Vitali Covering Lemma: \( \forall x \in E_\alpha = \{ Mf > \alpha \} \), choose \( B_x \) such that

\[
\frac{1}{\mu(B_x)} \int_{B_x} |f| \, dx > \alpha
\]

and \( \{ B_x \}_{x \in E_\alpha} \) covers \( E_\alpha \). Without loss of generality, \( \exists B_1, B_2, \ldots, B_k \), a sequence of disjoint open balls from the \( B_x \)’s such that

\[
\bigcup_k B_k \supset \bigcup_{x \in E_\alpha} B_x \supset E_\alpha
\]

So we’re going to show

\[
\mu \left( \{ M(f)(x) > \alpha \} \right) \leq \mu \left( \bigcup_{k=1}^N B_k \right) \leq \frac{C}{\alpha} \|f\|_1 \quad \forall N \) (\( C \) independent of \( N \))

The proof is now fairly straightforward:

\[
\mu \left( \bigcup_{k=1}^N B_k \right) = \sum_{k=1}^N \mu(B_k) < \sum_{k=1}^N \frac{1}{\alpha} \int_{B_k} |f| \, dx = \frac{1}{\alpha} \int_{\bigcup B_k} |f| \, dx \leq \frac{1}{\alpha} \|f\|_1
\]

The second step in this string of inequalities comes from the following observation: look where

\[
\frac{1}{\mu(B)} \int_B |f| \, dx > \alpha \quad \Rightarrow \quad \mu(B_k) < \frac{1}{\alpha} \int_{B_k} |f| \, dx \leq \frac{1}{\alpha} \|f\|_1
\]

This completes the proof of the Hardy-Littlewood Theorem for \( p = 1 \) (see section 8 for the other part). \( \square \)

7 The Generalized FTC Revisited

**Theorem 7.1 (Generalized FTC in \( \mathbb{R}^n \)).** If \( f \in L^1(\mathbb{R}^n) \) then

\[
\lim_{r \to 0} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} f \, dx = f \quad \text{outside a set of measure 0}
\]
Proof. Given any $f \in L^1(\mathbb{R}^n)$, $\exists$ a ‘beautiful’ function $C \in L^1(\mathbb{R}^n)$ such that $\int_{\mathbb{R}^n} |f - C| \, dx$ is as small as you wish. Define:

$$A_r(f)(x) = \frac{1}{\mu(B_r(x))} \int_{B_r(x)} f \, dx$$

We want to prove that

$$\lim_{r \to 0} A_r(f)(x) \longrightarrow f(x) \quad \text{outside a set of measure } 0$$

Note that $A_r$ is additive:

$$A_r(f_1 + f_2)(x) = A_r(f_1)(x) + A_r(f_2)(x)$$

We have $f \in L^1(\mathbb{R}^n)$. Let $f = C + \eta$ (C is the ‘beautiful’ function above) where $\eta = error = f - C$ and $\int_{\mathbb{R}^n} |\eta| \, dx < \varepsilon$. Since $A_r$ is additive:

$$A_r(f)(x) = A_r(C)(x) + A_r(\eta)(x)$$

Subtracting equations we get:

$$A_r(f)(x) - f(x) = \underbrace{[A_r(C)(x) - C(x)]}_{\text{this} \to 0} + \underbrace{[A_r(\eta)(x) - \eta(x)]}_{\text{this} \to 0}$$

Now we play some tricks with limsups:

$$\limsup_{r \to 0} |A_r(f)(x) - f(x)|$$

$$= \lim_{k \to 0} \sup_{0 < r < k} |A_r(f)(x) - f(x)|$$

$$\leq \limsup_{r \to 0} |A_r(C)(x) - C(x)| + \limsup_{r \to 0} |A_r(\eta)(x) - \eta(x)|$$

$$\leq \sup_{r > 0} |A_r(\eta)(x)| + |\eta(x)|$$

Remember that we want to prove that

$$\lim_{r \to 0} A_r(f)(x) \longrightarrow f(x) \quad \text{outside a set of measure } 0$$
so let's look at the following set:

\[
\mu \left\{ \limsup_{r \to 0} |A_r(f)(x) - f(x)| > \varepsilon \right\}
\]

\[
\leq \mu \left\{ M(\eta) > \frac{\varepsilon}{2} \right\} + \mu \left\{ |\eta(x)| > \frac{\varepsilon}{2} \right\}
\]

\[
\leq \frac{C}{\varepsilon} \|\eta\|_1 + \mu \left\{ |\eta(x)| > \frac{\varepsilon}{2} \right\} \quad \text{(by H-L)}
\]

\[
\leq \frac{2C'}{\varepsilon} \|\eta\|_1 + \frac{2}{\varepsilon} \|\eta\|_1 \quad \text{(by Chebychev)}
\]

\[
\]
Now $f \in L^p(\mathbb{R}^n)$, and define $f = f^{\alpha} + f_{\alpha}$ where
\[
f^{\alpha}(x) = \begin{cases} f(x) & \text{where } |f(x)| \geq \frac{\alpha}{2} \\ 0 & \text{where } |f(x)| < \frac{\alpha}{2} \end{cases}
\]
and
\[
f_{\alpha}(x) = \begin{cases} 0 & \text{where } |f(x)| \geq \frac{\alpha}{2} \\ f(x) & \text{where } |f(x)| < \frac{\alpha}{2} \end{cases}
\]
Since the maximum operator is a supremum of averages of values over balls, it is sub-additive:
\[
M(f) = M(f^{\alpha} + f_{\alpha}) \leq M(f^{\alpha}) + M(f_{\alpha})
\]
Now note that $f^{\alpha} \in L^1(\mathbb{R}^n)$ and $f_{\alpha} \in L^\infty(\mathbb{R}^n)$. We’re going to make some observations about the sets, and then look at their measures:
\[
\{M(f) > \alpha\} \subseteq \left\{M(f^{\alpha}) > \frac{\alpha}{2}\right\} \cup \left\{M(f_{\alpha}) > \frac{\alpha}{2}\right\}
\]
\[
\mu \{M(f) > \alpha\} \leq \mu \left\{M(f^{\alpha}) > \frac{\alpha}{2}\right\} + \mu \left\{M(f_{\alpha}) > \frac{\alpha}{2}\right\}
\]
\[
\leq \frac{C}{(\frac{\alpha}{2})} \|f^{\alpha}\|_1 \quad \text{(by part a) of H-L)}
\]
Now we’re going to use the distribution function of $Mf$:
\[
\lambda_{Mf}(\alpha) \leq \frac{2C}{\alpha} \int_{\mathbb{R}^n} |f^{\alpha}(x)| \, dx = \frac{2C}{\alpha} \int_{\{|f(x)| > \frac{\alpha}{2}\}} |f(x)| \, dx
\]
Using the same logic as before (and Fubini again):

\[
\int_{\mathbb{R}^n} |Mf(x)|^p \, dx = (\|Mf\|_p)^p
\]

\[
= p \int_0^{\infty} \alpha^{p-1} \lambda_{Mf}(\alpha) \, d\alpha
\]

\[
\leq p \int_0^{\infty} \frac{2C}{\alpha} \int_{\{|f(x)| > \frac{\alpha}{2}\}} |f(x)| \, dx \, \alpha^{p-1} \, d\alpha
\]

\[
= 2pC \int_{\mathbb{R}^n} \int_0^{2|f(x)|} \alpha^{p-2} \, d\alpha \cdot |f(x)| \, dx
\]

\[
= 2pC \int_{\mathbb{R}^n} \frac{(2|f(x)|)^{p-1}}{p-1} \cdot |f(x)| \, dx
\]

\[
= \frac{2^p pC}{p-1} \int_{\mathbb{R}^n} |f(x)|^p \, dx
\]

\[
= C_p(\|f\|_p)^p
\]  

\[\square\]

9 Some Closing Questions

Do you think the generalized FTC is still true if we replace \( B_r(x) \) with rectangles, and state the theorem the same way? If not, when is it true (by placing restrictions on \( f \))? When is it true in \( \mathbb{R}^2 \)? In \( \mathbb{R}^n \)?