1 Historical origins of the Hardy spaces $\mathcal{H}^p$

The Hardy spaces in one variable have their original setting in complex analysis. They first appeared as spaces of holomorphic functions and were introduced with the aim of characterizing boundary values of holomorphic functions on the unit disk $\mathbb{D} := \{|z| < 1\}$. Namely, let us look at the following problem: what are the possible functions $S^1 \to \mathbb{C}$ arising as boundary values of some holomorphic function $F : \mathbb{D} \to \mathbb{C}$?

This question, as just stated, is too vague: due to the lack of compactness of $\mathbb{D}$, holomorphic functions defined on $\mathbb{D}$ could exhibit a wild behaviour as we approach the boundary (for instance, we can prescribe arbitrary values of F on any discrete subset of $\mathbb{D}$). In order to obtain a meaningful notion of boundary value, it is natural to impose integrability conditions on our functions F.

As a motivation of the forthcoming definitions, let us make a *heuristic* remark: if the trace of F on $\partial\mathbb{D} = S^1$ is some complex function $f \in L^p(S^1)$, for some $1 \le p \le \infty$, then F (which is holomorphic and thus harmonic) is given by the Poisson integral of f. In polar coordinates we have the formula

$$F(re^{i\theta}) = \int_{S^1} P_r(e^{i(\theta-\eta)}) f(e^{i\eta})\, d\eta.$$

The Poisson kernel is everywhere positive and satisfies $\int_{S^1} P_r(e^{i\eta})\, d\eta = 1$ for any r, so (by Young's inequality on the group S^1)

$$\|F(r\cdot)\|_{L^p(S^1)} \le \|f\|_{L^p(S^1)}$$

and in particular all the norms in the left-hand side remain bounded as $r \uparrow 1$.

In 1915 Hardy observed that, for any holomorphic function $F : \mathbb{D} \to \mathbb{C}$, the map $r \mapsto \|F(r\cdot)\|_{L^p}$ is nondecreasing (for an arbitrary $0 < p \le \infty$). These observations lead us to define the space

$$H^p(\mathbb{D}) := \left\{ F : \mathbb{D} \to \mathbb{C} \text{ holomorphic with } \lim_{r\uparrow 1} \|F(r\cdot)\|_{L^p(S^1)} < +\infty \right\}.$$

When $p > 1$, using the weak* compactness of $L^p(S^1)$, it is not difficult to show that any $F \in H^p(\mathbb{D})$ is given by the Poisson integral of a complex-valued function $f \in L^p(S^1)$ satisfying $\widehat{f}(k) = 0$ for all $k < 0$. Conversely, given any such f, its Poisson integral lies in $H^p(\mathbb{D})$. Moreover, one can show that

$$F(r\cdot) \to f \text{ in } L^p(S^1) \quad \text{and} \quad \lim_{r\uparrow 1} F(re^{i\theta}) = f(e^{i\theta}) \text{ for a.e. } \theta,$$

so that f deserves to be regarded as the set of boundary values of F (we mention that for a.e. θ one has even a *nontangential* convergence of F to $f(e^{i\theta})$). This settles the problem for $1 < p \le \infty$. Let us also remark that the condition $\widehat{f}(k) = 0$ (for all $k < 0$) amounts to saying that $\Im(f)$, up to constants, equals the Hilbert-Riesz transform of $-\Re(f)$. When $1 < p < \infty$ the Hilbert-Riesz transform maps $L^p(S^1)$

into itself, so any function in $L^p(S^1)$ arises as the real part of the trace of some element of $H^p(\mathbb{D})$.

The case of $0 < p \leq 1$ is more difficult. F. Riesz, in a paper published in 1923, introduced the notation $H^p(\mathbb{D})$ for these spaces of holomorphic function (the letter H stands of course for Hardy) and proved many interesting properties, such as the following factorization theorem.

THEOREM 1. Any $F \in H^p(\mathbb{D})$ can be written as $F = BG$, for suitable holomorphic functions $B, G : \mathbb{D} \to \mathbb{C}$ such that $|B| \leq 1$, $G \neq 0$ everywhere and $G \in H^p(\mathbb{D})$ (B is the so-called *Blaschke product* associated to the zeros of F).

This theorem enabled him to prove the existence of a trace $f \in L^p(S^1)$ such that we have again all the convergence results mentioned before for the case $p > 1$: the trick is that, as $G \neq 0$ everywhere, one can take a k-th root of G (for an arbitrary $k > \frac{1}{p}$) and, observing that $G^{1/k} \in H^{kp}(\mathbb{D})$ and $B \in H^\infty(\mathbb{D})$, we get back to the previous case.

Again, when $p = 1$, it can be shown that the possible traces are precisely the complex-valued functions $f \in L^1(S^1)$ satisfying $\widehat{f}(k) = 0$ for all $k < 0$. As before, the possible real values of traces of functions form the set

$$\mathcal{H}^1(S^1) = \left\{ f \in L^1(S^1) : \mathcal{R}f \in L^1(S^1) \right\}$$

(here $\mathcal{R}$ denotes the Hilbert-Riesz transform). But $\mathcal{R}$ does not map $L^1(S^1)$ into itself any longer, so this set is a proper subspace of $L^1(S^1)$.

These results still hold in the upper half-plane $\mathbb{R}^2_+ := \{z = x + iy : y > 0\}$, replacing S^1 with $\mathbb{R}$ (see [2]), leading to the definition of $\mathcal{H}^1(\mathbb{R})$.

Later, in 1960, Stein and Weiss introduced the systems of conjugate harmonic functions in several variables, inspiring the correct definition of Hardy spaces in higher dimension.

The first characterization avoiding conjugate functions was provided by Burkholder, Gundy and Silverstein in 1971: they proved that a holomorphic function belongs to $\mathcal{H}^p$ if and only if the nontangential maximal function of its real part lies in L^p. The importance of this result lies in the fact that it allows to decide the membership of a function f to $\mathcal{H}^p$ by looking just at f itself.

In 1972 Fefferman and Stein, in a single pioneering paper, provided new real characterizations of the Hardy spaces, introducing different useful maximal functions and showing that the Poisson kernel can be replaced essentially by any other kernel. In this paper Fefferman and Stein also proved that singular integrals map Hardy spaces to themselves (and in particular $\mathcal{H}^1$ to L^1), as well as the duality $(\mathcal{H}^1)^* = BMO$.

The Littlewood–Paley characterization of these spaces was first given by Peetre, while the atomic decomposition was obtained by Coifman (in one dimension) and Latter (in arbitrary dimension).

2 Equivalent definitions and basic properties of $\mathcal{H}^1(\mathbb{R}^n)$

We now introduce the Hardy space $\mathcal{H}^1(\mathbb{R}^n)$, with a strong emphasis on the modern real-variable point of view outlined in the last part of [.]

This important space can be characterized in many useful ways: indeed, $\mathcal{H}^1(\mathbb{R}^n)$ is the space of all functions in $L^1(\mathbb{R}^n)$ satisfying one of the equivalent definitions provided by Theorem 2 ((8) being the closest to the historical one).

Before stating the theorem, let us recall that the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ is a Fréchet space, with the following increasing sequence of (semi)norms:

$$\|\psi\|_k := \sup_{x\in\mathbb{R}^n}(1+|x|^2)^{k/2}\sum_{|\alpha|\le k}\left|\frac{\partial^{|\alpha|}\psi}{\partial x^\alpha}\right|(x), \quad k\ge 0.$$

THEOREM 2. Fix any $\varphi \in \mathcal{S}(\mathbb{R}^n)$ with $\int \varphi(x)\,dx \neq 0$. There exists an $N \ge 0$ (independent of φ) such that the following are equivalent, for a function $f \in L^1(\mathbb{R}^n)$:

(1) the *vertical maximal function* $\mathcal{M}^v_\varphi f(x) := \sup_{t>0}|\varphi_t * f|(x)$ lies in $L^1(\mathbb{R}^n)$;

(2) the *conical maximal function*

$$\mathcal{M}^c_\varphi f(x) := \sup_{\substack{t>0,\\ y\in B_t(x)}} |\varphi_t * f|(y)$$

lies in $L^1(\mathbb{R}^n)$;

(3) the *tangential maximal function*

$$\mathcal{M}^t_\varphi f(x) := \sup_{\substack{t>0,\\ y\in\mathbb{R}^n}} |\varphi_t * f|(y)\left(1+\frac{|y-x|}{t}\right)^{-n-1}$$

lies in $L^1(\mathbb{R}^n)$;

(4) the *grand maximal function*

$$\mathcal{GM}f(x) := \sup\{|\psi_t * f|(x) \mid t>0, \psi\in\mathcal{S}(\mathbb{R}^n), \|\psi\|_N \le 1\}$$

lies in $L^1(\mathbb{R}^n)$;

(5) the similar grand maximal function

$$\mathcal{GM}'f(x) := \sup\{|\varphi_t * f|(x) \mid t>0, \psi\in C^\infty_c(B_1(0)), \|\nabla\psi\|_{L^\infty}\le 1\}$$

lies in $L^1(\mathbb{R}^n)$;

(6) there exists an *atomic decomposition*, namely there exist $\lambda_k \ge 0$ and ∞-atoms a_k (see Definition 4) such that

$$f = \sum_{k=0}^{\infty}\lambda_k a_k \text{ in } L^1(\mathbb{R}^n), \quad \sum_{k=0}^{\infty}\lambda_k < +\infty;$$

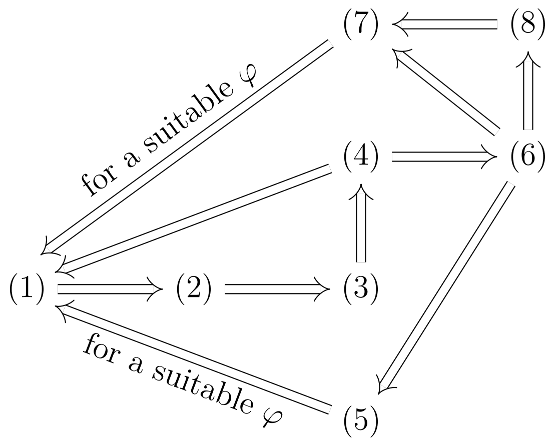
(7) the vertical maximal function with the Poisson kernel, i.e. $\mathcal{M}_P^v f$, lies in $L^1(\mathbb{R}^n)$ (notice that $P(x) = \mathcal{F}^{-1}(e^{-t|\xi|}) = \frac{c_n}{(1+|x|^2)^{(n+1)/2}} \notin \mathcal{S}(\mathbb{R}^n)$);

(8) the tempered distributions $\mathcal{R}_j f := \mathcal{F}^{-1}\left(-i\frac{\xi_j}{|\xi|}\widehat{f}(\xi)\right)$ belong to $L^1(\mathbb{R}^n)$;

(9) f belongs to the homogeneous Triebel-Lizorkin space $\dot{F}^0_{1,2}(\mathbb{R}^n)$.

Each of the preceding statements defines also a norm on $\mathcal{H}^1(\mathbb{R}^n)$: (1) defines the norm $\left\|\mathcal{M}_\varphi^v f\right\|_{L^1}$ (and similarly for (2), (3), (4), (5), (7)), (6) induces the norm $\inf \sum_k \lambda_k$ (the infimum ranging among all the possible decompositions), (8) provides the norm $\|f\|_{L^1} + \sum_{j=1}^n \|\mathcal{R}_j f\|_{L^1}$ and (9) defines $\|f\|_{\dot{F}^0_{1,2}}$ (which is a norm on $L^1(\mathbb{R}^n) \cap \dot{F}^0_{1,2}(\mathbb{R}^n)$).

The proof of this theorem is scattered across the next sections. We will show the following diagram of implications (with the corresponding bounds on the induced norms):



We left out (9) in the diagram, since its equivalence with the other definitions is slightly involved and invokes the vector-valued space $\mathcal{H}^1(\ell^2)$, which will be introduced in Section 8.

Similarly, as we will see in Section 4, the proof of (1) $\Rightarrow$ (2) is quite circuitous and uses a particular refinement of the proofs of (2) $\Rightarrow$ (3) $\Rightarrow$ (4).

Let us rely on definition (1) as for now, i.e. let $\mathcal{H}^1(\mathbb{R}^n)$ denote the space of functions $f \in L^1(\mathbb{R}^n)$ satisfying (1) and let $\|f\|_{\mathcal{H}^1} := \left\|\mathcal{M}_\varphi^v f\right\|_{L^1}$. A first basic question is whether $C_c^\infty(\mathbb{R}^n)$ is included in $\mathcal{H}^1(\mathbb{R}^n)$. Surprisingly, this property (which holds for most of the common functional spaces) fails for $\mathcal{H}^1(\mathbb{R}^n)$, as the next proposition shows.

Proposition 3. If $f \in \mathcal{H}^1(\mathbb{R}^n)$, then $\int f(x)\,dx = 0$.

Proof. Assume by contradiction that $m := \left|\int f(x)\,dx\right| \neq 0$. Choose any $x_0 \neq 0$ such that $c := |\varphi|(x_0) \neq 0$. Then we can find $R > 0$ such that

$$\left|\int_{B_R(0)} f(x)\,dx\right| \geq \frac{m}{2}, \quad \|\varphi\|_{L^\infty} \int_{\mathbb{R}^n \setminus B_R(0)} |f|(x)\,dx \leq \frac{cm}{4}.$$

For any $z \in \mathbb{R}^n$ close to 0 and any large $r > 0$ we have

$$\begin{aligned} r^n |\varphi_r * f|(r(x_0+z)) &\geq \left|\int_{B_R(0)} \varphi(x_0+z-r^{-1}y)f(y)\,dy\right| - \frac{cm}{4} \\ &\geq c\left|\int_{B_R(0)} f(y)\,dy\right| - \int_{B_R(0)} \left|\varphi(x_0+z-r^{-1}y)-\varphi(x_0)\right| |f|(y)\,dy - \frac{cm}{4} \\ &\geq \frac{cm}{4} - \|f\|_{L^1} \max_{y\in\overline{B}_R(0)} \left|\varphi(x_0+z-r^{-1}y)-\varphi(x_0)\right| - \frac{cm}{8} \geq \frac{cm}{8} \end{aligned}$$

provided that $\|f\|_{L^1} \max_{y\in\overline{B}_R(0)} |\varphi(x_0+z-r^{-1}y)-\varphi(x_0)| \leq \frac{cm}{8}$, which holds if $|z| < \epsilon$ and $r > \epsilon^{-1}$ for some small ϵ. We can assume that $\epsilon < \frac{|x_0|}{2}$. For such z, r it holds

$$\mathcal{M}^v_\varphi f(r(x_0+z)) \geq \frac{cm}{8} r^{-n} \gtrsim |r(x_0+z)|^{-n}.$$

But $E := \{r(x_0+z) \mid |z| < \epsilon, r > \epsilon^{-1}\}$ is an open cone minus a bounded set, so

$$\int \mathcal{M}^v_\varphi f(x)\,dx \geq \int_E \mathcal{M}^v_\varphi f(x)\,dx \gtrsim \int_E |x|^{-n}\,dx = +\infty,$$

contradicting the fact that $f \in \mathcal{H}^1(\mathbb{R}^n)$. □

As shown by the next proposition, the mean-zero property is the only requirement which a function in $C^\infty_c(\mathbb{R}^n)$ needs to satisfy in order to be in $\mathcal{H}^1(\mathbb{R}^n)$.

Definition 4. For any $1 < p \leq \infty$, a p-atom is a function $a \in L^p$ supported in some ball B, with zero mean and

$$\|a\|_{L^p} |B|^{1/p'} \leq 1.$$

We remark that (for $1 < p < \infty$) the last condition can be rewritten as

$$|B| \left(\fint_B |a|^p \right)^{1/p} \leq 1.$$

By Hölder's inequality, any q-atom a is also a p-atom for every $1 < p \leq q$ and $\|a\|_{L^1} \leq 1$. We now show that the $\mathcal{H}^1$-norm is bounded, as well.

Remark 5. $\mathcal{M}^v_\varphi f \lesssim Mf$ pointwise, since letting $h(x) := \max_{|x'|\geq|x|} |\varphi|(x')$ we have (noticing that h is radial and that the superlevel sets $\{h > s\}$ are either open balls or empty, for all $s > 0$)

$$\begin{aligned}|\varphi_t * f|(x) \leq& \int t^{-n} h(t^{-1}y)\,|f|(x-y)\,dy \\ =& \int h(y)\,|f|(x-ty)\,dy \\ =& \int_0^\infty \int_{\{h>s\}} |f|(x-ty)\,dy\,ds \\ \leq& Mf(x) \int_0^\infty |\{h>s\}| \\ =& Mf(x) \int h(y)\,dy \lesssim Mf(x),\end{aligned}$$

as $\int h(y)\,dy$ is finite. The same proof with P in place of f shows that $\mathcal{M}_P^v f \leq Mf$.

Proposition 6. If a is a p-atom supported in B, then $a \in \mathcal{H}^1(\mathbb{R}^n)$ and

$$\|a\|_{\mathcal{H}^1(\mathbb{R}^n)} \lesssim \|a\|_p\,|B|^{1/p'} \leq 1.$$

The implied constant depends on n, p and φ.

Proof. Let $B = B_r(x_0)$. For $x \in B_{2r}(x_0)$ we use the last remark to estimate

$$\mathcal{M}_\varphi^v a(x) \lesssim Ma(x),$$

which gives (by Hölder's inequality and Hardy-Littlewood maximal inequality)

$$\int_{B_{2r}(x_0)} \mathcal{M}_\varphi^v a(x)\,dx \lesssim \int_{B_{2r}(x_0)} Ma(x)\,dx \leq |B_{2r}(x_0)|^{1/p'}\,\|Ma\|_{L^p} \lesssim |B|^{1/p'}\,\|a\|_{L^p}.$$

For $x \notin B_{2r}(x_0)$ we use instead the mean-zero property:

$$\varphi_t * a(x) = \int (\varphi_t(x-y) - \varphi_t(x-x_0))a(y)\,dy.$$

By the mean value theorem, $|\varphi_t(x-y) - \varphi_t(x-x_0)| \leq r\,|\nabla\varphi_t|(x-z)$ for some z on the segment joining x_0 to y. So $|z-x_0| \leq r$, thus

$$|\nabla\varphi_t|(x-z) = t^{-n-1}\,|\nabla\varphi|\left(\frac{x-z}{t}\right) \lesssim t^{-n-1}\left(\frac{x-z}{t}\right)^{-n-1} \lesssim |x-x_0|^{-n-1}$$

since $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Hence,

$$\int_{\mathbb{R}^n\setminus B_{2r}(x_0)} \mathcal{M}_\varphi^v a(x)\,dx \lesssim r\,\|a\|_{L^1} \int_{\mathbb{R}^n\setminus B_{2r}(x_0)} |x-x_0|^{-n-1}\,dx \lesssim |B|^{1/p'}\,\|a\|_{L^p}.$$

□

Proposition 7. For any $f \in \mathcal{H}^1(\mathbb{R}^n)$ we have $\|f\|_{L^1} \lesssim \|f\|_{\mathcal{H}^1}$.

Proof. We assume (without loss of generality) that $\int \varphi(x)\,dx = 1$. Recall that $\lim_{h\to 0} \|f(\cdot + h) - f\|_{L^1} = 0$ for all functions $f \in L^1(\mathbb{R}^n)$. Thus,

$$\begin{aligned}\|\varphi_t * f - f\|_{L^1} &= \int \left| \int \varphi(y)\,(f(x-ty) - f(x))\,dy \right| dx \\ &\le \int \int |\varphi|\,(y)\,|f(x-ty) - f(x)|\,dx\,dy \to 0\end{aligned}$$

as $t \to 0$, by the dominated convergence theorem: indeed, the inner integral is bounded by $2\|f\|_{L^1} |\varphi|\,(y)$ and tends to 0 for all y, by the aforementioned property of functions in $L^1(\mathbb{R}^n)$. So $\|f\|_{L^1} = \lim_{t\to 0} \|\varphi_t * f\|_{L^1} \le \left\|\mathcal{M}_\varphi^v f\right\|_{\mathcal{H}^1}$. □

Proposition 8. The space $\mathcal{H}^1(\mathbb{R}^n)$ is a Banach space.

Proof. If (f_j) is a Cauchy sequence in $\mathcal{H}^1(\mathbb{R}^n)$, then we have

$$\|f_j - f_k\|_{L^1} \lesssim \|f_j - f_k\|_{\mathcal{H}^1} \to 0 \text{ as } j,k \to \infty$$

by Proposition 7, so (f_j) is a Cauchy sequence in $L^1(\mathbb{R}^n)$. Hence, $f_j \to f$ for some $f \in L^1(\mathbb{R}^n)$. For any $x \in \mathbb{R}^n$ we have

$$|\varphi_t * f|\,(x) = \lim_{j\to\infty} |\varphi_t * f_j|\,(x) \le \liminf_{j\to\infty} \mathcal{M}_\varphi^v f_j(x),$$

so $\mathcal{M}_\varphi^v f(x) \le \liminf_{j\to\infty} \mathcal{M}_\varphi^v f_j(x)$ and, by Fatou's lemma, we deduce

$$\|f\|_{\mathcal{H}^1} = \left\|\mathcal{M}_\varphi^v f\right\|_{L^1} \le \liminf_{j\to\infty} \left\|\mathcal{M}_\varphi^v f_j\right\|_{L^1} < +\infty.$$

So $f \in \mathcal{H}^1(\mathbb{R}^n)$. Moreover, since $f - f_j = \lim_{k\to\infty}(f_k - f_j)$ in $L^1(\mathbb{R}^n)$, the same argument shows that

$$\|f - f_j\|_{\mathcal{H}^1} \le \liminf_{k\to\infty} \|f_k - f_j\|_{\mathcal{H}^1}.$$

But the right-hand side can be made small at will, by taking j large enough (since (f_j) is a Cauchy sequence in $\mathcal{H}^1(\mathbb{R}^n)$). This proves that $f_j \to f$ in $\mathcal{H}^1(\mathbb{R}^n)$. □

Proposition 9. If $(f_j)_{j\in\mathbb{N}}$ is a bounded sequence in $\mathcal{H}^1(\mathbb{R}^n)$ then, up to extracting a subsequence, there exists $f \in \mathcal{H}^1(\mathbb{R}^n)$ such that $f_j \overset{*}{\rightharpoonup} f$ in $\mathcal{S}'(\mathbb{R}^n)$.

This result is related to the fact that $\mathcal{H}^1(\mathbb{R}^n)$ is a dual space. Notice that the same statement is false in $L^1(\mathbb{R}^n)$: for instance, it is easy to see that $\varphi_t \overset{*}{\rightharpoonup} (\int \varphi(x)\,dx)\delta$. In general, a distributional limit of functions in $L^1(\mathbb{R}^n)$ is a finite measure which can possess a singular part.

Proof. We assume (without loss of generality) that $\int \varphi(x)\,dx = 1$. Recall that the dual space of $C_0(\mathbb{R}^n)$, the space of continuous functions which are infinitesimal at infinity, is $C_0(\mathbb{R}^n)^* = \mathcal{M}(\mathbb{R}^n)$, the space of finite (signed) measures. $L^1(\mathbb{R}^n)$ is isometrically embedded into $\mathcal{M}(\mathbb{R}^n)$: a function $g \in L^1(\mathbb{R}^n)$ can be regarded as the finite measure $g\,dx$.

By Proposition 7, (f_j) is bounded in $L^1(\mathbb{R}^n)$ as well. Since $C_0(\mathbb{R}^n)$ is separable, by Banach-Alaoglu any closed ball in its dual is weakly* sequentially compact, so there exists a subsequence, which we still denote (f_j), and a measure $\mu \in \mathcal{M}(\mathbb{R}^n)$ such that

$$f_j\, dx \overset{*}{\rightharpoonup} \mu \text{ in } C_0(\mathbb{R}^n)^*.$$

We claim that μ is absolutely continuous with respect to the Lebesgue measure. Indeed, for any $x \in \mathbb{R}^n$ and any $t > 0$ we have

$$\varphi_t * f_j(x) = \int \varphi_t(x-y)\, f_j(y)\, dy \to \int \varphi_t(x-y)\, d\mu(y) =: \varphi_t * \mu(x).$$

Arguing as in the previous proof, we deduce

$$\sup_{t>0} |\varphi_t * \mu|(x) \le \liminf_{j\to\infty} \mathcal{M}^v_\varphi f_j =: g.$$

It is easy to check that $\varphi_t(-\cdot) * \rho \to \rho$ in $C_0(\mathbb{R}^n)$, for any $\rho \in C_0(\mathbb{R}^n)$. This implies, using Fubini's theorem, that $\varphi_t * \mu\, dx \overset{*}{\rightharpoonup} \mu$. Let now E be a Borel set with $|E| = 0$. We can find a decreasing sequence of open sets (V_k) such that $E \subseteq \cap_k V_k$ and $|V_k| = 0$. By weak* convergence we have

$$|\mu|(E) = |\mu|(V_k) \le \liminf_{t\to 0} \int_{V_k} |\varphi_t * \mu|(x)\, dx \le \int_{V_k} g(x)\, dx.$$

But $g \in L^1(\mathbb{R}^n)$ (by Fatou's lemma, since $\liminf_{j\to\infty} \left\|\mathcal{M}^v_\varphi f_j\right\|_{L^1} < +\infty$), so taking the limit as $k \to \infty$ we deduce

$$|\mu|(E) \le \lim_{k\to\infty} \int_{V_k} g(x)\, dx = 0.$$

Hence the claim is proved, i.e. $\mu = f\, dx$ for some $f \in L^1(\mathbb{R}^n)$. We deduce that $f \in \mathcal{H}^1(\mathbb{R}^n)$ as in the previous proof. The convergence $f_j \overset{*}{\rightharpoonup} f$ in $\mathcal{S}'(\mathbb{R}^n)$ follows from the fact that $\mathcal{S}(\mathbb{R}^n)$ injects continously into $C_0(\mathbb{R}^n)$. □

3 $\mathcal{H}^1 \to \mathcal{H}^1$ boundedness of Calderón-Zygmund operators

In this section we will take for granted Theorem 2 (with an abuse of notation, we will denote by $\|\cdot\|_{\mathcal{H}^1}$ any of the equivalent norms introduced above) and we will show why $\mathcal{H}^1(\mathbb{R}^n)$ is the good replacement of $L^1(\mathbb{R}^n)$ from the point of view of harmonic analysis.

Namely, its norm has the same behaviour as the L^1-norm: for any $\lambda > 0$,

$$\|f_\lambda\|_{\mathcal{H}^1} = \|f\|_{\mathcal{H}^1}$$

(to be precise, this identity becomes $\|f\|_{\mathcal{H}^1} \|f_\lambda\|_{\mathcal{H}^1} \lesssim \|f\|_{\mathcal{H}^1}$ if we use the norm given by (9), with an implied constant independent of f and λ). Furthermore, Calderón-Zygmund operators map $\mathcal{H}^1(\mathbb{R}^n)$ into itself: this property holds also for $L^p(\mathbb{R}^n)$ with $1 < p < \infty$, but it dramatically fails for $L^1(\mathbb{R}^n)$.

Let us mention that, as another confirmation of the appropriateness of Hardy spaces, if one carries over the theory into the general case of $\mathcal{H}^p(\mathbb{R}^n)$ spaces then, for $1 < p < \infty$, they collapse to $L^p(\mathbb{R}^n)$ (for which many important results in harmonic analysis already hold).

THEOREM 10. Let $K : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ be a Calderón-Zygmund kernel, i.e. a measurable function satisfying (for some finite constants $A, B > 0$)

- $|K|(x) \le A|x|^{-n}$ for all $x \in \mathbb{R}^n \setminus \{0\}$
- $\int_{|x|>2|y|} |K(x-y) - K(x)|\, dx^n \le B$ for all $y \in \mathbb{R}^n$
- $\int_{r<|x|<R} K(x)\, dx^n = 0$ for any $0 < r < R < +\infty$.

Let $K_\epsilon := K\mathbf{1}_{\mathbb{R}^n \setminus B_\epsilon(0)}$. Then, for any $f \in \mathcal{H}^1(\mathbb{R}^n)$, $K_\epsilon * f \in L^1(\mathbb{R}^n)$ and the limit

$$K * f := \lim_{\epsilon \to 0} K_\epsilon * f$$

exists in $L^1(\mathbb{R}^n)$. We have the estimate

$$\|K * f\|_{L^1} \le C(n, A, B)\, \|f\|_{\mathcal{H}^1}.$$

Proof. Recall that $K_\epsilon \in L^2$ still satisfies the above conditions (with B possibly replaced by $C(n)B$) and that $\left\|\widehat{K_\epsilon}\right\|_{L^\infty} \le C(n, A, B)$. Fix any $f \in \mathcal{H}^1(\mathbb{R}^n)$: by the characterization involving the atomic decomposition, we can find $\lambda_k \ge 0$ and ∞-atoms a_k with $f = \sum_k \lambda_k a_k$ and $\sum_k \lambda_k \lesssim \|f\|_{\mathcal{H}^1}$.

It suffices to prove the thesis for ∞-atoms: once this is done, for any $\epsilon > 0$

$$\|K_\epsilon * f\|_{L^1} \le \sum_k \lambda_k \|K_\epsilon * a_k\|_{L^1} \lesssim \sum_k \lambda_k \lesssim \|f\|_{\mathcal{H}^1}.$$

Moreover, $(K_\epsilon * f)$ is Cauchy in $\mathcal{H}^1(\mathbb{R}^n)$ as $\epsilon \to 0$: indeed, for an arbitrary N,

$$\begin{aligned}
&\limsup_{\epsilon,\epsilon' \to 0} \|K_\epsilon * f - K_{\epsilon'} * f\|_{L^1} \\
&\le \limsup_{\epsilon,\epsilon' \to 0} \sum_{k \le N} \lambda_k \|K_\epsilon * a_k - K_{\epsilon'} * a_k\|_{L^1} + \limsup_{\epsilon,\epsilon' \to 0} \sum_{k > N} \lambda_k \|K_\epsilon * a_k - K_{\epsilon'} * a_k\|_{L^1} \\
&\le 0 + \limsup_{\epsilon,\epsilon' \to 0} \sum_{k > N} \lambda_k \left(\|K_\epsilon * a_k\|_{L^1} + \|K_{\epsilon'} * a_k\|_{L^1}\right) \lesssim \sum_{k > N} \lambda_k,
\end{aligned}$$

which can be made arbitrarily small by letting $N \to +\infty$. Thus $K_\epsilon * f$ converges in $L^1(\mathbb{R}^n)$ and the limit satisfies the same estimate.

Let now a be an ∞-atom supported in $\overline{B}_R(x_0)$. Recall that

$$\|K_\epsilon * a\|_{L^2} \le C(n, A, B)\, \|a\|_{L^2} \le C(n, A, B)$$

and that $\lim_{\epsilon \to 0} K_\epsilon * a$ exists in L^2. Using Hölder's inequality we infer that $(K_\epsilon * a)\mathbf{1}_{\overline{B}_{2R}(x_0)}$ satisfies

$$\left\|(K_\epsilon * a)\mathbf{1}_{\overline{B}_{2R}(x_0)}\right\|_{L^1} \le C(n, A, B)$$

and converges in L^1 as $\epsilon \to 0$. Moreover, using the mean-zero property of a, for any $x \in \mathbb{R}^n \setminus \overline{B}_{2R}(x_0)$ we can write

$$K_\epsilon * a(x) = \int_{\mathbb{R}^n} (K_\epsilon(x-y) - K_\epsilon(x-x_0))\, a(y)\, dy^n,$$

so that

$$\begin{aligned}\int_{\mathbb{R}^n\setminus\overline{B}_{2R}(x_0)} |K_\epsilon * a|(x)\, dx^n \le& \int_{\mathbb{R}^n\setminus\overline{B}_{2R}(x_0)} \int_{\mathbb{R}^n} |K_\epsilon(x-y) - K_\epsilon(x-x_0)|\, |a(y)|\, dy^n\, dx^n \\ =& \int_{\mathbb{R}^n} \int_{\mathbb{R}^n\setminus\overline{B}_{2R}(x_0)} |K_\epsilon(x-y) - K_\epsilon(x-x_0)|\, |a(y)|\, dx^n\, dy^n \\ \le& B\, \|a\|_{L^1} \le B.\end{aligned}$$

Adding this to the preceding inequality we deduce that $\|K_\epsilon * a\|_{L^1} \le C(n, A, B)$. Finally, $(K_\epsilon * a)\mathbf{1}_{\mathbb{R}^n\setminus\overline{B}_{2R}(x_0)}$ is Cauchy in L^1 as well, since

$$(K_\epsilon * a)\mathbf{1}_{\mathbb{R}^n\setminus\overline{B}_{2R}(x_0)} = (K_{\epsilon'} * a)\mathbf{1}_{\mathbb{R}^n\setminus\overline{B}_{2R}(x_0)}$$

whenever $\epsilon, \epsilon' \le R$. □

The multiplier version of Calderón-Zygmund theorem holds as well, with the following statement.

Theorem 11. Assume $m \in C^\infty(\mathbb{R}^n \setminus \{0\})$ satisfies

$$\sup_{\xi\in\mathbb{R}^n\setminus\{0\}} |\xi|^{|\alpha|} \left| \frac{\partial^{|\alpha|} m}{\partial \xi^\alpha}(\xi) \right| < +\infty$$

for any $\alpha \in \mathbb{N}^n$. Then, for any $f \in \mathcal{H}^1(\mathbb{R}^n)$, the distribution $m\widehat{f} \in L^\infty(\mathbb{R}^n)$ lies in $\mathcal{F}(L^1(\mathbb{R}^n))$ and

$$\left\| \mathcal{F}^{-1}(m\widehat{f}) \right\|_{L^1} \lesssim \|f\|_{\mathcal{H}^1}.$$

Proof. Take an atomic decomposition $f = \sum \lambda_k a_k$ as in the preceding proof and fix a dyadic partition of unity $(\psi_\ell)_{\ell\in\mathbb{Z}}$ in $\mathbb{R}^n \setminus \{0\}$. Recall that the kernels $K_N := \mathcal{F}^{-1}\left(\sum_{\ell=-N}^{N} \psi_\ell m\right) \in \mathcal{S}(\mathbb{R}^n)$ satisfy the Hörmander condition for some constant B independent of N and have equibounded Fourier transforms. Thus we can argue as in the previous proof (without the need of truncating the kernel K_N, since it is a Schwartz function) and we get

$$\|K_N * a_k\|_{L^1} \lesssim 1.$$

But, by Plancherel's theorem, $K_N * a_k \to \mathcal{F}^{-1}(m\widehat{a}_k)$ in $L^2(\mathbb{R}^n)$ as $N \to \infty$, thus

$$\mathcal{F}^{-1}(m\widehat{a}_k) \in L^1(\mathbb{R}^n) \text{ and } \left\|\mathcal{F}^{-1}(m\widehat{a}_k)\right\|_{L^1} \lesssim 1$$

(by Fatou's lemma, since a subsequence $K_{N_j} * a_k$ converges a.e. to $\mathcal{F}^{-1}(m\widehat{a}_k)$). Thus the limit

$$g := \sum_k \lambda_k \mathcal{F}^{-1}(m\widehat{a}_k)$$

exists in $L^1(\mathbb{R}^n)$ and satisfies $\|g\|_{L^1} \lesssim \|f\|_{\mathcal{H}^1}$, as well as $\widehat{g} = \sum_k \lambda_k (m\widehat{a}_k) = m\widehat{f}$. □

Actually, in the preceding theorems we can easily upgrade the $\mathcal{H}^1 \to L^1$ boundedness to $\mathcal{H}^1 \to \mathcal{H}^1$.

Corollary 12. Under the hypotheses of Theorem 11, for any $f \in \mathcal{H}^1(\mathbb{R}^n)$ we have

$$\mathcal{F}^{-1} m\widehat{f} \in \mathcal{H}^1(\mathbb{R}^n), \quad \left\|\mathcal{F}^{-1} m\widehat{f}\right\|_{\mathcal{H}^1} \lesssim \|f\|_{\mathcal{H}^1}.$$

Proof. By the characterization of $\mathcal{H}^1(\mathbb{R}^n)$ using Riesz transforms, it suffices to show that $\mathcal{R}_j \mathcal{F}^{-1}(m\widehat{f}) \in L^1(\mathbb{R}^n)$ with an estimate on its L^1-norm (for any $1 \le j \le n$). But

$$\mathcal{R}_j \mathcal{F}^{-1}(m\widehat{f}) = \mathcal{F}^{-1}\left(-i\frac{\xi_j}{|\xi|} m(\xi)\widehat{f}(\xi)\right)$$

and the multiplier still satisfies the hypotheses of Theorem 11. □

Corollary 13. Under the hypotheses of Theorem 10, for any $f \in \mathcal{H}^1(\mathbb{R}^n)$ we have $K_\epsilon * f \in \mathcal{H}^1(\mathbb{R}^n)$ and the limit

$$K * f := \lim_{\epsilon \to 0} K_\epsilon * f$$

exists in $\mathcal{H}^1(\mathbb{R}^n)$, with the estimate

$$\|K * f\|_{\mathcal{H}^1} \le C(n, A, B) \|f\|_{\mathcal{H}^1}.$$

Proof. From Corollary 12 we know that, for any $1 \le j \le n$, $\mathcal{R}_j f \in \mathcal{H}^1(\mathbb{R}^n)$ with $\|\mathcal{R}_j f\|_{\mathcal{H}^1} \lesssim \|f\|_{\mathcal{H}^1}$. Moreover,

$$\mathcal{R}_j(K_\epsilon * f) = \mathcal{F}^{-1}\left(-i\frac{\xi_j}{|\xi|}\widehat{K}_\epsilon(\xi)\widehat{f}(\xi)\right) = K_\epsilon * (\mathcal{R}_j f),$$

so, by the conclusion of Theorem 10, $(\mathcal{R}_j(K_\epsilon * f))$ is Cauchy as $\epsilon \to 0$. As a consequence, $(K_\epsilon * f)$ is Cauchy in $\mathcal{H}^1(\mathbb{R}^n)$. □

4 Equivalence of some maximal functions

The goal of this section is to prove the equivalence among the norms defined by (1), (2), (3) and (4).

Trivially, we have

$$\mathcal{M}_\varphi^v f \lesssim \mathcal{G}\mathcal{M} f$$

pointwise (with the implied constant depending only on φ), so $\left\|\mathcal{M}_\varphi^v f \lesssim \mathcal{G}\mathcal{M} f\right\|_{L^1} \lesssim \|\mathcal{G}\mathcal{M} f\|_{L^1}$ and (4) $\Rightarrow$ (1) hold as well.

We also remark the following pointwise inequalities:

$$\mathcal{M}_\varphi^v f \le \mathcal{M}_\varphi^c f \le 2^{n+1}\mathcal{M}_\varphi^t f$$

pointwise (the second inequality follows from the fact that $2^{n+1}\left(1+\frac{|y-x|}{t}\right)^{-n-1}\geq 1$ whenever $y\in B_t(x)$).

Let us now turn to the first nontrivial inequality, namely the fact that $\left\|\mathcal{M}^t_\varphi f\right\|_{L^1}\lesssim \left\|\mathcal{M}^c_\varphi f\right\|_{L^1}$, which will give (2) $\Rightarrow$ (3).

LEMMA 14. For any $x\in\mathbb{R}^n$ we have

$$\mathcal{M}^t_\varphi f(x)\leq \left(M\left|\mathcal{M}^c_\varphi f\right|^{n/(n+1)}\right)^{(n+1)/n}(x).$$

Proof. The key observation is the fact that $|\varphi_t * f|(y)\leq \mathcal{M}^c_\varphi f(z)$ whenever $z\in B_t(y)$ (since $z\in B_t(y)$ is equivalent to $y\in B_t(z)$). From this it follows that

$$\begin{aligned}|\varphi_t * f|^{n/(n+1)}(y)\leq&\frac{1}{|B_t(y)|}\int_{B_t(y)}(\mathcal{M}^c_\varphi f)^{n/(n+1)}(z)\,dz\\ \leq&\frac{\left|B_{t+|y-x|}(x)\right|}{|B_t(y)|}\fint_{B_{t+|y-x|}(x)}(\mathcal{M}^c_\varphi f)^{n/(n+1)}(z)\,dz,\end{aligned}$$

which gives

$$|\varphi_t * f(y)|^{n/(n+1)}\left(1+\frac{|y-x|}{t}\right)^{-n}\leq M\left|\mathcal{M}^c_\varphi f\right|^{n/(n+1)}(x).$$

Raising both sides to the power $\frac{n+1}{n}$ we obtain the thesis. □

COROLLARY 15. Using the $L^{(n+1)/n}$-boundedness of the Hardy-Littlewood maximal function, we deduce

$$\left\|\mathcal{M}^t_\varphi f\right\|_{L^1}\leq\left\|M\left|\mathcal{M}^c_\varphi f\right|^{n/(n+1)}\right\|^{(n+1)/n}_{L^{(n+1)/n}}\lesssim\left\|\mathcal{M}^c_\varphi f\right\|_{L^1}.$$

Now we prove that the grand maximal function $\mathcal{GM}f$ is controlled pointwise by $\mathcal{M}^t_\varphi f$, which will trivially give (3) $\Rightarrow$ (4) and $\|\mathcal{GM}f\|_{L^1}\lesssim\left\|\mathcal{M}^t_\varphi f\right\|_{L^1}$. The choice of the seminorm $\|\cdot\|_N$ will be specified by the proof of the next lemma, which roughly says that every $\psi\in\mathcal{S}(\mathbb{R}^n)$ is a superposition of dilations of φ.

LEMMA 16. Any $\psi\in\mathcal{S}(\mathbb{R}^n)$ can be written as a series

$$\psi=\sum_{k=0}^{\infty}\eta^{(k)}*\varphi_{2^{-k}}$$

converging in $\mathcal{S}(\mathbb{R}^n)$, where the functions $\eta^{(k)}\in\mathcal{S}(\mathbb{R}^n)$ satisfy

$$\int(1+|y|)^{2(n+1)}\left|\eta^{(k)}\right|(y)\,dy\lesssim 2^{-k(n+2)}\|\psi\|_N$$

for a suitable seminorm $\|\cdot\|_N$ depending only on n (while the implied constant depends also on φ).

Proof. Let $(\rho_k)_{k\in\mathbb{N}}$ be a (inhomogeneous) dyadic partition of unity in $\mathbb{R}^n$, which can be obtained by taking $\rho_0 \in C_c^\infty(B_2)$, $\rho_0 \equiv 1$ in a neighbourhood of $\overline{B}_1$ and letting $\rho_k := \rho_0(2^{-k}\cdot) - \rho_0(2^{-(k-1)}\cdot)$ for $k > 0$ (so that, for $k > 0$, ρ_k is supported in the open annulus $B_{2^{k+1}} \setminus \overline{B}_{2^{k-1}}$).

Since $\int \varphi(x)\, dx \neq 0$, we have $\widehat{\varphi}(0) \neq 0$. By continuity we can find $k_0 \geq 0$ such that $\widehat{\varphi}(\xi) \neq 0$ for all $\xi \in \overline{B}_{2^{1-k_0}}$. For $k \geq k_0$ let $\eta^{(k)} \in \mathcal{S}(\mathbb{R}^n)$ be defined by

$$\widehat{\eta^{(k)}} := \frac{\rho_{k-k_0}\widehat{\psi}}{\widehat{\varphi}(2^{-k}\cdot)}$$

(notice that the right-hand side makes sense on $\overline{B}_{2^{k-k_0+1}}$ and vanishes near the boundary of this ball, so it can be smoothly extended by 0 on the complement). Let $\eta^{(k)} := 0$ for $k < k_0$. The series

$$\sum_{k=0}^{\infty} \widehat{\eta^{(k)}}\widehat{\varphi}(2^{-k}) = \sum_{k=k_0}^{\infty} \rho_{k-k_0}\widehat{\psi}$$

converges to $\widehat{\psi}$ in $\mathcal{S}(\mathbb{R}^n)$, so (by the continuity of $\mathcal{F}^{-1}$) we also have

$$\sum_{k=0}^{\infty} \eta^{(k)} * \varphi_{2^{-k}} = \psi$$

in $\mathcal{S}(\mathbb{R}^n)$. We can find a seminorm $\|\cdot\|_{N''}$ such that $\int (1+|y|)^{2(n+1)} |\eta|(y)\, dy \lesssim \|\widehat{\eta}\|_{N''}$, so that for $k \geq k_0$

$$\int (1+|y|)^{2(n+1)} \left|\eta^{(k)}\right|(y)\, dy \lesssim \left\| \frac{\rho_{k-k_0}\widehat{\psi}}{\widehat{\varphi}(2^{-k}\cdot)} \right\|_{N''}.$$

Using the Leibniz rule it is easy to see that, for a suitable bigger seminorm $\|\cdot\|_{N'}$ independent of φ,

$$\left\| \frac{\rho_{k-k_0}\widehat{\psi}}{\widehat{\varphi}(2^{-k}\cdot)} \right\|_{N''} \lesssim 2^{-k(n+2)} \left\|\widehat{\psi}\right\|_{N'}$$

(the implied constant, however, will depend on φ and k_0, i.e. on φ). We can finally find $\|\cdot\|_N$ such that $\left\|\widehat{\psi}\right\|_{N'} \lesssim \|\psi\|_N$. □

Corollary 17. For any $x \in \mathbb{R}^n$ we have $\mathcal{GM}f(x) \lesssim \mathcal{M}_\varphi^t f(x)$.

Proof. Fix $\psi \in \mathcal{S}(\mathbb{R}^n)$ such that $\|\psi\|_N \leq 1$. Since $\psi_t = \sum_{k=0}^\infty \varphi_{2^{-k}t} * (\eta^{(k)})_t$,

$$\begin{aligned}
|\psi_t * f|(x) &\leq \sum_{k=0}^{\infty} \left|\varphi_{2^{-k}t} * (\eta^{(k)})_t * f\right|(x) \\
&\leq \sum_{k=0}^{\infty} \int |\varphi_{2^{-k}t} * f|(x-y)\, t^{-n} \left|\eta^{(k)}\right|(t^{-1}y)\, dy \\
&\leq \mathcal{M}_\varphi^t f(x) \sum_{k=0}^{\infty} \int \left(1 + \frac{|y|}{2^{-k}t}\right)^{n+1} t^{-n} \left|\eta^{(k)}\right|(t^{-1}y)\, dy.
\end{aligned}$$

But the last integral is bounded by

$$2^{k(n+1)}\int\left(1+\frac{|y|}{t}\right)^{n+1}t^{-n}\left|\eta^{(k)}\right|(t^{-1}y)\,dy=2^{k(n+1)}\int(1+|y|)^{n+1}\left|\eta^{(k)}\right|(y)\,dy\lesssim 2^{-k}$$

for all $k\geq 0$. So we obtain $|\psi_t * f|(x)\lesssim \mathcal{M}_\varphi^t f(x)$ and the thesis follows by taking the supremum over $t>0$ and over $\psi\in\mathcal{S}(\mathbb{R}^n)$, $\|\psi\|_N\leq 1$. □

In order to prove the implication (1) ⇒ (2), we need some technical preliminaries. Fix $0<\epsilon<1$ and define the following modified maximal functions:

$$\widetilde{\mathcal{M}}_\varphi^c f(x):=\sup_{0<t<\epsilon^{-1},y\in B_t(x)}|\varphi_t * f|(y)\left(\frac{t}{t+\epsilon+\epsilon|y|}\right)^{n+1},$$

$$\widetilde{\mathcal{M}}_\varphi^t f(x):=\sup_{0<t<\epsilon^{-1},y\in\mathbb{R}^n}|\varphi_t * f|(y)\left(1+\frac{|y-x|}{t}\right)^{-n-1}\left(\frac{t}{t+\epsilon+\epsilon|y|}\right)^{n+1},$$

$$\widetilde{\mathcal{GM}}f(x):=\sup\left\{|\psi_t * f|(x)\left(\frac{t}{t+\epsilon+\epsilon|x|}\right)^{n+1}\ \middle|\ 0<t<\epsilon^{-1},\psi\in\mathcal{S}(\mathbb{R}^n),\|\psi\|_N\leq 1\right\}.$$

Clearly $\widetilde{\mathcal{M}}_\varphi^c f$ converges to $\mathcal{M}_\varphi^c f$ pointwise from below, as $\epsilon\to 0$, and most importantly it always lies in $L^1(\mathbb{R}^n)$: from $t+\epsilon|y|\geq\epsilon t+\epsilon|y|\geq\epsilon|x|$ we infer

$$\begin{aligned}|\varphi_t * f|(y)\left(\frac{t}{t+\epsilon+\epsilon|y|}\right)^{n+1}&\leq\|\varphi_t\|_{L^\infty}\|f\|_{L^1}\frac{t^{n+1}}{(\epsilon+\epsilon|x|)^{n+1}}\\&\lesssim t^{-n}\|f\|_{L^1}t^{n+1}\epsilon^{-n-1}(1+|x|)^{-n-1}\\&\leq\epsilon^{-n-2}(1+|x|)^{-n-1}\in L^1(\mathbb{R}^n).\end{aligned}$$

Lemma 18. We still have

$$\left\|\widetilde{\mathcal{GM}}f\right\|_{L^1}\lesssim\left\|\widetilde{\mathcal{M}}_\varphi^t f\right\|_{L^1}\lesssim\left\|\widetilde{\mathcal{M}}_\varphi^c f\right\|_{L^1},$$

the implied constants being independent of ϵ and f.

Proof. The second inequality is proved exactly as we did for the original maximal functions: again we have, whenever $0<t<\epsilon^{-1}$ and $z\in B_t(y)$,

$$|\varphi_t * f|(y)\left(\frac{t}{t+\epsilon+\epsilon|y|}\right)^{n+1}\leq\widetilde{\mathcal{M}}_\varphi^c f(z)$$

and, raising this inequality to the power $\frac{n}{n+1}$, averaging as z varies in $B_t(y)$ and then raising to the power $\frac{n+1}{n}$, we get again

$$\widetilde{\mathcal{M}}_\varphi^t f(x)\leq\left(M\left|\widetilde{\mathcal{M}}_\varphi^c f\right|^{n/(n+1)}\right)^{(n+1)/n}(x)$$

for any $x\in\mathbb{R}^n$, from which the second inequality follows (using the $L^{(n+1)/n}$-boundedness of the Hardy-Littlewood maximal function).

Let us turn to the first inequality. Using the decomposition $\psi_t = \sum_{k=0}^{\infty} \varphi_{2^{-k}t} * \left(\eta^{(k)}\right)_t$ we obtain again (for any $x \in \mathbb{R}^n$ and any $0 < t < \epsilon^{-1}$)

$$|\psi_t * f|(x) \leq \sum_{k=0}^{\infty} \int |\varphi_{2^{-k}t} * f|(x-y)\, t^{-n} \left|\eta^{(k)}\right|(t^{-1}y)\, dy,$$

but now we estimate (using $0 < 2^{-k}t < \epsilon^{-1}$)

$$|\varphi_{2^{-k}t} * f|(x-y) \leq \widetilde{\mathcal{M}}_\varphi^t f(x) \left(1 + \frac{|y|}{2^{-k}t}\right)^{n+1} \left(\frac{2^{-k}t + \epsilon + \epsilon|x-y|}{t}\right)^{n+1}.$$

Inserting this into the preceding inequality and multiplying both sides by $\left(\frac{t}{t+\epsilon+\epsilon|x|}\right)^{n+1}$ we arrive at

$$|\psi_t * f|(x) \left(\frac{t}{t+\epsilon+\epsilon|x|}\right)^{n+1} \leq \widetilde{\mathcal{M}}_\varphi^t f(x) \sum_{k=0}^{\infty} I_k,$$

where I_k denotes the following integral:

$$I_k := \int \left(1 + \frac{|y|}{2^{-k}t}\right)^{n+1} \left(\frac{2^{-k}t + \epsilon + \epsilon|x-y|}{t+\epsilon+\epsilon|x|}\right)^{n+1} t^{-n} \left|\eta^{(k)}\right|(t^{-1}y)\, dy.$$

The second factor in the definition of I_k is bounded by

$$\left(\frac{t+\epsilon+\epsilon|x|+\epsilon|y|}{t+\epsilon+\epsilon|x|}\right)^{n+1} = \left(1 + \frac{\epsilon|y|}{t+\epsilon+\epsilon|x|}\right)^{n+1} \leq \left(1 + \frac{|y|}{t}\right)^{n+1},$$

where we used the assumption $\epsilon < 1$, while the first factor is again bounded by $2^{k(n+1)}\left(1+\frac{|y|}{t}\right)^{n+1}$. Thus,

$$\begin{aligned} I_k \leq & 2^{k(n+1)} \int \left(1 + \frac{|y|}{t}\right)^{2(n+1)} t^{-n} \left|\eta^{(k)}\right|(t^{-1}y)\, dy \\ = & 2^{k(n+1)} \int (1+|y|)^{2(n+1)} \left|\eta^{(k)}\right|(y)\, dy \lesssim 2^{-k}, \end{aligned}$$

in view of the statement of Lemma 16. So we get

$$|\psi_t * f|(x) \left(\frac{t}{t+\epsilon+\epsilon|x|}\right)^{n+1} \lesssim \widetilde{\mathcal{M}}_\varphi^t f(x)$$

and taking the supremum over $0 < t < \epsilon^{-1}$ we obtain the pointwise bound $\widetilde{\mathcal{GM}} f(x) \lesssim \widetilde{\mathcal{M}}_\varphi^t f(x)$, from which we infer the first inequality of the thesis. □

Theorem 19. For any $0 < \epsilon < 1$ we have $\left\|\widetilde{\mathcal{M}}_\varphi^c f\right\|_{L^1} \lesssim \left\|\mathcal{M}_\varphi^v f\right\|_{L^1}$ (the implied constant is independent of ϵ).

Proof. We claim that it suffices to bound the integral $\int_E \widetilde{\mathcal{M}}^c_\varphi f(x)\,dx$ on the 'bad' set

$$E := \left\{ \widetilde{\mathcal{G}\mathcal{M}} f \le \lambda \widetilde{\mathcal{M}}^c_\varphi f \right\},$$

for some large enough λ. Indeed,

$$\int_{\mathbb{R}^n\setminus E} \widetilde{\mathcal{M}}^c_\varphi f(x)\,dx \le \lambda^{-1} \int_{\mathbb{R}^n\setminus E} \widetilde{\mathcal{G}\mathcal{M}} f(x)\,dx \le C\lambda^{-1} \int \widetilde{\mathcal{M}}^c_\varphi f(x)\,dx,$$

since by the preceding lemma $\left\|\widetilde{\mathcal{G}\mathcal{M}} f\right\|_{L^1} \le C \left\|\widetilde{\mathcal{M}}^c_\varphi f\right\|_{L^1}$ (for some C depending only on n and φ). Choosing $\lambda := 2C$ we get

$$\int_{\mathbb{R}^n\setminus E} \widetilde{\mathcal{M}}^c_\varphi f(x)\,dx \le \frac{1}{2} \int \widetilde{\mathcal{M}}^c_\varphi f(x)\,dx.$$

We can now subtract the finite quantity $\frac{1}{2}\int_{\mathbb{R}^n\setminus E} \widetilde{\mathcal{M}}^c_\varphi f(x)\,dx$ from both sides (this step is the reason why we needed to introduce these modified maximal functions: the same integral with $\mathcal{M}^c_\varphi f$ could a priori be infinite) and obtain

$$\frac{1}{2}\int_{\mathbb{R}^n\setminus E} \widetilde{\mathcal{M}}^c_\varphi f(x)\,dx \le \frac{1}{2}\int_E \widetilde{\mathcal{M}}^c_\varphi f(x)\,dx,$$

so that

$$\int \widetilde{\mathcal{M}}^c_\varphi f(x)\,dx = \int_{\mathbb{R}^n\setminus E} \widetilde{\mathcal{M}}^c_\varphi f(x)\,dx + \int_E \widetilde{\mathcal{M}}^c_\varphi f(x)\,dx \le 2\int_E \widetilde{\mathcal{M}}^c_\varphi f(x)\,dx.$$

Fix now $x \in E$ and let (y,t) such that $0 < t < \epsilon^{-1}$, $y \in B_t(x)$ and

$$|\varphi_t * f|(y) \left(\frac{t}{t+\epsilon+\epsilon|y|} \right)^{n+1} \ge \frac{1}{2} \widetilde{\mathcal{M}}^c_\varphi f(x).$$

We aim at showing that the same inequality holds, with $\frac{1}{4}$ in place of $\frac{1}{2}$, for all y' in a small ball $B_{\eta t}(y)$ ($0 < \eta < 1$ will be specified later). Once this is achieved, we will have

$$\begin{aligned}
\frac{1}{4}\widetilde{\mathcal{M}}^c_\varphi f(x) &\le \left(\fint_{B_{\eta t}} (\mathcal{M}^v_\varphi f)^{1/2}(y')\,dy' \right)^2 \\
&\le \left(\left(\frac{t+\eta t}{\eta t}\right)^n \fint_{B_{t+\eta t}(x)} (\mathcal{M}^v_\varphi f)^{1/2}(y')\,dy' \right)^2 \\
&\le \left(\frac{1+\eta}{\eta}\right)^{2n} \left(M(\mathcal{M}^v_\varphi f)^{1/2} \right)^2 (x),
\end{aligned}$$

from which the thesis follows as usual (integrating over E and using the L^2-boundedness of the Hardy-Littlewood maximal function).

Let $g(y') := \varphi_t * f(y') \left(\frac{t}{t+\epsilon+\epsilon|y|}\right)^{n+1}$. The function g is locally Lipschitz and is smooth on $\mathbb{R}^n \setminus \{0\}$, so for $y' \in B_{\eta t}(y)$

$$|g(y') - g(y)| \leq \eta t \sup_{z \in B_{\eta t}(y)\setminus\{0\}} |\nabla g|(z).$$

We compute

$$\nabla g(z) = t^{-1}(\nabla\varphi)_t * f(z) \left(\frac{t}{t+\epsilon+\epsilon|y|}\right)^{n+1} - \varphi_t * f(z) \frac{(n+1)t^{n+1}}{(t+\epsilon+\epsilon|y|)^{n+2}} \cdot \epsilon \frac{z}{|z|}.$$

But, writing $z = x + th$ (with $|h| < 1 + \eta < 2$),

$$\varphi_t * f(z) = \int t^{-n}\varphi\left(\frac{x+th-u}{t}\right) f(u)\,du = \int t^{-n}\varphi\left(\frac{x-u}{t} + h\right) f(u)\,du = (\varphi(\cdot + h))_t * f(x)$$

and similarly $(\nabla\varphi) * f(z) = (\nabla\varphi(\cdot + h))_t * f(x)$. Assuming without loss of generality $\epsilon < \frac{1}{4}$, we also have

$$t + \epsilon + \epsilon|z| \geq t + \epsilon + \epsilon(|x| - (1+\eta)t) \geq \frac{1}{2}(t + \epsilon + \epsilon|x|)$$

(as $t - \epsilon(1+\eta)t \geq \frac{t}{2}$). Putting everything together,

$$\begin{aligned}
|\nabla g|(z) \lesssim& t^{-1}\left(|\nabla\varphi * f|(x) + |\varphi * f|(z)\frac{\epsilon t}{t+\epsilon+\epsilon|z|}\right)\left(\frac{t}{t+\epsilon+\epsilon|z|}\right)^{n+1} \\
\lesssim& t^{-1}\big(|(\nabla\varphi(\cdot+h))_t * f|(x) + |(\varphi(\cdot+h))_t * f(x)|\big)\left(\frac{t}{t+\epsilon+\epsilon|x|}\right)^{n+1} \\
\lesssim& t^{-1}\widetilde{\mathcal{GM}}f(x),
\end{aligned}$$

thanks to the fact that the quantities $\sup_{|h|<2}\|\varphi(\cdot+h)\|_N$ and $\sup_{|h|<2}\left\|\frac{\partial\varphi}{\partial x_i}(\cdot+h)\right\|_N$ are finite (for $i = 1, \ldots, n$). Hence,

$$|g(y') - g(y)| \leq \eta t \cdot C' t^{-1}\widetilde{\mathcal{GM}}f(x) \leq \eta C'\lambda \widetilde{\mathcal{M}}^c_\varphi f(x)$$

(for some C' depending only on n and φ), as $x \in E$. Choosing $\eta := \min\left(\frac{1}{2}, \frac{1}{4C'\lambda}\right)$ we arrive at

$$g(y') \geq g(y) - |g(y) - g(y')| \geq \frac{1}{2}\widetilde{\mathcal{M}}^c_\varphi f(x) - \frac{1}{4}\widetilde{\mathcal{M}}^c_\varphi f(x) = \frac{1}{4}\widetilde{\mathcal{M}}^c_\varphi f(x),$$

which is what we wanted to obtain. □

5 Further remarks

We collect in this section the proofs of some easier parts of Theorem 2. By what we proved in the previous section, given $\varphi' \in \mathcal{S}(\mathbb{R}^n)(\mathbb{R}^n)$ with $\int \varphi'(x)\,dx \neq 0$, we have

$$\left\|\mathcal{M}^v_{\varphi'} f\right\|_{L^1} \lesssim \|\mathcal{GM}f\|_{L^1} \lesssim \|\mathcal{M}_\varphi f\|_{L^1}$$

and similarly $\left\|\mathcal{M}^v_\varphi f\right\|_{L^1} \lesssim \left\|\mathcal{M}^v_{\varphi'} f\right\|_{L^1}$. So $\mathcal{M}^v_\varphi f$ and $\mathcal{M}^v_{\varphi'} f$ have comparable L^1-norms. This shows that, in order to prove (5) $\Rightarrow$ (1) and (7) $\Rightarrow$ (1), we are free to choose φ at will (provided it satisfies $\int \varphi(x)\,dx \neq 0$).

Proof of (5) $\Rightarrow$ (1). As just remarked, we can assume $\varphi \in C^\infty_c(B_1)$ and $\|\nabla\varphi\|_{L^\infty} \leq 1$. The thesis follows from the trivial pointwise inequality $\mathcal{M}^v_\varphi f \leq \mathcal{G}\mathcal{M}' f$. □

Proof of (7) $\Rightarrow$ (1). First of all, we claim that there exists a continuous function $\rho : [1,+\infty) \to \mathbb{R}$ such that ρ is rapidly decreasing at infinity (i.e. $\sup_t t^k\,|\rho|\,(t) < +\infty$ for every $k \geq 0$) and

$$\int_1^\infty \rho(t)\,dt = 1, \quad \int_1^{+\infty} t^k \rho(t)\,dt = 0 \text{ for } k = 1, 2, \ldots$$

(these integrals make sense by the rapid decrease assumption on ρ).

An explicit example is the following:

$$\rho(t) := \frac{e}{\pi t}\Im\left(\exp\left(e^{3\pi i/4}(t-1)^{1/4}\right)\right).$$

The rapid decrease at infinity is clear since $|\rho|\,(t) \leq \frac{e}{\pi t}\exp\left(\Re\left(e^{3\pi i/4}(t-1)^{1/4}\right)\right) = \frac{e}{\pi t}\exp\left(-\frac{1}{\sqrt{2}}(t-1)^{1/4}\right)$. Let

$$g : \Omega := \mathbb{C}\setminus\{t\in\mathbb{R}, t\geq 1\} \to \mathbb{C}, \quad g(z) := \frac{e}{\pi}\exp\left(e^{3\pi i/4}(t-1)^{1/4}\right),$$

where $(z-1)^{1/4}$ means the unique holomorphic function $h : \Omega \to \mathbb{C}$ such that $h^4(z) = z-1$ and $\lim_{\epsilon\to 0^+} h(t+\epsilon i) = (t-1)^{1/4}$ for every real $t > 1$. We remark that $z \mapsto e^{3\pi i/4}(t-1)^{1/4}$ maps Ω into $\left\{re^{i\theta} \mid r > 0, \theta \in \left(\frac{3}{4}\pi, \frac{5}{4}\pi\right)\right\}$ and so

$$|g|\,(z) \leq \frac{e}{\pi}\exp\left(\Re\left(e^{3\pi i/4}(z-1)^{1/4}\right)\right) \leq \frac{e}{\pi}\exp\left(-\frac{1}{\sqrt{2}}\,|z-1|^{1/4}\right) \lesssim |z|^{-k}.$$

Let γ be the loop (in Ω) obtained by concatenating the parametrized paths

$$\begin{aligned} &t+\epsilon i\ (t\in[1,R]), \quad \sqrt{R^2+\epsilon^2}e^{i\alpha}\ (\alpha\in[\alpha_0, 2\pi-\alpha_0]),\\ &R-t-\epsilon i\ (t\in[0,R-1]), \quad 1+\epsilon e^{-i\alpha}\ (\alpha\in\left[\frac{\pi}{2},\frac{3\pi}{2}\right]), \end{aligned}$$

for arbitrary $\epsilon, R > 0$. By Cauchy's theorem we have $\int_\gamma z^{k-1} g(z)\,dz = 0$ for $k = 1, 2, \ldots$ and

$$\int_\gamma z^{-1} g(z)\,dz = 2\pi i g(0) = 2ei\exp\left(e^{3\pi i/4}e^{\pi i/4}\right) = 2i.$$

Taking the imaginary part of both identities, sending $\epsilon \to 0$ and then $R \to \infty$ (and noticing that the contributions of the two circular arcs are infinitesimal), we get precisely

$$2\int_1^\infty t^k\rho(t)\,dt = 0 \text{ for } k = 1, 2, \ldots, \quad 2\int_1^\infty \rho(t)\,dt = 2,$$

which is the claim.

We now build a Schwartz function out of the Poisson kernel: let

$$\varphi(x) := \int_1^\infty \rho(t) P_t(x)\, dt.$$

This integral converges (as $|P_t|(x) = t^{-n}P(t^{-1}x) \le t^{-n}P(0) \le P(0)$) and defines a function in $L^1(\mathbb{R}^n)$, since

$$\int_{\mathbb{R}^n} |\varphi|(x)\, dx \le \int_1^\infty \int_{\mathbb{R}^n} |\rho|(t) P_t(x)\, dx\, dt = \int_1^\infty |\rho|(t)\, dt < +\infty.$$

Moreover, using Fubini's theorem, $\widehat{\varphi}(\xi) = \int_1^\infty \rho(t) e^{-t|\xi|}\, dt$. It is easy to show inductively that, for $\xi \neq 0$ and any multiindex $\alpha \neq 0$,

$$\frac{\partial^{|\alpha|}\widehat{\varphi}(\xi)}{\partial \xi^\alpha} = \int_1^\infty \rho(t) \cdot t\, Q_\alpha(t, \xi, |\xi|^{-1}) e^{-t|\xi|}\, dt$$

for a suitable polynomial Q_α. In particular, $\widehat{\varphi}$ is smooth on $\mathbb{R}^n \setminus \{0\}$ and all its derivatives are rapidly decreasing at infinity. Moreover, $\widehat{\varphi}$ is clearly continuous. Given any $\alpha \neq 0$, we write

$$e^{-s} = \sum_{k<K} \frac{(-s)^k}{k!} + R_K(s)$$

and notice that $s^{-K}|R_K|(s)$ is bounded for $s \in \mathbb{R} \setminus \{0\}$ close to the origin, while it is infinitesimal as $|s| \to \infty$; thus $|R_K|(s) \lesssim |s|^K$. This implies

$$\frac{\partial^{|\alpha|}\widehat{\varphi}(\xi)}{\partial \xi^\alpha} = \int_1^\infty \rho(t) \cdot t\, Q_\alpha(t, \xi, |\xi|^{-1}) \left(\sum_{k<K} \frac{(-t|\xi|)^k}{k!} + R_K(t|\xi|) \right) dt.$$

Calling d and d' the degrees of Q_α with respect to its first and last argument, respectively, we obtain that for every $K > d'$

$$\int_1^\infty \left|\rho(t) \cdot t\, Q_\alpha(t, \xi, |\xi|^{-1})\right| |R_K|(t|\xi|)\, dt \lesssim \int_1^\infty t^{1+d+K} |\xi|\, dt \lesssim |\xi|$$

(whenever $0 < |\xi| \le 1$), while

$$\int_1^\infty \rho(t) \cdot t\, Q_\alpha(t, \xi, |\xi|^{-1}) \sum_{k<K} \frac{(-t|\xi|)^k}{k!}\, dt = 0$$

by the special properties satisfied by η. This shows that all the derivatives of $\widehat{\varphi}$ extend continuously up to the origin, hence $\widehat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$ and we deduce that $\varphi \in \mathcal{S}(\mathbb{R}^n)$, as well. Finally, $\int_{\mathbb{R}^n} \varphi(x)\, dx = \int_1^\infty \int_{\mathbb{R}^n} \rho(t) P_t(x)\, dx = \int_1^\infty \rho(t)\, dt = 1$ and

$$\mathcal{M}^v_\varphi f(x) \le \int_1^\infty |\rho|(t) \mathcal{M}^v_P f(x)\, dt \lesssim \mathcal{M}^v_P f(x)$$

for any $f \in L^1(\mathbb{R}^n)$, showing that (7) $\Rightarrow$ (1) for this particular φ. □

Proof of (6) ⇒ (5). Let $\psi \in C_c^\infty(B_1(0))$ with $\|\nabla\psi\|_{L^\infty} \le 1$. Given an ∞-atom supported in $B_r(x_0)$, we have

$$|\psi_t * a|(x) \le \|\psi_t\|_{L^1}\|a\|_{L^\infty} = \|\psi\|_{L^1}\|a\|_{L^\infty} \lesssim \|a\|_{L^\infty}$$

for any $x \in B_{2r}(x_0)$. Fix now $x \in \mathbb{R}^n \setminus B_{2r}(x_0)$ and notice that $\psi_t * a(x) = 0$ if $t < |x - x_0| - r$ (since in this case $\psi_t(x - \cdot)$ and a are supported in the disjoint balls $\overline{B}_t(x)$ and $\overline{B}_r(x_0)$). Assume instead that $t \ge |x - x_0| - r$: in this case we get $t \ge \frac{|x-x_0|}{2}$, so

$$\begin{aligned}|\psi_t * a|(x) \le& \int |\psi_t(x-y) - \psi_t(x-x_0)|\,|a|(y)\,dy\\ \le& r\,\|\nabla\psi_t\|_{L^\infty}\|a\|_{L^1}\\ \lesssim& rt^{-n-1} \lesssim \frac{r}{|x-x_0|^{n+1}}\end{aligned}$$

(as $\nabla\psi_t(x) = t^{-n-1}\nabla\psi(t^{-1}x)$). Thus,

$$\begin{aligned}\|\mathcal{G}\mathcal{M}'f\|_{L^1} \lesssim& \int_{B_{2r}(x_0)} \|a\|_{L^\infty}\,dx + \int_{\mathbb{R}^n\setminus B_{2r}(x_0)} \frac{r}{|x-x_0|^{n+1}}\,dx\\ \lesssim& \|a\|_{L^\infty}|B_r(x_0)| + r\int_{2r}^\infty \frac{\rho^{n-1}\,d\rho}{\rho^{n+1}} \lesssim 1.\end{aligned}$$

Hence, if $f = \sum_k \lambda_k a_k$ is an atomic decomposition,

$$\|\mathcal{G}\mathcal{M}'f\|_{L^1} \le \left\|\sum_k \lambda_k \mathcal{G}\mathcal{M}'a_k\right\|_{L^1} \lesssim \sum_k \lambda_k.$$

□

Proof of (6) ⇒ (7). The proof of Proposition 6 can be repeated verbatim, with φ replaced by P, to show that

$$\|\mathcal{M}_P^v a\|_{L^1} \lesssim 1$$

for any ∞-atom a. Hence, if $f = \sum_k \lambda_k a_k$ is an atomic decomposition,

$$\|\mathcal{M}_P^v f\|_{L^1} \le \left\|\sum_k \lambda_k \mathcal{M}_P^v a_k\right\|_{L^1} \lesssim \sum_k \lambda_k.$$

□

Proof of (6) ⇒ (8). It suffices to notice that the proof of Theorem 11 used only the atomic decomposition of f. So, choosing $m(\xi) := -i\frac{\xi_j}{|\xi|}$, we deduce

$$\|\mathcal{R}_j f\|_{L^1} \lesssim \inf \sum \lambda_k$$

(the infimum ranging over all the possible atomic decompositions). Moreover, for any decomposition

$$\|f\|_{L^1} \le \sum_k \lambda_k \|a_k\|_{L^1} \le \sum_k \lambda_k.$$

Thus, $\|f\|_{L^1} + \sum_{j=1}^n \|\mathcal{R}_j f\|_{L^1} \lesssim \inf \sum_k \lambda_k$. □

6 Characterization with the Riesz transforms

We now show the implication (8) $\Rightarrow$ (7) among the equivalent definitions of $\mathcal{H}^1(\mathbb{R}^n)$. The proof will implicitly show the corresponding inequality on the norms, namely

$$\|\mathcal{M}_P^v f\|_{L^1} \lesssim \|f\|_{L^1} + \|\mathcal{R}_1 f\|_{L^1} + \cdots + \|\mathcal{R}_n f\|_{L^1}.$$

Assume that f and all its Riesz transforms are in $L^1(\mathbb{R}^n)$. So far we have tacitly allowed any function to be either real or complex valued, but now it is convenient to assume f real valued (without loss of generality, as $\mathcal{R}_j$ maps real functions to real distributions). The functions

$$u_j(x,t) := P_t * \mathcal{R}_j f(x) \text{ for } 1 \leq j \leq n, \quad u_{n+1}(x,t) := P_t * f(x)$$

form a *system of conjugate harmonic functions* on $\mathbb{H}^{n+1} := \{(x,t) \in \mathbb{R}^{n+1} : t > 0\}$, i.e. they satisfy the following system of *generalized Cauchy-Riemann equations*:

$$\sum_{j=1}^{n+1} \frac{\partial u_j}{\partial x_j} = 0, \quad \frac{\partial u_j}{\partial x_k} = \frac{\partial u_k}{\partial x_j} \text{ for any } 1 \leq j, k \leq n+1,$$

where x_{n+1} is an alias for the auxiliary variable t. This can be checked using the formulas

$$\mathcal{F}(P_t * f)(\xi) = (2\pi)^{-n/2} e^{-t|\xi|} \widehat{f}(\xi), \quad \mathcal{F}(P_t * \mathcal{R}_j f)(\xi) = -(2\pi)^{-n/2} i \frac{\xi_j}{|\xi|} e^{-t|\xi|} \widehat{f}(\xi).$$

Clearly it suffices to prove that

$$\sup_{t>0} |u|(\cdot, t) \in L^1(\mathbb{R}^n),$$

where $u := (u_1, \ldots, u_{n+1})$. We could bound $|u|(x,t)$ by the Hardy-Littlewood maximal function of $(f, \mathcal{R}_1 f, \ldots, \mathcal{R}_n f)$ at x, but this would be useless (as the Hardy-Littlewood maximal function satisfies only a weak (1,1) bound). Rather, we aim at showing that $|u|^q (x,t) \lesssim Mg(x)$ for some $q < 1$ and some $g \in L^{1/q}(\mathbb{R}^n)$ with

$$\|g\|_{L^{1/q}}^{1/q} \lesssim \|f\|_{L^1} + \|\mathcal{R}_1 f\|_{L^1} + \cdots + \|\mathcal{R}_n f\|_{L^1},$$

from which the thesis will follow since

$$\left\| \sup_{t>0} |u|(\cdot, t) \right\|_{L^1} \lesssim \|Mg\|_{L^{1/q}}^{1/q} \lesssim \|g\|_{L^{1/q}}^{1/q}.$$

Pick now $\frac{n-1}{n} < q < 1$ (so that in particular $q > 0$). From Lemma 20 below, we know that $(|u|^2 + \epsilon^2)^{q/}2$ is subharmonic (for any $\epsilon > 0$). Thus it satisfies the following version of the maximum principle: for any $\Omega \Subset \mathbb{H}^{n+1}$ and any $h \in C^0(\overline{\Omega})$ harmonic in Ω, the implication

$$(|u|^2 + \epsilon^2)^{q/2} \leq h \text{ on } \partial\Omega \ \Rightarrow \ (|u|^2 + \epsilon^2)^{q/2} \leq h \text{ on } \overline{\Omega}$$

holds. Sending $\epsilon \to 0$, it is easy to check that $|u|^q$ satisfies the same property. Lemma 21 below tells us that this property applies also with the harmonic function $h(x,t) := P_{t-\delta} * |u|^q (x,\delta)$ on the unbounded domain $\{(x,t) : t > \delta\} \subseteq \mathbb{H}^{n+1}$, for any $\delta > 0$.

Notice that

$$\sup_{\delta>0} \||u|^q (\cdot,\delta)\|_{L^{1/q}}^{1/q} = \sup_{\delta>0} \|u(\cdot,\delta)\|_{L^1} \leq \|f\|_{L^1} + \|\mathcal{R}_1 f\|_{L^1} + \cdots + \|\mathcal{R}_n f\|_{L^1}.$$

Since any closed ball in $L^{1/q}$ is weakly sequentially compact, we can find a sequence $\delta_k \to 0$ and a function $g \in L^{1/q}(\mathbb{R}^n)$ (whose $L^{1/q}$-norm satisfies the same bound) such that $|u|^q (\cdot,\delta_k) \rightharpoonup g$. Since $P_{t-\delta_k} \to P_t$ in $L^{(1/q)'}$, we deduce that

$$|u|^q (x,t) \leq \lim_{k\to\infty} (P_{t-\delta_k} * |u|^q (\cdot,\delta_k))(x) = P_t * g(x).$$

Finally, by Remark 5, we have $P_t * g \leq Mg$, which was our goal. It remains to prove the two lemmas.

Lemma 20. For any $q \geq \frac{n-1}{n}$ the function $(|u|^2 + \epsilon^2)^{q/2}$ is subharmonic, i.e.

$$\Delta\left((|u|^2 + \epsilon^2)^{q/2}\right) \geq 0.$$

Proof. We will use the shorthand notation $\partial_j := \frac{\partial}{\partial x_j}$. We compute

$$\partial_j (|u|^2 + \epsilon^2)^{q/2} = q(|u|^2 + \epsilon^2)^{(q/2)-1} u \cdot \partial_j u,$$

$$\sum_j \partial_{jj}^2 (|u|^2+\epsilon^2)^{q/2} = \sum_j q(q-2)(|u|^2+\epsilon^2)^{(q/2)-2}(u\cdot\partial_j u)^2 + \sum_j q(|u|^2+\epsilon^2)^{(q/2)-1} |\partial_j u|^2$$

(using $\Delta u = 0$). The thesis follows immediately if $q \geq 2$, so we can assume $\frac{n-1}{n} \leq q < 2$, i.e. $0 < 2 - q \leq \frac{n+1}{n}$. It suffices to prove that

$$\frac{n+1}{n} \sum_j (u \cdot \partial_j u)^2 \leq |u|^2 \sum_j |\partial_j u|^2.$$

This inequality is a consequence of the generalized Cauchy-Riemann equations: indeed, the matrix $A := (\partial_j u_k(x))_{jk}$ is symmetric, so (by the spectral theorem) we can find $P \in \mathbb{O}(n+1)$ and a diagonal matrix D such that $A = P^t D P$. The coefficients on the diagonal of D are the eigenvalues $\lambda_1, \ldots, \lambda_{n+1}$ of A. We remark that

$$\sum_j \lambda_j = \mathrm{tr}(D) = \mathrm{tr}(A) = 0.$$

We pick j_0 such that $|\lambda_{j_0}| = \max_j |\lambda_j|$. By Cauchy-Schwarz we have

$$(n+1) |\lambda_{j_0}|^2 = n |\lambda_{j_0}| + \left|\sum_{j\neq j_0} \lambda_j\right|^2 \leq n \sum_j |\lambda_j|^2,$$

so, letting $v := P \begin{pmatrix} u_1(x) \\ \vdots \end{pmatrix}$, we can estimate

$$\begin{aligned} \frac{n+1}{n} \sum_j |u \cdot \partial_j u|^2 =& \frac{n+1}{n} \sum_j \left| A \begin{pmatrix} u_1(x) \\ \vdots \end{pmatrix} \right|^2 = \frac{n+1}{n} |Dv|^2 \leq \frac{n+1}{n} |\lambda_{j_0}|^2 |v|^2 \\ \leq& \sum_j |\lambda_j|^2 |u|^2 (x). \end{aligned}$$

We finally observe that

$$\sum_j |\lambda_j|^2 = \operatorname{tr}(D^t D) = \operatorname{tr}(PA^t P^t PAP^t) = \operatorname{tr}(A^t A) = \sum_j |\partial_j u|^2 .$$

□

Lemma 21. For $t > \delta$ we have $|u|^q (x,t) \leq (P_{t-\delta} * |u|^q (\cdot, \delta))(x)$.

Proof. Let us first prove that, for every $\eta > 0$, there exists an arbitrarily large radius $R > 0$ such that $|u| \leq \eta$ on the set $\{(x,t) : t \geq \delta, |(x,t)| \geq R\}$. From the mean-value property of harmonic functions, for any $(x,t))$ in this set we have

$$|u| (x,t) \leq \frac{1}{\left|B_{t/2}(x,t)\right|} \int_{B_{t/2}(x,t)} |u| (y,s) \, dy \, ds.$$

If $|x| \leq t$ then $t \geq \frac{R}{\sqrt{2}}$ and we can estimate

$$|u| (x,t) \leq \frac{1}{\left|B_{t/2}(x,t)\right|} \int_{\mathbb{R}^n \times \left(\frac{t}{2}, \frac{3}{2}t\right)} |u| (y,s) \, dy \, ds \lesssim At^{-n} \lesssim AR^{-n}$$

(where $A := \sup_{s>0} \|u(\cdot,s)\|_{L^1} < +\infty$), which becomes small at will taking R large enough. Otherwise, if $|x| > t$, then $|x| \geq \frac{R}{\sqrt{2}}$ and any point $(y,s) \in B_{t/2}(x,t)$ satisfies $|y| > \frac{|x|}{2}$, so

$$|u| (x,t) \lesssim t^{-n-1} \int_{t/2}^{3t/2} \int_{|y|>|x|/2} |u| (y,s) \, dy \, ds \lesssim \int_{t/2}^{3t/2} s^{-n-1} \int_{|y|>R/\sqrt{8}} |u| (y,s) \, dy \, ds.$$

But the latter quantity can be uniformly estimated by

$$\int_{\delta/2}^{\infty} s^{-n-1} \int_{|y|>R/\sqrt{8}} |u| (y,s) \, dy \, ds,$$

which can be made arbitrarily small taking R large enough, thanks to the dominated convergence theorem (since the inner integral is bounded by A and tends to 0 as $R \to +\infty$).

Now $h(x,t) := (P_{t-\delta} * |u|^q (\cdot, \delta))(x)$ is harmonic on $\{(x,t) : t > \delta\}$ and extends continuously to the boundary $\mathbb{R}^n \times \{\delta\}$, where it coincides with $|u|^q$. So we have proved that

$$|u|^q (x,t) \leq (P_{t-\delta} * |u|^q (\cdot, \delta))(x) + \eta^q$$

on the boundary of $S_R := \{(x,t) : t > \delta, |(x,t)| < R\}$ for any R large enough. We deduce that this inequality is also true on S_R itself. Thus, letting $R \to +\infty$, we infer that it holds on $\{(x,t) : t > \delta\}$. The thesis follows as we let $\eta \to 0$. □

7 Existence of the atomic decomposition

In this section we show that any function $f \in L^1(\mathbb{R}^n)$ with $\mathcal{GM}f \in L^1(\mathbb{R}^n)$ admits an atomic decomposition

$$f = \sum_{k=0} \lambda_k a_k$$

with $\lambda_k \geq 0$, (a_k) a collection of ∞-atoms and $\sum_k \lambda_k \lesssim \|\mathcal{GM}f\|_{L^1}$, thereby proving the implication (4) $\Rightarrow$ (6) and the bound on the corresponding norms.

[$\cdots$ work in progress $\cdots$]

8 Littlewood-Paley characterization

In this section we are going to prove that $\mathcal{H}^1(\mathbb{R}^n) = \dot{F}^0_{1,2}(\mathbb{R}^n)$, in the sense specified by Theorem 23, denoting by $\mathcal{H}^1(\mathbb{R}^n)$ the space of functions satisfying (any of) the definitions (1) – (8), whose equivalence has been established in the previous sections.

We fix a function $\psi \in C^\infty_c(B_2 \setminus \overline{B}_{1/2})$ such that $\sum_{j\in\mathbb{Z}} \psi(2^{-j}\xi) = 1$ for any $\xi \in \mathbb{R}^n \setminus \{0\}$. Recall that such a ψ can be produced by taking any $\phi \in C^\infty_c(B_2)$ such that $\phi \equiv 1$ in a neighbourhood of $\overline{B}_1$ and letting $\psi := \phi - \phi(2\cdot)$.

We let $\varphi \in \mathcal{S}(\mathbb{R}^n)$ be any function such that $\int_{\mathbb{R}^n} \varphi \, dx^n \neq 0$ and $\operatorname{supp}(\varphi) \subseteq B_2$. For instance we can take $\varphi := \mathcal{F}^{-1}(\phi)$ for any ϕ as above (using $\int_{\mathbb{R}^n} \varphi \, dx^n = (2\pi)^{-n/2}\phi(0) \neq 0$).

Lemma 22. For any $f \in \mathcal{S}'(\mathbb{R}^n)$ and any $r \in (0, \infty)$ we have

$$\sup_{t>0} |\varphi_t * P_j f|(x) \leq C(n,r) M |P_j f|^r (x)^{1/r}.$$

Proof. Recall that, whenever $v \in \mathcal{S}'(\mathbb{R}^n)$ has its Fourier transform supported in B_1, we have the inequality

$$\sup_z \frac{|v(x-z)|}{(1+|z|)^{n/r}} \lesssim (M|v|^r)^{1/r}(x).$$

More generally, if $\widehat{u}$ is supported in B_s, letting $v := u(s^{-1}\cdot)$ we obtain

$$\sup_z \frac{|u(x-z)|}{(1+s|z|)^{n/r}} = \sup_z \frac{|v(sx-z)|}{(1+|z|)^{n/r}} \lesssim (M|v|^r)^{1/r}(sx) = (M|u|^r)^{1/r}(x).$$

If $\frac{2}{t} \leq 2^{j-1}$ (i.e. if $t \geq 2^{2-j}$) we have $\varphi_t * P_j f \equiv 0$, since the supports of $\widehat{\varphi}_t$ and $\psi(2^{-j}\cdot)$ are disjoint in this case. Assume now that $t \leq 2^{2-j}$: in this case $\psi(2^{-j}\cdot)$ is supported in $B_{8/t}$, so choosing any $N > \frac{n}{r} + n$ and estimating $|\varphi(z)| \lesssim (1+|z|)^{-N}$ we get

$$|\varphi_t * P_j f|(x) \lesssim \int t^{-n} \frac{|P_j f|(x-z)}{(1+t^{-1}|z|)^N} \, dz \leq \sup_z \frac{|P_j f|(x-z)}{(1+t^{-1}|z|)^{n/r}} \int \frac{t^{-n}}{(1+t^{-1}|z|)^{N-n/r}} \, dz.$$

The last integral is a finite constant independent of t, while

$$\sup_z \frac{|P_j f|(x-z)}{(1+t^{-1}|z|)^{n/r}} \lesssim \sup_z \frac{|P_j f|(x-z)}{(1+\frac{8}{t}|z|)^{n/r}} \lesssim (M|P_j f|^r)^{1/r}(x).$$

□

Before stating and proving the next theorem, we introduce the vector-valued Hardy space $\mathcal{H}^1(\mathbb{R}^n, \ell^2)$: it is the subspace of

$$L^1(\mathbb{R}^n, \ell^2) := \left\{ (f_j)_{j\in\mathbb{Z}} \subseteq L^1(\mathbb{R}^n) : \int \left(\sum_j |f_j|^2 \right)^{1/2} < +\infty \right\}$$

made of elements (f_j) satisfying one of the equivalent definitions $(1)-(7)$ in vectorized form. For instance, (1) amounts to ask that

$$\sup_{t>0} \|\varphi_t * (f_j)\|_{\ell^2} = \sup_{t>0} \|(\varphi_t * f_j)\|_{\ell^2} \in L^1(\mathbb{R}^n).$$

Their equivalence comes from the fact that the proofs for the scalar case can be repeated verbatim for the vectorial case (we exclude definition (8) since its equivalence with the other definitions uses real numbers in an essential way, due to the appeal to the spectral theorem).

Theorem 23. For any $f \in \mathcal{H}^1(\mathbb{R}^n)$ we have

$$\|(P_j f)_{j\in\mathbb{Z}}\|_{L^1(\ell^2)} \lesssim \|f\|_{\mathcal{H}^1}.$$

Conversely, if for some $f \in \mathcal{S}'(\mathbb{R}^n)$ we have $\|(P_j f)_{j\in\mathbb{Z}}\|_{L^1(\ell^2)} < \infty$, then there exists a unique polynomial Q such that $f - Q \in \mathcal{H}^1(\mathbb{R}^n)$; moreover

$$\|f - Q\|_{\mathcal{H}^1(\mathbb{R}^n)} \lesssim \|(P_j f)_{j\in\mathbb{Z}}\|_{L^1(\ell^2)}.$$

Proof. The first statement follows immediately from the $\mathcal{H}^1 \to L^1(\ell^2)$ version of Theorem 10, applied with $K_j := \mathcal{F}^{-1}(\eta(2^{-j}\cdot))$, with assumptions (1) and (3) replaced by the validity of the $L^2 \to L^2(\ell^2)$ bound (which holds as a consequence of Plancherel's theorem). This variant of Theorem 10 is simply obtained by vectorizing the same proof (and in this case there is no need of truncating the kernel). Recall that this ℓ^2-valued kernel satisfies the Hörmander condition

$$\int_{|x|>2|y|} \|(K_j(x-y) - K_j(x))\|_{\ell^2} \, dx^n \lesssim 1.$$

We now turn to the converse. Pick $\eta := \psi(2\cdot) + \psi + \psi\left(\frac{\cdot}{2}\right)$ and notice that $\eta \equiv 1$ near the support of ψ. Let

$$\tilde{P}_j : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n), \quad \tilde{P}_j(g) := \mathcal{F}^{-1}(\eta(2^{-j}\cdot)\mathcal{F}g)$$

and remark that $\tilde{P}_j P_j = P_j$. Applying the $\mathcal{H}^1(\ell^2) \to L^1$ version of Theorem 10 with $f := (P_j f)_{j\in\mathbb{Z}}$ and $K_j := \mathcal{F}^{-1}(\eta(2^{-j}\cdot))$, we can estimate

$$\left\| \sum_{j=-N}^{N} P_j f \right\|_{L^1} = \left\| \sum_{j=-N}^{N} \tilde{P}_j P_j f \right\|_{L^1} \lesssim \left\| \sum_{j=-N}^{N} P_j f \right\|_{\mathcal{H}^1(\ell^2)} .$$

We can similarly estimate the L^1-norm of $\mathcal{R}_k \sum_{j=-N}^{N} P_j f$, using the ℓ^2-valued kernel $K_j := \mathcal{F}^{-1}\left(-i\frac{\xi_k}{|\xi|}\eta(2^{-j}\cdot)\right)$. Thus,

$$\left\| \sum_{j=-N}^{N} P_j f \right\|_{\mathcal{H}^1} \lesssim \left\| \sum_{j=-N}^{N} P_j f \right\|_{\mathcal{H}^1(\ell^2)} = \left\| \sup_{t>0} \left(\sum_{j=-N}^{N} |\varphi_t * P_j f|^2 \right)^{1/2} \right\|_{L^1} .$$

Using Lemma 22 with any $0 < r < 1$, as well as the Hardy-Littlewood maximal estimate for $L^{1/r}(\ell^{2/r})$, the last quantity is bounded up to constants by

$$\begin{aligned} \left\| \left(\sum_{j=-N}^{N} (M\, |P_j f|^r)^{2/r} \right)^{1/2} \right\|_{L^1} &= \|(M\, |P_j f|^r)\|_{L^{1/r}\left(\ell_N^{2/r}\right)}^{1/r} \lesssim \|(|P_j f|^r)\|_{L^{1/r}\left(\ell_N^{2/r}\right)}^{1/r} \\ &= \|(P_j f)\|_{L^1(\ell^2)} , \end{aligned}$$

where ℓ_N^p denotes the truncated space of sequences $a = (a_{-N}, \ldots, a_N)$ with the norm $\|a\|_{\ell_N^p} := \left(\sum_{j=-N}^{N} |a_j|^p\right)^{1/p}$. The same argument shows that the partial sums $\sum_{j=-N}^{N} P_j f$ form a Cauchy sequence in $\mathcal{H}^1(\mathbb{R}^n)$ and thus, by Proposition 8, converge to some $g \in \mathcal{H}^1(\mathbb{R}^n)$.

But $\mathcal{F}\left(\sum_{j=-N}^{N} P_j f\right) \to \widehat{f}$ in $\mathcal{D}'(\mathbb{R}^n \setminus \{0\})$, so the tempered distribution $\widehat{f} - \widehat{g}$ is supported in $\{0\}$. This means that

$$Q := f - g = \mathcal{F}^{-1}\left(\widehat{f} - \widehat{g}\right)$$

is a polynomial. So $f - Q = g \in \mathcal{H}^1(\mathbb{R}^n)$ and, letting $N \to \infty$ in the above estimate, we also have

$$\|f - Q\|_{\mathcal{H}^1} = \|g\|_{\mathcal{H}^1} \lesssim \|(P_j f)\|_{L^1(\ell^2)} .$$

□

References

[1] D.L. Burkholder, R.F. Gundy, and M. L. Silverstein. “A maximal function characterization of the class H^p”. In: *Transactions of the American Mathematical Society* 157 (1971), pp. 137–153.

[2] P.L. Duren. *Theory of H^p spaces*. Courier Corporation, 2000.

[3] C. Fefferman and E.M. Stein. "H^p spaces of several variables". In: *Acta mathematica* 129.1 (1972), pp. 137–193.
[4] L. Grafakos. *Modern Fourier Analysis*. Vol. 250. Springer-Verlag New York, 2009. doi: 10.1007/978-0-387-09434-2.
[5] G.H. Hardy. "The mean value of the modulus of an analytic function". In: *Proceedings of the London Mathematical Society* 2.1 (1915), pp. 269–277.
[6] F. Riesz. "Über die Randwerte einer analytischen Funktion". In: *Mathematische Zeitschrift* 18.1 (1923), pp. 87–95.
[7] W. Rudin. *Real and complex analysis*. McGraw-Hill, 1987.
[8] E.M. Stein. *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*. Vol. 43. Princeton University Press, 1993.
[9] E.M. Stein and G. Weiss. "On the theory of harmonic functions of several variables". In: *Acta Mathematica* 103.1-2 (1960), pp. 25–62.
[10] H. Triebel. *Theory of function spaces*. Birkhäuser Basel, 1983. doi: 10.1007/978-3-0346-0416-1.