Lecture 1: One Parameter Maximal Functions and Covering Lemmas

In this first lecture we start studying one of the basic and fundamental operators in harmonic analysis, the Hardy-Littlewood maximal function. We will focus primarily on the standard maximal function, but at the end will make some comments about the dyadic maximal operator and the standard results that one can deduce from the boundedness of the maximal function.

1. Definitions and Basic Properties of the Maximal Function

For a function $f \in L^1_{loc}(\mathbb{R}^n)$ we define the Hardy-Littlewood maximal function in the following manner:

$$f^*(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y)|\, dy.$$

Here the supremum is taken over all cubes Q that have the point x as its center. Based solely on the definition, we can see some elementary properties of the maximal function of f.

Trivially, we have that $f^*(x) \geq 0$. Even more precisely, when x is large we have that

$$f^*(x) \geq \frac{C}{|x|^n} \tag{1.1}$$

where C is an absolute constant. This implies that the maximal function can never be in $L^1(\mathbb{R}^n)$, and so we will need a replacement for this space of functions.

Next, note that the operator $\cdot^*$ is a sub-linear operator. Namely,

$$(f+g)^*(x) \leq f^*(x) + g^*(x) \quad \forall f, g \in L^1_{loc}(\mathbb{R}^n)$$

and

$$(cf)^*(x) = |c|\, f^*(x) \quad \forall c \in \mathbb{R} \quad \forall f \in L^1_{loc}(\mathbb{R}^n).$$

Both these computations follow immediately (and trivially) from the definitions.

Finally, observe that if $f \in L^\infty(\mathbb{R}^n)$ then we have that

$$\|f^*\|_\infty \leq \|f\|_\infty .$$

Our first question to answer is what happens for the other $L^p(\mathbb{R}^n)$ when $1 < p < \infty$. By (1.1) we easily see that $f^*(x) \notin L^1(\mathbb{R}^n)$, and so we also need to address the appropriate replacement for the estimate

$$\|f^*\|_1 \lesssim \|f\|_1 .$$

Exercise 1.1. *Define an alternate maximal function by*

$$f^{**}(x) = \sup_{r>0} \frac{1}{c_n r^n} \int_{B_r(x)} |f(y)|\, dy$$

whre $B_r(x)$ is the ball of radius r centered at the point x. Here c_n is the volume of the unit ball in $\mathbb{R}^n$. Show that the maximal function $f^*(x)$ is pointwise equivalent to the function $f^{**}(x)$, namely there exists constants $c_1(n)$ and $c_2(n)$ such that

$$c_2 f^{**}(x) \leq f^*(x) \leq c_1(n) f^{**}(x).$$

2. Marcinkiewicz Interpolation

We now turn to proving a general theorem about interpolation of sub-linear operators. Before, we do this, we recall a fact about how to compute the norm of a function in L^p from its distribution function.

Lemma 2.1. *Let (X,μ) be a measure space. If f is measurable and $0<p<\infty$ then*

$$\int_X |f(x)|^p \, d\mu(x) = \int_0^\infty p\lambda^{p-1}\mu\left(\{x\in X : |f(x)|>\lambda\}\right) d\lambda$$

Proof. Simply observe that

$$\begin{aligned}\int_X |f(x)|^p \, d\mu(x) &= \int_X\int_0^{|f(x)|} p\lambda^{p-1} d\lambda d\mu(x)\\ &= \int_0^\infty p\lambda^{p-1}\int_X \chi_{\{f(x)>\lambda\}}(x) d\mu(x) d\lambda\\ &= \int_0^\infty p\lambda^{p-1}\mu\left(\{x\in X : |f(x)|>\lambda\}\right) d\lambda.\end{aligned}$$

□

Theorem 2.2 (Marcinkiewicz Interpolation)**.** *Suppose that (X,μ) and (Y,ν) are measure spaces. Let $1\le p_0<p_1\le\infty$ and suppose that $T: L^{p_0}(X;\mu)+L^{p_1}(X;\mu)$ is a mapping to the ν-measurable functions on Y such that*

(a) $|T(f+g)(y)|\le |Tf(y)|+|Tg(y)|$;
(b) $\nu\left(\{y\in Y : |Tf(y)|>\lambda\}\right)\le \frac{A_0^{p_0}}{\lambda^{p_0}}\|f\|_{L^{p_0}(X;\mu)}^{p_0}$ *for all* $f\in L^{p_0}(X;\mu)$;
(c) $\nu\left(\{y\in Y : |Tf(y)|>\lambda\}\right)\le \frac{A_1^{p_1}}{\lambda^{p_1}}\|f\|_{L^{p_1}(X;\mu)}^{p_1}$ *for all* $f\in L^{p_1}(X;\mu)$.

Then for any $p_0<p<p_1$ we have that the operator $T: L^p(X;\mu)\to L^p(Y;\nu)$ is bounded by a constant C depending on p_0,p_1,A_0,A_1. Namely, when $p_0<p<p_1$

$$\|Tf\|_{L^p(Y;\nu)}\le C\|f\|_{L^p(X;\mu)}.$$

The condition that T maps into the measurable functions is assumed just so we can define Tf when $f\in L^p(X\ mu)$. The properties in estimates in (b) and (c) are important enough for an operator, that they deserve their own definition going forward.

Definition 2.3. *Let $0<p<\infty$. Then a operator T from $L^p(X;\nu)$ to the measurable functions on Y is said to be weak-type (p,p) if there exists a constant A such that*

$$\nu\left(\{y\in Y : |Tf(y)|>\lambda\}\right)\le \frac{A^p}{\lambda^p}\|f\|_{L^p(X;\mu)}^p$$

for all $\lambda>0$ and all $f\in L^p(X;\mu)$. When this estimate holds, we say that $T: L^p(X;\mu)\to L^{p,\infty}(Y;\nu)$.

Proof. We first decompose and arbitrary $f\in L^p(X;\mu)$ into $f_j\in L^{p_j}(X;\mu)$ for $j=0,1$ for which $f=f_0+f_1$. Let K be a constant to be chosen later, and for a fixed α define

$$f_0(x)=\begin{cases} f(x) & \text{if} \quad |f(x)|>K\alpha\\ 0 & \text{else}\end{cases}$$

and

$$f_1(x) = \begin{cases} f(x) & \text{if} \quad |f(x)| \leq K\alpha \\ 0 & \text{else.} \end{cases}$$

Then we have that $f = f_0 + f_1$. Now observe that $f_j \in L^{p_j}(X;\mu)$ for $j = 0, 1$. Indeed, we have

$$\begin{aligned}
\|f_0\|_{L^{p_0}(X;\mu)}^{p_0} &= \int_X |f_0(x)|^{p_0}\, d\mu(x) \\
&= \int_0^\infty p_0 \lambda^{p_0-1} \mu\left(\{x \in X : |f_0(x)| > \lambda\}\right) d\lambda \\
&= \int_{K\alpha}^\infty p_0 \lambda^{p_0-1} \mu\left(\{x \in X : |f(x)| > \lambda\}\right) d\lambda \\
&= \frac{p_0}{p} \int_{K\alpha}^\infty \lambda^{p_0-p} p \lambda^{p-1} \mu\left(\{x \in X : |f(x)| > \lambda\}\right) d\lambda \\
&\leq \frac{p_0}{p} (K\alpha)^{p_0-p} \int_0^\infty p\lambda^{p-1} \mu\left(\{x \in X : |f(x)| > \lambda\}\right) d\lambda \\
&= \frac{p_0}{p} (K\alpha)^{p_0-p} \|f\|_{L^p(X;\mu)}^p < \infty.
\end{aligned}$$

In this computation, we have used Lemma 2.1 twice, and the fact that $p_0 < p$ so that on $[0, \alpha]$ we have that $\lambda^{p_0-p} \leq (K\alpha)^{p_0-p}$. The conclusion that $f_1 \in L^{p_1}(X;\mu)$ is similar, and one concludes that

$$\|f_1\|_{L^{p_1}(X;\mu)}^{p_1} \leq \frac{p_1}{p} (K\alpha)^{p_1-p} \|f\|_{L^p(X;\mu)}^p.$$

Now for fixed α we have that

$$\{y \in Y\, |Tf(y)| > \alpha\} \subset \left\{y \in Y : |Tf_0(y)| > \frac{\alpha}{2}\right\} \cup \left\{y \in Y : |Tf_1(y)| > \frac{\alpha}{2}\right\}$$

by condition (a) from the hypotheses of the Theorem. Thus, we have that

$$\begin{aligned}
\nu\left(\{y \in Y\, |Tf(y)| > \alpha\}\right) &\leq \nu\left(\left\{y \in Y : |Tf_0(y)| > \frac{\alpha}{2}\right\}\right) + \nu\left(\left\{y \in Y : |Tf_1(y)| > \frac{\alpha}{2}\right\}\right) \\
&\leq \frac{(2A_0)^{p_0}}{\alpha^{p_0}} \|f_0\|_{L^{p_0}(X;\mu)}^{p_0} + \frac{(2A_1)^{p_1}}{\alpha^{p_1}} \|f_1\|_{L^{p_1}(X;\mu)}^{p_1},
\end{aligned}$$

where we have now used conditions (b) and (c) in the hypotheses of the Theorem. We will apply this for each α when computing the norm of Tf by the distribution function.

Now we again use Lemma 2.1 to find that

$$\begin{aligned}
\|Tf\|_{L^p(Y;\nu)}^p &= \int_Y |Tf(y)|^p\, d\nu(y) \\
&= p\int_0^\infty \alpha^{p-1} \nu\left(\{y \in Y\, |Tf(y)| > \alpha\}\right) d\alpha \\
&\leq p\int_0^\infty \alpha^{p-1} \left(\nu\left(\left\{y \in Y : |Tf_0(y)| > \frac{\alpha}{2}\right\}\right) + \nu\left(\left\{y \in Y : |Tf_1(y)| > \frac{\alpha}{2}\right\}\right)\right) d\alpha \\
&\leq p\int_0^\infty \alpha^{p-1} \left(\frac{A_0^{p_0} 2^{p_0}}{\alpha^{p_0}} \|f_0\|_{L^{p_0}(X;\mu)}^{p_0} + \frac{A_1^{p_1} 2^{p_1}}{\alpha^{p_1}} \|f_1\|_{L^{p_1}(X;\mu)}^{p_1}\right) d\alpha \\
&= pA_0^{p_0} 2^{p_0} \int_0^\infty \alpha^{p-p_0-1} \|f_0\|_{L^{p_0}(X;\mu)}^{p_0} d\alpha + pA_1^{p_1} 2^{p_1} \int_0^\infty \alpha^{p-p_1-1} \|f_1\|_{L^{p_1}(X;\mu)}^{p_1} d\alpha.
\end{aligned}$$

Now, for the first integral we have

$$\begin{aligned} p(2A_0)^{p_0}\int_0^\infty \alpha^{p-p_0-1}\left\|f_0\right\|_{L^{p_0}(X;\mu)}^{p_0}d\alpha &= p(2A_0)^{p_0}\int_0^\infty \alpha^{p-p_0-1}\left(\int_{\{|f(x)|>K\alpha\}}|f(x)|^{p_0}\,d\mu(x)\right)d\alpha\\ &= p(2A_0)^{p_0}\int_X |f(x)|^{p_0}\left(\int_0^{\frac{|f(x)|}{K}}\alpha^{p-p_0-1}d\alpha\right)d\mu(x)\\ &= p(2A_0)^{p_0}\frac{1}{K^{p-p_0}(p-p_0)}\int_X |f(x)|^p\,d\mu(x). \end{aligned}$$

Similarly, for the second integral, we have

$$\begin{aligned} p(2A_1)^{p_1}\int_0^\infty \alpha^{p-p_1-1}\left\|f_1\right\|_{L^{p_1}(X;\mu)}^{p_1}d\alpha &= p(2A_1)^{p_1}\int_0^\infty \alpha^{p-p_1-1}\left(\int_{\{|f(x)|\leq K\alpha\}}|f(x)|^{p_1}\,d\mu(x)\right)d\alpha\\ &= p(2A_1)^{p_1}\int_X |f(x)|^{p_1}\left(\int_{\frac{|f(x)|}{K}}^\infty \alpha^{p-p_1-1}d\alpha\right)d\mu(x)\\ &= p(2A_1)^{p_1}\frac{1}{K^{p-p_1}(p_1-p)}\int_X |f(x)|^p\,d\mu(x). \end{aligned}$$

Combining things, we see that

$$\begin{aligned} \left\|Tf\right\|_{L^p(Y;\nu)}^p &\leq \left(p(2A_0)^{p_0}\frac{1}{K^{p-p_0}(p-p_0)}+p(2A_1)^{p_1}\frac{1}{K^{p-p_1}(p_1-p)}\right)\left\|f\right\|_{L^p(X;\mu)}^p\\ &= \frac{p}{K^p}\left(\frac{1}{p-p_0}(2A_0K)^{p_0}+\frac{1}{p_1-p}(2A_1K)^{p_1}\right)\left\|f\right\|_{L^p(X;\mu)}^p. \end{aligned}$$

Now, choose the constant K so that $(2A_0K)^{p_0}=(2A_1K)^{p_1}$, and then one can recognize this last estimate as

$$\left\|Tf\right\|_{L^p(Y;\nu)}^p\leq 2^pp\left(\frac{1}{p-p_0}+\frac{1}{p_1-p}\right)A_0^{1-\theta}A_1^\theta\left\|f\right\|_{L^p(X;\mu)}^p$$

where $0<\theta<1$ and $\frac{1}{p}=\frac{\theta}{p_1}+\frac{1-\theta}{p_0}$. □

One can observe that the estimate for the operator T blows up as both $p\to p_0$ and p_1.

Exercise 2.4. *In the proof we assumed that $p_1<\infty$. Note that in the case when $p_1=\infty$ the proof can be simplified some.*

3. Weak-type Estimates for the Maximal Function

We now turn to proving the following Theorem

Theorem 3.1. *The Hardy-Littlewood maximal function is weak-type $(1,1)$. Namely, for any $f\in L^1(\mathbb{R}^n)$ and for any $\lambda>0$, there exists a constant $C(n)$, depending only on n, such that*

$$|\{x\in\mathbb{R}^n : f^*(x)>\lambda\}|\leq\frac{C(n)}{\lambda}\left\|f\right\|_{L^1(\mathbb{R}^n)}.$$

This Theorem is enough to already deduce an interesting Theorem,

Theorem 3.2. *Let $1<p<\infty$, then*

$$\left\|f^*\right\|_{L^p(\mathbb{R}^n)}\leq C(n,p)\left\|f\right\|_{L^p(\mathbb{R}^n)}.$$

Moreover, $C(n,p)\to 1$ as $p\to\infty$ and $C(n,p)\to\infty$ as $p\to 1$.

Proof. Observe that $f^*(x)$ is sub-linear. By trivial estimates we have

$$\|f^*\|_\infty \leq \|f\|_\infty .$$

Theorem 3.1 implies that $\cdot^*$ is weak-type $(1,1)$. Then apply the Marcikiewicz Interpolation Theorem, Theorem 2.2 to conclude the Theorem. ☐

Before we prove Theorem 3.1 we first have to prove an useful fact in its own right. Namely, we have to prove a covering lemma. It turns out that the behavior of maximal functions are closely connected to covering lemmas, and knowing the appropriate weak-type estimates is equivalent to knowing an appropriate covering lemma.

Lemma 3.3 (Simple Vitali Covering Lemma)**.** *Let $E \subset \mathbb{R}^n$ and let K be a collection of cubes covering E. There there exists a positive constant $C(n)$, and a finite number of cubes $Q_1, \ldots, Q_n$ in K such that*

$$|E| \leq C(n) \sum_{j=1}^{N} |Q_j| .$$

Let's assume the lemma for the moment and show how we can use this to prove Theorem 3.1.

Proof of Theorem 3.1. Let $E = \{f^* > \lambda\}$. For each $x \in E$, by definition of f^* there must exist a cube Q_x such that x is the center and

$$\lambda < \frac{1}{|Q_x|} \int_{Q_x} |f(y)|\, dy.$$

Now we have that $K = \{Q_x : x \in E\}$ where Q_x is the cube selected above is a collection of cubes covering E. Apply Lemma 3.3 to find a finite collection of cubes $Q_1, \ldots, Q_N$ so that

$$\begin{aligned} |E| &\leq C(n) \sum_{j=1}^{N} |Q_j| \\ &\leq \frac{C(n)}{\lambda} \int_{\cup_k Q_k} |f(y)|\, dy \\ &\leq \frac{C(n)}{\lambda} \int_{\mathbb{R}^n} |f(y)|\, dy. \end{aligned}$$

This then proves the desired result. ☐

We now turn to the proof of the Simple Vitali Covering Lemma. The idea behind the proof will be, at each stage, select the largest possible cube, and then remove from the collection the cubes that overlap with this large cube. We now turn to giving a proof of this result.

Proof of Lemma 3.3. Index the size of each cube $Q \in K$ by writing $Q = Q(t)$, where t denotes the corresponding side length. Let $K_1 = K$ and set

$$t_1^* = \sup\{t : Q = Q(t) \in K_1\}.$$

We will assume that case that $t_1^* < \infty$ since if not, the desired conclusion is an easy exercise. Select a cube in K_1 so that $Q_1 = Q(t_1) \in K_1$ and $t_1 > \frac{1}{2} t_1^*$. Divide $K_1 = K_2 \cup K_2'$ where K_2 are the cubes that are disjoint from Q_1 and K_2' are the cubes that intersect Q_1. Let Q_1^* denote the cube concentric with Q_1 by sidelength increased by a factor of 5. Then, it is easy to see that every cube in K_2' is contained in Q_1^* and $|Q_1^*| = 5^n |Q_1|$. This follows from the geometric observation that if two cubes intersect, the an appropriately large dilation of one will completely contain the other.

We continue this selection process for all j. Namely, let

$$t_j^* = \sup\{t : Q = Q(t) \in K_j\}$$

and select a cube $Q_j = Q(t_j)$ with $t_j > \frac{1}{2}t_j^*$ and split $K_j = K_{j+1} \cup K'_{j+1}$, with the set K_{j+1} the collection of cubes in K_j that are disjoint from Q_j, and K'_{j+1} the cubes that intersection Q_j. Let Q_j^* denote the cube concentric with Q_j by sidelength increased by a factor of 5. Then, it is easy to see that every cube in K'_{j+1} is contained in Q_j^* and $\left|Q_j^*\right| = 5^n \left|Q_j\right|$.

If K_{j+1} is empty, then this process stops. Note that the construction produces $t_j^* \geq t_{j+1}^*$. Also, note that by construction, the cubes $Q_1, Q_2, \ldots, Q_j$ are disjoint from each other and from every cube in K_{j+1}.

Suppose now that there is a K_{N+1} that is empty. Then $t_j^* = 0$ for all $j \geq N+1$, and it is easy to see that E is covered by the cubes in $K'_2 \cup \cdots \cup K'_N$, hence covered by the cubes Q_j^*. To see that E is covered by this set, note that E is covered by $K_1 = K$, and that $K_1 = K_2 \cup K'_2$, then note that $K_2 = K_3 \cup K'_3$, and repeat. Thus,

$$|E| \leq \sum_j^N \left|Q_j^*\right| = C(n) \sum_j \left|Q_j\right|,$$

which proves the lemma.

If no t_j^* is zero, then there are two possibilities. First, it could happen that $t_j^* \geq \delta$ for all j and for some $\delta > 0$, or else $t_j^* \to 0$. The case when $t_j^* \geq \delta$ is easy since in this case $t_j > \frac{1}{2}\delta$, so the sides of the selected cubes are always large. Hence, as $N \to \infty$ we have that

$$\sum_{j=1}^N \left|Q_j\right| \to \infty.$$

So, by selecting N sufficiently large, we will have the result for any choice of $C(n)$.

On the other hand, if $t_j^* \to 0$, then simply observe that E is covered by $\{Q_j^*\}$ (this is observed by a simple contradiction argument). Then we have

$$|E| \leq C(n) \sum_j \left|Q_j\right|.$$

□

3.1. **Applications of the Maximal Function.** The standard application of the maximal function is the following Theorem due to Lebesgue.

Theorem 3.4 (Lebesgue's Differentiation Theorem)**.** *If $f \in L^1(\mathbb{R}^n)$ then for almost every $x \in \mathbb{R}^n$ it is true that*

$$\lim_{Q \searrow x} \frac{1}{|Q|} \int_Q f(y)dy = f(x).$$

Exercise 3.5. *Prove this theorem. Hint: It is obviously true if f is continuous. Attempt to approximate f by continuous functions c_k. Study the set*

$$\left\{x \in \mathbb{R}^n : \limsup_{Q \searrow x} \left|\frac{1}{|Q|} \int_Q f(y)dy - f(x)\right| > \epsilon\right\}$$

and use the weak-type $(1,1)$ estimate to help control the size of this set.

4. Dyadic Maximal Functions

Let $\mathcal{D}$ denote the collection of dyadic cubes in $\mathbb{R}^n$. Namely the collection of cubes of the form

$$2^{-k}\prod_{l=1}^{n}[j_l, j_l+1]$$

where $k \in \mathbb{Z}$ and $j \in \mathbb{Z}^n$. We can then form a maximal function with respect to this family of cubes, namely,

$$M_d f(x) = \sup_{Q\in\mathcal{D}} \frac{1}{|Q|}\int_Q |f(y)|\, dy.$$

One can then ask about the behvior of this operator on the spaces $L^p(\mathbb{R}^n)$ when $1 < p < \infty$. It turns out that one can prove the following results, in an analogous manner, to what appeared in this lecture. In particular one can prove that

$$M_d : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n) \quad 1 < p < \infty$$

and

$$\{x \in \mathbb{R}^n : M_d f(x) > \lambda\} \leq \frac{C}{\lambda}\left\Vert f\right\Vert_{L^1(\mathbb{R}^n)}.$$

These proofs are mostly repeats of what appears in this lecture, so we omit them but encourage the interested reader to think more about these facts.

References

The following are excellent books for an introduction to much of the harmonic analysis that we will be learning.

[1] Javier Duoandikoetxea, *Fourier analysis*, Graduate Studies in Mathematics, vol. 29, American Mathematical Society, Providence, RI, 2001. Translated and revised from the 1995 Spanish original by David Cruz-Uribe. ↑

[2] John B. Garnett, *Bounded analytic functions*, 1st ed., Graduate Texts in Mathematics, vol. 236, Springer, New York, 2007. ↑

[3] Loukas Grafakos, *Modern Fourier analysis*, 2nd ed., Graduate Texts in Mathematics, vol. 250, Springer, New York, 2009. ↑

[4] Elias M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy; Monographs in Harmonic Analysis, III. ↑

[5] ———, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970. ↑