LESSON 25 - STUDY GUIDE

Abstract. This is the second of two lessons focused on the machinery required to prove pointwise almost everywhere convergence of approximate identities. For this lesson we will see the definition of a maximal operator of a family of linear operators and the importance of establishing weak type inequalities for it in order to prove pointwise almost everywhere convergence. We will then present and study the Hardy-Littlewood maximal operator as the first and fundamental such operator, but also as a universal majorant for a large class of maximal operators of approximate identities.

1. Maximal operators, the Hardy-Littlewood maximal function and pointwise almost everywhere convergence of approximate identities.

Study material: Just like in the previous lesson, this lesson continues to be a personal summary and selection of the material on maximal operators and the Hardy-Littlewood maximal function found in Javier Duoandikoetxea's book [1] in chapter **2 The Hardy-Littlewood Maximal Function**, Grafakos' [3] section **2.1 Maximal Functions** and Folland's [2] section **3.4 Differentiation on Euclidean Space**.

For the particular subject of maximal operators and the Hardy-Littlewood maximal function I cannot help strongly recommending Elias Stein's masterpiece [5] that begins exactly with this subject.

After having presented the fundamental definitions and results concerning weak type spaces and inequalities, in the previous lesson, we now can now develop the main tool used to prove pointwise convergence almost everywhere of sequences or families of measurable functions and operators: the maximal operator.

Definition 1.1. *Let $\{T_t\}$ be a family of operators $T_t : L^p(X,\mu) \to L^p(X,\mu)$. The* maximal operator, *associated with this family, denoted by T^*, is defined as*

$$T^*f(x) = \sup_t |T_t f(x)|.$$

The following is the central result that relates all these concepts.

Theorem 1.2. *Let T^* be the maximal operator associated with a family of linear operators $\{T_t\}$ defined on $L^P(X,\mu)$. If, for every $f \in L^p(X,\mu)$, the maximal function T^*f is measurable[1] and T^* is of weak type (p,q) for some $1 \le p, q < \infty$ then the set*

$$F = \{f \in L^p(X,\mu) : \lim_{t \to t_0} T_t f(x) = f(x) \quad \mu\text{-almost everywhere } x \in X\}$$

is closed in $L^p(X,\mu)$.

Proof. Let $f \in L^p(X,\mu)$ be in the closure of F, $f \in \overline{F}$. We will show that

$$\mu\big(\{x \in X : \limsup_{t \to t_0} |T_t f(x) - f(x)| > 0\}\big) = 0,$$

Date: May 25, 2020.

[1]Of course, if the family of operators is a sequence, then the maximal function is always measurable. There might be issues, though, if the family of operators is uncountable and for those cases measurability of the maximal function must be separately verified for each case.

thus concluding that $\lim_{t\to t_0} T_t f(x) = f(x)$ for μ-almost every $x \in X$ and therefore that $f \in F$. As $f \in \overline{F}$ we can pick a sequence $f_n \in F$ such that $f_n \to f$ in the $L^p(X,\mu)$ norm (of course, if F is empty the conclusion of the theorem is trivial). Then

$$\begin{aligned} |T_t f(x) - f(x)| &\leq |T_t f(x) - T_t f_n(x)| + |T_t f_n(x) - f_n(x)| + |f_n(x) - f(x)| \\ &= |T_t(f - f_n)(x)| + |T_t f_n(x) - f_n(x)| + |f_n(x) - f(x)| \end{aligned}$$

so that

$$\begin{aligned} \limsup_{t\to t_0} |T_t f(x) - f(x)| &\leq \limsup_{t\to t_0} |T_t(f - f_n)(x)| + \limsup_{t\to t_0} |T_t f_n(x) - f_n(x)| + |f_n(x) - f(x)| \\ &\leq |T^*(f - f_n)(x)| + \limsup_{t\to t_0} |T_t f_n(x) - f_n(x)| + |f_n(x) - f(x)| \end{aligned}$$

and thus

$$\begin{aligned} \mu\big(\{x \in X : \limsup_{t\to t_0} |T_t f(x) - f(x)| > \varepsilon\}\big) &\leq \mu\big(\{x \in X : |T^*(f - f_n)(x)| > \varepsilon/3\}\big) + \\ &+ \mu\big(\{x \in X : \limsup_{t\to t_0} |T_t f_n(x) - f_n(x)| > \varepsilon/3\}\big) + \mu\big(\{x \in X : |f_n(x) - f(x)| > \varepsilon/3\}\big) \\ &= \mu\big(\{x \in X : |T^*(f - f_n)(x)| > \varepsilon/3\}\big) + \mu\big(\{x \in X : |f_n(x) - f(x)| > \varepsilon/3\}\big), \end{aligned}$$

where we use the fact that the middle term is zero, because for the $f_n \in F$ the pointwise convergence holds except for a set of zero measure. Finally, using the weak type bounds for the maximal operator and for $L^p \subset L^p_w$ functions (the Chebyshev inequality), we obtain

$$\begin{aligned} &\mu\big(\{x \in X : \limsup_{t\to t_0} |T_t f(x) - f(x)| > \varepsilon\}\big) \\ &\quad \leq \mu\big(\{x \in X : |T^*(f - f_n)(x)| > \varepsilon/3\}\big) + \mu\big(\{x \in X : |f_n(x) - f(x)| > \varepsilon/3\}\big) \\ &\quad \leq \left(C\frac{3\|f_n - f\|_{L^p(X,\mu)}}{\varepsilon}\right)^q + \left(\frac{3\|f_n - f\|_{L^p(X,\mu)}}{\varepsilon}\right)^p, \end{aligned}$$

and because $\|f_n - f\|_{L^p(X,\mu)} \to 0$ we conclude that, for every $\varepsilon > 0$,

$$\mu\big(\{x \in X : \limsup_{t\to t_0} |T_t f(x) - f(x)| > \varepsilon\}\big) = 0.$$

Therefore, as

$$\{x \in X : \limsup_{t\to t_0} |T_t f(x) - f(x)| > 0\} = \cup_{k=1}^{\infty} \{x \in X : \limsup_{t\to t_0} |T_t f(x) - f(x)| > 1/k\},$$

we conclude

$$\mu\big(\{x \in X : \limsup_{t\to t_0} |T_t f(x) - f(x)| > 0\}\big) \leq \sum_{k=1}^{\infty} \mu\big(\{x \in X : \limsup_{t\to t_0} |T_t f(x) - f(x)| > 1/k\}\big) = 0.$$

□

A simple variation of the previous proof also yields the following.

Corollary 1.3. *With the same conditions as the previous theorem, the set*

$$F = \{f \in L^p(X,\mu) : \lim_{t\to t_0} T_t f(x) \quad \text{exists } \mu\text{-almost everywhere } x \in X\}$$

is closed in $L^p(X,\mu)$.

Proof. One just repeats the exact same proof as before, but proving instead that

$$\mu\big(\{x \in X : \limsup_{t\to t_0} T_t f(x) - \liminf_{t\to t_0} T_t f(x) > 0\}\big) = 0,$$

if $T_t f$ is real valued. If it is complex valued, one just separates into its real and imaginary parts. □

Now, this theorem clearly shows what a powerful tool the concept of maximal operator is in order to prove pointwise almost everywhere convergence of sequences, or families, of linear operators in L^p spaces. Very often, one already knows that pointwise convergence holds everywhere for some dense set of functions, typically C_c^∞. Then, one "just" needs to establish a weak type estimate for the maximal operator in order to conclude that pointwise convergence holds almost everywhere for all functions in L^p because it is the closure of the dense subset.

Here is an example. We already know that Fourier series converge at every point $t \in \mathbb{T}$ for $f \in C^\infty(\mathbb{T})$. And $C^\infty(\mathbb{T})$ is dense in $L^2(\mathbb{T})$. So, to prove pointwise convergence almost everywhere of Fourier series in $L^2(\mathbb{T})$ what Carleson actually proved in his famous result of 1965 was that the maximal operator of the partial sums

$$S^* f(t) = \sup_{N\geq 1} \left| \sum_{n=-N}^{N} \hat{f}(n) e^{int} \right|,$$

today called the Carleson maximal operator and usually denoted by $\mathcal{C}f$, is weak type $(2,2)$.

The first, and arguably the most important maximal operator is the Hardy-Littlewood maximal operator. It was introduced in 1930 in a paper [4] written jointly by the two famous British mathematicians, titled "A Maximal Theorem With Function-Theoretic Applications", in order to try to relate L^p norms of the supremum at fixed argument over all radii $0 \leq r \leq 1$ of analytic functions on the unit disk D, with the L^p norms of its boundary values at $r = 1$. They start their motivation of the definition of the maximal function with a famous sentence:

> "The problem is most easily grasped when stated in the language of cricket, or any other game in which a player compiles a series of scores of which an average is recorded."

Basically, the Hardy-Littlewood maximal function of a locally integrable function[2] $f \in L^1_{loc}(\mathbb{R}^n)$ is the supremum of the averages of the absolute value of f over balls centered at a point.

Definition 1.4. *Let $f \in L^1_{loc}(\mathbb{R}^n)$. Then, the (centered) Hardy-Littlewood maximal function of f, denoted by Mf, is defined as*

$$Mf(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)| dy, \tag{1.1}$$

where $B_r(x)$ is the Euclidean ball of radius r centered at x and $|B_r(x)|$ its Lebesgue measure, and this definition includes the possibility that $Mf(x) = +\infty$.

Many other variants of the Hardy-Littlewood maximal operator exist: with cubes instead of balls, taking the supremum over all balls that contain x and not just the ones centered at x, etc. But they are all comparable and we will restrict our definition to the case (1.1) above.

If one considers averages of Lebesgue integrable function over balls centered at x

$$A_r f(x) = \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y) dy, \tag{1.2}$$

[2]Recall that a locally integrable function $f \in L^1_{loc}(\mathbb{R}^n)$ is a measurable function $f : \mathbb{R}^n \to \mathbb{C}$ such that $f \in L^1(K)$ for every compact $K \subset \mathbb{R}^n$. Due to the L^p space inclusions for sets of finite measure, it is obvious that, for all $1 \leq p \leq \infty$, $L^p(\mathbb{R}^n) \subset L^1_{loc}(\mathbb{R}^n)$.

and looks at possible pointwise convergence almost everywhere of these averages as $r \to 0$ then, from Theorem 1.2, we are led to the maximal operator for the averages

$$A^* f(x) = \sup_{r>0} \left| \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y)dy \right|. \tag{1.3}$$

So the Hardy-Littlewood maximal function clearly is related to pointwise almost everywhere convergence of averages of Lebesgue integrable function over balls centered at x as the radius r shrinks to zero. What is required, therefore, are weak type estimates for it. But that is the content of the following fundamental theorem of this theory.

Theorem 1.5. *Let $f \in L^1_{loc}(\mathbb{R}^n)$. Then, the Hardy-Littlewood maximal function Mf is measurable and*

(1) *M is weak type $(1,1)$, i.e. if $f \in L^1(\mathbb{R}^n)$ then*
$$\big|\{x \in X : Mf(x) > \alpha\}\big| \le \frac{C_n}{\alpha} \int_{\mathbb{R}^n} |f(x)|dx,$$
where the constant C_n depends only on the dimension.

(2) *For $1 < p \le \infty$, M is strong type (p,p), i.e. if $f \in L^p(\mathbb{R}^n)$, then $\|Mf\|_{L^p(\mathbb{R}^n)} \le C_{n,p}\|f\|_{L^p(\mathbb{R}^n)}$, where the constant $C_{n,p}$ depends only on the dimension and p.*

Proof. It is a simple exercise to show that, for $f \in L^1_{loc}(\mathbb{R}^n)$, the Hardy-Littlewood maximal function Mf is lower semicontinuous so that, for every $\alpha \in \mathbb{R}$ the set $\{x \in X : Mf(x) > \alpha\}$ is open and, therefore, it is measurable.

(1) If $f \in L^1(\mathbb{R}^n)$ let us denote by E_α the set $\{x \in X : Mf(x) > \alpha\}$ and let $x \in E_\alpha$ so that $Mf(x) > \alpha$. Then, there exists a ball $B_r(x)$ centered at x for which
$$\alpha < \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|dy \Rightarrow |B_r(x)| < \frac{1}{\alpha} \int_{B_r(x)} |f(y)|dy,$$
which already looks like a weak type estimate for the ball. The set $E_\alpha = \{x \in X : Mf(x) > \alpha\}$ is then covered by such balls, centered at each of its points. So, if we could extract a countable subcover $\cup_{x_j \in E_\alpha} B_{r_{x_j}}(x_j) \supset E_\alpha$, then
$$\big|\{x \in X : Mf(x) > \alpha\}\big| \le \sum_{j=1}^{\infty} |B_{r_{x_j}}(x_j)| < \frac{1}{\alpha} \sum_{j=1}^{\infty} \int_{B_{r_{x_j}}(x_j)} |f(y)|dy,$$
which is almost what we want, except that the balls overlap each other and therefore the sum of the integrals on the right hand side would not be comparable to $\int_{\mathbb{R}^n} |f(y)|dy$. What we need is a wiser way to extract from the original cover by balls, not a subcover, but a disjoint collection of balls - so that they do not overlap and thus the sum of integrals can be bounded by $\int_{\mathbb{R}^n} |f(y)|dy$ - but whose measures can be compared to the measures of larger balls that would actually cover the whole of E_α. So we appeal to a Vitali-type covering lemma, important in its own right, that we will present after the proof of the current theorem. There are a few variants of this type of covering lemma that can be used, but we are going to apply the formulation in which, for every $c < \big|\{x \in X : Mf(x) > \alpha\}\big|$ we can pick a finite disjoint subcollection of the balls, let us denote them by $B_1, \dots, B_k$ such that $3^n \sum_{j=1}^{j} |B_j| > c$. Therefore, for every such $c < \big|\{x \in X : Mf(x) > \alpha\}\big|$ we have
$$c < 3^n \sum_{j=1}^{j} |B_j| < \frac{3^n}{\alpha} \sum_{j=1}^{j} \int_{B_j} |f(y)|dy \le \frac{3^n}{\alpha} \int_{\mathbb{R}^n} |f(y)|dy.$$

The right hand side of this inequality is independent of the finite collection of balls chosen for each c, and as c can be made as close as desired to $\big|\{x \in X : Mf(x) > \alpha\}\big|$ we conclude that

$$\big|\{x \in X : Mf(x) > \alpha\}\big| \leq \frac{3^n}{\alpha} \int_{\mathbb{R}^n} |f(y)|dy.$$

(2) From part (1), for $p = 1$, and the obvious inequality $\|Mf\|_{L^\infty(\mathbb{R}^n)} \leq \|f\|_{L^\infty(\mathbb{R}^n)}$, for $p = \infty$, we have that the Hardy-Littlewood maximal operator is weak type $(1, 1)$ and (∞, ∞). So we can use the Marcinkiewicz interpolation theorem - the Hardy-Littlewood maximal operator is clearly sublinear - to conclude that it is strong type (p, p) for all $1 < p < \infty$.

□

Let us now prove the Vitali-type covering lemma.

Lemma 1.6. *Let $\Omega \subset \mathbb{R}^n$ be a Lebesgue measurable set covered by a collection of balls. Then, for any $c < |\Omega|$ a finite subcollection $B_1, \dots, B_k$ of these balls can be chosen so that*

$$c < 3^n \sum_{j=1}^{k} |B_j|.$$

Proof. Let us assume that the measure of Ω is positive and that $0 < c < |\Omega|$, otherwise the result is totally obvious. As the Lebesgue measure is regular, we can choose a compact $K \subset \Omega$ such that $c < |K| \leq \Omega$ and, for that compact, pick a finite subcover of the balls, say $A_1, \dots, A_m$ such that $K \subset \cup_{j=1}^m A_j$.

Now, chose B_1 to be the ball, among the $\{A_j\}$, with the largest radius. If there are more than one with the maximum radius, just pick any one of them. Then, all the remaining balls A_j will have radii which are not bigger than B_1's. And any one of these balls that intersects B_1 will be contained in a ball with the same center as B_1 but with three times the radius. Therefore $3^n|B_1|$ is the measure of that ball with radius three times as big as B_1's that contains B_1 and all the remaining A_j that intersect it.

Now, keep B_1 and discard all the A_j that intersect it. Those are covered by $3^n|B_1|$. For the remaining A_j that do not intersect B_1 choose the one with the largest radius, and call it B_2. Again, a ball with the same center and three times its radius will cover B_2 and all the remaining A_j that intersect it because they have smaller or equal radii.

We can repeat this process for a finite number of steps, until we get to the end with a finite subcollection of balls with decreasing radii $B_1, B_2, \dots, B_k$ such that all the previous A_j intersect one of them, and therefore

$$c < |K| \leq 3^n \sum_{j=1}^{k} |B_j|.$$

□

Some observations are in order. The first one is that, from Theorem 1.5 we conclude that, for $f \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$, its Hardy-Littlewood maximal function Mf is finite almost everywhere.

A second observation is that, for $f \in L^1(\mathbb{R}^n)$ we can do no better than the weak type $(1, 1)$ estimate of part (1) of the theorem. In fact, except for $f = 0$, $Mf \notin L^1(\mathbb{R}^n)$ when $f \in L^1(\mathbb{R})$, a simple fact which we will leave as an exercise to be proved.

So, it is now clear that we have all the ingredients to prove a generalized version of the Fundamental Theorem of Calculus, for $\mathbb{R}^n$, where one shows that the averages of Lebesgue integrable functions over balls (1.2) converge pointwise almost everywhere to the functions, when the balls shrink to zero radius. The maximal operator for the averages (1.3) obviously satisfies $A^*f(x) \leq Mf(x)$ and Theorem 1.5 yields the required weak type estimates. We therefore have the important following result as a consequence of our study of the Hardy-Littlewood maximal function.

Theorem 1.7. (Lebesgue Differentiation Theorem) *Let $f \in L^1_{loc}(\mathbb{R}^n)$. Then, for almost every $x \in \mathbb{R}^n$,*

$$\lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y)\, dy = f(x). \tag{1.4}$$

Proof. We just need to prove that the theorem holds for functions in $L^1(\mathbb{R}^n)$ because, given $f \in L^1_{loc}$, then one just cuts off f at arbitrarily large radii, say $\tilde{f} = f\chi_{\{|x|\le N+1\}}$ to be in $L^1(\mathbb{R}^n)$, and the limit $r \to 0$ of the averages does not distinguish between f or $\tilde{f}$ when $r < 1$ for the points $|x| \le N$. Of course the balls centered at the origin with radius N form a countable collection whose union is $\mathbb{R}^n$ so if this theorem holds almost everywhere for every cutoff $f\chi_{\{|x|\le N+1\}}$ when $|x| \le N$, then it also holds almost everywhere for f on $\mathbb{R}^n$.

Part (1) of Theorem 1.5 yields the weak type $(1,1)$ estimate for the Hardy-Littlewood maximal operator, and the inequality $A^*f(x) \le Mf(x)$ implies that it holds for the averages maximal operator too. Finally one just needs to recall that, obviously, (1.4) is true for every $x \in \mathbb{R}^n$ if $f \in C^\infty_c(\mathbb{R}^n)$, for example, which is dense in $L^1(\mathbb{R}^n)$. And from Theorem 1.2 we obtain the pointwise almost everywhere convergence of the averages for any function in $L^1(\mathbb{R}^n)$. □

We can rewrite (1.4) as

$$\lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} \big(f(y) - f(x)\big)\, dy = 0. \tag{1.5}$$

but, actually, a stronger result is true, with absolute values inside the integral.

Definition 1.8. *Let $f \in L^1_{loc}(\mathbb{R}^n)$. Then, a point $x \in \mathbb{R}^n$ is called a Lebesgue point of f if*

$$\lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} \big|f(y) - f(x)\big|\, dy = 0. \tag{1.6}$$

The set of Lebesgue points of f is called the Lebesgue set of f.

Clearly (1.6) is stronger than (1.5) so that whenever x is a Lebesgue point of f, we have (1.4). The interesting fact is that almost every point in $\mathbb{R}^n$ is a Lebesgue point so the stronger (1.6) actually holds almost everywhere as well.

Proposition 1.9. *Let $f \in L^1_{loc}(\mathbb{R}^n)$. Then, the set of points $x \in \mathbb{R}^n$ which are not Lebesgue points of f has measure zero.*

Proof. Let $c \in \mathbb{C}$ be an arbitrary constant. Then, applying Theorem 1.7 to the function $|f(x) - c|$ we conclude that, except for x in a set $E_c \subset \mathbb{R}^n$ of measure zero, we have

$$\lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - c|\, dy = |f(x) - c|.$$

Taking now a countable dense set of points $D \in \mathbb{C}$ we conclude that $E = \cup_{c\in D} E_c$ also has zero measure. So if $x \notin E$, for arbitrary $\varepsilon > 0$ there exists $c \in D$ such that $|f(x) - c| < \varepsilon$ and

$$\limsup_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|\, dy \le \lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - c|\, dy + |c - f(x)| = 2|f(x) - c| \le 2\varepsilon.$$

And as ε is arbitrarily small, this shows that for all $x \notin E$

$$\lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|\, dy = 0,$$

concluding the proof. □

Finally, the Hardy-Littlewood is not only important because it leads to the proof of the fundamental Lebesgue differentiation theorem about almost everywhere pointwise convergence of averages of $L^1_{loc}(\mathbb{R}^n)$ functions, but because it also serves as quite a universal majorant for a very general collection of approximate identities. To motivate it, notice that we can rewrite the averages over balls (1.2) as

$$A_r f(x) = \sup_{r>0} \frac{1}{|B_r(0)|} \int_{B_r(0)} f(x-y)dy,$$

which can therefore be interpreted as convolution operators $\Phi_r * f$ where

$$\Phi_r(x) = \frac{1}{|B_r(0)|} \chi_{B_r(0)}(x),$$

and these are rescaled versions of $\Phi_1(x)$, which is a nonnegative function with $\int \Phi_1 = 1$. So the averages really are a classical example of approximate identity obtained by rescaling, as seen in Lesson 11,

$$A_r f(x) = \Phi_r * f(x) = \frac{1}{r^n} \Phi_1\left(\frac{\cdot}{r}\right) * f(x) \quad \text{as} \quad r \to 0.$$

The same way that the Hardy-Littlewood maximal operator is a majorant for the maximal operator of the approximate identity of the averages over balls, the following result shows that it is also the majorant of maximal operators for a large family of other approximate identities obtained by rescaling.

Theorem 1.10. *Let Φ be a nonnegative, radial, decreasing (in radius) and integrable function. Then*

$$\sup_{t>0} \left| \frac{1}{t^n} \Phi\left(\frac{\cdot}{t}\right) * f(x) \right| \leq \|\Phi\|_{L^1(\mathbb{R}^n)} Mf(x).$$

Proof. Let ϕ be a simple function satisfying the same hypotheses. Then

$$\phi(x) = \sum_j a_j \chi_{B_{r_j}(0)}(x),$$

with $a_j > 0$. Then

$$|\phi * f(x)| \leq \sum_j a_j \chi_{B_{r_j}(0)} * |f|(x) = \sum_j a_j |B_{r_j}(0)| \frac{1}{|B_{r_j}(0)|} \chi_{B_{r_j}(0)} * |f|(x) \leq \|\phi\|_{L^1(\mathbb{R}^n)} Mf(x),$$

where $\|\phi\|_{L^1(\mathbb{R}^n)} = \sum_j a_j |B_{r_j}(0)|$. The rescaled versions $\phi_t = 1/t^n \phi(\cdot/t)$ are also simple functions with the same properties and $L^1(\mathbb{R}^n)$ norm, so the estimate holds for them as well.

An arbitrary function satisfying the hypotheses of the theorem can then be obtained as the limit of an increasing sequence of these simple functions, and the general result therefore follows. □

Even if a function does not satisfy all the hypotheses of the previous theorem, it might nevertheless have a majorant that does. So we still have the final corollary.

Corollary 1.11. *If $\psi : \mathbb{R}^n \to \mathbb{C}$ is a measurable function with a nonnegative, radial, decreasing and integrable majorant, i.e. a function Φ with those properties such that $|\psi(x)| \leq \Phi(x)$ almost everywhere, then*

$$\sup_{t>0} \left| \frac{1}{t^n} \psi\left(\frac{\cdot}{t}\right) * f(x) \right| \leq \|\Phi\|_{L^1(\mathbb{R}^n)} Mf(x),$$

so that the maximal operator $\sup_t |\psi_t * f|$ *is weak type* $(1,1)$ *and strong type* (p,p)*, for* $1 < p \leq \infty$.

And finally.

Corollary 1.12. *Let ϕ be a positive measurable function defined on $\mathbb{R}^n$ such that $\int \phi = 1$ and with a radial, decreasing and integrable majorant. Then, denoting by $\phi_t(x) = 1/t^n \phi(x/t)$ its rescalings, we have, for $f \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$,*

$$\lim_{t \to 0} \phi_t * f(x) = f(x) \quad \textit{almost everywhere } x \in \mathbb{R}^n.$$

Proof. That ϕ_t is an approximate identity was seen in Lesson 11. Which implies that for $f \in C_c^\infty(\mathbb{R}^n)$ we have $\phi_t * f(x) \to f(x)$ uniformly. As $C_c^\infty(\mathbb{R}^n)$ is dense in all of the $L^p(\mathbb{R}^n)$, for $1 \leq p < \infty$, the result follows from the weak type estimates of the maximal operator associated to this approximate identity, from the previous corollary. □

References

[1] Javier Duoandikoetxea, *Fourier Analysis,* American Mathematical Society, Graduate Studies in Mathematics 29, 2001.
[2] Gerald B. Folland, *Real Analysis, Modern Techniques and Applications,* 2nd Edition, John Wiley & Sons, 1999.
[3] Loukas Grafakos, *Classical Fourier Analysis,* 3rd Edition, Springer, Graduate Texts in Mathematics 249, 2014.
[4] G. H. Hardy & J. E. Littlewood, *A Maximal Theorem With Function-Theoretic Applications,* Acta Mathematica, volume 54, 81–116, 1930.
[5] Elias Stein, *Singular Integrals and Differentiabilty Properties of Functions,* Princeton University Press, 1970.