The Hardy-Littlewood Maximal Function

Jacob Lawrence

April 2022, Revised: October 2022

1 Introduction

In their 1930 paper "A Maximal Theorem with Function-Theoretic Applications" (reference [1]), G.H. Hardy and J.E. Littlewood consider the following question:

> Suppose that $\lambda > 0$, that
>
> $$f(z) = f\left(re^{i\theta}\right)$$
>
> is an analytic function regular for $r \leq 1$, and that
>
> $$F(\theta) = \max_{0 \leq r \leq 1} \left| f\left(re^{i\theta}\right) \right|$$
>
> is the maximum of $|f|$ on the radius θ. Is it true that
>
> $$\frac{1}{2\pi} \int_{-\pi}^{\pi} F^{\lambda}(\theta) d\theta \leq A(\lambda) \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| f\left(e^{i\theta}\right) \right|^{\lambda} d\theta,$$
>
> where $A(\lambda)$ is a function of λ only?

The actual question itself is not pertinent to the focus of this paper; rather, we will discuss the two following major ideas of Hardy and Littlewood: the Hardy-Littlewood maximal function and the bounds on the norm of this function.

On $\mathbb{R}$, we define four versions of the Hardy-Littlewood maximal function in the following ways. If $f : \mathbb{R} \to \mathbb{R}$ is a locally integrable function, we define the right Hardy-Littlewood maximal function $Mf^+ : \mathbb{R} \to \mathbb{R}$ as

$$Mf^+(x) := \sup_{h>0} \frac{1}{h} \int_{x}^{x+h} |f(t)| dt.$$

Similarly, we define the left Hardy-Littlewood maximal function $Mf^- : \mathbb{R} \to \mathbb{R}$ as

$$Mf^-(x) := \sup_{h>0} \frac{1}{h} \int_{x-h}^{x} |f(t)| dt.$$

The centered Hardy-Littlewood maximal function $M^c f : \mathbb{R} \to \mathbb{R}$ is defined by

$$M^c f(x) \coloneqq \sup_{h>0} \frac{1}{2h} \int_{x-h}^{x+h} |f(t)| dt.$$

Lastly, the uncentered Hardy-Littlewood maximal function $Mf : \mathbb{R} \to \mathbb{R}$ is defined by

$$Mf(x) \coloneqq \sup_{I \ni x} \frac{1}{|I|} \int_I |f(t)| dt,$$

where $|I|$ denotes the Lebesgue measure of the set I. For clarity, $Mf(x)$ does not take the supremum over all x in some interval I; rather, $Mf(x)$ takes the supremum over all intervals that contain x. Notice, each version of the Hardy-Littlewood maximal function is the maximum average value of $f(x)$ at some x over their respective intervals.

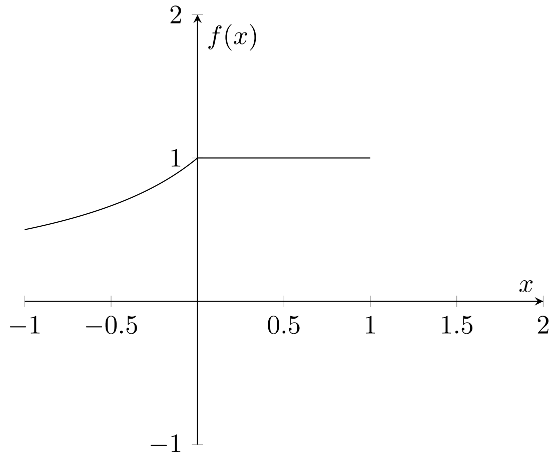
To help illustrate the Hardy-Littlewood maximal function, we turn to the following example.

Example 1.1. Let $f(x) = \chi_{[0,1]}(x)$, the indicator function over $[0,1]$. Note, in order to find any version of the Hardy-Littlewood maximal function, we will choose the interval that contains the greatest proportion of $[0,1]$. Also, note that the average value of $f(x)$ is at most 1.

First we will compute $M^+ f(x)$. For $x \in [0,1)$, the maximum average value of $f(x)$ over the interval $[0,1)$ is 1, and hence, $M^+ f(x) = 1$. For $x \geq 1$, every interval that contains x is of the form $[x, b]$ where $b > x$, and every interval of this form contains no non-zero value of $f(x)$. Hence, $M^+ f(x) = 0$. For $x < 0$, the interval that contains the highest proportion of $[0,1]$ is clearly $[x, 1]$, and so, $M^+ f(x) = 1/(1-x)$. Thus,

$$M^+ f(x) = \begin{cases} \frac{1}{1-x}, & \text{if } x < 0 \\ 1, & \text{if } 0 \leq x < 1 \\ 0, & \text{if } x \geq 1. \end{cases}$$

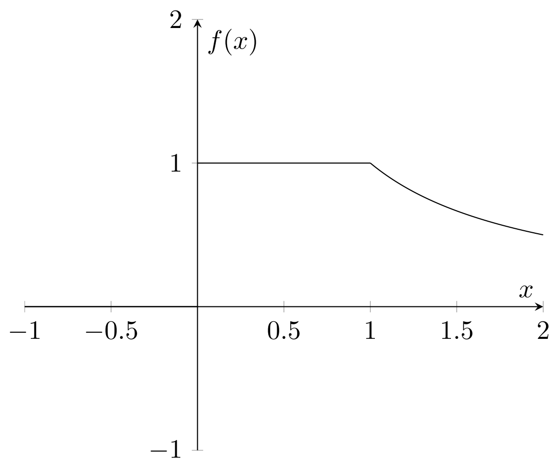
The graph of $M^+ f(x)$ is



Similarly, we compute the left Hardy-Littlewood maximal function as

$$M^{-}f(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } 0 \leq x < 1 \\ \frac{1}{x}, & \text{if } x \geq 1. \end{cases}$$

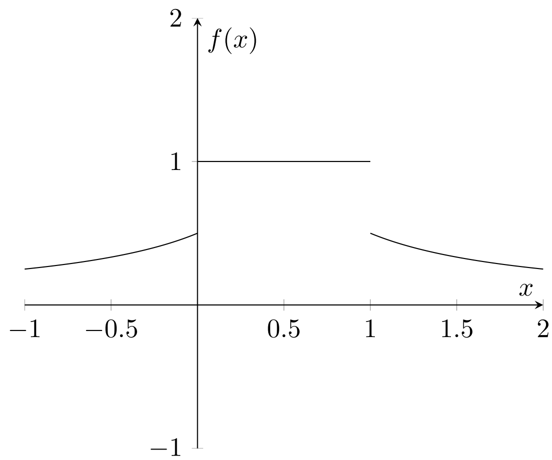
The graph is



Similar to the previous two examples, for $x \in (0,1)$, $M^c f(x) = 1$. For $x < 0$, the interval will be $[2x-1, 1]$ since it contains the highest proportion of $[0,1]$. Similarly, for $x > 1$, the interval will be $[0, 2x]$. Thus, $M^c f(x) = 1/2(1-x)$ and $M^c f(x) = 1/2x$ respectively. Thus,

$$M^c f(x) = \begin{cases} \frac{1}{2(1-x)}, & \text{if } x < 0 \\ 1, & \text{if } 0 \leq x < 1 \\ \frac{1}{2x}, & \text{if } x \geq 1. \end{cases}$$

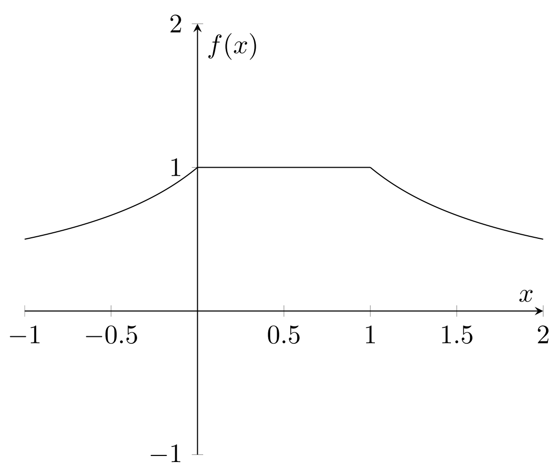
The graph of this is



Lastly, the uncentered Hardy-Littlewood maximal function will be the maximum of these three previous maximum functions above. Thus,

$$M^c f(x) = \begin{cases} \frac{1}{1-x}, & \text{if } x < 0 \\ 1, & \text{if } 0 \le x < 1 \\ \frac{1}{x}, & \text{if } x \ge 1. \end{cases}$$

The graph of this is



There are two things to notice with this example. First, for a given x, $M^+f(x)$, $M^-f(x)$, and $M^c f(x)$ are all less than or equal to $Mf(x)$. Second, for a given x, the uncentered Hardy-Littlewood maximal function is the maximum between $M^+f(x)$ and $M^-f(x)$. Moreover, these facts are true for all locally integrable functions.

Theorem 1.1. *Given any locally integrable function $f : \mathbb{R} \to \mathbb{R}$, then the following are true:*

1. $M^+f(x), M^-f(x), M^cf(x) \leq Mf(x)$, *and*
2. $Mf(x) = \max\{M^f(x), M^-f(x)\}$.

Proof. Fix some $x \in \mathbb{R}$. Let I denote the set of all intervals that contains x. Notice, I has three distinct subsets. The first subset, I_1 is the set of all intervals of the form $[x, b]$ where $b > x$. The second subset, I_2, is the set of all intervals of the form $[a, x]$, where $a < x$. The third subset, I_3 is the set of all intervals of the form $[a, b]$ where $a < x < b$ and for some $h > 0$, $a = x - h$ and $b = x + h$. Note that $M^+f(x)$ takes the supremum over I_1, $M^-f(x)$ takes the supremum over I_2, $M^cf(x)$ takes the supremum over I_3, and $Mf(x)$ is the supremum over I. Since the supremum of a subset is less than or equal to supremum over the entire set, it follows that $M^+f(x), M^-f(x), M^cf(x) \leq Mf(x)$.

Now, we want to show that $Mf(x) = \max\{M^+(x), M^-f(x)\}$. It follows from the previous result that $\max\{M^f(x), M^-f(x)\} \leq Mf(x)$. Consider the interval $[c, d]$ where $c < x < d$, and consider the following

$$\begin{aligned}
\frac{1}{d-c}\int_c^d |f(t)|dt &= \frac{1}{d-c}\int_c^x |f(t)|dt + \frac{1}{d-c}\int_x^d |f(t)|dt \\
&= \frac{x-c}{d-c}\frac{1}{x-c}\int_c^x |f(t)|dt + \frac{d-x}{d-c}\frac{1}{d-x}\int_x^d |f(t)|dt \\
&\leq \frac{x-c}{d-c}M^-f(x) + \frac{d-x}{d-c}M^+f(x)
\end{aligned}$$

Notice,

$$\frac{x-c}{d-c} + \frac{d-x}{d-c} = 1.$$

Hence,

$$\frac{x-c}{d-c}M^-f(x) + \frac{d-x}{d-c}M^+f(x) \leq \max\{M^-f(x), M^+f(x)\}.$$

Taking the supremum over all intervals that contain x,

$$Mf(x) \leq \max\{M^f(x), M^-f(x)\}.$$

Thus, the result follows. $\square$

1.1 Higher Dimensions

The previous definitions of the Hardy-Littlewood maximal function were only defined on $\mathbb{R}$; however, we can extend the definitions of the uncentered and centered Hardy-Littlewood maximal functions to be defined on $\mathbb{R}^n$ for $n \in \mathbb{N}$.

If $f : \mathbb{R}^n \to \mathbb{C}$ is a locally integrable function and $x \in \mathbb{R}^n$, then we define the centered Hardy-Littlewood maximal function over cubes as

$$M^c f(x) = \sup_{r>0} \frac{1}{|Q_r(x)|} \int_{Q_r(x)} |f(t)| dt,$$

where $Q_r(x)$ is the cube with sides parallel to the axes centered at x with sidelength $2r$. We also define the uncentered Hardy-Littlewood maximal function over cubes as

$$Mf(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(t)| dt,$$

where Q is a cube containing x with sides parallel to the axes. Similarly, we define the centered Hardy-Littlewood maximal function (over balls) as

$$\mathcal{M}^c f(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(t)| dt,$$

where $B_r(x)$ is the ball centered at x with radius r. We also define the uncentered Hardy-Littlewood maximal function over balls as

$$\mathcal{M} f = \sup_{B \ni x} \frac{1}{|B|} \int_B |f(t)| dt,$$

where B is a ball containing x. Clearly, these four variations do not necessarily equal; however, we can show they are equivalent up to some constant.

Theorem 1.2. *$M^c f(x)$, $Mf(x)$, $\mathcal{M}^c f(x)$, and $\mathcal{M} f(x)$ are equivalent up to a constant. We will denote this by $\simeq$.*

Proof. For this proof, fix $x \in \mathbb{R}$. To show that two functions, $g(x)$ and $h(x)$ say, are equivalent up to some constant, i.e., $g(x) \simeq h(x)$, we will find constants c and C so that

$$ch(x) \leq g(x) \leq Ch(x).$$

For this proof, we will need the two following facts. First, given any cube Q with sidelength $\ell(Q)$, then $|Q| = \ell(Q)^n$, where n is the dimension of $\mathbb{R}^n$. Second, given any ball B with radius r, then

$$|B| = \frac{\pi^{n/2}}{\Gamma((n/2)+1)} r^n,$$

where Γ is the gamma function. For simplicity, let

$$\omega_n = \frac{\pi^{n/2}}{\Gamma((n/2)+1)}.$$

First, we will show that $Mf(x) \simeq M^c f(x)$. Notice, the set of all cubes containing x contains the set of all cubes with x center. Since $Mf(x)$ takes the supremum over all cubes and $M^c f(x)$ takes the supremum over all cubes with x center, $M^c f(x) \leq Mf(x)$. Now, let Q be a cube that contains x with sidelength $\ell(Q)$. We can construct a new cube, Q' say, such that $Q \subseteq Q'$ and the sidelength of Q' is $2\ell(Q)$. Thus,

$$\frac{1}{|Q|}\int_Q |f(t)|dt \leq \frac{1}{|Q|}\int_{Q'} |f(t)|dt = \frac{2^n}{|Q'|}\int_{Q'} |f(t)|dt \leq 2^n M^c f(x).$$

Taking the supremum overall cubes that contain x, $Mf(x) \leq 2^n M^c f(x)$. Hence, $Mf(x) \simeq M^c f(x)$.

Next, we will show that $\mathcal{M}f(x) \simeq \mathcal{M}^c f(x)$. Notice, the set of all balls containing x contains the set of all balls with x center. Since $\mathcal{M}f(x)$ takes the supremum over all balls while $\mathcal{M}^c f(x)$ takes the supremum over all balls with x center, $\mathcal{M}^c f(x) \leq \mathcal{M}f(x)$. Now, let B be a ball that contains x with radius r. We can construct a new ball, B' say, such that $B \subseteq B'$ and B has radius $2r$. Thus,

$$\frac{1}{|B|}\int_B |f(t)|dt \leq \frac{1}{|B|}\int_{B'} |f(t)|dt = \frac{2^n}{|B'|}\int_{B'} |f(t)|dt \leq 2^n \mathcal{M}^c f(x).$$

Taking the supremum over all balls that contain x, $\mathcal{M}f(x) \leq 2^n \mathcal{M}^c f(x)$. Hence, $\mathcal{M}f(x) \simeq \mathcal{M}^c f(x)$.

Lastly, we will show that $Mf(x) \simeq \mathcal{M}f(x)$. Note that this and the two previous results will imply that $M^c f(x) \simeq \mathcal{M}^c f(x)$. Let Q be a cube that contains x with sidelength $\ell(Q)$. We can circumscribe a ball B around Q that will have radius $r = (\sqrt{n}/2)\ell(Q)$. Clearly, $x \in B$ and $Q \subseteq B$. Then,

$$\begin{aligned}\frac{1}{|Q|}\int_Q |f(t)|dt &\leq \frac{1}{|Q|}\int_B |f(t)|dt \\ &= \frac{\omega_n(\sqrt{n}/2)^n}{|B|}\int_B |f(t)|dt \\ &\leq \omega_n(\sqrt{n}/2)^n \mathcal{M}f(x).\end{aligned}$$

Taking the supremum over all cubes that contain x,

$$Mf(x) \leq \omega_n(\sqrt{n}/2)^n \mathcal{M}f(x).$$

Next, let B be a ball that contains x with radius r. We can circumscribe a cube Q around B that will have sidelength $\ell(Q) = 2r$. Thus,

$$\frac{1}{|B|}\int_B |f(t)|dt \leq \frac{1}{|B|}\int_Q |f(t)|dt = \frac{2^n}{\omega_n|Q|}\int_Q |f(t)|dt \leq \frac{2^n}{\omega_n} Mf(x).$$

Taking the supremum over all balls that contain x, $\mathcal{M}f(x) \leq (2^n/\omega_n)\mathcal{M}f(x)$. Rearranging, $(\omega_n/2^n)\mathcal{M}f(x) \leq Mf(x)$. Thus, $Mf(x) \simeq \mathcal{M}f(x)$. □

Another variation of the Hardy-Littlewood Maximal function we can consider is defined over the dyadic cubes. Recall, $\mathcal{D}^k$ is the set of all cubes with sidelength 2^{-k} and corners in the set

$$\{2^{-k}(m_1, m_2, \ldots, m_n) : m_i \in \mathbb{Z}\}.$$

The dyadic cubes is the union of all $\mathcal{D}^k$ for $k \in \mathbb{Z}$, i.e., $\mathcal{D} = \bigcup_{k\in\mathbb{Z}} \mathcal{D}^k$. Some key properties of the dyadic cubes include

1. Every $Q \in \mathcal{D}^k$ has sidelength 2^{-k},
2. For each k, $\mathcal{D}^k$ partitions $\mathbb{R}^n$, and
3. Given $P, Q \subseteq \mathcal{D}$, $P \subseteq Q$, $Q \subseteq P$, or $P \cap Q = \varnothing$.

Now, if f is a locally integrable function, we define the dyadic Hardy-Littlewood maximal function as

$$M^d f(x) := \sup_{\substack{Q\in\mathcal{D} \\ Q\ni x}} \frac{1}{|Q|} \int_Q |f(t)|dt,$$

which is to say we are calculating the maximum average value of $f(x)$ over the dyadic cubes that contain x. In the same way as Theorem 1.2, it is clear that $M^d f(x)$ is equivalent to the other variations up to some constant.

1.2 Application of the Hardy-Littlewood Maximal Function

We will now move on to a key application of the Hardy-Littlewood maximal function. Recall that the fundamental theorem of calculus states the following:

Theorem 1.3. *If f is continuous on $[a, b]$, and the function $F(x)$ is defined by*

$$F(x) = \int_a^x f(t)dt$$

then $F'(x) = f(x)$.

This theorem is a property of the Riemann integral, so it naturally follows to attempt to find a similar theorem for the Lebesgue integral. First, by the definition of the derivative,

$$\begin{aligned} f(x) = F'(x) &= \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} \\ &= \lim_{h\to 0} \frac{1}{h} \left(\int_a^{x+h} f(t)dt - \int_a^x f(t)dt \right) \\ &= \lim_{h\to 0} \frac{1}{h} \int_x^{x+h} f(t)dt. \end{aligned}$$

Notice, for any given h, we are finding the average value of f on the interval $[x, x+h]$. This is similar to the definition of the right Hardy-Littlewood maximal function. Using this idea, the following theorem is able to accomplish our goal:

Theorem 1.4 (Lebesgue Differentiation Theorem)**.** *If $f : \mathbb{R}^n \to \mathbb{R}^n$ is a locally integrable function, then for almost everywhere $x \in \mathbb{R}^n$,*

$$\lim_{r \to 0^+} \left[\frac{1}{|B_r(x)|} \int_{B_r(x)} f(t)dt \right] = f(x).$$

Moreover, for almost everywhere $x \in \mathbb{R}^n$,

$$\lim_{r \to 0^+} \left[\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(t) - f(x)|dt \right] = 0.$$

Notice, this theorem says that for any given $x \in \mathbb{R}^n$, f(x) is equal to the limit of the average value of f on balls centered at x with a given radius. While not the purpose of this paper, the Lebesgue differentiation theorem utilizes the Hardy-Littlewood maximal function and results from Section 2, and as such, a proof is presented in Section 4.

2 L^p Boundedness

In the original question by Hardy and Littlewood presented at the beginning of this paper, they are concerned with the boundedness of a given Maximal function. From this, we can start to construct a bound on the norm of the Hardy-Littlewood maximal function. One possible option is given any $f \in L^p(\mathbb{R}^n)$,

$$\|Mf\|_p \leq C_p \|f\|_p, \quad \text{for } 1 \leq p \leq \infty,$$

where C_p is a constant determined only be p. This is false for $p = 1$. Consider any $f \in L^1(\mathbb{R}^n)$ and let $c = \int_{Q_R(x)} |f(t)|dt$. Then, for $|x| > R$,

$$Mf(x) \geq \frac{1}{|Q_{R+|x|}(x)|} \int_{Q_R(x)} |f(t)|dt \geq \frac{c}{(R + |x|)^n}.$$

We cannot integrate $c/(R + |x|)^n$ over $\mathbb{R}^n$, and so, $Mf(x)$ cannot be integrated over $\mathbb{R}^n$, i.e., $Mf(x) \notin L^1(\mathbb{R}^n)$. Nevertheless, the bound above holds for $p > 1$, and for $p = 1$, there exists a weak bound of $Mf(x)$. In order to prove these bounds, we need the following lemma:

Lemma 2.1 (Vitali Covering Lemma)**.** *Given $\{Q_1, Q_2, \ldots, Q_n\}$, a set of cubes in $\mathbb{R}$, there exists $S \subseteq \{1, 2, \ldots, n\}$ such that*

1. *$\{Q_i\}_{i \in S}$ is disjoint, and*
2. *$\bigcup_{i=1}^n Q_i \subseteq \bigcup_{i \in S} 3Q_i$.*

Proof. Without loss of generality, we may assume that the collection of cubes is not empty. Let Q_{j_i} be the cube with the largest side length. Inductively assume that $Q_{j_1}, Q_{j_2}, \ldots, Q_{j_k}$ have been chosen. If there is some cube in $Q_1, Q_2, \ldots, Q_n$

that is disjoint from $Q_{j_1} \cup Q_{j_2} \cup \cdots \cup Q_{j_k}$, let $Q_{j_{k+1}}$ be such cube with largest size (if there are multiple of said cubes, choose the cube arbitrarily); otherwise, set $m = k$ and terminate the inductive step.

Now, set $X = \bigcup_{k=1}^{m} 3Q_{j_k}$. It is left to show that $Q_i \subseteq X$ for every $i \in \{1, 2, \ldots, n\}$. This is clear if $i \in \{j_1, j_2, \ldots, j_m\}$. Otherwise, there necessarily is some $k \in \{1, 2, \ldots, m\}$ such that $Q_i \cap Q_{j_k} \neq \varnothing$ and the sidelength of Q_{j_k} is at least as large as that of Q_i. Thus, for every $i \in \{1, 2, \ldots, n\}$, $Q_i \subseteq 3Q_{j_k} \subseteq X$, completing the proof. $\square$

We can now prove that $Mf(x)$ is weakly bounded; this is also know as Weak (1,1) bounded.

Theorem 2.1 (Weak (1,1) Inequality)**.** *For $n \geq 1$, there is a constant $C_1 > 0$ such that for all $\lambda > 0$ and $f \in L^1(\mathbb{R}^n)$, we have*

$$\sup_{\lambda > 0} \lambda |\{x : Mf(x) > \lambda\}| \leq C_1 \|f\|_1.$$

Proof. Fix $\lambda > 0$, where $\{x : Mf(x) > \lambda\} \neq \varnothing$. Let $K \subseteq \{x : Mf(x) > \lambda\}$ be a non-empty compact set. For each $x \in K$, there exists a cube Q_x such that $x \in Q_x$ and

$$\frac{1}{|Q_x|} \int_{Q_x} |f(t)| dt > \lambda.$$

So, $K \subseteq \bigcup_{x \in K} Q_x$. Thus, there exists a set of cubes $\{Q_1, Q_2, \ldots, Q_N\}$ such that $K \subseteq \bigcup_{i=1}^{N} Q_i$, and

$$\frac{1}{|Q_i|} \int_{Q_i} |f(t)| dt > \lambda \quad \forall i \in \{1, 2, \ldots, N\}.$$

Let $S \subseteq \{1, 2, \ldots, N\}$ be the set from Lemma 2.1 so that $K \subseteq \bigcup_{i=1}^{N} Q_i \subseteq \bigcup_{i \in S} 3Q_i$. Then,

$$\begin{aligned} |K| &\leq \sum_{i \in S} |3Q_i| \\ &= \sum_{i \in S} 3^n |Q_i| \\ &< 3^n \sum_{i \in S} \frac{1}{\lambda} \int_{Q_i} |f(t)| dt \\ &= \frac{3^n}{\lambda} \int_{\mathbb{R}^n} |f(t)| dt. \end{aligned}$$

Taking the supremum over all compact sets $K \subseteq \{x : Mf(x) > \lambda\}$ yields

$$|\{x : Mf(x) > \lambda\}| \leq \frac{3^n}{\lambda} \int_{\mathbb{R}^n} |f(t)| dt$$

Thus, the result clearly follows. $\square$

Theorem 2.2 (L^p Inequality)**.** *For $n \geq 1$, $1 < p \leq \infty$, and $f \in L^p(\mathbb{R}^n)$, there exists a constant $C_p > 0$ such that*

$$\|Mf\|_p \leq C_p \|f\|_p.$$

Proof. Suppose $n \geq 1$, $1 < p \leq \infty$, and $f \in L^p(\mathbb{R}^n)$. Let $p' = p/(p-1)$. By definition,

$$\|Mf\|_p^p = \int_{\mathbb{R}^n} |Mf|^p dx,$$

and by the "layer" cake representation,

$$\int_{\mathbb{R}^n} |Mf|^p dx = p \int_0^\infty t^{p-1} |\{x : Mf > t\}| dt.$$

By Theorem 2.1,

$$p \int_0^\infty t^{p-1} |\{x : Mf > t\}| dt \leq 3^n p \int_0^\infty t^{p-2} \int_{\{x : Mf > t\}} |f| dx dt$$

It follows then

$$\begin{aligned} 3^n p \int_0^\infty t^{p-2} \int_{\{x:Mf>t\}} |f| dx dt &= 3^n p \int_{\mathbb{R}^n} |f| \int_0^{Mf(x)} t^{p-2} dt dx \\ &= 3^n p \int_{\mathbb{R}^n} |f| \left(\frac{t^{p-1}}{p-1} \Bigg|_{t=0}^{t=Mf(x)} \right) dx \\ &= 3^n \frac{p}{p-1} \int_{\mathbb{R}^n} |f| Mf(x)^{p-1} dx. \end{aligned}$$

Using Hölder's Inequality,

$$\begin{aligned} 3^n \frac{p}{p-1} \int_{\mathbb{R}^n} |f| Mf(x)^{p-1} dx &\leq 3^n p' \|f\|_p \left(\int_{\mathbb{R}^n} \left(Mf(x)^{p-1} \right)^{p'} \right)^{1/p'} \\ &= 3^n p' \|f\|_p \left(\int_{\mathbb{R}^n} Mf(x)^p dx \right)^{1/p'} \\ &= 3^n p' \|f\|_p \|Mf\|_p^{p-1}. \end{aligned}$$

Thus,

$$\|Mf\|_p \leq 3^n p' \|f\|_p.$$

Hence, the result follows. □

3 Best Constants

What follows from Theorems 2.1 and 2.2 is an investigation into sharp constants for these inequalities. For this section, we will denote $\|M\|_{L^{1,\infty}(\mathbb{R}^n)}$ to be the sharp constant for Theorem 2.1 and $\|M\|_{L^p(\mathbb{R}^n)}$ be the sharp constant for Theorem 2.2.

3.1 Sharp constants for the Weak (1,1) Inequality

3.1.1 Uncentered Maximal Function on $\mathbb{R}$

In order to determine what $\|M\|_{L^{1,\infty}(\mathbb{R})}$ is, we will need the following lemma:

Lemma 3.1. *Given a finite collection of intervals $\{I_1, I_2, \ldots, I_N\}$, there exists two subcollections $\{J_1, J_2, \ldots, J_m\}$ and $\{L_1, L_2, \ldots, L_n\}$ such that*

1. *Each subcollection is pairwise disjoint, and*
2. *The union of all elements in $\{J_1, J_2, \ldots, J_m\}$ with all the elements of $\{L_1, L_2, \ldots, L_n\}$ will equal to the union of the elements in $\{I_1, I_2, \ldots, I_N\}$, i.e.,*

$$\bigcup_{i=1}^{N} I_i = \left(\bigcup_{j=1}^{m} J_i \right) \cup \left(\bigcup_{k=1}^{n} L_k \right).$$

This proof is an adaption of Garnett's proof found in [2].

Proof. By induction, $\{I_1, I_2, \ldots, I_N\}$ can be replaced by a subfamily of intervals such that no interval I_j is contained in the union of the others and such that the refined family has the same union as the original family. Write each I_j in the refined family as the open interval (α_j, β_j) and index them so that

$$\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n.$$

Notice, $\beta_{j+1} > \beta_j$ since otherwise $I_{j+1} \subseteq I_j$, and $\alpha_{j+1} > \beta_{j-1}$ since otherwise $I_j \subseteq I_{j-1} \cup I_{j+1}$. This implies that the even-indexed intervals and and the odd- indexed intervals form two subcollections of $\{I_1, I_2, \ldots, I_N\}$ such that each subcollection is pairwise disjoint. Thus, the result follows. $\square$

Now, we can prove the following:

Theorem 3.1. *If $f \in L^1(\mathbb{R})$ and $\lambda > 0$, then*

$$|\{x : Mf(x) > \lambda\}| \leq \frac{2}{\lambda} \int_{\mathbb{R}} |f(t)| dt.$$

Hence, $\|M\|_{L^{1,\infty}(\mathbb{R})} = 2$.

Proof. Suppose $f \in L^1(\mathbb{R})$ and fix $\lambda > 0$. Let $K \subseteq \{x : Mf(x) > \lambda\}$ be a compact set. Notice, $\{x : Mf(x) > \lambda\} \subseteq \mathbb{R}$, then we can cover this set by intervals such that

$$\{x : Mf(x) > \lambda\} \subseteq \bigcup_{i=1}^{N} I_i.$$

Then,

$$K \subseteq \bigcup_{i=1}^{N} I_i,$$

where

$$\frac{1}{|I_i|}\int_{I_i}|f(t)|dt > \lambda.$$

By Lemma 3.1, we can write

$$\bigcup_{i=1}^{N} I_i = \left(\bigcup_{j=1}^{m} J_j\right) \cup \left(\bigcup_{k=1}^{n} L_k\right),$$

where J_j and L_k come from the subcollections found in Lemma 3.1. Then,

$$\begin{aligned}
|K| &\leq \sum_{j=1}^{m}|J_j| + \sum_{k=1}^{n}|L_k| \\
&\leq \frac{1}{\lambda}\left(\sum_{j=1}^{m}\int_{J_j}|f(t)|dt + \sum_{k=1}^{n}\int_{L_k}|f(t)|dt\right) \\
&\leq \frac{2}{\lambda}\int_{\mathbb{R}}|f(t)|dt.
\end{aligned}$$

Taking the supremum over all compact sets $K \subseteq \{x : Mf(x) > \lambda\}$ results in

$$|\{x : Mf(x) > \lambda\}| \leq \frac{2}{\lambda}\|f\|_1,$$

i.e., $\|M\|_{L^{1,\infty}(\mathbb{R})} \leq 2$.

Now, to show that this constant is sharp, consider $f(x) = \chi_{[0,1]}(x)$. For $\lambda > 1$, $|\{x : Mf(x) > \lambda\}| = 0$; so, suppose $\lambda < 1$. Notice,

$$\{x : Mf(x) > \lambda\} = (1 - 1/\lambda, 1/\lambda),$$

and hence,

$$|\{x : Mf(x) > \lambda\}| = |(1 - 1/\lambda, 1/\lambda)| = 1/\lambda - (1 - 1\lambda) = 2\lambda - 1.$$

Multiplying my λ,

$$\lambda|\{x : Mf(x) > \lambda\}| = 2 - \lambda.$$

Thus,

$$2 = \sup_{\lambda>0}\lambda|\{x : Mf(x) > \lambda\}| \leq \|Mf\|_{L^{1,\infty}(\mathbb{R})}\|f\|_1 = \|Mf\|_{L^{1,\infty}(\mathbb{R})} \leq 2.$$

Thus, $\|M\|_{L^{1,\infty}(\mathbb{R})} = 2$ is sharp. $\square$

3.1.2 Dyadic on $\mathbb{R}^n$

Next, we can determine $\|M^d\|_{L^{1,\infty}(\mathbb{R}^n)}$. Note that if $f \in L^1(\mathbb{R}^n)$ and $\lambda > 0$ is such that

$$\{x \in \mathbb{R}^n : M^d f(x) > \lambda\} \neq \varnothing,$$

then

$$\{x \in \mathbb{R}^n : M^d f(x) > \lambda\} = \bigcup_{k \in \mathbb{N}} Q_k$$

where $\{Q_k\}$ is a countable collection of disjoint dyadic cubes such that

$$\frac{1}{|Q_k|} \int_{Q_k} |f(t)| dt > \lambda.$$

Theorem 3.2. *If $f \in L^1(\mathbb{R}^n)$ and $\lambda > 0$, then*

$$|\{x : M^d f(x) > \lambda\}| \leq \frac{1}{\lambda} \int_{\{x : M^d f(x) > \lambda\}} |f(t)| dt$$

Hence, $\|M^d\|_{L^{1,\infty}(\mathbb{R}^n)} = 1$.

Proof. Suppose $f \in L^1(\mathbb{R}^n)$ and $\lambda > 0$. Then, using the fact stated above,

$$\begin{aligned}
|\{x : M^d f(x) > \lambda\}| &= \sum_{k \in \mathbb{N}} |Q_k| \\
&\leq \sum_{k \in \mathbb{N}} \frac{1}{\lambda} \int_{Q_k} |f(t)| dt \\
&= \frac{1}{\lambda} \int_{\bigcup_{k \in \mathbb{N}} Q_k} |f(t)| dt \\
&= \frac{1}{\lambda} \int_{\{x : M^d f(x) > \lambda\}} |f(t)| dt.
\end{aligned}$$

It follows that $\|M^d\|_{L^{1,\infty}(\mathbb{R}^n)} \leq 1$.

For $\frac{1}{2} < \lambda < 1$, we have $\{x : M^d(\chi_{[0,1]}(x) > \lambda\} = [0,1]$. Thus,

$$\begin{aligned}
\lambda = \lambda |\{x : M^d(\chi_{[0,1]})(x) > \lambda\}| &\leq \|M^d(\chi_{[0,1]})\|_{L^{1,\infty}(\mathbb{R}^n)} \\
&\leq \|M^d\|_{L^{1,\infty}(\mathbb{R}^n)} \|\chi_{[0,1]}\|_1 \\
&= \|M^d\|_{L^{1,\infty}(\mathbb{R}^n)}.
\end{aligned}$$

Thus, $\|M^d\|_{L^{1,\infty}(\mathbb{R}^n)} = 1$ is sharp. ∎

3.1.3 Centered Maximal Function on $\mathbb{R}$

As a result of Antonios Melas (reference [3]), the following theorem states the sharp constant for the centered Hardy-Littlewood maximal function on $\mathbb{R}$ in the Weak (1,1) Inequality.

Theorem 3.3. *For the centered Hardy-Littlewood maximal function $M^c f(x)$ on the real line, the best constant in the weak (1,1) inequality is*

$$\|M^c\|_{L^{1,\infty}(\mathbb{R})} = \frac{11 + \sqrt{61}}{12} \approx 1.5675,$$

which is the largest root of $12x^2 - 22x + 5$.

3.2 Best Constants in the L^p Inequality

For simplicity, we will simplify $\{x : f(x) > \lambda\}$ to $\{f > \lambda\}$ and $\{x : Mf(x) > \lambda\}$ to $\{Mf > \lambda\}$.

3.2.1 Dyadic Maximal Function on $\mathbb{R}^n$

Theorem 3.4. *If $f \in L^p(\mathbb{R}^n)$ for $1 < p \leq \infty$, then*

$$\|Mf\|_p \leq p'\|f\|_p.$$

Hence, $\|M\|_{L^p(\mathbb{R}^n)} = p'$.

Proof. Suppose $f \in L^p(\mathbb{R}^n)$ for $1 < p \leq \infty$. By definition,

$$\|M^d f\|_p^p = \int_{\mathbb{R}^n} |M^d f(t)|^p dt.$$

By the "layer" cake representation,

$$\int_{\mathbb{R}^n} |M^d f(t)|^p dt = p \int_0^\infty \lambda^{p-1} |\{Mf > \lambda\}| d\lambda.$$

Using Theorem 3.2 and Hölder's Inequality,

$$\begin{aligned} p \int_0^\infty \lambda^{p-1} |\{Mf(t) > \lambda\} d\lambda &\leq p \int_0^\infty \lambda^{p-2} \int_{\{M^d f > \lambda\}} |f(t)| dt d\lambda \\ &= p \int_{\mathbb{R}^n} |f(t)| \int_0^{M^d f(t)} \lambda^{p-2} d\lambda dt \\ &= p' \int_{\mathbb{R}^n} |f(t)| M^d f(t)^{p-1} dt \\ &\leq p' \|f\|_p \|M^d f\|_p^{p-1}. \end{aligned}$$

Thus, $\|Mf\|_p \leq p'\|f\|_p$, and hence, $\|M\|_{L^p(\mathbb{R}^n)} \leq p'$. The proof of the sharpness can be found in [4] and [5]. □

3.2.2 Uncentered Maximal Function on $\mathbb{R}$

In order to determine $\|M\|_{L^p(\mathbb{R})}$, we will need the two following lemmas:

Lemma 3.2. *For a set E with finite measure,*

$$\frac{1}{\lambda} \int_E f(x) dx + |\{x : f(x) > \lambda\}| \leq \frac{1}{\lambda} \int_{\{f > \lambda\}} f(x) dx + |E|.$$

Proof. Suppose $|E| < \infty$. Then,

$$\begin{aligned}\int_E f(x)dx - \lambda|E| &= \int_E (f(x) - \lambda)dx \\ &= \int_{E\cap\{f\leq\lambda\}} (f(x) - \lambda)dx + \int_{E\cap\{f>\lambda\}} (f(x) - \lambda) \\ &\leq \int_{\{f>\lambda\}} (f(x) - \lambda)dx \\ &= \int_{\{f>\lambda\}} f(x)dx - \lambda|\{x : f(x) > \lambda\}|.\end{aligned}$$

Thus,

$$\frac{1}{\lambda}\int_E f(x)dx + |\{x : f(x) > \lambda\}| \leq \frac{1}{\lambda}\int_{\{f>\lambda\}} f(x)dx + |E|. \qquad \square$$

Lemma 3.3. *For $f \geq 0$ and $f \in L^1(\mathbb{R})$, then*

$$|\{x : Mf(x) > \lambda\}| + |\{x : f(x) > \lambda\}| \leq \frac{1}{\lambda}\int_{\{Mf>\lambda\}} f(x)dx + \frac{1}{\lambda}\int_{\{f>\lambda\}} f(x)dx.$$

Proof. Let $K \subseteq \{x : Mf(x) > \lambda\}$ be a compact set. Then $K \subseteq J \cup L$ where

$$|J| \leq \frac{1}{\lambda}\int_J f(t)dt \quad \text{and} \quad |K| \leq \frac{1}{\lambda}\int_L f(t)dt.$$

Thus,

$$\begin{aligned}|J \cup L| + |J \cap L| &= |J| + |L| \\ &\leq \frac{1}{\lambda}\int_J f(t)dt + \frac{1}{\lambda}\int_L f(t)dt \\ &= \frac{1}{\lambda}\int_{L\cup J} f(t)dt + \frac{1}{\lambda}\int_{L\cap J} f(t)dt.\end{aligned}$$

By Lemma 3.2,

$$\begin{aligned}|K| + |\{f > \lambda\}| + |J \cap L| &\leq |J \cup L| + |J \cap L| + |\{f > \lambda\}| \\ &\leq \frac{1}{\lambda}\int_{J\cup L} f(t)dt + \frac{1}{\lambda}\int_{J\cap L} f(t)dt + |\{f > \lambda\}| \\ &\leq \frac{1}{\lambda}\int_{J\cup L} f(t)dt + \frac{1}{\lambda}\int_{\{f>\lambda\}} f(t)dt + |J \cap L|.\end{aligned}$$

Taking the supremum over compact sets K results in

$$|\{x : Mf(x) > \lambda\}| + |\{x : f(x) > \lambda\}| \leq \frac{1}{\lambda}\int_{\{Mf>\lambda\}} f(t)dt + \frac{1}{\lambda}\int_{\{f>\lambda\}} f(t)dt. \qquad \square$$

Now, we move onto the sharp constant $\|M\|_{L^p(\mathbb{R})}$.

Theorem 3.5. *For $p > 1$, $\|M\|_{L^p(\mathbb{R})}$ is the largest root of $(p-1)x^p - px^{p-1} - 1$.*

The following proof is an adaptation of L. Grafakos and S. Montgomery Smith proof found in [6].

Proof. Using Lemma 3.3 and Hölder's inequality,

$$\begin{aligned}
\int_{\mathbb{R}} Mf(t)^p dt + \int_{\mathbb{R}} f(t)^p dt &= p \int_{\mathbb{R}} \lambda^{p-1} \left(|\{Mf > \lambda\}| + |\{f > \lambda\}| \right) d\lambda \\
&\leq p \int_0^\infty \lambda^{p-2} \int_{\{Mf<\lambda\}} f(t) dt d\lambda + p \int_0^\infty \lambda^{p-2} \int_{\{f>\lambda\}} f(t) dt d\lambda \\
&= p' \int_{\mathbb{R}} f(t) Mf(t)^{p-1} dt + p' \int_{\mathbb{R}} f(t)^p dt \\
&\leq p' \|f\|_p \|Mf\|_p^{p-1} + p' \|f\|_p^p.
\end{aligned}$$

Reorganizing,

$$(p-1) \left(\frac{\|Mf\|_p}{\|f\|_p} \right)^p - p \left(\frac{\|Mf\|_p}{\|f\|_p} \right)^{p-1} - 1 \leq 0.$$

Hence, the root of the equation $(p-1)x^p - px^{p-1} - 1 = 0$ is an upper bound to $\|M\|_{L^p(\mathbb{R})}$. We will denote this root by A_p.

In order to show that constant is sharp, we will consider the function $f_\varepsilon(x) = |x|^{-1/p} \min(|x|^{-\varepsilon}, |x|^{\varepsilon})$ for $\varepsilon > 0$. Let γ be the solution to

$$\frac{p}{p-1} \frac{\gamma^{1/p'+1} + 1}{\gamma + 1} = \gamma^{-1/p}.$$

It is easy to verify that this previous equation is a positive solution to $(p-1)x^p - px^{p-1} - 1 = 0$, i.e.,

$$\frac{p}{p-1} \frac{\gamma^{1/p'+1} + 1}{\gamma + 1} = A_p.$$

Now, consider any $0 < x < 1$. Then,

$$\begin{aligned}
Mf_e(x) &\geq \frac{1}{x + \gamma x} \int_{-\gamma x}^{x} |t|^{-1/p+\varepsilon} dt \\
&= \frac{1}{x + \gamma x} \left[\int_{-\gamma x}^{0} (-t)^{-1/p+\varepsilon} dt + \int_0^x t^{-1/p+\varepsilon} \right] \\
&= \frac{1}{x + \gamma x} \left[\frac{-(-t)^{-1/p+\varepsilon+1}}{1/p' + \varepsilon} \Bigg|_{-\gamma x}^{0} + \frac{t^{-1/p+\varepsilon+1}}{1/p' + \varepsilon} \Bigg|_0^x \right] \\
&= \frac{1}{x + \gamma x} \left[\frac{\gamma^{1/p'+\varepsilon} x^{-1/p+\varepsilon+1}}{1/p' + \varepsilon} + \frac{x^{-1/p+\varepsilon+1}}{1/p' + \varepsilon} \right] \\
&= \frac{\gamma^{1/p'+\varepsilon} + 1}{(1/p' + \varepsilon)(1 + \gamma)} x^{-1/p+\varepsilon}.
\end{aligned}$$

Now, since $f_\varepsilon(x)$ is an even function, then $Mf_\varepsilon(x)$ is an even function as well. Thus, the previous inequality holds for $-1 < x < 0$. For $x > 1$, we find through a similar calculation that

$$M_\varepsilon(x) \geq \frac{\gamma^{1/p'-\varepsilon}+1}{(1/p'-\varepsilon)(1+\gamma)} x^{-1/p-\varepsilon}.$$

Since $M_\varepsilon f(x)$ is even, this inequality holds for $x < -1$. Notice, for sufficiently small ε,

$$\frac{\gamma^{1/p'-\varepsilon}+1}{(1/p'-\varepsilon)(1+\gamma)} x^{-1/p-\varepsilon} \geq \frac{\gamma^{1/p'+\varepsilon}+1}{(1/p'+\varepsilon)(1+\gamma)} x^{-1/p-\varepsilon}.$$

Combining these facts,

$$M_\varepsilon(x) \geq \frac{\gamma^{1/p'+\varepsilon}+1}{(1/p'+\varepsilon)(1+\gamma)} f_e(x).$$

Thus,

$$\|Mf_\varepsilon\|_p \geq \frac{\gamma^{1/p'+\varepsilon}+1}{(1/p'+\varepsilon)(1+\gamma)} \|f\|_p.$$

As ε approaches 0, then

$$\frac{\gamma^{1/p'+\varepsilon}+1}{(1/p'+\varepsilon)(1+\gamma)} \to \frac{\gamma^{1/p'}+1}{(1/p')(1+\gamma)} = A_p.$$

Thus, $\|Mf_\varepsilon\|_p \geq A_p\|f_\varepsilon\|_p$. Thus, $\|M\|_{L^p(\mathbb{R})} = A_p$. $\square$

3.3 Higher Dimensions: Dimensional Dependence

Beyond what has already been presented, little is know about the best constant other than a few facts. For $p = 1$, Stein and Strömberg showed in [7] that

$$\|M\|_{L^{1,\infty}(\mathbb{R}^n)} \leq O(n \log n),$$

and

$$\|\mathcal{M}\|_{L^{1,\infty}(\mathbb{R}^n)} \leq O(n),$$

where O denotes big O notation. Aldaz showed in [8] that the sharp constant in the weak (1,1) inequality for the centered maximal function with respect to cubes approaches infinity approaches infinity, i.e.,

$$\|M^c\|_{L^{1,\infty}(\mathbb{R}^n)} \to \infty, \quad \text{as } n \to \infty.$$

For $p > 1$, Grafakos and Montgomery-Smith show in [6] that the uncentered maximal function with respect to balls is equivalent to a power of some constant, i.e.,

$$\|\mathcal{M}\|_{L^p(\mathbb{R}^n)} \equiv c^n.$$

Stein and Strömberg show in [7] that the centered maximal function with respect to balls is independent of dimension, i.e.,

$$\|\mathcal{M}^c\|_{L^p(\mathbb{R}^n)} = B(p),$$

where $B(p)$ is a function dependent only on p. Similarly, Bourgain shows in [9] that the centered maximal function with respect to cubes is independent of dimension for $p > 3/2$.

4 Lebesgue Differentiation Theorem

Now, we move onto the proof of the Lebesgue differentiation theorem. For this proof we will need Chebyshev's inequality.

Theorem 4.1 (Chebyshev's Inequality)**.** *If f is an integrable function and $0 < t < \infty$, then*

$$|\{|f(x)| > t\}| \leq \frac{1}{t}\|f\|_1.$$

Proof. Let $E_t = \{x : |f(x)| > t\}$. Then

$$\int |f(x)|dx \geq \int_{E_t} |f(x)|dx \geq t|E_t|.$$

Thus, the result follows. $\square$

Proof of the Lebesgue Differentiation Theorem. Notice, we only have to prove that

$$\lim_{r \to 0^+} \left[\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(t) - f(x)|dt \right] = 0.$$

since

$$\left| \frac{1}{|B_r(x)|} \int_{B_r(x)} (f(t) - f(x))dt \right| \leq \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(t) - f(x)|.$$

We will start by defining $f^* : \mathbb{R}^n \to [0, \infty]$ by

$$f^*(x) = \limsup_{r \to 0^+} \left[\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(t) - f(x)|dt \right].$$

Thus, we want to show that $f^* = 0$ pointwise almost everywhere.

Consider any compactly supported continuous function $g(x)$. Then, for $\varepsilon > 0$, there exists a $\delta > 0$ such that $|g(x) - g(t)| < \varepsilon$ whenever $|x - t| < \delta$. Hence, for $r < \delta$,

$$\frac{1}{|B_r(x)|} \int_{B_r(x)} |g(t) - g(x)|dt < \varepsilon,$$

i.e., $g^* = 0$. We can use the fact that g has this property to show that any locally integrable function f also has this property. We may assume that $f \in L^1(\mathbb{R}^n)$ since if $f\chi_{E_k} \in L^1(\mathbb{R}^n)$ except on a set of measure zero for each k, then $E = \bigcup_{k\in\mathbb{N}} E_k$ has measure zero, and so, f is a locally integrable $L^1(\mathbb{R}^n)$ function except on E. Notice,

$$|f(t) + g(t) - [f(x) + g(x)]| \leq |f(t) - f(x)| + |g(t) - g(x)|,$$

and hence,

$$(f + g)^* \leq f^* + g^*.$$

So, if $f \in L^1(\mathbb{R}^n)$ and g is a compactly supported continuous function, then

$$(f - g)^* \leq f^* + g^* \quad \text{and} \quad f^* = (f - g + g)^* \leq (f - g)^* + g^* = (f - g)^*$$

Thus, $(f - g)^* = f^*$. Now, we estimate f^* by

$$\begin{aligned} f^*(x) &\leq \sup_{r>0} \left[\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(t) - f(x)| dt \right] \\ &\leq \sup_{r>0} \left[\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(t)| dt \right] + |f(x)| \\ &\leq \mathcal{M}f(x) + |f(x)|. \end{aligned}$$

It follows then,

$$\{f^* > t\} \subseteq \{\mathcal{M}f + |f| > t\} \subseteq \{Mf > t/2\} \cup \{|f| > t/2\}.$$

By results in Section 2,

$$|\{x : \mathcal{M}f(x) > t/2\}| \leq \frac{2 \cdot 3^n}{t} \|f\|_1$$

By Chebyshev's inequality,

$$|\{x : |f(x)| > t/2\}| \leq \frac{2}{t} \|f\|_1.$$

Hence,

$$|\{x : f^*(x) > t\}| \leq \frac{2(3^n + 1)}{t} \|f\|_1.$$

Let $C = 2(3^n + 1)/t$. Finally, suppose $f \in L^1(\mathbb{R}^n)$ and $0 < t < \infty$. Since the space of compactly supported continuous functions is dense, then for $\varepsilon > 0$, there exists a compactly supported continuous function g such that $\|f - g\|_1 < \varepsilon$. Then

$$|\{x : f^*(x) > t\}| = |\{x : (f - g)^*(x) > t\}| \leq C_t \|f - g\|_1 \leq \frac{C\varepsilon}{t}.$$

Since ε is arbitrary,

$$|\{x : f^*(x) > t\}| = 0.$$

Notice,

$$\{x : f^*(x) > 0\} = \bigcup_{k=1}^{\infty} \{x : f^*(x) > 1/k\},$$

and hence,

$$|\{x : f^*(x) > 0\}| = 0,$$

proving the result. $\square$

References

[1] G. H. Hardy and J. E. Littlewood, “A maximal theorem with function-theoretic applications,” *Acta Mathematica*, vol. 54, pp. 81–116, 1930.

[2] J. B. Garnett, *Bounded Analytic Functions*, vol. 236 of *Graduate Texts in Mathematics*. New York: Springer, revised first edition ed., 1981.

[3] A. D. Melas, “The best constant for the centered hardy-littlewood maximal inequality,” *Annals of Mathematics*, vol. 157, no. 2, pp. 647–688, 2003.

[4] A. D. Melas, “The bellman functions of dyadic-like maximal operators and related inequalities,” *Advances in Mathematics*, vol. 192, no. 2, pp. 310–340, 2005.

[5] L. Slavin, A. Stokolos, and V. Vasyunin, “Monge-ampère equations and bellman functions: The dyadic maximal operator,” *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 585–588, 2008.

[6] L. Grafakos and S. Montgomery-Smith, “Best constants for uncentered maximal functions,” *Bulletin of the London Mathematical Society*, vol. 29, no. 1, pp. 60–64, 1997.

[7] E. M. Stein and J. O. Strömberg, “Behavior of maximal functios in rn for large n,” *Arkiv för Matematik*, vol. 21, no. 1-2, pp. 259 – 269, 1983.

[8] J. Aldaz, “The weak type (1,1) bounds for the maximal function associated to cubes grow to infinity with the dimension,” *Annals of Mathematics*, vol. 173, no. 2, pp. 1013–1023, 2011.

[9] J. Bourgain, “On dimension free maximal inequalities for convex symmetric bodies in $\mathbb{R}^n$,” in *Geometrical Aspects of Functional Analysis* (J. Lindenstrauss and V. D. Milman, eds.), (Berlin, Heidelberg), pp. 168–176, Springer Berlin Heidelberg, 1987.