It is quite easy to see that the Dirac mass δ_0 does not belong in any Hardy space; indeed, $\delta_0 * P_t = P_t$ and $\sup_{t>0} P_t(x)$ is comparable to $|x|^{-n}$ which does not lie in $L^p(\mathbf{R}^n)$ for any p. However, the difference of Dirac masses $\delta_1 - \delta_{-1}$ lies in $H^p(\mathbf{R})$ for $1/2 < p < 1$. To see this, notice that

$$\sup_{t>0}\Big|(\delta_1 * P_t)(x) - (\delta_{-1} * P_t)(x)\Big| = \sup_{t>0} \frac{4|x|}{\pi} \frac{t}{(t^2+|x-1|^2)(t^2+|x+1|^2)}. \tag{2.1.5}$$

Suppose that $|x+1| < |x-1|$, i.e., $x < 0$. Then we have

$$\sup_{t\le|x+1|} \frac{t\,|x|}{(t^2+|x-1|^2)(t^2+|x+1|^2)} \approx \sup_{t\le|x+1|} \frac{t\,|x|}{|x-1|^2|x+1|^2} = \frac{|x|}{|x-1|^2|x+1|}.$$

Also,

$$\sup_{|x+1|\le t\le|x-1|} \frac{t\,|x|}{(t^2+|x-1|^2)(t^2+|x+1|^2)} \approx \sup_{|x+1|\le t\le|x-1|} \frac{t\,|x|}{|x-1|^2t^2} = \frac{|x|}{|x-1|^2|x+1|},$$

while

$$\sup_{t\ge|x-1|} \frac{t\,|x|}{(t^2+|x-1|^2)(t^2+|x+1|^2)} \approx \sup_{t\ge|x-1|} \frac{t\,|x|}{t^4} = \frac{|x|}{|x-1|^3}.$$

Thus (2.1.5) is comparable to $\frac{|x|}{|x-1|^2|x+1|}$ for $x<0$ and analogously to $\frac{|x|}{|x+1|^2|x-1|}$ for $x>0$. Consequently, (2.1.5) lies in $L^p(\mathbf{R})$ if and only if $1/2 < p < 1$.

At this point we don't know whether the H^p spaces coincide with any other known spaces for some values of p. In the next theorem we show that this is the case when $1 < p < \infty$.

Theorem 2.1.2. *(a) Let $1 < p < \infty$. Then every bounded tempered distribution f in H^p is an element of L^p. Moreover, there is a constant $C_{n,p}$ such that for all such f we have*

$$\big\|f\big\|_{L^p} \le \big\|f\big\|_{H^p} \le C_{n,p}\big\|f\big\|_{L^p},$$

and therefore $H^p(\mathbf{R}^n)$ coincides with $L^p(\mathbf{R}^n)$.
(b) When $p=1$, every element of H^1 is an integrable function. In other words, $H^1(\mathbf{R}^n) \subseteq L^1(\mathbf{R}^n)$ and for all $f \in H^1$ we have

$$\big\|f\big\|_{L^1} \le \big\|f\big\|_{H^1}. \tag{2.1.6}$$

Proof. (a) Let $f \in H^p(\mathbf{R}^n)$ for some $1<p<\infty$. The set $\{P_t * f:\ t>0\}$ lies in a multiple of the unit ball of $L^p(\mathbf{R}^n)$, which is the dual space of the separable Banach space $L^{p'}(\mathbf{R}^n)$, and hence it is weak* sequentially compact by the Banach–Alaoglu theorem. Therefore, there exists a sequence $t_j \to 0$ such that $P_{t_j} * f$ converges to some L^p function f_0 in the weak* topology of L^p. On the other hand, in view of (2.1.3), $P_{t_j} * f \to f$ in $\mathscr{S}'(\mathbf{R}^n)$ as $t_j \to 0$, and thus the bounded tempered distribution f coincides with the L^p function f_0. Since the family $\{P_t\}_{t>0}$ is an approximate identity, Theorem 1.2.19 in [156] gives that

$$\big\|P_t * f - f\big\|_{L^p} \to 0 \qquad \text{as } t \to 0,$$

from which it follows that

$$\|f\|_{L^p} \le \big\| \sup_{t>0} |P_t * f| \big\|_{L^p} = \|f\|_{H^p}. \tag{2.1.7}$$

The converse inequality is a consequence of the fact that

$$\sup_{t>0} |P_t * f| \le M(f),$$

where M is the Hardy–Littlewood maximal operator. (See Corollary 2.1.12 in [156].)

(b) The case $p = 1$ requires only a small modification of the case $p > 1$. We embed L^1 in the space of finite Borel measures $\mathscr{M}$ which is the dual of the separable space $\mathscr{C}_{00}(\mathbf{R}^n)$ of all continuous functions on $\mathbf{R}^n$ that vanish at infinity. By the Banach-Alaoglu theorem, the unit ball of $\mathscr{M}$ is weak* sequentially compact, and we can extract a sequence $t_j \to 0$ such that $P_{t_j} * f$ converges to some measure μ in the topology of measures. In view of (2.1.3), it follows that the distribution f can be identified with the measure μ.

It remains to show that μ is absolutely continuous with respect to Lebesgue measure, which would imply that it coincides with some L^1 function. We show that μ is absolutely continuous with respect to Lebesgue measure by showing that for all subsets E of $\mathbf{R}^n$ we have $|E| = 0 \implies |\mu(E)| = 0$. Since $\sup_{t>0} |P_t * f|$ lies in $L^1(\mathbf{R}^n)$, given $\varepsilon > 0$, there exists a $\delta > 0$ such that for any measurable subset F of $\mathbf{R}^n$ we have

$$|F| < \delta \implies \int_F \sup_{t>0} |P_t * f|\, dx < \varepsilon.$$

Given E with $|E| = 0$, we can find an open set U such that $E \subseteq U$ and $|U| < \delta$. Let us denote by $\mathscr{C}_{00}(U)$ the space of continuous functions $g(x)$ that are supported in U and tend to zero as $|x| \to \infty$. Then for any g in $\mathscr{C}_{00}(U)$ we have

$$\begin{aligned} \left| \int_{\mathbf{R}^n} g\, d\mu \right| &= \lim_{j\to\infty} \left| \int_{\mathbf{R}^n} g(x)\,(P_{t_j} * f)(x)\, dx \right| \\ &\le \|g\|_{L^\infty} \int_U \sup_{t>0} |(P_t * f)(x)|\, dx \\ &< \varepsilon \|g\|_{L^\infty}. \end{aligned}$$

Let $|\mu|$ be the ~~total~~ absolute variation of μ. Then we have (see [190] (20.49))

$$|\mu|(U) = \int_U 1\, d|\mu| = \sup\left\{ \left| \int_{\mathbf{R}^n} g\, d\mu \right| :\ g \in \mathscr{C}_{00}(U), \quad \|g\|_{L^\infty} \le 1 \right\},$$

which implies $|\mu|(U) < \varepsilon$. Since ε was arbitrary, it follows that $|\mu|(E) = 0$ and thus $\mu(E) = 0$; hence μ is absolutely continuous with respect to Lebesgue measure. Finally, (2.1.6) is a consequence of (2.1.7), which is also valid for $p = 1$. □

We may wonder whether H^1 coincides with L^1. We show in Corollary 2.4.8 that elements of H^1 have integral zero; thus H^1 is a proper subspace of L^1.

2.1.2 Quasi-norm Equivalence of Several Maximal Functions

We now obtain some characterizations of these spaces.

Definition 2.1.3. Let $a,b>0$. Let Φ be a Schwartz function and let f be a tempered distribution on $\mathbf{R}^n$. We define the *smooth maximal function of f with respect to Φ* as

$$M(f;\Phi)(x)=\sup_{t>0}|(\Phi_t*f)(x)|\,.$$

We define the *nontangential maximal function (with aperture a) of f with respect to Φ* as

$$M_a^*(f;\Phi)(x)=\sup_{t>0}\sup_{\substack{y\in\mathbf{R}^n\\ |y-x|<at}}|(\Phi_t*f)(y)|\,.$$

We also define the *auxiliary maximal function*

$$M_b^{**}(f;\Phi)(x)=\sup_{t>0}\sup_{y\in\mathbf{R}^n}\frac{|(\Phi_t*f)(x-y)|}{(1+t^{-1}|y|)^b}\,, \tag{2.1.8}$$

and we observe that

$$M(f;\Phi)\le M_a^*(f;\Phi)\le(1+a)^b M_b^{**}(f;\Phi) \tag{2.1.9}$$

for all $a,b>0$. We note that if Φ is merely integrable, for example, if Φ is the Poisson kernel, the maximal functions $M(f;\Phi)$, $M_a^*(f;\Phi)$, and $M_b^{**}(f;\Phi)$ are well defined only for bounded tempered distributions f on $\mathbf{R}^n$.

For a fixed positive integer N and a Schwartz function φ we define the quantity

$$\mathfrak{N}_N(\varphi)=\int_{\mathbf{R}^n}(1+|x|)^N\sum_{|\alpha|\le N+1}|\partial^\alpha\varphi(x)|\,dx\,. \tag{2.1.10}$$

We now define

$$\mathscr{F}_N=\Big\{\varphi\in\mathscr{S}(\mathbf{R}^n):\ \mathfrak{N}_N(\varphi)\le 1\Big\}\,, \tag{2.1.11}$$

and we also define the *grand maximal function of f (with respect to N)* as

$$\mathscr{M}_N(f)(x)=\sup_{\varphi\in\mathscr{F}_N}M_1^*(f;\varphi)(x)\,.$$

It is a fact that all the maximal functions of the preceding subsection have comparable L^p quasi-norms for all $0<p<\infty$. This is the essence of the following theorem.

Theorem 2.1.4. *Let $0 < p < \infty$. Then the following statements are valid:*
(a) There exists a Schwartz function Φ^o with $\int_{\mathbf{R}^n} \Phi^o(x)\,dx = 1$ such that

$$\big\|M(f;\Phi^o)\big\|_{L^p} \le 500\,\big\|f\big\|_{H^p} \tag{2.1.12}$$

for all bounded distributions $f \in \mathscr{S}'(\mathbf{R}^n)$.
(b) For every $a > 0$ and every Φ in $\mathscr{S}(\mathbf{R}^n)$ with $\int_{\mathbf{R}^n} \Phi(x)dx = 1$ one has

$$\big\|M_a^*(f;\Phi)\big\|_{L^p} \le C_2(n,p,a,\Phi)\big\|M(f;\Phi)\big\|_{L^p} \tag{2.1.13}$$

for some constant $C_2(n,p,a,\Phi) < \infty$ and for all distributions $f \in \mathscr{S}'(\mathbf{R}^n)$.
(c) For every $a > 0$, $b > n/p$, and every Φ in $\mathscr{S}(\mathbf{R}^n)$ there exists a constant $C_3(n,p,a,b) < \infty$ such that

$$\big\|M_b^{**}(f;\Phi)\big\|_{L^p} \le C_3(n,p,a,b)\big\|M_a^*(f;\Phi)\big\|_{L^p} \tag{2.1.14}$$

for all distributions $f \in \mathscr{S}'(\mathbf{R}^n)$.
(d) For every $b > 0$ and every Φ in $\mathscr{S}(\mathbf{R}^n)$ with $\int_{\mathbf{R}^n} \Phi(x)\,dx = 1$ there exists a constant $C_4(b,\Phi) < \infty$ such that if $N = [b]+1$ we have

$$\big\|\mathscr{M}_N(f)\big\|_{L^p} \le C_4(b,\Phi)\big\|M_b^{**}(f;\Phi)\big\|_{L^p} \tag{2.1.15}$$

for all distributions $f \in \mathscr{S}'(\mathbf{R}^n)$.
(e) For every positive integer N there exists a constant $C_5(n,N)$ such that every tempered distribution f with $\big\|\mathscr{M}_N(f)\big\|_{L^p} < \infty$ is a bounded distribution and satisfies

$$\big\|f\big\|_{H^p} \le C_5(n,N)\big\|\mathscr{M}_N(f)\big\|_{L^p}\,, \tag{2.1.16}$$

that is, it lies in the Hardy space H^p.

Choosing $\Phi = \Phi^o$ in parts (b), (c), and (d), $\frac{n}{p} < b < [\frac{n}{p}]+1$, and $N = [\frac{n}{p}]+1$, we conclude that for bounded distributions f we have

$$\big\|f\big\|_{H^p} \approx \big\|\mathscr{M}_N(f)\big\|_{L^p}\,.$$

Moreover, for any Schwartz function Φ with $\int_{\mathbf{R}^n} \Phi(x)\,dx = 1$ and any bounded distribution f in $\mathscr{S}'(\mathbf{R}^n)$, the following quasi-norms are equivalent

$$\big\|f\big\|_{H^p} \approx \big\|M(f;\Phi)\big\|_{L^p}\,,$$

with constants that depend only on Φ, n, p.

Before we begin the proof of Theorem 2.1.4, we state and prove a useful lemma.

Lemma 2.1.5. *Let $m \in \mathbf{Z}^+$ and let Φ in $\mathscr{S}(\mathbf{R}^n)$ satisfy $\int_{\mathbf{R}^n} \Phi(x)\,dx = 1$. Then there exists a constant $C_0(\Phi,m)$ such that for any Ψ in $\mathscr{S}(\mathbf{R}^n)$, there are Schwartz functions $\Theta^{(s)}$, $0 < s \le 1$, with the properties*

$$\Psi(x) = \int_0^1 (\Theta^{(s)} * \Phi_s)(x)\,ds \tag{2.1.17}$$

and

$$\int_{\mathbf{R}^n}(1+|x|)^m|\Theta^{(s)}(x)|\,dx \le C_0(\Phi,m)\,s^m\,\mathfrak{N}_m(\Psi). \tag{2.1.18}$$

Proof. We start with a smooth function ζ supported in $[0,1]$ that satisfies

$$\begin{aligned} 0 \le \zeta(s) &\le \tfrac{s^m}{m!} && \text{for all } 0 \le s \le 1\,, \\ \zeta(s) &= \tfrac{s^m}{m!} && \text{for all } 0 \le s \le \tfrac{1}{2}\,, \\ \zeta(s) &= 0 && \text{for all } \tfrac{7}{8} \le s \le 1. \end{aligned}$$

We define

$$\Theta^{(s)} = (-1)^{m+1}\zeta(s)\;\Xi^{(s)} * \Psi - \frac{d^{m+1}\zeta}{ds^{m+1}}(s)\;\overbrace{\Phi_s * \cdots * \Phi_s}^{m+1 \text{ terms}} * \Psi\,, \tag{2.1.19}$$

where $\Xi^{(s)}$ is chosen so that $\frac{d^{m+1}}{ds^{m+1}}\big(\overbrace{\Phi_s * \cdots * \Phi_s}^{m+2 \text{ terms}}\big) = \Xi^{(s)} * \Phi_s$. Notice that

$$\Xi^{(s)} = \sum_{j_1,\dots,j_m \ge 0:\, j_1+\cdots+j_m = m+1} c_{j_1,\dots,j_m} \frac{d^{j_1}}{ds^{j_1}}\Phi_s * \cdots * \frac{d^{j_{m+1}}}{ds^{j_{m+1}}}\Phi_s\,,$$

for some constants $c_{j_1,\dots,j_m}$. We claim that (2.1.17) holds for this choice of $\Theta^{(s)}$. To verify this assertion, we apply $m+1$ integration by parts to write

$$\begin{aligned} \int_0^1 \Theta^{(s)} * \Phi_s\,ds = \int_0^1 (-1)^{m+1}\zeta(s)\Xi^{(s)} * \Phi_s * \Psi\,ds + \frac{d^m\zeta}{ds^m}(0)\lim_{s\to 0+}\big(\overbrace{\Phi * \ldots * \Phi}^{m+2 \text{ terms}}\big)_s * \Psi \\ -(-1)^{m+1}\int_0^1 \zeta(s)\,\frac{d^{m+1}}{ds^{m+1}}\Big(\overbrace{\Phi_s * \cdots * \Phi_s}^{m+2 \text{ terms}}\Big) * \Psi\,ds\,, \end{aligned}$$

noting that all the boundary terms vanish except for the term at $s=0$ in the first integration by parts. The first and the third terms on the right above add up to zero, while the second term equals Ψ, since Φ has integral one, which implies that the family $\{(\Phi * \cdots * \Phi)_s\}_{s>0}$ is an approximate identity as $s \to 0$. Thus (2.1.17) holds.

We now prove estimate (2.1.18). Let Ω be the $(m+1)$-fold convolution of Φ. For the second term on the right in (2.1.19), we note that the $(m+1)$st derivative of $\zeta(s)$ vanishes on $[0,\frac{1}{2}]$, so that we may write

$$\begin{aligned} &\int_{\mathbf{R}^n}(1+|x|)^m\Big|\frac{d^{m+1}\zeta(s)}{ds^{m+1}}\Big|\,|\Omega_s * \Psi(x)|\,dx \\ &\le C_m\,\chi_{[\frac{1}{2},1]}(s)\int_{\mathbf{R}^n}(1+|x|)^m\bigg[\int_{\mathbf{R}^n}\tfrac{1}{s^n}\big|\Omega\big(\tfrac{x-y}{s}\big)\big|\,|\Psi(y)|\,dy\bigg]dx \\ &\le C_m\,\chi_{[\frac{1}{2},1]}(s)\int_{\mathbf{R}^n}\int_{\mathbf{R}^n}(1+|y+sx|)^m|\Omega(x)|\,|\Psi(y)|\,dy\,dx \\ &\le C_m\,\chi_{[\frac{1}{2},1]}(s)\int_{\mathbf{R}^n}\int_{\mathbf{R}^n}(1+|sx|)^m|\Omega(x)|\,(1+|y|)^m|\Psi(y)|\,dy\,dx \\ &\le C_m\,\chi_{[\frac{1}{2},1]}(s)\Big(\int_{\mathbf{R}^n}(1+|x|)^m|\Omega(x)|\,dx\Big)\Big(\int_{\mathbf{R}^n}(1+|y|)^m|\Psi(y)|\,dy\Big) \end{aligned}$$

$$\leq C_0'(\Phi,m)\, s^m\, \mathfrak{N}_m(\Psi)\,,$$

since $\chi_{[\frac{1}{2},1]}(s) \leq 2^m s^m$. For the first term on the right in (2.1.19) we have

$$\begin{aligned}
&\int_{\mathbf{R}^n} (1+|x|)^m \left|\frac{d^{j_1}\Phi_s}{ds^{j_1}} * \Psi(x)\right| dx \\
&= \int_{\mathbf{R}^n} (1+|x|)^m \left|\frac{d^{j_1}}{ds^{j_1}} \int_{\mathbf{R}^n} \Phi(y)\Psi(x-sy)\,dy\right| dx \\
&= \int_{\mathbf{R}^n} (1+|x|)^m \left|\int_{\mathbf{R}^n} \Phi(y)\frac{d^{j_1}}{ds^{j_1}}\Psi(x-sy)\,dy\right| dx \\
&\leq \int_{\mathbf{R}^n} (1+|x|)^m \int_{\mathbf{R}^n} |\Phi(y)| \Big[\sum_{|\alpha|\leq j_1} |\partial^\alpha \Psi(x-sy)|\,|y|^{|\alpha|}\Big]\,dy\,dx \\
&\leq \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} (1+|x+sy|)^m |\Phi(y)| \sum_{|\alpha|\leq j_1} |\partial^\alpha \Psi(x)|\,(1+|y|)^{j_1}\,dy\,dx \\
&\leq \int_{\mathbf{R}^n} (1+|y|)^{j_1}\,|\Phi(y)|\,(1+|y|)^m\,dy \int_{\mathbf{R}^n} (1+|x|)^m \sum_{|\alpha|\leq j_1} |\partial^\alpha \Psi(x)|\,dx \\
&\leq C'_{m,\Phi} \int_{\mathbf{R}^n} (1+|x|)^m \sum_{|\alpha|\leq j_1} |\partial^\alpha \Psi(x)|\,dx,
\end{aligned}$$

using $j_1 \leq m+1$. Applying this estimate to the function $\frac{d^{j_2}\Phi_s}{ds^{j_2}} * \Psi$ in place of Ψ we obtain a similar estimate for $\frac{d^{j_1}\Phi_s}{ds^{j_1}} * \frac{d^{j_2}\Phi_s}{ds^{j_2}} * \Psi$ where the last displayed sum is taken over $|\alpha| \leq j_1 + j_2$. Continuing in this way we obtain the desired estimate for every term that appears in the sum defining $\Xi^{(s)}$, and consequently for $\Xi^{(s)}$ itself. Keeping in mind that the function $\zeta(s)$ is pointwise bounded by s^m for $0 < s \leq 1$, yields the desired estimate. This concludes the proof of (2.1.18). □

Next, we discuss the proof of Theorem 2.1.4.

Proof. (a) We pick a continuous and integrable function $\psi(s)$ on the interval $[1,\infty)$ that decays faster than any negative power of s (i.e., $|\psi(s)| \leq C_N s^{-N}$ for all $N > 0$) and such that

$$\int_1^\infty s^k\,\psi(s)\,ds = \begin{cases} 1 & \text{if } k = 0, \\ 0 & \text{if } k = 1,2,3,\dots. \end{cases} \tag{2.1.20}$$

Such a function exists; see Exercise 2.1.3. In fact, we may take

$$\psi(s) = \frac{e}{\pi}\,\frac{1}{s}\,e^{-\frac{\sqrt{2}}{2}(s-1)^{\frac14}} \sin\Big(\frac{\sqrt{2}}{2}(s-1)^{\frac14}\Big). \tag{2.1.21}$$

We now define the function

$$\Phi^o(x) = \int_1^\infty \psi(s) P_s(x)\,ds\,, \tag{2.1.22}$$

where P_s is the Poisson kernel. Note that the double integral

$$\int_{\mathbf{R}^n}\int_1^\infty \frac{s}{(s^2+|x|^2)^{\frac{n+1}{2}}}\, s^{-N}\,ds\,dx$$

converges and so it follows from (2.1.20) and (2.1.22) via Fubini's theorem that

$$\int_{\mathbf{R}^n} \Phi^o(x)\,dx = 1\,.$$

Moreover, another application of Fubini's theorem yields that

$$\widehat{\Phi^o}(\xi) = \int_1^\infty \psi(s)\widehat{P_s}(\xi)\,ds = \int_1^\infty \psi(s)e^{-2\pi s|\xi|}\,ds$$

using that $\widehat{P_s}(\xi) = e^{-2\pi s|\xi|}$ (cf. Exercise 2.2.11 in [156]). This function is rapidly decreasing as $|\xi| \to \infty$ and the same is true for all the derivatives

$$\partial^\alpha \widehat{\Phi^o}(\xi) = \int_1^\infty \psi(s)\partial_\xi^\alpha \big(e^{-2\pi s|\xi|}\big)\,ds\,. \tag{2.1.23}$$

Moreover, the function $\widehat{\Phi^o}$ is clearly smooth on $\mathbf{R}^n \setminus \{0\}$ and we will show that it is also smooth at the origin. Notice that for all multi-indices α we have

$$\partial_\xi^\alpha (e^{-2\pi s|\xi|}) = s^{|\alpha|} p_\alpha(\xi)|\xi|^{-m_\alpha} e^{-2\pi s|\xi|}$$

for some $m_\alpha \in \mathbf{Z}^+$ and some polynomial $p_\alpha(\xi)$. By Taylor's theorem, for some function $v(s,|\xi|)$ with $0 \le v(s,|\xi|) \le 2\pi s|\xi|$, we have

$$e^{-2\pi s|\xi|} = \sum_{k=0}^{L} (-2\pi)^k \frac{|\xi|^k}{k!} s^k + \frac{(-2\pi s|\xi|)^{L+1}}{(L+1)!} e^{-v(s,|\xi|)}\,.$$

Choosing $L > m_\alpha$ gives

$$\partial_\xi^\alpha (e^{-2\pi s|\xi|}) = \sum_{k=0}^{L} (-2\pi)^k \frac{|\xi|^k}{k!} s^{k+|\alpha|} \frac{p_\alpha(\xi)}{|\xi|^{m_\alpha}} + s^{|\alpha|} \frac{p_\alpha(\xi)}{|\xi|^{m_\alpha}} \frac{(-2\pi s|\xi|)^{L+1}}{(L+1)!} e^{-v(s,|\xi|)},$$

which, inserted in (2.1.23) and in view of (2.1.20), yields that when $|\alpha| > 0$, the derivative $\partial^\alpha \widehat{\Phi^o}(\xi)$ tends to zero as $\xi \to 0$ and when $\alpha = 0$, $\widehat{\Phi^o}(\xi) \to 1$ as $\xi \to 0$. We conclude that $\widehat{\Phi^o}$ is continuously differentiable and hence smooth at the origin (cf. Exercise 1.1.2); hence it lies in the Schwartz class, and thus so does Φ^o.

Finally, we have the estimate

$$\begin{aligned} M(f;\Phi^o)(x) &= \sup_{t>0} |(\Phi_t^o * f)(x)| \\ &= \sup_{t>0} \left| \int_1^\infty \psi(s)(f * P_{ts})(x)\,ds \right| \\ &\le \int_1^\infty |\psi(s)|\,ds\, M(f;P)(x)\,, \end{aligned}$$

and the required conclusion follows since $\int_1^\infty |\psi(s)|\,ds \le 500$. Note that we actually obtained the stronger pointwise estimate

$$M(f;\Phi^o) \le 500\,M(f;P)$$

rather than (2.1.12).

(b) The control of the nontagential maximal function $M_a^*(\cdot\,;\Phi)$ in terms of the vertical maximal function $M(\cdot\,;\Phi)$ is the hardest and most technical part of the proof. For matters of exposition, we present the proof only in the case that $a=1$ and we note that the case of general $a>0$ presents only notational differences. We derive (2.1.13) as a consequence of the estimate

$$\big\|M_1^*(f;\Phi)\big\|_{L^p}^p \le C_2''(n,p,\Phi)^p\big\|M(f;\Phi)\big\|_{L^p}^p + \frac{1}{2}\big\|M_1^*(f;\Phi)\big\|_{L^p}^p\,, \tag{2.1.24}$$

which is useful only if we know that $\|M_1^*(f;\Phi)\|_{L^p}<\infty$. This presents a significant hurdle that needs to be overcome by an approximation. For this reason we introduce a family of maximal functions $M_1^*(f;\Phi)^{\varepsilon,N}$ for $0\le\varepsilon,N<\infty$ such that $\|M_1^*(f;\Phi)^{\varepsilon,N}\|_{L^p}<\infty$ and such that $M_1^*(f;\Phi)^{\varepsilon,N}\uparrow M_1^*(f;\Phi)$ as $\varepsilon\downarrow 0$ and we prove (2.1.24) with $M_1^*(f;\Phi)^{\varepsilon,N}$ in place of $M_1^*(f;\Phi)$. In other words we prove

$$\big\|\,M_1^*(f;\Phi)^{\varepsilon,N}\big\|_{L^p}^p \le C_2'(n,p,\Phi,N)^p\big\|M(f;\Phi)\big\|_{L^p}^p + \frac{1}{2}\big\|M_1^*(f;\Phi)^{\varepsilon,N}\big\|_{L^p}^p\,, \tag{2.1.25}$$

where there is an additional dependence on N in the constant $C_2'(n,p,\Phi,N)$, but there is no dependence on ε. The $M_1^*(f;\Phi)^{\varepsilon,N}$ are defined as follows: for a bounded distribution f in $\mathscr{S}'(\mathbf{R}^n)$ such that $M(f;\Phi)\in L^p$ we define

$$M_1^*(f;\Phi)^{\varepsilon,N}(x) = \sup_{0<t<\frac{1}{\varepsilon}}\ \sup_{|y-x|<t} \big|(\Phi_t * f)(y)\big|\Big(\frac{t}{t+\varepsilon}\Big)^N \frac{1}{(1+\varepsilon|y|)^N}\,.$$

We first show that $M_1^*(f;\Phi)^{\varepsilon,N}$ lies in $L^p(\mathbf{R}^n)\cap L^\infty(\mathbf{R}^n)$ if N is large enough depending on f. Indeed, using that $(\Phi_t * f)(x) = \langle f, \Phi_t(x-\cdot)\rangle$ and the fact that f is in $\mathscr{S}'(\mathbf{R}^n)$, we obtain constants C_f and $m=m_f$ such that:

$$\begin{aligned}
|(\Phi_t * f)(y)| &\le C_f \sum_{|\gamma|\le m,|\beta|\le m}\ \sup_{w\in\mathbf{R}^n} |w^\gamma(\partial^\beta\Phi_t)(y-w)| \\
&\le C_f \sum_{|\beta|\le m}\ \sup_{z\in\mathbf{R}^n}(1+|y|^m+|z|^m)|(\partial^\beta\Phi_t)(z)| \\
&\le C_f\,(1+|y|^m) \sum_{|\beta|\le m}\ \sup_{z\in\mathbf{R}^n}(1+|z|^m)|(\partial^\beta\Phi_t)(z)| \\
&\le C_f\,\frac{(1+|y|^m)}{\min(t^n,t^{n+m})} \sum_{|\beta|\le m}\ \sup_{z\in\mathbf{R}^n}(1+|z|^m)|(\partial^\beta\Phi)(z/t)|
\end{aligned}$$

$$\le C_f \frac{(1+|y|)^m}{\min(t^n,t^{n+m})}(1+t^m)\sum_{|\beta|\le m}\sup_{z\in\mathbf{R}^n}(1+|z/t|^m)|(\partial^\beta\Phi)(z/t)|$$
$$\le C_{f,\Phi}(1+\varepsilon|y|)^m\varepsilon^{-m}(1+t^m)(t^{-n}+t^{-n-m}).$$

Multiplying by $(\frac{t}{t+\varepsilon})^N(1+\varepsilon|y|)^{-N}$ for some $0<t<\frac{1}{\varepsilon}$ and $|y-x|<t$ yields

$$\big|(\Phi_t * f)(y)\big|\Big(\frac{t}{t+\varepsilon}\Big)^N\frac{1}{(1+\varepsilon|y|)^N}\le C_{f,\Phi}\frac{\varepsilon^{-m-N}(1+\varepsilon^{-m})(\varepsilon^{n-N}+\varepsilon^{n+m-N})}{(1+\varepsilon|y|)^{N-m}},$$

and using that $1+\varepsilon|y|\ge\frac{1}{2}(1+\varepsilon|x|)$, we obtain for some $C''(f,\Phi,\varepsilon,n,m,N)<\infty$,

$$M_1^*(f;\Phi)^{\varepsilon,N}(x)\le\frac{C''(f,\Phi,\varepsilon,n,m,N)}{(1+\varepsilon|x|)^{N-m}}.$$

Taking $N>m+n/p$, we have that $M_1^*(f;\Phi)^{\varepsilon,N}$ lies in $L^p(\mathbf{R}^n)$. This choice of N depends on m and hence on the distribution f.

We now introduce functions

$$U(f;\Phi)^{\varepsilon,N}(x)=\sup_{0<t<\frac{1}{\varepsilon}}\sup_{|y-x|<t}t\big|\nabla(\Phi_t * f)(y)\big|\Big(\frac{t}{t+\varepsilon}\Big)^N\frac{1}{(1+\varepsilon|y|)^N}$$

and

$$V(f;\Phi)^{\varepsilon,N}(x)=\sup_{0<t<\frac{1}{\varepsilon}}\sup_{y\in\mathbf{R}^n}\big|(\Phi_t * f)(y)\big|\Big(\frac{t}{t+\varepsilon}\Big)^N\frac{1}{(1+\varepsilon|y|)^N}\Big(\frac{t}{t+|x-y|}\Big)^{[\frac{2n}{p}]+1}.$$

Let $C(n)=\|M\|_{L^2(\mathbf{R}^n)\to L^2(\mathbf{R}^n)}$, where M is the Hardy–Littlewood maximal operator. We need the norm estimate

$$\big\|V(f;\Phi)^{\varepsilon,N}\big\|_{L^p}\le C(n)^{\frac{2}{p}}\big\|M_1^*(f;\Phi)^{\varepsilon,N}\big\|_{L^p} \tag{2.1.26}$$

and the pointwise estimate

$$U(f;\Phi)^{\varepsilon,N}\le A(n,p,\Phi,N)\,V(f;\Phi)^{\varepsilon,N}, \tag{2.1.27}$$

where

$$A(\Phi,N,n,p)=2^{[\frac{2n}{p}]+1}\sum_{j=1}^{n}C_0(\partial_j\Phi,N+[\tfrac{2n}{p}]+1)\,\mathfrak{N}_{N+[\frac{2n}{p}]+1}(\partial_j\Phi).$$

To prove (2.1.26) we observe that when $z\in B(y,t)\subseteq B(x,|x-y|+t)$ we have

$$\big|(\Phi_t * f)(y)\big|\Big(\frac{t}{t+\varepsilon}\Big)^N\frac{1}{(1+\varepsilon|y|)^N}\le M_1^*(f;\Phi)^{\varepsilon,N}(z),$$