Proof. The sketch of the proof: By a density of continuous functions in $L^{p}$, we can choose $g \in C_{0}\left(\mathbf{R}^{n}\right)$ so that $\|f-g\|_{p}$ is small. By adding and subtracting $g$, we can estimate

$$
\begin{align*}
\left|\left(f * \phi_{\varepsilon}\right)(x)-a f(x)\right| & \leq\left|\phi_{\varepsilon} *(f-g)(x)-a(f-g)(x)\right| \\
& +\left|\left(g * \phi_{\varepsilon}\right)(x)-a g(x)\right| . \tag{3.13}
\end{align*}
$$

Since $g \in C_{0}\left(\mathbf{R}^{n}\right)$, the second term tends to zero as $\varepsilon \rightarrow 0$. Thus we can focus attention on the first term on the right hand side. By Theorem 3.10, we can estimate

$$
\begin{aligned}
\left|\left(f * \phi_{\varepsilon}\right)(x)-a f(x)\right| & \leq\left|\phi_{\varepsilon} *(f-g)(x)-a(f-g)(x)\right| \\
& \leq M(f-g)(x)+a|(f-g)(x)|
\end{aligned}
$$

Finally, we can show by using the weak type estimates that the quantities on the right hand side get small almost everywhere.

Details: Case $1 \leq p<\infty$ :
As sketched above the weak type estimates play a key role. Theorem Hardy-Littlewood I (Theorem 2.12) implies

$$
\begin{equation*}
m\left(\left\{x \in \mathbf{R}^{n}: M f(x)>\lambda\right\}\right) \leq \frac{C}{\lambda}\|f\|_{1} \tag{3.14}
\end{equation*}
$$

for $\lambda>0$, and Hardy-Littlewood II (Theorem 2.19) imply

$$
\begin{equation*}
m\left(\left\{x \in \mathbf{R}^{n}: M f(x)>\lambda\right\}\right) \stackrel{\text { Chebyshev }}{\leq} \frac{C}{\lambda^{p}}\|M f\|_{p}^{p} \stackrel{\text { H-L II }}{\leq} C\|f\|_{p}^{p} \text {. } \tag{3.15}
\end{equation*}
$$

As $g$ is continuous at $x \in \mathbf{R}^{n}$ it follows that for every $\eta>0$ there exists $\delta>0$ such that

$$
|g(x-y)-g(x)|<\eta \quad \text { whenever } \quad|y|<\delta
$$

Thus

$$
\begin{aligned}
\left|\left(g * \phi_{\varepsilon}\right)(x)-a g(x)\right| & \leq \int_{\mathbf{R}^{n}}|g(x-y)-g(x)| \phi_{\varepsilon}(y) \mathrm{d} y \\
& \leq \eta \underbrace{\int_{B(0, \delta)} \phi_{\varepsilon}(y) \mathrm{d} y}_{\leq\|\phi\|_{1}}+2\|g\|_{\infty} \underbrace{\int_{\mathbf{R}^{n} \backslash B(0, \delta)} \phi_{\varepsilon}(x) \mathrm{d} y}_{\rightarrow 0 \text { as } \varepsilon \rightarrow 0 \text { by Lemma } 3.6}
\end{aligned}
$$

Since $\eta$ was arbitrary, it follows that

$$
\lim _{\varepsilon \rightarrow 0}\left|\left(g * \phi_{\varepsilon}\right)(x)-a g(x)\right|=0
$$

for all $x \in \mathbf{R}^{n}$.

This in mind we can estimate

$$
\begin{align*}
\limsup _{\varepsilon \rightarrow 0} & \left|\left(f * \phi_{\varepsilon}\right)(x)-a f(x)\right| \\
& \leq \limsup _{\varepsilon \rightarrow 0}^{\lim \sup }\left|\left((f-g) * \phi_{\varepsilon}\right)(x)-a(f-g)(x)\right| \\
& +\underbrace{\limsup _{\varepsilon \rightarrow 0}\left|\left(g * \phi_{\varepsilon}\right)(x)-a g(x)\right|}_{=0}  \tag{3.16}\\
\leq & \sup _{\varepsilon>0}\left|\left((f-g) * \phi_{\varepsilon}\right)(x)\right|+a|(f-g)(x)| \\
& \text { Theorem } 3.10 \\
\leq & C M(f-g)(x)+a|(f-g)(x)| .
\end{align*}
$$

Next we define

$$
A_{i}=\left\{x \in \mathbf{R}^{n}: \limsup _{\varepsilon \rightarrow 0}\left|\left(f * \phi_{\varepsilon}\right)(x)-a f(x)\right|>\frac{1}{i}\right\} .
$$

By the previous estimate,
$A_{i} \subset\left\{x \in \mathbf{R}^{n}: C M(f-g)(x)>\frac{1}{2 i}\right\} \cup\left\{x \in \mathbf{R}^{n}: a|f(x)-g(x)|>\frac{1}{2 i}\right\}$,
for $i=1,2, \ldots$ Let $\eta>0$, and let $g \in C_{0}\left(\mathbf{R}^{n}\right)$ be such that (density)

$$
\|f-g\|_{p} \leq \eta
$$

This and the previous inclusion imply

$$
\begin{aligned}
m\left(A_{i}\right) & \leq m\left(\left\{x \in \mathbf{R}^{n}: C M(f-g)(x)>\frac{1}{2 i}\right\}\right)+m\left(\left\{x \in \mathbf{R}^{n}: a|f(x)-g(x)|>\frac{1}{2 i}\right\}\right) \\
& \stackrel{(3.14),(3.15)}{\leq} C i^{p}\|f-g\|_{p}^{p}+C i^{p}\|f-g\|_{p}^{p} \\
& \leq C i^{p}\|f-g\|_{p}^{p} \leq C i^{p} \eta^{p}
\end{aligned}
$$

for every $\eta, i=1,2, \ldots$. Thus

$$
m\left(A_{i}\right)=0
$$

and

$$
m\left(\cup_{i=1}^{\infty} A_{i}\right) \leq \sum_{i=1}^{\infty} m\left(A_{i}\right)=0 .
$$

This gives us

$$
m\left(\left\{x \in \mathbf{R}^{n}: \limsup _{\varepsilon \rightarrow 0}\left|\left(f * \phi_{\varepsilon}\right)(x)-a f(x)\right|>0\right\}\right)=0
$$

which proofs the claim

$$
\lim _{\varepsilon \rightarrow 0}\left|\left(f * \phi_{\varepsilon}\right)(x)-a f(x)\right|=0 \quad \text { a.e. } x \in \mathbf{R}^{n} .
$$

Case $p=\infty$ : Now $f \in L^{\infty}\left(\mathbf{R}^{n}\right)$. We show that

$$
\lim _{\varepsilon \rightarrow 0}\left(f * \phi_{\varepsilon}\right)(x)=a f(x)
$$

for almost every $x \in B(0, r), r>0$. Let

$$
f_{1}(x)=f \chi_{B(0, r+1)}(x)= \begin{cases}f(x), & x \in B(0, r+1) \\ 0, & \text { otherwise }\end{cases}
$$

and $f_{2}=f-f_{1}$. Now $f_{1} \in L^{1}\left(\mathbf{R}^{n}\right)$ and by the previous case

$$
\lim _{\varepsilon \rightarrow 0}\left(f_{1} * \phi_{\varepsilon}\right)(x)=a f_{1}(x)
$$

for almost every $x \in \mathbf{R}^{n}$. By utilizing this, we obtain for almost every $x \in B(0, r)$ that

$$
\begin{aligned}
\lim _{\varepsilon \rightarrow 0}\left(f * \phi_{\varepsilon}\right)(x) & =\lim _{\varepsilon \rightarrow 0}\left(f_{1} * \phi_{\varepsilon}\right)(x)+\lim _{\varepsilon \rightarrow 0}\left(f_{2} * \phi_{\varepsilon}\right)(x) \\
& =a f(x)+\lim _{\varepsilon \rightarrow 0}\left(f_{2} * \phi_{\varepsilon}\right)(x)
\end{aligned}
$$

and it remains to show that $\lim _{\varepsilon \rightarrow 0}\left(f_{2} * \phi_{\varepsilon}\right)(x)=0$ for almost all $x \in$ $B(0, r)$. To this end, let $x \in B(0, r)$ so that $f_{2}(x-y)=0$ for $y \in B(0,1)$ and calculate
as $\varepsilon \rightarrow 0$.
By choosing

$$
\phi(x)=\chi_{B(0,1)}(x) / m(B(0,1)),
$$

so that

$$
\phi_{\varepsilon}(x)=\chi_{B(0, \varepsilon)} /\left(\varepsilon^{n} m(B(0,1))\right)=\chi_{B(0, \varepsilon)} / m(B(0, \varepsilon))
$$

we immediately obtain
Theorem 3.17 (Lebesgue density theorem). If $f \in L_{l o c}^{1}\left(\mathbf{R}^{n}\right)$, then

$$
\lim _{r \rightarrow 0} f_{B(x, R)} f(y) \mathrm{d} y=f(x)
$$

for almost every $x \in \mathbf{R}^{n}$.
Example 3.18. Let

$$
\phi(x)=P(x)=\frac{C(n)}{\left(1+|x|^{2}\right)^{(n+1) / 2}}
$$

where the constant is chosen so that

$$
\int_{\mathbf{R}^{n}} P(x) \mathrm{d} x=1 .
$$

Next we define

$$
P_{t}(x)=\frac{1}{t^{n}} P\left(\frac{x}{t}\right)=C(n) \frac{t}{\left(|x|^{2}+t^{2}\right)^{(n+1) / 2}}, t>0
$$

and

$$
u(x, t)=\left(f * P_{t}\right)(x)=\int_{\mathbf{R}^{n}} P_{t}(x-y) f(y) \mathrm{d} y
$$

This is called the Poisson integral for $f$. It has the following properties
(i) $\Delta u=\frac{\partial^{2} u}{\partial t^{2}}+\frac{\partial^{2} u}{\partial x_{1}^{2}}+\ldots+\frac{\partial^{2} u}{\partial x_{2}^{2}}=0$ and
(ii) $\lim _{t \rightarrow 0} u(x, t)=f(x)$ for almost every $x \in \mathbf{R}^{n}$ by Theorem 3.12. Let

$$
\mathbf{R}_{+}^{n+1}=\left\{\left(x_{1}, x_{2}, \ldots, t\right) \in \mathbf{R}^{n+1}: t>0\right\}
$$

denote the upper half space. As stated above $u$ is harmonic in $\mathbf{R}_{+}^{n+1}$ so that $u(x, t)=\int_{\mathbf{R}^{n}} P_{t}(x-y) f(y) \mathrm{d} y$ solves

$$
\begin{cases}\Delta u(x, t)=0, & (x, t) \in \mathbf{R}_{+}^{n+1} \\ u(x, 0)=f(x), & , x \in \partial \mathbf{R}_{+}^{n+1}=\mathbf{R}^{n}\end{cases}
$$

where the boundary condition is obtained in the sense

$$
\lim _{t \rightarrow 0} u(x, t)=f(x)
$$

almost everywhere on $\mathbf{R}^{n}$. As $(x, t) \rightarrow(x, 0)$ along a perpendicular axis, we call this radial convergence.

Question Does the Poisson integral converge better than radially?
Definition 3.19. Let $x \in \mathbf{R}^{n}$ and $\alpha>0$. Then
(i) We define a cone

$$
\Gamma_{\alpha}(x)=\left\{(y, t) \in \mathbf{R}_{+}^{n+1}:|x-y|<\alpha t\right\} .
$$

(ii) Function $u(x, t)$ converges nontangentially, if $u(y, t) \rightarrow f(x)$ and $(y, t) \rightarrow(x, 0)$ so that $(y, t)$ remains inside the cone $\Gamma_{\alpha}(x)$.

Theorem 3.20. Let $f \in L^{p}\left(\mathbf{R}^{n}\right), 1 \leq p \leq \infty$, and $u(x, t)=\left(f * P_{t}\right)(x)$. Then for every $\alpha>0$, there exists $C=C(n, \alpha)$ such that

$$
u_{\alpha}^{*}(x):=\sup _{(y, t) \in \Gamma_{\alpha}(x)}|u(y, t)| \leq C M f(x)
$$

for every $x \in \mathbf{R}^{n}$.
$u^{*}$ is called a nontangential maximal function.
Proof. First we show that

$$
P_{t}(y-z) \leq C(\alpha, n) P_{t}(x-z) \quad \text { for every }(y, t) \in \Gamma_{\alpha}(x), z \in \mathbf{R}^{n}
$$

