

§3: Maximal functions

3.1: The Hardy-Littlewood maximal function

We denote by B_r the Euclidean ball $B(0, r)$ centered at the origin with radius $r > 0$.

Given $f \in L^1_{loc}(\mathbb{R}^n)$, we define its Hardy-Littlewood maximal function Mf by:

$$Mf(x) := \sup_{r>0} \frac{1}{|B_r|} \int_{B_r} |f(x-y)| dy. \quad \text{Note that } M \text{ is sublinear.}$$

We can write this as:

$$Mf(x) = \sup_{r>0} \int_{B_r(x)} |f(y)| dy.$$

Here $\int_Q (\dots) dy$ denotes the average over Q .

• Some related notions are given as follows:

$$M'f(x) := \sup_{r>0} \frac{1}{|Q_r|} \int_{Q_r} |f(x-y)| dy$$

where $Q_r = [-r, r]^n$ is a cube.

We note that when $n=1$, M and M' coincide.

More generally, one can consider:

$$M''f(x) := \sup_{\substack{Q \ni x \\ Q \text{ is a cube}}} \frac{1}{|Q|} \int_Q |f(y)| dy$$

It can be shown that $Mf(x) \sim M'f(x) \sim M''f(x)$.

• A related concept can be associated to a family of linear operators.

Namely, let $\{T_t\}$ be a family of linear operators on $L^p(X, \mu)$

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

Then, T^* is called the maximal operator associated with the family $\{T_t\}$. Note that T^* is sublinear.

• These notions can be related as follows:

$$\text{Let } \varphi := \frac{1}{|B_1|} \chi_{B_1}$$

$$\varphi_r := \frac{1}{r^m} \varphi\left(\frac{\cdot}{r}\right) = \frac{1}{|B_r|} \chi_{B_r} \Rightarrow \frac{1}{|B_r|} \int |f(x-y)| dy = \varphi_r * |f|(x)$$

$$\text{Then } Mf(x) = \sup_{r>0} \varphi_r * |f|$$

Hence, for $f \geq 0$, $f \in L^1_{loc}(\mathbb{R}^m)$, the Hardy-Littlewood maximal operator coincides with the maximal operator associated to the approximation of the identity corresponding to φ .

A fundamental result is the following:

Theorem 3.1: The Hardy-Littlewood maximal operator is weak (1,1) and strong (p,p) for $1 < p \leq \infty$.

Proof: By construction, it follows that

$$\|Mf\|_{L^\infty} \leq \|f\|_{L^\infty}.$$

The operator M is sublinear.

Hence, the claim will follow from the Marcinkiewicz Interpolation theorem if we show that M is weak (1,1).

In order to prove this claim, we can use the following result:

Lemma 3.2: (Wiener's Vitali-type Covering Lemma)

Let $\{B_1, \dots, B_N\}$ be a collection of balls in \mathbb{R}^m .

Then, there exists a subcollection $\{B_{j_1}, \dots, B_{j_m}\}$ of disjoint balls such that:

$$\mu\left(\bigcup_{k=1}^m B_{j_k}\right) \geq 3^{-m} \cdot \mu\left(\bigcup_{k=1}^N B_k\right)$$

Proof of Lemma:

We use a greedy algorithm.

• Let B_{j_1} be the ball of largest radius. (If there are more, choose one.)

We then choose B_{j_2} to be the ball disjoint from B_{j_1} which has the largest radius. In general, given B_{j_1}, \dots, B_{j_k} , we let $B_{j_{k+1}}$ be the ball in the collection which is disjoint from B_{j_1}, \dots, B_{j_k} and has maximal radius. (62)

• This procedure terminates after finitely many steps, when we obtain $\{B_{j_1}, \dots, B_{j_m}\}$, since the total number of balls is finite, this procedure terminates after a finite number of steps. By construction, $\{B_{j_1}, \dots, B_{j_m}\}$ is a disjoint collection.

• If $l \in \{1, \dots, N\}$, then by construction B_l intersects B_{j_k} for some $k \in \{1, 2, \dots, m\}$. Otherwise, we could continue the procedure.

We choose k such that B_{j_k} has maximal radius, i.e. we choose k to be minimal.

• By construction, we get that $r(B_l) \leq r(B_{j_k})$, for $l > k$.

Namely, if $r(B_l) > r(B_{j_k})$, then we would have had to choose B_l as an element of the subcollection at some previous step of the procedure.

• From the triangle inequality and the fact that $B_l \cap B_{j_k} \neq \emptyset$, it follows that:

$$\Rightarrow \bigcup_{l=1}^N B_l \subseteq \bigcup_{k=1}^m 3B_{j_k}$$

$$\Rightarrow \mu\left(\bigcup_{k=1}^N B_k\right) \leq \mu\left(\bigcup_{k=1}^m 3B_{j_k}\right) \leq \sum_{k=1}^m \mu(3B_{j_k})$$

$$= \sum_{k=1}^m 3^m \cdot \mu(B_{j_k}) = 3^m \cdot \mu\left(\bigcup_{k=1}^m B_{j_k}\right)$$

$$\Rightarrow \mu\left(\bigcup_{k=1}^m B_{j_k}\right) \geq 3^{-m} \cdot \mu\left(\bigcup_{k=1}^N B_k\right). \quad \square$$

We can now finish the proof of Theorem 3.1.

Namely, let $\lambda > 0$ be given.

Let $A := \{x \in \mathbb{R}^n, Mf(x) > \lambda\}$.

We know that, $\forall x \in A, \exists B_x = B_x(r_x)$ s.t. $\int_{B_x} |f| > \lambda$.

Let $K \subseteq A$ be a compact set.

Then, by compactness $\exists x_1, \dots, x_N$ s.t. $K \subseteq B_{x_1} \cup \dots \cup B_{x_N}$.

Let $B_i := B_{x_i}$ for $i=1,2,\dots,N$.

Hence, $K \subseteq B_1 \cup \dots \cup B_N$ and $\int_{B_i} |f| dy > \lambda$ for all $i=1,2,\dots,N$.

By Wiener's Lemma, we can find a disjoint subcollection $\{B_{j_1}, \dots, B_{j_m}\}$ such that:

$$\mu\left(\bigcup_{k=1}^m B_{j_k}\right) \geq 3^{-m} \cdot \mu\left(\bigcup_{k=1}^N B_k\right) \geq 3^{-m} \cdot \mu(K)$$

$$\Rightarrow \mu(K) \leq 3^m \cdot \mu\left(\bigcup_{k=1}^m B_{j_k}\right) = 3^m \cdot \sum_{k=1}^m \mu(B_{j_k})$$

$$< 3^m \cdot \sum_{k=1}^m \frac{1}{\lambda} \cdot \int_{B_{j_k}} |f| dy \leq \frac{3^m}{\lambda} \cdot \int_{\mathbb{R}^m} |f| dy = \frac{3^m \|f\|_{L^1}}{\lambda}$$

We take suprema over $K \subseteq a$ and we obtain that:

$$\mu(a) = \mu(\{x \in \mathbb{R}^m : Mf(x) > \lambda\}) \leq \frac{3^m \|f\|_{L^1}}{\lambda}$$

$\Rightarrow M$ is weak $(1,1)$.

Theorem 3.1 now follows. \square

Proposition 3.2: Suppose that $f \in L^1(\mathbb{R}^m)$, $f \neq 0$. Then $Mf \notin L^1(\mathbb{R}^m)$.

In particular, M is not strong $(1,1)$.

Proof: We let f be as above.

Then, there exists $R > 0$ s.t. $\int_{B_R(0)} |f| dy = \varepsilon > 0$

If $|x| > R$, then $B_R(0) \subseteq B_{2|x|}(x)$ by the triangle inequality.

In particular:

$$Mf(x) \geq \frac{1}{c_m |x|^m} \cdot \int_{B_{2|x|}(x)} |f| dy \geq \frac{1}{c_m |x|^m} \cdot \int_{B_R(0)} |f| dy = \frac{\varepsilon}{c_m |x|^m}$$

So, $Mf \notin L^1(\mathbb{R}^m)$. \square

The following fact concerning the maximal operator associated to a family $\{T_t\}$ is useful:

Proposition 3.3: Let $\{T_t\}$ be a family of linear operators on $L^p(X, \mu)$.

Let $T^* := \sup |T_t|$ (we suppose that T^* is well-defined)

Suppose that T^* is weak (p, q) , for some $1 \leq p, q < \infty$.

Then, $A := \{f \in L^p(X, \mu), \lim_{t \rightarrow t_0} T_t f = f \text{ a.e.}\}$ is closed in $L^p(X, \mu)$, for all t_0 .

Proof: Suppose that (f_n) is a sequence in A such that

$$f_n \xrightarrow{L^p} f.$$

Let $\lambda > 0$ be given. Then, since $\lim_{t \rightarrow t_0} T_t f_n = f_n \text{ a.e.}$, it

follows that:

$$\mu(\{x: \limsup_{t \rightarrow t_0} |T_t f - f| > \lambda\})$$

$$= \mu(\{x: \limsup_{t \rightarrow t_0} |T_t(f - f_n) - (f - f_n)| > \lambda\})$$

We note that we are using the linearity of T_t at this step.

This expression is:

$$\leq \mu(\{x: \limsup_{t \rightarrow t_0} |T_t(f - f_n)| > \frac{\lambda}{2}\}) + \mu(\{x: |f - f_n| > \frac{\lambda}{2}\})$$

$$\leq \mu(\{x: |T^*(f - f_n)| > \frac{\lambda}{2}\}) + \mu(\{x: |f - f_n| > \frac{\lambda}{2}\})$$

$$\leq \left(\frac{2C \|f - f_n\|_{L^p}}{\lambda}\right)^q + \left(\frac{2\|f - f_n\|_{L^p}}{\lambda}\right)^p \rightarrow 0 \text{ as } n \rightarrow \infty$$

since $f_n \rightarrow f$ in L^p .

It follows that:

$$\mu(\{x: \limsup_{t \rightarrow t_0} |T_t f - f| > 0\}) \leq \sum_{n=1}^{\infty} \mu(\{x: \limsup_{t \rightarrow t_0} |T_t f - f| > \frac{1}{n}\}),$$

which equals to zero

Hence $\lim_{t \rightarrow t_0} T_t f = f$ a.e.

Thus, $f \in \mathcal{A}$. So, \mathcal{A} is closed. \square

3.2: The dyadic maximal function

Given $k \in \mathbb{Z}$, let \mathcal{Q}_k denote the set of all cubes of the form

$$\left[2^{-k} \cdot m_1, 2^{-k} \cdot (m_1+1) \right) \times \left[2^{-k} \cdot m_2, 2^{-k} \cdot (m_2+1) \right) \times \dots \times \left[2^{-k} \cdot m_n, 2^{-k} \cdot (m_n+1) \right)$$

for $m_1, m_2, \dots, m_n \in \mathbb{Z}$.

These are cubes of sidelength 2^{-k} in \mathbb{R}^n .

Given $f \in L^1_{loc}(\mathbb{R}^n)$, we define:

$$E_k f := \sum_{Q \in \mathcal{Q}_k} (f)_Q \cdot \chi_Q$$

This can be interpreted as the conditional expectation of f with respect to the σ -algebra generated by \mathcal{Q}_k .

The dyadic maximal function M_d is defined by:

$$M_d f := \sup |E_k f|.$$

• We will now study the properties of M_d .

Given a cube $Q \subseteq \mathbb{R}^n$, we say that it is a dyadic cube if $Q \in \mathcal{Q}_k$ for some $k \in \mathbb{Z}$.

Theorem 3.4:

1) M_d is weak (1,1) and strong (p,p) for all $1 < p \leq \infty$.

2) $E_k f \xrightarrow[k \rightarrow \infty]{} f$ a.e. for all $f \in L^1_{loc}(\mathbb{R}^n)$.

Proof: 1) By construction $\|M_d f\|_{L^\infty} \leq \|f\|_{L^\infty}$.

Hence, by the Marcinkiewicz interpolation theorem, it suffices to show that M_d is weak $(1,1)$.

By decomposing f into its real and imaginary part and then by decomposing each of these into its non-negative and negative part and by using the sublinearity of M_d , it suffices to show the claim when $f \in L^1(\mathbb{R}^n)$ and $f \geq 0$.

Let $\lambda > 0$.

Since $f \in L^1(\mathbb{R}^n)$, we know that $\lim_{k \rightarrow \infty} E_k f = 0$ a.e.

Now, we deduce that:

$$\{x \in \mathbb{R}^n, M_d f(x) > \lambda\} = \bigcup_{k \in \mathbb{Z}} \Omega_k$$

where:

$$\Omega_k := \{x \in \mathbb{R}^n : E_k f(x) > \lambda, E_j f(x) \leq \lambda \text{ for all } j < k\}$$

In other words, we find the largest dyadic cube over which the average is $> \lambda$.

The Ω_k are then disjoint sets which are unions of elements of \mathcal{Q}_k .

Hence: (throughout this discussion, μ denotes the Lebesgue measure on \mathbb{R}^n)

$$\mu(|M_d f| > \lambda) = \mu(M_d f > \lambda) = \sum_k \mu(\Omega_k)$$

We note that, for each k :

$$\mu(\Omega_k) < \frac{1}{\lambda} \int_{\Omega_k} E_k f = \frac{1}{\lambda} \int_{\Omega_k} f, \text{ by construction.}$$

So:

$$\mu(|M_d f| > \lambda) < \sum_k \frac{1}{\lambda} \int_{\Omega_k} f \leq \frac{1}{\lambda} \|f\|_{L^1}.$$

2) By replacing f with $f \cdot \chi_Q$ for Q a cube of sidelength 2^m centered at 0 it suffices to consider the case $f \in L^1(\mathbb{R}^n)$.

We know that $E_k f \xrightarrow[k \rightarrow \infty]{} f$ pointwise if $f \in C_0(\mathbb{R}^n)$,

We know from part a) that $M_d = \sup |E_k|$ is weak (1,1).

Hence, by Proposition 3.3, it follows that

$\mathcal{A} := \{ f \in L^1(\mathbb{R}^n), E_k f \rightarrow f \text{ a.e.} \}$ is closed in $L^1(\mathbb{R}^n)$.

Since $C_0(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$, it follows that

$\mathcal{A} = L^1(\mathbb{R}^n)$. The claim now follows. \square

The following result can be deduced by using the methods of the above proof:

Theorem 3.5. (Calderón - Zygmund Decomposition)

Let $f \in L^1(\mathbb{R}^n)$ with $f \geq 0$ and let $\lambda > 0$ be given.

Then, there exists a countable collection (Q_j) of disjoint dyadic cubes such that:

$$1) \quad f \leq \lambda \text{ a.e. on } \left(\bigcup_j Q_j \right)^c$$

$$2) \quad \mu \left(\bigcup_j Q_j \right) \leq \frac{1}{\lambda} \|f\|_{L^1}$$

$$3) \quad \lambda < \int_{Q_j} f \leq 2^n \lambda \quad \forall j$$

\rightarrow Such a decomposition of \mathbb{R}^n is called a Calderón-Zygmund Decomposition of f at height λ .

Proof: We use the proof of Theorem 3.4.

The Q_j are obtained from decomposing each Ω_k as a disjoint union of dyadic cubes, (and taking the union over k)

• By construction, we know that $E_k f \leq \lambda$ a.e. on $\left(\bigcup_j Q_j \right)^c$.

Hence, from Part 2) of Theorem 3.4, it follows that

$f \leq \lambda$ a.e. on $(\bigcup_j Q_j)^c$, which proves 1).

- Part 2) follows from the construction of the family (Q_j) .
- Let us take an element Q_j of the family.

Then, we know that:

$$\lambda < \int_{Q_j} f = \frac{1}{\mu(Q_j)} \cdot \int_{Q_j} f$$

• Let \tilde{Q}_j be the dyadic cube containing Q_j of twice the sidelength.

Then, by construction:

$$\int_{\tilde{Q}_j} f = \frac{1}{|\tilde{Q}_j|} \cdot \int_{\tilde{Q}_j} f \leq \lambda$$

$$\text{So: } \frac{1}{\mu(Q_j)} \cdot \int_{Q_j} f = \underbrace{\frac{\mu(\tilde{Q}_j)}{\mu(Q_j)}}_{=2^m} \cdot \frac{1}{\mu(\tilde{Q}_j)} \cdot \int_{Q_j} f \leq 2^m \cdot \frac{1}{\mu(\tilde{Q}_j)} \cdot \int_{\tilde{Q}_j} f \leq 2^m \lambda$$

□

Note: We know by Hölder's inequality that for $1 \leq p < \infty$, for $Q \subseteq \mathbb{R}^n$ a bounded set that:

$$\int_Q |f| \leq \left(\int_Q |f|^p \right)^{\frac{1}{p}}$$

So, if $f \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$ and $f \geq 0$, by applying the previous construction, we get a decomposition (Q_j) of \mathbb{R}^n that satisfies 1) and 3). Namely, we know that $f \leq \lambda$ on $(\bigcup Q_j)^c$ because M_d is strong (p, p) .

• We can use Theorem 3.4 to give an alternative proof of Theorem 3.1.

This is done by relating M' and M_d :

Proposition 3.5: Let $f \in L^1(\mathbb{R}^m)$ with $f \geq 0$ and $\lambda > 0$ be given. Then:

$$\mu(\{x \in \mathbb{R}^m : M'f(x) > 4^m \lambda\}) \leq 2^m \mu(\{x \in \mathbb{R}^m : M_d f(x) > \lambda\})$$

Proof: We use the same notation as in the previous arguments to write:

$$\{x \in \mathbb{R}^m, M_d f(x) > \lambda\} = \bigcup_j Q_j, \quad \text{a disjoint union of dyadic cubes}$$

We will now show that: (with Q_j satisfying the assumptions given earlier.)

$$\{x \in \mathbb{R}^m, M'f(x) > 4^m \lambda\} \subseteq \bigcup_j \widehat{Q}_j \quad (\Delta)$$

Here, each \widehat{Q}_j is the cube with same center as Q_j of twice the sidelength.

Let us note that (Δ) indeed implies that:

$$\begin{aligned} \mu(\{x \in \mathbb{R}^m, M'f(x) > 4^m \lambda\}) &\leq \sum_j \mu(\widehat{Q}_j) = \\ &= \sum_j 2^m \mu(Q_j) = 2^m \cdot \mu\left(\bigcup_j Q_j\right) = 2^m \cdot \mu(\{x \in \mathbb{R}^m, M_d f(x) > \lambda\}) \end{aligned}$$

(This step is reminiscent of the Vitali covering method used earlier.)

• Let us now show (Δ) .

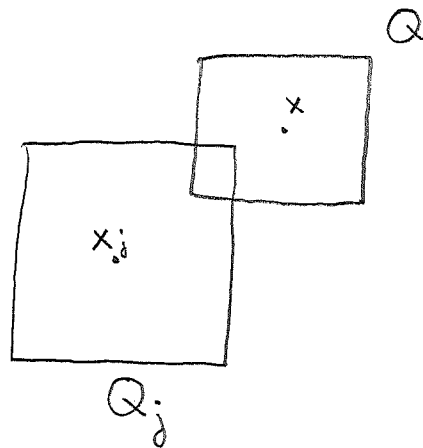
Suppose that $x \notin \bigcup_j \widehat{Q}_j$.

Let Q be a cube centered at x . We find $k \in \mathbb{Z}$ such that $2^{-k-1} \leq l(Q) < 2^{-k}$, where $l(Q) :=$ sidelength of Q .

The cube Q intersects cubes R_1, \dots, R_m that belong to \mathcal{Q}_k . Since $l(Q) < 2^{-k}$, it follows that $m \leq 2^m$. Let us fix one such R_i .

• Let us observe that R_i cannot be contained in any of the Q_j .
 Namely, if this were the case, then one would obtain ^{by the decomposition} that $x \in \widehat{Q}_j$, which is a contradiction.

More precisely: If $R_i \subseteq Q_j$, then $l(Q_j) \geq 2^{-k}$ and $Q \cap Q_j \neq \emptyset$



We observe that each coordinate of $x - x_j$ is bounded in absolute value by $\frac{1}{2}(l(Q_j) + l(Q)) \leq l(Q_j)$
 $\Rightarrow x \in \widehat{Q}_j$, which is the desired contradiction.

• In particular, since $R_i \in \mathcal{Q}_k$, it follows that $\int_{R_i} f \leq \lambda$.

So:

$$\int_Q f = \frac{1}{\mu(Q)} \cdot \sum_{i=1}^m \int_{Q \cap R_i} f \leq \sum_{i=1}^m \underbrace{\frac{\mu(R_i)}{\mu(Q)}}_{\leq 2^m} \cdot \underbrace{\left(\frac{1}{\mu(R_i)} \int_{R_i} f \right)}_{\leq \lambda}$$

$$\leq 2^m m \lambda \leq 4^m \lambda.$$

Hence, $M'f(x) \leq 4^m \lambda$ for $x \notin \bigcup_j \widehat{Q}_j$, i.e.

$$\{x : M'f(x) > 4^m \lambda\} \subseteq \bigcup_j \widehat{Q}_j.$$

This is (Δ) . The claim now follows. \square

• It is now possible to give an alternative proof of Theorem 3.1 from Theorem 3.4, Proposition 3.5 and the fact that $Mf \sim M'f$. (71)

3.3: Examples

• We can use maximal functions to study approximations of the identity.

Proposition 3.6: Let $\varphi \in L^1(\mathbb{R}^n)$ be a non-negative, radial, decreasing (wrt. radius) function. Then:

$$\sup_{t>0} |\varphi_t * f| \leq \|\varphi\|_{L^1} \cdot Mf, \text{ for all } f \in L^p(\mathbb{R}^n) \quad 1 \leq p \leq \infty.$$

Proof: By rescaling, it suffices to prove $|\varphi * f| \leq \|\varphi\|_{L^1} \cdot Mf$ (i.e. we use the fact that $\|\varphi_t\|_{L^1} = \|\varphi\|_{L^1}$).

Since $|\varphi * f| \leq \varphi * |f|$ and $M|f| = Mf$, we may just consider the case when $f \geq 0$.

So, we must show that, for $f \geq 0$:

$$\varphi * f \leq \|\varphi\|_{L^1} \cdot Mf.$$

• If φ is a simple function, we can write:

$$\varphi = \sum_{i=1}^k a_i \chi_{B_i}, \text{ where } a_i > 0 \text{ and } B_i \text{ are balls centered at the origin.}$$

(note that these are all the simple functions of this form).

$$\text{Then } \varphi * f = \sum_{i=1}^k (a_i \chi_{B_i}) * f = \sum_{i=1}^k a_i \mu(B_i) \cdot \underbrace{\left(\frac{1}{\mu(B_i)} \cdot (\chi_{B_i} * f) \right)}_{\leq Mf}$$

$$\leq \sum_{i=1}^k a_i \mu(B_i) \cdot Mf = \|\varphi\|_{L^1} \cdot Mf.$$

We observe that every non-negative, radial, decreasing (wrt. radius) integrable function φ can be written as a monotone increasing limit of simple functions $\varphi^{(k)}$ as above. Note that $\varphi^{(k)} \rightarrow \varphi$ in L^1 .

Therefore, by Young's inequality $\varphi^{(k)} * f \rightarrow \varphi * f$ in L^p .

\Rightarrow up to a subsequence (which we also denote by $\varphi^{(k)}$), we then obtain that $\varphi * f = \lim_k \varphi^{(k)} * f$, almost everywhere. (72)

$$\Rightarrow \Psi * f = \lim_k \Psi^{(k)} * f \leq \lim_k \|\Psi^{(k)}\|_{L^1} \cdot Mf = \|\Psi\|_{L^1} \cdot Mf$$

$$\text{so: } \sup_{t>0} |\Psi_t * f| \leq \|\Psi\|_{L^1} \cdot Mf. \quad \square$$

• We can apply this result to study pointwise convergence of approximations of the identity:

Proposition 3.7: Let Ψ satisfy the assumptions as above.

Let $1 \leq p < \infty$ and $f \in L^p(\mathbb{R}^n)$ be given. Then:

$$\Psi_t * f \xrightarrow{t \rightarrow 0} \left(\int_{\mathbb{R}^n} \Psi \right) \cdot f \quad \text{a.e.}$$

Proof: From Proposition 3.6 and Theorem 3.1, it follows that

$T^* f := \sup_{t>0} |\Psi_t * f|$ is a weak (p, p) operator.

If $f \in \mathcal{S}(\mathbb{R}^n)$, then $\Psi_t * f \xrightarrow{t \rightarrow 0} f$ pointwise a.e.

The claim follows for general $f \in L^p(\mathbb{R}^n)$

by Proposition 3.3. \square (Note that here, we consider $p < \infty$)

Remark: The same conclusions hold if $|\Psi| \leq \Psi$ where Ψ is a non-negative, radial, decreasing (wrt. radius) and integrable function. This is because $|\Psi * f| \leq |\Psi| * |f| \leq \Psi * |f|$ and we then argue as before.