GRADIENT BOUNDS FOR
RADIAL MAXIMAL FUNCTIONS

Emanuel Carneiro and Cristian González-Riquelme

ICTP - The Abdus Salam International Centre for Theoretical Physics
Strada Costiera, 11, I - 34151, Trieste, Italy; carneiro@ictp.it, and
IMPA - Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina, 110
Jardim Botânico, Rio de Janeiro - RJ, Brazil, 22460-320; carneiro@impa.br

IMPA - Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina, 110
Jardim Botânico, Rio de Janeiro - RJ, Brazil, 22460-320; cristian@impa.br

Abstract. In this paper we study the regularity properties of certain maximal operators of convolution type at the endpoint $p = 1$, when acting on radial data. In particular, for the heat flow maximal operator and the Poisson maximal operator, when the initial datum $u_0 \in W^{1,1}(\mathbf{R}^d)$ is a radial function, we show that the associated maximal function u^* is weakly differentiable and

$$\|\nabla u^*\|_{L^1(\mathbf{R}^d)} \lesssim_d \|\nabla u_0\|_{L^1(\mathbf{R}^d)}.$$

This establishes the analogue of a recent result of Luiro for the uncentered Hardy–Littlewood maximal operator, now in a centered setting with smooth kernels. In a second part of the paper, we establish similar gradient bounds for maximal operators on the sphere $\mathbf{S}^d$, when acting on polar functions. Our study includes the uncentered Hardy–Littlewood maximal operator, the heat flow maximal operator and the Poisson maximal operator on $\mathbf{S}^d$.

1. Introduction

1.1. A brief historical perspective. Maximal operators are central objects of study in harmonic analysis. One of the most basic examples is the centered Hardy–Littlewood maximal operator, denoted here by M. For $f \in L^1_{\text{loc}}(\mathbf{R}^d)$ it is defined as

$$Mf(x) = \sup_{r>0} \frac{1}{m(B_r(x))} \int_{B_r(x)} |f(y)| \, \mathrm{d}y = \sup_{r>0} \fint_{B_r(x)} |f(y)| \, \mathrm{d}y,$$

where $B_r(x) \subset \mathbf{R}^d$ is the open ball centered at x with radius r, and $m(B_r(x))$ denotes its d-dimensional Lebesgue measure. The uncentered Hardy–Littlewood maximal operator, denoted here by $\widetilde{M}$, is defined analogously, taking the supremum over open balls that simply contain the point x but that are not necessarily centered at x. The fundamental theorem of Hardy, Littlewood and Wiener states that $M \colon L^1(\mathbf{R}^d) \to L^{1,\infty}(\mathbf{R}^d)$ and $M \colon L^p(\mathbf{R}^d) \to L^p(\mathbf{R}^d)$, for $1 < p \leq \infty$, are bounded operators. The same holds for $\widetilde{M}$.

In the seminal paper [14], Kinnunen studied the action of the Hardy–Littlewood maximal operator on Sobolev functions, giving an elegant proof that $M \colon W^{1,p}(\mathbf{R}^d) \to W^{1,p}(\mathbf{R}^d)$ is bounded for $1 < p \leq \infty$. This work paved the way for several interesting contributions to the regularity theory of maximal operators over the past two decades, with interesting connections to potential theory and partial differential equations, see for instance [1, 3, 4, 6, 7, 8, 11, 13, 15, 16, 18, 19, 22, 23, 21, 24, 25, 26]. One of the longstanding problems in this field is concerned with the regularity at the endpoint

$p = 1$. This is the $W^{1,1}$-problem, formally posed by Hajłasz and Onninen in [12]: if $f \in W^{1,1}(\mathbf{R}^d)$, do we have that Mf is weakly differentiable and

$$\|\nabla Mf\|_{L^1(\mathbf{R}^d)} \lesssim_d \|\nabla f\|_{L^1(\mathbf{R}^d)}?$$

This problem has been settled affirmatively in dimension $d = 1$, in the uncentered case by Tanaka [29] and Aldaz and Pérez Lázaro [2], and in the centered case by Kurka [17]. The higher dimensional version is generally open, having been settled affirmatively only for the *uncentered* Hardy–Littlewood maximal operator $\widetilde{M}$ in the important case of *radial datum* f, by Luiro in [20]. This beautiful work of Luiro [20] is fundamental for the present paper, for we aim to extend it to different contexts.

1.2. Maximal operators of convolution type on $\mathbf{R}^d$. We start by investigating the higher dimensional $W^{1,1}$-problem for certain centered maximal operators of convolution type associated to partial differential equations, *in the case of radial data*, establishing a result analogous to that of Luiro [20]. To our knowledge, this is the first instance of an affirmative result for centered maximal operators, in what concerns the boundedness of the variation, in the higher dimensional setting.

We borrow the basic setup from [5]. Let $\varphi \colon \mathbf{R}^d \times (0, \infty) \to \mathbf{R}$ be a nonnegative function such that

$$\int_{\mathbf{R}^d} \varphi(x, t)\, \mathrm{d}x = 1$$

for each $t > 0$. Assume also that, when $t \to 0$, the family $\varphi(\cdot, t)$ is an approximation of the identity, in the sense that $\lim_{t \to 0} \varphi(\cdot, t) * u_0(x) = u_0(x)$ for a.e. $x \in \mathbf{R}^d$, if $u_0 \in L^p(\mathbf{R}^d)$ for some $1 \le p \le \infty$. For an initial datum $u_0 \colon \mathbf{R}^d \to \mathbf{R}$ we consider the evolution

$$(1.1) \qquad u(x, t) = \big(|u_0| * \varphi(\cdot, t)\big)(x)$$

and the associated maximal function

$$(1.2) \qquad u^*(x) := \sup_{t > 0} u(x, t).$$

Notice the use of the shorter notation u^* for simplicity. One could also refer to (1.2) as $M_\varphi u_0$. In this setting, note that the centered Hardy–Littlewood maximal operator corresponds to the kernel $\varphi(x, t) = \frac{1}{t^d m(B_1)} \chi_{B_1}(x/t)$. We consider here kernels $\varphi_{a,b}$ that are fundamental solutions of

$$au_{tt} - bu_t + \Delta u = 0 \quad \text{in } \mathbf{R}^d \times (0, \infty),$$

with $a, b \ge 0$ and $(a, b) \ne (0, 0)$. That is, the function $u(x, t)$ defined in (1.1) solves this equation in the upper half-space with initial datum $u(x, 0) = |u_0(x)|$. By appropriate space-time dilations it suffices to consider the following three nonnegative and radial decreasing kernels as basic profiles:

$$(1.3) \qquad \varphi_{1,0}(x, t) = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\pi^{(d+1)/2}} \frac{t}{(|x|^2 + t^2)^{(d+1)/2}} \quad \text{(Poisson kernel)},$$

$$(1.4) \qquad \varphi_{0,1}(x, t) = \frac{1}{(4\pi t)^{d/2}} e^{-|x|^2/4t} \quad \text{(Heat kernel)},$$

$$(1.5) \qquad \varphi_{1,1}(x, t) = \int_{\mathbf{R}^d} e^{-t\left(\frac{-1+\sqrt{1+16\pi^2|\xi|^2}}{2}\right)} e^{2\pi i x \cdot \xi}\, \mathrm{d}\xi.$$

The fact that (1.5) is nonnegative and radial decreasing was proved in [5]. The *Poisson maximal operator* and the *heat flow maximal operator*, given by the kernels (1.3) and (1.4) respectively, are the classical and most important examples we want

to keep in mind, but our methods could be adapted to treat other maximal operators associated to differential equations. Our first result is the following.

Theorem 1. *Let φ be given by* (1.3), (1.4) *or* (1.5). *If $u_0 \in W^{1,1}(\mathbf{R}^d)$ is radial, then u^* is weakly differentiable and*

$$\|\nabla u^*\|_{L^1(\mathbf{R}^d)} \lesssim_d \|\nabla u_0\|_{L^1(\mathbf{R}^d)}.$$

The intuitive idea behind the proof of this result is as follows. First we reduce matters to the study of nonnegative functions u_0 with some degree of smoothness, say Lipschitz. We are then able to invoke one of the main results of [5, 9], that in the *detachment set* $\{u^* > |u_0|\}$ the function u^* is *subharmonic*. The proof of this fact relies on some of the qualitative properties of the underlying partial differential equations (e.g. maximum principles and semigroup property). As observed in [5, Theorem 1 (iv)], this subharmonicity implies a control on the L^2-norm of ∇u^* by the L^2-norm of ∇u_0. To arrive at the L^1-control we use the fact that u^* is pointwise smaller than $\widetilde{M}u_0$. Hence, in the case of radial functions, we have a relatively well-behaved (i.e. subharmonic in the detachment set) function, namely u^*, that is trapped between u_0 and $\widetilde{M}u_0$, and the latter comes with an L^1-control of the gradient by the result of Luiro [20]. As we shall see, these pieces together will ultimately imply the control of the L^1-norm of ∇u^* as well.

1.3. The Hardy–Littlewood maximal operator on $\mathbf{S}^d$. We now move our discussion to consider maximal operators acting on functions defined on the sphere $\mathbf{S}^d \subset \mathbf{R}^{d+1}$, in order to develop an analogous theory. First, let us establish the basic notation to be used in this context. We let $d(\zeta, \eta)$ denote the geodesic distance between two points $\zeta, \eta \in \mathbf{S}^d$. Let $\mathcal{B}_r(\zeta) \subset \mathbf{S}^d$ be the open geodesic ball of center $\zeta \in \mathbf{S}^d$ and radius $r > 0$, that is

$$\mathcal{B}_r(\zeta) = \{\eta \in \mathbf{S}^d \colon d(\zeta, \eta) < r\},$$

and let $\overline{\mathcal{B}_r(\zeta)}$ be the corresponding closed ball. Let $\widetilde{\mathcal{M}}$ denote the uncentered Hardy–Littlewood maximal operator on the sphere $\mathbf{S}^d$, that is, for $f \in L^1_{\mathrm{loc}}(\mathbf{S}^d)$,

$$\widetilde{\mathcal{M}}f(\xi) = \sup_{\{\mathcal{B}_r(\zeta) \colon \xi \in \overline{\mathcal{B}_r(\zeta)}\}} \frac{1}{\sigma(\mathcal{B}_r(\zeta))} \int_{\mathcal{B}_r(\zeta)} |f(\eta)| \, \mathrm{d}\sigma(\eta) = \sup_{\{\overline{\mathcal{B}_r(\zeta)} \colon \xi \in \overline{\mathcal{B}_r(\zeta)}\}} \fint_{\mathcal{B}_r(\zeta)} |f(\eta)| \, \mathrm{d}\sigma(\eta),$$

where $\sigma = \sigma_d$ denotes the usual surface measure on the sphere $\mathbf{S}^d$. The centered version $\mathcal{M}$ would be defined with centered geodesic balls. Fix $\mathbf{e} = (1, 0, 0, \dots, 0) \in \mathbf{R}^{d+1}$ to be our north pole. We say that a function $f \colon \mathbf{S}^d \to \mathbf{C}$ is *polar* if for every $\xi, \eta \in \mathbf{S}^d$ with $\xi \cdot \mathbf{e} = \eta \cdot \mathbf{e}$ we have $f(\xi) = f(\eta)$. This will be the analogue, in the spherical setting, of a radial function in the Euclidean setting.

When working on the circle $\mathbf{S}^1$, an adaptation of the proof of Aldaz and Pérez Lázaro [2] yields $\mathrm{Var}(\widetilde{\mathcal{M}}f) \le \mathrm{Var}(f)$, where $\mathrm{Var}(f)$ denotes the total variation of the function f. This follows from the fact that $\widetilde{\mathcal{M}}f$ has no local maxima in the detachment set $\{\widetilde{\mathcal{M}}f > |f|\}$ (say, for f Lipschitz). Our second result is the extension of this statement to the multidimensional setting, in the case of polar functions. For the basic theory of Sobolev spaces on the sphere $\mathbf{S}^d$ we refer the reader to [10].

Theorem 2. *If $f \in W^{1,1}(\mathbf{S}^d)$ is a polar function, then $\widetilde{\mathcal{M}}f$ is weakly differentiable and*

$$\|\nabla \widetilde{\mathcal{M}}f\|_{L^1(\mathbf{S}^d)} \lesssim_d \|\nabla f\|_{L^1(\mathbf{S}^d)}.$$

This is the analogue on the sphere $\mathbf{S}^d$ of Luiro's result [20] for radial functions in the Euclidean space. The proof we present below follows broadly the strategy outlined by Luiro [20]. However, due to the different geometry, several nontrivial technical points arise along the proof and must be considered carefully. A good example that such difficulties cannot be underestimated is Lemma 12 below, one of the core results used in our proof of Theorem 2. As in the case of $\mathbf{R}^d$, the analogue of Theorem 2 for the centered Hardy–Littlewood maximal operator $\mathcal{M}$ on $\mathbf{S}^d$ is an open problem.

1.4. Maximal operators of convolution type on $\mathbf{S}^d$. We now treat two important cases of maximal operators of convolution type on the sphere: the Poisson maximal operator and the heat flow maximal operator. We briefly recall the basic definitions and refer the reader to [5, Section 1.4] for additional details.

1.4.1. Poisson maximal function on $\mathbf{S}^d$. Let $0 \leq \rho < 1$ and let $\xi, \eta \in \mathbf{S}^d$. We define the Poisson kernel $\mathcal{P}$ on the sphere by

$$\mathcal{P}(\xi, \eta, \rho) = \frac{1 - \rho^2}{\kappa_d \, |\rho\xi - \eta|^d} = \frac{1 - \rho^2}{\kappa_d \, (\rho^2 - 2\rho \, \xi \cdot \eta + 1)^{d/2}},$$

with $\kappa_d = \sigma(\mathbf{S}^d)$ being the total surface area of $\mathbf{S}^d$. If $u_0 \in L^1(\mathbf{S}^d)$ we let $u(\xi, \rho) = u(\rho\xi)$ be the function defined on the unit $(d+1)$-dimensional open ball $B_1 \subset \mathbf{R}^{d+1}$ by

$$u(\xi, \rho) = \int_{\mathbf{S}^d} \mathcal{P}(\xi, \eta, \rho) \, |u_0(\eta)| \, \mathrm{d}\sigma(\eta),$$

and consider the associated maximal function

$$(1.6) \qquad\qquad u^*(\xi) = \sup_{0 \leq \rho < 1} u(\xi, \rho).$$

Observe that $u \in C^\infty(B_1)$ and solves the Dirichlet problem

$$\begin{cases} \Delta u = 0 & \text{in } B_1; \\ \lim_{\rho \to 1^-} u(\xi, \rho) = |u_0(\xi)| & \text{for a.e. } \xi \in \mathbf{S}^d. \end{cases}$$

1.4.2. Heat flow maximal function on $\mathbf{S}^d$. Let $\{Y_n^\ell\}$, $\ell = 1, 2, \ldots, \dim H_n^{d+1}$, be an orthonormal basis of the space H_n^{d+1} of spherical harmonics of degree n in the sphere $\mathbf{S}^d$. For $t \in (0, \infty)$ and $\xi, \eta \in \mathbf{S}^d$ we define the heat kernel $\mathcal{K}$ on the sphere (see [10, Lemma 1.2.3, Theorem 1.2.6 and Eq. 7.5.5]) by

$$\mathcal{K}(\xi, \eta, t) = \sum_{n=0}^\infty e^{-tn(n+d-1)} \sum_{\ell=1}^{\dim H_n^{d+1}} Y_n^\ell(\xi) Y_n^\ell(\eta) = \sum_{n=0}^\infty e^{-tn(n+d-1)} \frac{(n+\lambda)}{\lambda} C_n^\lambda(\xi \cdot \eta),$$

where $\lambda = \frac{d-1}{2}$ and $t \mapsto C_n^\beta(t)$, for $\beta > 0$, are the *Gegenbauer polynomials* defined in terms of the generating function

$$(1 - 2rt + r^2)^{-\beta} = \sum_{n=0}^\infty C_n^\beta(t) \, r^n.$$

If $u_0 \in L^1(\mathbf{S}^d)$ we consider

$$u(\xi, t) = \int_{\mathbf{S}^d} \mathcal{K}(\xi, \eta, t) \, |u_0(\eta)| \, \mathrm{d}\sigma(\eta),$$

and consider the associated maximal function

$$(1.7) \qquad u^*(\xi) = \sup_{t>0} u(\xi, t).$$

Note that u is a smooth function on $\mathbf{S}^d \times (0, \infty)$ and solves the heat equation

$$\begin{cases} \partial_t u - \Delta u = 0 & \text{in } \mathbf{S}^d \times (0, \infty); \\ \lim_{t \to 0^+} u(\xi, t) = |u_0(\xi)| & \text{for a.e. } \xi \in \mathbf{S}^d. \end{cases}$$

1.4.3. Gradient bounds. We note that the smooth kernels $\mathcal{P}$ and $\mathcal{K}$ depend only on $d(\xi, \eta)$ and are decreasing with respect to this distance. If we fix one of these two parameters, they have integral 1 on $\mathbf{S}^d$ and are approximate identities as $\rho \to 1^-$ and $t \to 0^+$, respectively. The discussion on the heat kernel can be found in [27, Chapter III, Section 2]. Also, from [10, Chapter 2, Theorem 2.3.6], note that the associated maximal functions u^* are dominated by the Hardy–Littlewood maximal function, that is

$$(1.8) \qquad u^*(\xi) \le \mathcal{M}u_0(\xi) \le \widetilde{\mathcal{M}}u_0(\xi).$$

Our third result establishes the following.

Theorem 3. *Let u^* be the Poisson maximal function given by (1.6) or the heat flow maximal function given by (1.7). If $u_0 \in W^{1,1}(\mathbf{S}^d)$ is a polar function, then u^* is weakly differentiable and*

$$\|\nabla u^*\|_{L^1(\mathbf{S}^d)} \lesssim_d \|\nabla u_0\|_{L^1(\mathbf{S}^d)}.$$

1.5. A word on notation. In what follows we write $A \lesssim_d B$ if $A \le CB$ for a certain constant $C > 0$ that may depend on the dimension d. We say that $A \simeq_d B$ if $A \lesssim_d B$ and $B \lesssim_d A$. If there are other parameters of dependence, they will also be indicated. The characteristic function of a generic set H is denoted by χ_H. In the few occasions that we write universal constants C_d in Section 4, these may change from line to line.

2. Proof of Theorem 1

In this section we prove Theorem 1. Without loss of generality we may assume that u_0 is *real-valued and nonnegative* (or $+\infty$). Assume also that $d \ge 2$, since the result is already known for dimension $d = 1$ from [5, Theorem 1]. Throughout the proof below, with a slight abuse of notation, we identify radial functions of the variable $x \in \mathbf{R}^d$ with their one-dimensional versions of the variable $r \in (0, \infty)$, with the understanding that $r = |x|$. Naturally, if u_0 is radial, the maximal function u^* is also radial. In what follows, variables r, s, t, τ, a, b will be one-dimensional, whereas the variable x is always reserved for $\mathbf{R}^d$. We recall the fact [28, Chapter III, Theorem 2] that

$$(2.1) \qquad u^*(x) \le Mu_0(x) \le \widetilde{M}u_0(x)$$

for every $x \in \mathbf{R}^d$.

2.1. Lipschitz case. Let us first assume that our initial datum u_0 is a Lipschitz function. In this case u^* is also Lipschitz. Reducing matters to radial variables, we claim the following:

$$(2.2) \qquad \int_0^\infty \left|(u^*)'(r)\right| r^{d-1}\,\mathrm{d}r \le \int_0^\infty \left|u_0'(r)\right| r^{d-1}\,\mathrm{d}r + \int_0^\infty \left|\big(\widetilde{M}u_0\big)'(r)\right| r^{d-1}\,\mathrm{d}r.$$

Once we have established (2.2), the theorem follows easily by Luiro's result [20], that bounds the third integral in terms of the second.

Step 1. Partial control by the uncentered maximal function. Let us define the radial detachment set (excluding the origin)

$$A_d = \big\{ x \in \mathbf{R}^d \setminus \{0\} \colon u^*(x) > u_0(x) \big\}. \tag{2.3}$$

The one-dimensional radial version of this set will be denoted by

$$A_1 = \big\{ |x| \colon x \in A_d \big\}.$$

These are open sets and from [5, Lemma 7] we know that u^* is subharmonic on A_d. Let us write

$$A_1 = \bigcup_{i=1}^{\infty} (a_i, b_i) \tag{2.4}$$

as a countable union of disjoint open intervals. Let (a, b) denote a generic interval (a_i, b_i) of this union. If u^* had a strict local maximum in (a, b) (that is, a point $t_0 \in (a, b)$ for which there exist c and d with $a < c < t_0 < d < b$ such that $u^*(r) \le u^*(t_0)$ for $r \in (c, d)$ and $u^*(c), u^*(d) < u^*(t_0)$), we could then take the average of u^* over the ball in $\mathbf{R}^d$ centered at x_0, with $|x_0| = t_0$, and radius $\min\{|t_0 - c|, |t_0 - d|\}$ to reach a contradiction to the subharmonicity of u^* in A_d. Therefore u^* has no strict local maximum in (a, b) and there exists τ with $a \le \tau \le b$ such that u^* is non-increasing in $[a, \tau]$ and non-decreasing in $[\tau, b]$. We then have $(u^*)'(t) \le 0$ a.e. in $a < t < \tau$, and $(u^*)'(t) \ge 0$ a.e. in $\tau < t < b$.

Let us first consider the case $0 < a < b < \infty$. Using (2.1) and integration by parts we get

$$\int_a^b \big| (u^*)'(r) \big| \, r^{d-1} \, \mathrm{d}r = -\int_a^\tau (u^*)'(r) \, r^{d-1} \, \mathrm{d}r + \int_\tau^b (u^*)'(r) \, r^{d-1} \, \mathrm{d}r$$

$$= u^*(a) \, a^{d-1} + u^*(b) \, b^{d-1} - 2 \, u^*(\tau) \, \tau^{d-1}$$

$$+ (d-1) \int_a^\tau u^*(r) \, r^{d-2} \, \mathrm{d}r - (d-1) \int_\tau^b u^*(r) \, r^{d-2} \, \mathrm{d}r$$

$$\le u_0(a) \, a^{d-1} + u_0(b) \, b^{d-1} - 2 \, u_0(\tau) \, \tau^{d-1}$$

$$+ (d-1) \int_a^\tau \widetilde{M}u_0(r) \, r^{d-2} \, \mathrm{d}r - (d-1) \int_\tau^b u_0(r) \, r^{d-2} \, \mathrm{d}r$$

$$= u_0(a) \, a^{d-1} - u_0(\tau) \, \tau^{d-1}$$

$$+ (d-1) \int_a^\tau \widetilde{M}u_0(r) \, r^{d-2} \, \mathrm{d}r + \int_\tau^b u_0'(r) \, r^{d-1} \, \mathrm{d}r$$

$$\le \int_a^b \big| u_0'(r) \big| \, r^{d-1} \, \mathrm{d}r + (d-1) \int_a^\tau \widetilde{M}u_0(r) \, r^{d-2} \, \mathrm{d}r. \tag{2.5}$$

The last inequality holds since

$$u_0(a) \, a^{d-1} - u_0(\tau) \, \tau^{d-1} \le -\int_a^\tau u_0'(r) \, r^{d-1} \, \mathrm{d}r \le \int_a^\tau \big| u_0'(r) \big| \, r^{d-1} \, \mathrm{d}r.$$

If $b = \infty$, since $u^* \in L^{1,\infty}(\mathbf{R}^d)$ we must have $\tau = \infty$ as well (i.e. u^* non-increasing in the interval (a, ∞)) and a simple limiting argument leads to inequality (2.5) again. Note that $\lim_{r \to \infty} u_0(r) \, r^{d-1} = 0$ since $r \mapsto u_0(r) \, r^{d-1}$ is locally Lipschitz with integrable derivative in $(0, \infty)$.

Finally, if $a = 0$, the proof of (2.5) follows as above noting that $\lim_{r \to 0} u^*(r) \, r^{d-1} = 0$ (for $d \geq 2$).

If we add up (2.5) over all the intervals (a_i, b_i) of the disjoint union (2.4) we find

$$\int_{A_1} \left|(u^*)'(r)\right| r^{d-1} \, \mathrm{d}r \leq \int_{A_1} \left|u_0'(r)\right| r^{d-1} \, \mathrm{d}r + (d-1) \int_0^\infty \widetilde{M}u_0(r) \, r^{d-2} \, \mathrm{d}r,$$

which then leads to (note that in A_1^c we have $u^* = u_0$, and hence $(u^*)' = u_0'$ a.e. in A_1^c)

$$(2.6) \qquad \int_0^\infty \left|(u^*)'(r)\right| r^{d-1} \, \mathrm{d}r \leq \int_0^\infty \left|u_0'(r)\right| r^{d-1} \, \mathrm{d}r + (d-1) \int_0^\infty \widetilde{M}u_0(r) \, r^{d-2} \, \mathrm{d}r.$$

Step 2. Control of weighted norms. As $r \mapsto \widetilde{M}u_0(r)$ is Lipschitz and its derivative is integrable (in fact $\left(\widetilde{M}u_0\right)'(r) \, r^{d-1} \in L^1(0, \infty)$ from Luiro's work [20]) we have that $\lim_{r \to \infty} \widetilde{M}u_0(r)$ exists and it is equal to 0 since $\widetilde{M}u_0 \in L^{1,\infty}(\mathbf{R}^d)$. Then

$$\widetilde{M}u_0(r) = -\int_r^\infty \left(\widetilde{M}u_0\right)'(t) \, \mathrm{d}t$$

and

$$
\begin{aligned}
(2.7) \qquad (d-1) \int_0^\infty \widetilde{M}u_0(r) \, r^{d-2} \, \mathrm{d}r &= (d-1) \int_0^\infty \left(\int_r^\infty -\left(\widetilde{M}u_0\right)'(t) \, \mathrm{d}t \right) r^{d-2} \, \mathrm{d}r \\
&\leq (d-1) \int_0^\infty \left(\int_r^\infty \left|\left(\widetilde{M}u_0\right)'(t)\right| \, \mathrm{d}t \right) r^{d-2} \, \mathrm{d}r \\
&= (d-1) \int_0^\infty \int_0^t r^{d-2} \left|\left(\widetilde{M}u_0\right)'(t)\right| \, \mathrm{d}r \, \mathrm{d}t \\
&= \int_0^\infty \left|\left(\widetilde{M}u_0\right)'(t)\right| t^{d-1} \, \mathrm{d}t.
\end{aligned}
$$

Finally, we combine (2.6) and (2.7) to arrive at (2.2), concluding the proof in this case.

2.2. General case. Let us first record a basic lemma about radial functions and weak derivatives. In what follows, when we say that a function f is weakly differentiable in a certain domain $\Omega \subset \mathbf{R}^d$, it is naturally understood that f and its weak derivatives are locally integrable in such a domain.

Lemma 4. .

(i) *A radial function $f(x)$ is weakly differentiable in $\mathbf{R}^d \setminus \{0\}$ if and only if its radial restriction $f(r)$ is weakly differentiable in $(0, \infty)$. In this case, the weak gradient ∇f of $f(x)$ and the weak derivative f' of $f(r)$ are related by $\nabla f(x) = f'(|x|)\frac{x}{|x|}$.*

(ii) *In the situation above, if $f(x)$ and $\nabla f(x)$ are locally integrable in a neighborhood of the origin, then f is weakly differentiable in $\mathbf{R}^d$.*

Proof. This result is most certainly standard but we could not find an exact explicit reference. We then provide a brief proof for completeness.

Part (i). Assume that $f(x)$ is weakly differentiable in $\mathbf{R}^d \setminus \{0\}$ and let ∇f be its weak gradient. Let $\varphi \in C_c^\infty(\mathbf{R}^d \setminus \{0\})$ be a radial test function. Letting $r = |x|$ we

have, by definition,

$$\int_{\mathbf{R}^d\setminus\{0\}} f(x)\left(\frac{(d-1)}{|x|}\varphi(x) + \frac{\partial\varphi}{\partial r}(x)\right)\mathrm{d}x$$

$$(2.8)\qquad = \int_{\mathbf{R}^d\setminus\{0\}} f(x)\left(\sum_{i=1}^d \frac{\partial}{\partial x_i}\left(\frac{x_i}{|x|}\varphi(x)\right)\right)\mathrm{d}x$$

$$= -\int_{\mathbf{R}^d\setminus\{0\}}\left(\sum_{i=1}^d \frac{\partial f}{\partial x_i}\frac{x_i}{|x|}\varphi(x)\right)\mathrm{d}x = -\int_{\mathbf{R}^d\setminus\{0\}} \nabla f(x)\cdot\frac{x}{|x|}\,\varphi(x)\,\mathrm{d}x.$$

Write $x = r\omega$, with $\omega\in\mathbf{S}^{d-1}$. Letting $\Phi(r) = \varphi(r)\,r^{d-1}$, rewrite (2.8) in polar coordinates to get

$$\sigma_{d-1}\big(\mathbf{S}^{d-1}\big)\int_0^\infty f(r)\,\Phi'(r)\,\mathrm{d}r = -\int_0^\infty\left(\int_{\mathbf{S}^{d-1}}(\nabla f(r\omega))\cdot\omega\,\mathrm{d}\sigma_{d-1}(\omega)\right)\Phi(r)\,\mathrm{d}r.$$

This is the required integration by parts in $(0,\infty)$ for the generic test function Φ.

   Assume now that $f(r)$ is weakly differentiable in $(0,\infty)$. If g is its weak derivative, then $f(r) - \int_1^r g(t)\,\mathrm{d}t$ has weak derivative zero and hence is constant a.e. in $(0,\infty)$. We can then modify f on a set of measure zero so that f is continuous in $(0,\infty)$; in fact absolutely continuous in each interval $[a,b]\subset(0,\infty)$. In particular, f is differentiable a.e. and $g = f'$. The radial extension $f(x)$ is then continuous in $\mathbf{R}^d\setminus\{0\}$ and differentiable almost everywhere. Let us show that integration by parts holds, say, with respect to the first coordinate x_1. Write $x = (x_1, x_1,\ldots,x_d) = r\omega = (r\cos\theta, r(\sin\theta)\xi)$, with $r\in(0,\infty)$, $\omega\in\mathbf{S}^{d-1}\subset\mathbf{R}^d$, $0\le\theta\le\pi$ and $\xi\in\mathbf{S}^{d-2}\subset\mathbf{R}^{d-1}$. Let $\psi\in C_c^\infty(\mathbf{R}^d\setminus\{0\})$ be a generic test function and consider

$$\Psi(r) = \left(\int_{\mathbf{S}^{d-1}}\psi\,\frac{x_1}{|x|}\,\mathrm{d}\sigma_{d-1}(\omega)\right)r^{d-1}$$

$$= \left(\int_0^\pi\left(\int_{\mathbf{S}^{d-2}}\psi\,\mathrm{d}\sigma_{d-2}(\xi)\right)\cos\theta\,(\sin\theta)^{d-2}\,\mathrm{d}\theta\right)r^{d-1}.$$

Then

$$\Psi'(r) = \left(\int_0^\pi\int_{\mathbf{S}^{d-2}}\left(\frac{\partial\psi}{\partial r}\cos\theta - \frac{\partial\psi}{\partial\theta}\frac{\sin\theta}{r}\right)(\sin\theta)^{d-2}\,\mathrm{d}\sigma_{d-2}(\xi)\,\mathrm{d}\theta\right)r^{d-1},$$

where an integration by parts in the variable θ was used. Using polar coordinates one now sees that

$$\int_{\mathbf{R}^d\setminus\{0\}} f(x)\,\frac{\partial\psi}{\partial x_1}\,\mathrm{d}x = \int_0^\infty f(r)\,\Psi'(r)\,\mathrm{d}r = -\int_0^\infty f'(r)\,\Psi(r)\,\mathrm{d}r$$

$$= -\int_{\mathbf{R}^d\setminus\{0\}}\left(f'(|x|)\frac{x_1}{|x|}\right)\psi(x)\,\mathrm{d}x.$$

This shows that $f(x)$ is weakly differentiable with weak gradient given by $\nabla f(x) = f'(|x|)\frac{x}{|x|}$.

   Part (ii). Let $\psi\colon\mathbf{R}^d\to\mathbf{R}$ be a smooth radial non-increasing function with $\psi\equiv 1$ on $\{|x|\le 1\}$ and $\psi\equiv 0$ on $\{|x|\ge 2\}$. Let $\Psi_\alpha(x) = 1 - \psi(x/\alpha)$. Let $\phi\in C_c^\infty(\mathbf{R}^d)$ be any test function. Since we know that f is weakly differentiable in $\mathbf{R}^d\setminus\{0\}$ we have,

for any direction $i = 1, 2, \ldots, d$ (here we denote $\partial f / \partial x_i$ simply by f_{x_i}),

$$(2.9) \quad \begin{aligned} -\int_{\mathbf{R}^d} f_{x_i}(x)\,(\phi\Psi_\alpha)(x)\,\mathrm{d}x &= \int_{\mathbf{R}^d} f(x)\,(\phi\Psi_\alpha)_{x_i}(x)\,\mathrm{d}x \\ &= \int_{\mathbf{R}^d} f(x)\,\phi_{x_i}(x)\,\Psi_\alpha(x)\,\mathrm{d}x + \int_{\mathbf{R}^d} f(x)\,\phi(x)\,(\Psi_\alpha)_{x_i}(x)\,\mathrm{d}x. \end{aligned}$$

Note that the last integral takes place inside the ball of radius 2α. In this ball we have $\phi(x) = \phi(0) + R(x)$ with $|R(x)| \le C\alpha$. Since $f(x)$ is even in the variable x_i and $(\Psi_\alpha)_{x_i}(x)$ is odd in the variable x_i we get

$$(2.10) \qquad \int_{\mathbf{R}^d} f(x)(\Psi_\alpha)_{x_i}(x)\,\mathrm{d}x = 0,$$

and since $(\Psi_\alpha)_{x_i}(x) = -\frac{1}{\alpha}\psi_{x_i}(x/\alpha)$ we find

$$(2.11) \qquad \int_{\mathbf{R}^d} f(x) R(x) (\Psi_\alpha)_{x_i}(x)\,\mathrm{d}x \to 0$$

as $\alpha \to 0$, since f is locally integrable. Using (2.10) and (2.11) and the fact that ∇f is also locally integrable we may pass the limit as $\alpha \to 0$ in (2.9) to find

$$-\int_{\mathbf{R}^d} f_{x_i}(x)\,\phi(x)\,\mathrm{d}x = \int_{\mathbf{R}^d} f(x)\,\phi_{x_i}(x)\,\mathrm{d}x,$$

as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We now consider the case of general $u_0 \in W^{1,1}(\mathbf{R}^d)$ radial. We have seen in Lemma 4 that its radial version $u_0(r)$ is weakly differentiable in $(0, \infty)$ and

$$\int_0^\infty |u_0'(r)|\, r^{d-1}\,\mathrm{d}r < \infty.$$

In particular, after a possible redefinition on a set of measure zero, one can take $u_0(r)$ continuous in $(0, \infty)$ (in fact, absolutely continuous in each interval $[a, \infty)$ for $a > 0$). This is equivalent to assuming that $u_0(x)$ is continuous in $\mathbf{R}^d \setminus \{0\}$.

Step 3. u^* is continuous in $\mathbf{R}^d \setminus \{0\}$. With $u_0(x)$ continuous in $\mathbf{R}^d \setminus \{0\}$, the detachment set A_d defined in (2.3) is open. Throughout the rest of this section let us write

$$u_\varepsilon(x) := u(x, \varepsilon) = \big(u_0 * \varphi(\cdot, \varepsilon)\big)(x), \quad x \in \mathbf{R}^d,\ \varepsilon > 0.$$

We claim that u^* is locally Lipschitz in A_d. In fact, if $x_0 \in A_d$, there exists $t_0 > 0$ such that

$$u^*(x_0) = u(x_0, t_0) > u(x_0).$$

From the continuity of $u(x, t)$, there exist a neighborhood V of x_0 and an $\varepsilon_0 > 0$ such that

$$(2.12) \qquad u^*(x) = \sup_{t>0} u(x, t) = \sup_{t \ge \varepsilon_0} u(x, t) = \sup_{t>0} \big(u_{\varepsilon_0} * \varphi(\cdot, t)\big)(x) =: u^*_{\varepsilon_0}(x)$$

for all $x \in V$. Note that in the third equality above we used the semigroup property of the family $\varphi(\cdot, t)$ (i.e. the fact that $\varphi(\cdot, t_1) * \varphi(\cdot, t_2) = \varphi(\cdot, t_1 + t_2)$). Since u_{ε_0} is Lipschitz, we have that $u^* = u^*_{\varepsilon_0}$ is Lipschitz on V, which proves our claim.

Writing $\mathbf{R}^d \setminus \{0\} = A_d \cup A_d^c$, we now need so show that u^* is continuous at the points of A_d^c. Let $x_0 \in A_d^c$. If $x_0 \in \mathrm{int}(A_d^c)$ we are done since $u^* = u_0$ is continuous in a neighborhood of x_0. Assume now that $x_0 \in A_d^c \setminus \mathrm{int}(A_d^c)$ and that there exists a sequence $\{x_n\}_{n \in \mathbf{N}} \subset A_d$ such that $x_n \to x_0$ but $u^*(x_n) \not\to u^*(x_0) = u_0(x_0)$. Then there exist $t_n > 0$ and $\delta > 0$ such that $u(x_n, t_n) \ge u_0(x_0) + \delta$ for all n. From the

integrability of u_0, the t_n are bounded, and passing to a subsequence we may assume that $t_n \to t \geq 0$. Then $u(x_n, t_n) \to u(x_0, t) \geq u_0(x_0) + \delta$, and we get that $t > 0$ and $x_0 \in A_d$, a contradiction. This establishes that u^* is continuous in $\mathbf{R}^d \setminus \{0\}$.

Step 4. Weak differentiability and conclusion. In the previous step we showed that $u^*(r)$ is continuous on $(0, \infty)$ and locally Lipschitz in A_1. For almost every $r \in A_1$, from (2.12) we have

$$(u^*)'(r) = \lim_{\varepsilon \to 0} (u_\varepsilon^*)'(r).$$

From Minkowski's inequality we recall that

$$(2.13) \qquad \|\nabla u_\varepsilon\|_{L^1(\mathbf{R}^d)} \leq \|\nabla u_0\|_{L^1(\mathbf{R}^d)}$$

for any $\varepsilon > 0$. Using Fatou's lemma, the bound in Theorem 1 already proved for Lipschitz functions, and (2.13), we arrive at

$$(2.14) \qquad \begin{aligned} \int_{A_1} \left| (u^*)'(r) \right| r^{d-1} \, \mathrm{d}r &\leq \liminf_{\varepsilon \to 0} \int_{A_1} \left| (u_\varepsilon^*)'(r) \right| r^{d-1} \, \mathrm{d}r \\ &\lesssim_d \liminf_{\varepsilon \to 0} \|\nabla u_\varepsilon\|_{L^1(\mathbf{R}^d)} \leq \|\nabla u_0\|_{L^1(\mathbf{R}^d)}. \end{aligned}$$

With this in hand, an adaptation of the argument in [9, Section 5.4] shows that $u^*(r)$ is weakly differentiable in $(0, \infty)$ with weak derivative given by $\chi_{A_1^c} u_0'(r) + \chi_{A_1}(u^*)'(r)$. This in turn implies that $u^*(x)$ is weakly differentiable in $\mathbf{R}^d \setminus \{0\}$ by Lemma 4. From (2.14), its weak gradient ∇u^* on $\mathbf{R}^d \setminus \{0\}$ verifies

$$(2.15) \qquad \begin{aligned} \|\nabla u^*\|_{L^1(\mathbf{R}^d)} &= \kappa_{d-1} \int_0^\infty \left| (u^*)'(r) \right| r^{d-1} \, \mathrm{d}r \\ &= \kappa_{d-1} \left(\int_{A_1} \left| (u^*)'(r) \right| r^{d-1} \, \mathrm{d}r + \int_{A_1^c} \left| u_0'(r) \right| r^{d-1} \, \mathrm{d}r \right) \\ &\lesssim_d \|\nabla u_0\|_{L^1(\mathbf{R}^d)}, \end{aligned}$$

with κ_{d-1} being the total surface measure of $\mathbf{S}^{d-1}$. This is our desired bound. As a final remark note that, from the Sobolev embedding, $u_0 \in L^{d/(d-1)}(\mathbf{R}^d)$ and hence so does u^*. In particular, u^* is locally integrable in $\mathbf{R}^d$. Since we already know from (2.15) that $\nabla u^* \in L^1(\mathbf{R}^d)$, an application of Lemma 4 (ii) gives us that u^* is in fact weakly differentiable in $\mathbf{R}^d$. This completes the proof of Theorem 1.

Remark. A crucial insight in the proof above was to relate the variation of u^* with the variation of the uncentered Hardy–Littlewood maximal operator $\widetilde{M}u_0$, expressed in inequality (2.2). Since $\widetilde{M}u_0(x) \lesssim_d u^*(x)$, uniformly for all $x \in \mathbf{R}^d$, we could just run the exact same proof to obtain the gradient bound for $\widetilde{M}u_0$ starting from the gradient bound for u^*, showing that these two bounds are actually equivalent to each other.

3. Proof of Theorem 2

Recall that σ denotes the usual surface measure on the sphere $\mathbf{S}^d$. We denote by $\kappa_d = \sigma(\mathbf{S}^d) = 2\pi^{(d+1)/2}/\Gamma((d+1)/2)$ the total surface area of $\mathbf{S}^d$. With a slight abuse of notation, we shall also write

$$(3.1) \qquad \sigma(r) := \sigma\big(\mathcal{B}_r(\zeta)\big) = \kappa_{d-1} \int_0^r (\sin t)^{d-1} \, \mathrm{d}t.$$

Throughout this section we assume, without loss of generality, that f *is real-valued and nonnegative* (or $+\infty$).

3.1. Preliminaries. If $f \in L^1(\mathbf{S}^d)$, by Lebesgue differentiation we may modify it in a set of measure zero so that

$$(3.2) \qquad f(\xi) = \limsup_{\{r \to 0^+ \,:\, \xi \in \overline{\mathcal{B}_r(\zeta)}\}} \fint_{\mathcal{B}_r(\zeta)} f(\eta) \, d\sigma(\eta)$$

holds everywhere. Let us assume that is the case. For $f \in L^1(\mathbf{S}^d)$ and $\xi \in \mathbf{S}^d$ let us define the set $\mathbf{B}_\xi$ as the set of closed balls that realize the supremum in the definition of the maximal function, that is

$$(3.3) \qquad \mathbf{B}_\xi = \left\{ \overline{\mathcal{B}_r(\zeta)}; \ \zeta \in \mathbf{S}^d, \ r \geq 0, \ \xi \in \overline{\mathcal{B}_r(\zeta)} \colon \widetilde{\mathcal{M}}f(\xi) = \fint_{\overline{\mathcal{B}_r(\zeta)}} f(\eta) \, d\sigma(\eta) \right\}.$$

Here we consider the slight abuse of notation

$$(3.4) \qquad \overline{\mathcal{B}_0(\xi)} := \{\xi\} \quad \text{and} \quad \fint_{\{\xi\}} f(\eta) \, d\sigma(\eta) := f(\xi),$$

in order to include the closed ball of radius zero as a potential candidate in the definition of $\mathbf{B}_\xi$. In light of (3.2) we always have that $\mathbf{B}_\xi$ is non-empty. Our first lemma holds for general Sobolev functions in $W^{1,1}(\mathbf{S}^d)$ (not necessarily polar functions).

Lemma 5. *Let $f \in W^{1,1}(\mathbf{S}^d)$ be a nonnegative function that verifies (3.2) and let $\xi \in \mathbf{S}^d$ be a point such that $\widetilde{\mathcal{M}}f(\xi) > f(\xi)$. Assume that $\widetilde{\mathcal{M}}f$ is differentiable at ξ and that $\overline{\mathcal{B}} \in \mathbf{B}_\xi$. Then*

$$\nabla \widetilde{\mathcal{M}}f(\xi)v = \fint_{\mathcal{B}} \nabla f(\eta)\big(-(\eta \cdot v)\xi + (\eta \cdot \xi)v\big) \, d\sigma(\eta)$$

for every $v \in \mathbf{R}^{d+1}$ with $v \perp \xi$. In particular,

$$\big|\nabla \widetilde{\mathcal{M}}f(\xi)\big| \leq \fint_{\mathcal{B}} |\nabla f(\eta)| \, d\sigma(\eta).$$

Proof. Observe first that the condition $\widetilde{\mathcal{M}}f(\xi) > f(\xi)$ implies that the ball $\mathcal{B}$ has positive radius. Without loss of generality let us assume that $|v| = 1$. Let $R_t = R_{t,\xi,v}$ be the rotation of angle t over the plane spanned by ξ and v that leaves the orthogonal complement invariant, i.e.

$$R_t(\eta) = \big((\cos t)(\eta \cdot \xi) - (\sin t)(\eta \cdot v)\big)\xi + \big((\sin t)(\eta \cdot \xi) + (\cos t)(\eta \cdot v)\big)v + z(\eta),$$

where $z(\eta)$ is the component of the vector η that is orthogonal to the plane generated by ξ and v. Then

$$\nabla \widetilde{\mathcal{M}}f(\xi)v = \lim_{t \to 0+} \frac{\widetilde{\mathcal{M}}f(R_t\xi) - \widetilde{\mathcal{M}}f(\xi)}{t} \geq \lim_{t \to 0+} \frac{1}{t}\left(\fint_{R_t(\mathcal{B})} f - \fint_{\mathcal{B}} f\right)$$

$$(3.5) \qquad = \lim_{t \to 0+} \fint_{\mathcal{B}} \frac{f(R_t\eta) - f(\eta)}{t} \, d\sigma(\eta)$$

$$= \fint_{\mathcal{B}} \nabla f(\eta)\big(-(\eta \cdot v)\xi + (\eta \cdot \xi)v\big) \, d\sigma(\eta).$$

The reverse inequality is obtained similarly by considering the limit as $t \to 0^-$. $\quad\square$

Remark. The passage to the limit in (3.5) uses the fact that the difference quotients are bounded in L^1 by a multiple of L^1-norm of the gradient of f, uniformly in

t. With such a uniform bound one can establish the required limit by approximating f by smooth g.

3.2. Lipschitz case. Throughout this subsection we assume that our polar $f \in W^{1,1}(\mathbf{S}^d)$ is a Lipschitz function. Recalling that $\mathbf{e} = (1, 0, 0, \ldots, 0) \in \mathbf{R}^{d+1}$, for $\xi \in \mathbf{S}^d$ we write

$$\cos \theta = \xi \cdot \mathbf{e}$$

with $\theta \in [0, \pi]$. Note that $\theta = \theta(\xi) = d(\mathbf{e}, \xi)$ is the polar angle. We generally write $f(\xi)$ for the function on $\mathbf{S}^d$, and $f(\theta)$ for its polar version on $(0, \pi)$. We then have

$$|\nabla f(\xi)| = |f'(\theta)|$$

for a.e. $\xi \in \mathbf{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$, and

$$\|\nabla f\|_{L^1(\mathbf{S}^d)} = \kappa_{d-1} \int_0^\pi |f'(\theta)| \, (\sin \theta)^{d-1} \, \mathrm{d}\theta.$$

3.2.1. Estimates for small radii. For $\zeta \in \mathbf{S}^d$ let us define

$$w(\zeta) = \min \left\{ \theta(\zeta), \pi - \theta(\zeta) \right\} = \min\{d(\mathbf{e}, \zeta), d(-\mathbf{e}, \zeta)\}.$$

Let us define the auxiliary maximal operator $\widetilde{\mathcal{M}}^I$ by (recall convention (3.4))

$$(3.6) \qquad \widetilde{\mathcal{M}}^I f(\xi) = \sup_{\{\xi \in \overline{\mathcal{B}_r(\zeta)} : \, 0 \leq r \leq w(\zeta)/4\}} \fint_{\overline{\mathcal{B}_r(\zeta)}} f(\eta) \, \mathrm{d}\sigma(\eta).$$

In analogy to (3.3), for each $\xi \in \mathbf{S}^d$ we define the set of good balls

$$\mathbf{B}_\xi^I = \left\{ \overline{\mathcal{B}_r(\zeta)}; \; \zeta \in \mathbf{S}^d, \; 0 \leq r \leq \frac{w(\zeta)}{4}; \; \xi \in \overline{\mathcal{B}_r(\zeta)} : \widetilde{\mathcal{M}}^I f(\xi) = \fint_{\overline{\mathcal{B}_r(\zeta)}} f(\eta) \, \mathrm{d}\sigma(\eta) \right\}.$$

Notice that $\widetilde{\mathcal{M}}^I f$ is also a polar function. We consider the detachment set

$$\mathcal{E}_d := \left\{ \xi \in \mathbf{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\} : \widetilde{\mathcal{M}}^I f(\xi) > f(\xi) \right\},$$

and its polar version, denoted by

$$\mathcal{E}_1 = \{\theta(\xi) = d(\mathbf{e}, \xi) : \xi \in \mathcal{E}_d\}.$$

One can check that $\widetilde{\mathcal{M}}^I f$ is a continuous function in $\mathbf{S}^d$. Further qualitative properties of $\widetilde{\mathcal{M}}^I f$ are described in the next two results.

Lemma 6. $\widetilde{\mathcal{M}}^I f$ *does not have a strict local maximum in* $\mathcal{E}_1$.

Proof. The proof is identical to [20, Lemma 3.10]. $\qquad\qquad\qquad\qquad\qquad \square$

Lemma 7. $\widetilde{\mathcal{M}}^I f$ *is locally Lipschitz in* $\mathcal{E}_d$.

Proof. Let $\xi \in \mathcal{E}_d$. Let $\overline{\mathcal{B}_r(\zeta)} \in \mathbf{B}_\xi^I$ with r minimal. Then $r > 0$ and it is possible to find a neighborhood V of ξ of the form $V = \{\eta \in \mathbf{S}^d : \theta(\xi) - \varepsilon < \theta(\eta) < \theta(\xi) + \varepsilon\}$ such that: (i) $\varepsilon < r/100$ and (ii) if $\eta \in V$ and $\overline{\mathcal{B}_s(\omega)} \in \mathbf{B}_\eta^I$ then $s > 99r/100$.

Let $\eta_1, \omega_2 \in V$. Let S be the half great circle connecting $\mathbf{e}, \eta_1, -\mathbf{e}$. If $\eta_2 \in S$ is such that $d(\mathbf{e}, \eta_2) = d(\mathbf{e}, \omega_2)$ then we have $d(\eta_1, \eta_2) \leq d(\eta_1, \omega_2)$. Since $\widetilde{\mathcal{M}}^I f(\eta_2) = \widetilde{\mathcal{M}}^I f(\omega_2)$, for the purposes of proving Lipschitz continuity it suffices to work with $\eta_1, \eta_2 \in S$. Assume without loss of generality that $\widetilde{\mathcal{M}}^I f(\eta_1) > \widetilde{\mathcal{M}}^I f(\eta_2)$. Let $\overline{\mathcal{B}_{r_1}(\zeta_1)} \in \mathbf{B}_{\eta_1}^I$ with $\zeta_1 \in S$. Then $\eta_2 \notin \overline{\mathcal{B}_{r_1}(\zeta_1)}$, and hence η_2 is not between ζ_1

and η_1. It is also easy to see that we cannot have ζ_1 between η_1 and η_2 due to conditions (i) and (ii) above. Hence we must have η_1 between ζ_1 and η_2. We now choose a ball $\mathcal{B}_{r_2}(\zeta_2)$, with $\zeta_2 \in S$ lying between ζ_1 and η_2, such that $\eta_2 \in \partial\overline{\mathcal{B}_{r_2}(\zeta_2)}$ and

$$(3.7) \qquad\qquad r_2 = d(\zeta_2, \eta_2) = \min\left\{r_1, \frac{w(\zeta_2)}{4}\right\}$$

(one may think of moving the center ζ_1 along S in the direction of η_2 until finding the unique choice of ζ_2). Note that ζ_2 is in fact between ζ_1 and η_1 and hence

$$(3.8) \quad r_2 = d(\zeta_2, \eta_2) = d(\zeta_1, \eta_1) - d(\zeta_1, \zeta_2) + d(\eta_1, \eta_2) \le r_1 - d(\zeta_1, \zeta_2) + d(\eta_1, \eta_2).$$

If $r_2 = r_1$ in (3.7) then we have $d(\zeta_1, \zeta_2) \le d(\eta_1, \eta_2)$. In the other case we have

$$r_2 = \frac{w(\zeta_2)}{4} \ge \frac{w(\zeta_1)}{4} - \frac{d(\zeta_1, \zeta_2)}{4} \ge r_1 - \frac{d(\zeta_1, \zeta_2)}{4},$$

and combining with (3.8) we obtain $d(\zeta_1, \zeta_2) \le \frac{4}{3}d(\eta_1, \eta_2)$, which yields $r_1 - r_2 \le \frac{1}{3}d(\eta_1, \eta_2)$. We conclude by observing that

$$\widetilde{\mathcal{M}}^I f(\eta_1) - \widetilde{\mathcal{M}}^I f(\eta_2) \le \fint_{\mathcal{B}_{r_1}(\zeta_1)} f - \fint_{\mathcal{B}_{r_2}(\zeta_2)} f$$

$$\le \left(\fint_{\mathcal{B}_{r_1}(\zeta_1)} f - \fint_{\mathcal{B}_{r_2}(\zeta_1)} f\right) + \left(\fint_{\mathcal{B}_{r_2}(\zeta_1)} f - \fint_{\mathcal{B}_{r_2}(\zeta_2)} f\right)$$

$$\lesssim_{d,r,f} d(\eta_1, \eta_2). \qquad\qquad \square$$

An adaptation of the argument in [9, Section 5.4] then shows that $\widetilde{\mathcal{M}}^I f(\theta)$ is weakly differentiable in $(0, \pi)$, with weak derivative given by $\chi_{\mathcal{E}_1^c} f'(\theta) + \chi_{\mathcal{E}_1}(\widetilde{\mathcal{M}}^I f)'(\theta)$. In fact, if $\theta \in \mathcal{E}_1^c$ is a point of differentiability of f (which are almost all points of $\mathcal{E}_1^c$) one can plainly see that $f'(\theta) = 0$, otherwise one could do better than $f(\theta)$ in the maximal function (3.6) and θ would belong to $\mathcal{E}_1$ instead. The weak derivative of $\widetilde{\mathcal{M}}^I f(\theta)$ is then simply $\chi_{\mathcal{E}_1}(\widetilde{\mathcal{M}}^I f)'(\theta)$. From Lemma 13 below we have that $\widetilde{\mathcal{M}}^I f(\xi)$ is weakly differentiable in $\mathbf{S}^d$. The next proposition establishes the desired control of the variation.

Proposition 8. *The following inequality holds*

$$\left\|\nabla\widetilde{\mathcal{M}}^I f\right\|_{L^1(\mathbf{S}^d)} \lesssim_d \|\nabla f\|_{L^1(\mathbf{S}^d)}.$$

Proof. The proof follows the outline of [20, Lemma 3.5] with minor changes. We need to prove that

$$\int_{\mathcal{E}_1} \left|(\widetilde{\mathcal{M}}^I f)'(\theta)\right| (\sin\theta)^{d-1}\, d\theta \lesssim_d \int_0^\pi \left|f'(\theta)\right| (\sin\theta)^{d-1}\, d\theta.$$

We shall prove that

$$(3.9) \qquad \int_{\mathcal{E}_1 \cap [0, \pi/2]} \left|(\widetilde{\mathcal{M}}^I f)'(\theta)\right| (\sin\theta)^{d-1}\, d\theta \lesssim_d \int_0^\pi \left|f'(\theta)\right| (\sin\theta)^{d-1}\, d\theta$$

and the proposition follows by symmetry. For $k \ge 1$, we define $\mathcal{E}_1^k = \mathcal{E}_1 \cap \left[\frac{\pi}{2^{k+1}}, \frac{\pi}{2^k}\right]$, and since $\mathcal{E}_1$ is open we may write $\text{int}(\mathcal{E}_1^k) = \bigcup_{i=1}^\infty (a_i^k, b_i^k)$. We observe that $\frac{(\sin 2\theta)^{d-1}}{(\sin\theta)^{d-1}}$

$\simeq_d 1$ for $\theta \le \frac{\pi}{4}$. When $a_i^k = \frac{\pi}{2^{k+1}}$ or $b_i^k = \frac{\pi}{2^k}$ we observe, from the definition of the auxiliary operator in (3.6), that

$$\widetilde{\mathcal{M}}^I f(\pi/2^{k+1}) \, , \, \widetilde{\mathcal{M}}^I f(\pi/2^k) \le \sup_{\theta(\xi) \in [\pi/2^{k+2}, \pi/2^{k-1}]} f(\xi)$$

for $k \ge 2$. These are the ingredients needed to run the argument in [20, Lemma 3.5] in order to get

$$(3.10) \qquad \int_{\mathcal{E}_1^k} \left| (\widetilde{\mathcal{M}}^I f)'(\theta) \right| (\sin \theta)^{d-1} \, d\theta \lesssim_d \int_{\pi/2^{k+2}}^{\pi/2^{k-1}} \left| f'(\theta) \right| (\sin \theta)^{d-1} \, d\theta$$

for $k \ge 2$. In the case $k = 1$ we must be a bit more careful when $b_i^1 = \pi/2$ by using the bound

$$\widetilde{\mathcal{M}}^I f(\pi/2) \le \sup_{\theta(\xi) \in [\pi/4, 3\pi/4]} f(\xi),$$

which then yields

$$(3.11) \qquad \begin{aligned} \int_{\mathcal{E}_1^1} \left| (\widetilde{\mathcal{M}}^I f)'(\theta) \right| (\sin \theta)^{d-1} \, d\theta &\le \int_{\mathcal{E}_1^1} \left| (\widetilde{\mathcal{M}}^I f)'(\theta) \right| d\theta \lesssim \int_{\pi/8}^{3\pi/4} \left| f'(\theta) \right| d\theta \\ &\lesssim_d \int_{\pi/8}^{3\pi/4} \left| f'(\theta) \right| (\sin \theta)^{d-1} \, d\theta. \end{aligned}$$

Finally, we add up (3.10) and (3.11) to get (3.9). $\qquad\qquad\square$

3.2.2. Estimates for large radii – preliminary lemmas. The other crucial ingredient in the proof of Luiro [20, Lemma 2.2 (v)] is the bound

$$\left| \nabla \widetilde{M} f(x) \right| \le \frac{1}{|x|} \fint_B |\nabla f(y)| \, |y| \, dy,$$

where $\overline{B} \ni x$ is a ball in which the maximal function is realized. The main difficulty in the case of $\mathbf{S}^d$ is in establishing a bound that will serve a similar purpose. This is accomplished in Lemma 12 below but before we actually get there we need a few preliminary lemmas. Recall the definition of $\sigma(r)$ in (3.1), and observe that $\sigma'(r) = \kappa_{d-1}(\sin r)^{d-1}$ is equal to the $(d-1)$-dimensional area of $\partial \mathcal{B}_r(\zeta)$.

Lemma 9. *Let $\xi \in \mathbf{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ and let $\overline{\mathcal{B}_r(\zeta)} \in \mathbf{B}_\xi$, with ζ in the half great circle determined by $\mathbf{e}$, ξ and $-\mathbf{e}$. Assume that $0 \le \theta(\zeta) < \theta(\xi)$, that $\xi \in \partial \mathcal{B}_r(\zeta)$, that $\widetilde{\mathcal{M}} f(\xi) > f(\xi)$ and that $\widetilde{\mathcal{M}} f$ is differentiable at ξ. Then*

$$\nabla \widetilde{\mathcal{M}} f(\xi)(v(\xi, \mathbf{e})) = \frac{\sigma'(r)}{\sigma(r)} \fint_{\mathcal{B}_r(\zeta)} \nabla f(\eta)(v(\eta, \zeta)) \frac{\sigma(d(\zeta, \eta))}{\sigma'(d(\zeta, \eta))} \, d\sigma(\eta),$$

where

$$v(\eta, \zeta) = \frac{\zeta - (\eta \cdot \zeta)\eta}{|\zeta - (\eta \cdot \zeta)\eta|}$$

is the unit vector, tangent to η, in the direction of the geodesic that goes from η to ζ.

Proof. Since $\widetilde{\mathcal{M}} f(\xi) > f(\xi)$ we have $r > 0$. Let S be the great circle determined by $\mathbf{e}$ and ξ. For small $h \in \mathbf{R}$ we consider a rotation R_h of angle h in this circle (in

the direction from ξ to $\mathbf{e}$) leaving the orthogonal complement in $\mathbf{R}^{d+1}$ invariant, and write $\zeta - h := R_h(\zeta)$. The idea is to look at the following quantity

$$(3.12) \quad \lim_{h \to 0} \frac{\fint_{\mathcal{B}_{r+h}(\zeta-h)} f - \fint_{\mathcal{B}_r(\zeta)} f}{h}$$

$$= \lim_{h \to 0} \frac{\fint_{\mathcal{B}_{r+h}(\zeta-h)} f - \fint_{\mathcal{B}_r(\zeta-h)} f + \fint_{\mathcal{B}_r(\zeta-h)} f - \fint_{\mathcal{B}_r(\zeta)} f}{h}.$$

In principle we do not know that the limit above exists. We shall prove that it in fact exists using the right-hand side of (3.12). Once this is established, the left-hand side of (3.12) tells us that this limit must be zero, since the numerator is always nonpositive regardless of the sign of h.

From Lemma 5 (in particular, see computation (3.5)) we note that

$$(3.13) \quad \lim_{h \to 0} \frac{\fint_{\mathcal{B}_r(\zeta-h)} f - \fint_{\mathcal{B}_r(\zeta)} f}{h} = \nabla \widetilde{\mathcal{M}} f(\xi)(v(\xi, \mathbf{e})).$$

Note also that

$$\frac{\fint_{\mathcal{B}_{r+h}(\zeta-h)} f - \fint_{\mathcal{B}_r(\zeta-h)} f}{h}$$

$$(3.14) \quad = \frac{\frac{1}{\sigma(r+h)} - \frac{1}{\sigma(r)}}{h} \int_{\mathcal{B}_{r+h}(\zeta-h)} f + \frac{1}{\sigma(r)} \frac{\int_{\mathcal{B}_{r+h}(\zeta-h)} f - \int_{\mathcal{B}_r(\zeta-h)} f}{h}$$

$$\to -\frac{\sigma'(r)}{\sigma(r)^2} \int_{\mathcal{B}_r(\zeta)} f + \frac{1}{\sigma(r)} \int_{\partial \mathcal{B}_r(\zeta)} f$$

as $h \to 0$. Hence the limit in (3.12) exists and is zero. Now we consider momentarily ζ as the north pole in the computation below and proceed with the standard polar coordinates on the sphere. Writing $\eta = (\cos\theta, \omega \sin\theta)$, with $\omega \in \mathbf{S}^{d-1}$ we use integration by parts to get

$$\int_{\mathcal{B}_r(\zeta)} \nabla f(\eta) \, (-v(\eta, \zeta)) \frac{\sigma(d(\zeta, \eta))}{\sigma'(d(\zeta, \eta))} \, d\sigma(\eta)$$

$$= \int_{\mathbf{S}^{d-1}} \int_0^r \frac{\partial f}{\partial \theta}(\theta, \omega) \left(\int_0^\theta (\sin t)^{d-1} dt \right) d\theta \, d\sigma_{d-1}(\omega)$$

$$(3.15) \quad = \int_{\mathbf{S}^{d-1}} f(r, \omega) \left(\int_0^r (\sin t)^{d-1} dt \right) d\sigma_{d-1}(\omega)$$

$$- \int_{\mathbf{S}^{d-1}} \int_0^r f(\theta, \omega)(\sin\theta)^{d-1} d\theta \, d\sigma_{d-1}(\omega)$$

$$= \frac{\sigma(r)}{\sigma'(r)} \int_{\partial \mathcal{B}_r(\zeta)} f - \int_{\mathcal{B}_r(\zeta)} f.$$

The lemma then plainly follows from (3.12), (3.13), (3.14) and (3.15). $\qquad \square$

We now state a basic geometric lemma.

Lemma 10. *Denote by $\triangle ABC$ a geodesic triangle with vertices A, B, C, opposite geodesic side lengths a, b, c, and (geodesic) angles $\hat{A}, \hat{B}, \hat{C}$.*

(i) *There exist universal constants $\gamma > 1$ and $\rho > 0$ such that for every $\triangle ABC \subset \overline{\mathcal{B}_\rho(\mathbf{e})}$ we have*

$$a \sin \hat{B} \le \gamma \, b.$$

(ii) *Under the same hypotheses, if $\hat{B} \le \frac{\pi}{2}$ we have*

$$\left| c - a \cos \hat{B} \right| \le b.$$

Proof. Part (i). By the triangle inequality we have $a \le 2\rho$. Then, for any $\gamma > 1$ we can choose ρ small so that $\sin \theta \le \theta \le \gamma \sin \theta$ for $0 \le \theta \le 2\rho$. Using the spherical law of sines we have

$$a \sin \hat{B} \le \gamma \sin a \sin \hat{B} = \gamma \sin b \sin \hat{A} \le \gamma \sin b \le \gamma b.$$

Part (ii). Assume that ρ is small. We shall prove that $\cos(c - a \cos \hat{B}) \ge \cos b$, which shall imply that $|c - a \cos \hat{B}| \le b$. By the spherical law of cosines we have

$$\cos b = \cos c \cos a + \sin c \sin a \cos \hat{B}.$$

Note that

$$\cos(c - a \cos \hat{B}) = \cos c \cos(a \cos \hat{B}) + \sin c \sin(a \cos \hat{B}).$$

Since $0 \le a \cos \hat{B} \le a$ we have that $\cos(a \cos \hat{B}) \ge \cos a$. Also, by elementary calculus we have $\sin(a \cos \hat{B}) \ge \sin a \cos \hat{B}$, and the result plainly follows from these estimates. $\square$

We conclude this part with another elementary fact.

Lemma 11. *We have*

$$u(t) := \frac{\int_0^t (\sin s)^{d-1} \, \mathrm{d}s}{t \, (\sin t)^{d-1}} = \frac{\sigma(t)}{t \, \sigma'(t)} \simeq_d 1$$

for $0 \le t \le 1/4$. Moreover, u is a C^∞-function in this range.

Proof. Note that

$$\frac{\int_0^t (\sin s)^{d-1} \, \mathrm{d}s}{t(\sin t)^{d-1}} = \frac{1}{t} \int_0^t \left(\frac{\sin s}{\sin t} \right)^{d-1} \, \mathrm{d}s = \frac{\sin t}{t} \int_0^1 a^{d-1} \frac{1}{(1 - a^2 (\sin t)^2)^{1/2}} \, \mathrm{d}a,$$

and both $t \mapsto \frac{\sin t}{t}$ and $t \mapsto \int_0^1 a^{d-1} \frac{1}{(1 - a^2 (\sin t)^2)^{1/2}} \, \mathrm{d}a$ are smooth functions bounded above and below in the proposed range. $\square$

3.2.3. Estimates for large radii – main lemma. We are now in position to prove the key result of this subsection.

Lemma 12. *Let $\xi \in \mathbf{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ and let $\overline{\mathcal{B}_r(\zeta)} \in \mathbf{B}_\xi$, with ζ in the half great circle determined by $\mathbf{e}$, ξ and $-\mathbf{e}$. Assume that $0 \le \theta(\zeta) < \theta(\xi)$, that $\xi \in \partial \mathcal{B}_r(\zeta)$, that $\widetilde{\mathcal{M}} f(\xi) > f(\xi)$ and that $\widetilde{\mathcal{M}} f$ is differentiable at ξ. There is a universal constant $\rho > 0$ such that if $\mathcal{B} = \mathcal{B}_r(\zeta) \subset \overline{\mathcal{B}_\rho(\mathbf{e})}$ then*

$$(3.16) \qquad \left| \nabla \widetilde{\mathcal{M}} f(\xi) \right| \lesssim_d \frac{1}{\theta(\xi)} \fint_{\mathcal{B}} |\nabla f(\eta)| \, \theta(\eta) \, \mathrm{d}\sigma(\eta) + \frac{r \, \theta(\zeta)}{\theta(\xi)} \fint_{\mathcal{B}} |\nabla f(\eta)| \, \mathrm{d}\sigma(\eta).$$

Proof. From Lemma 9 we have

$$(3.17) \qquad \nabla \widetilde{\mathcal{M}} f(\xi)(-v(\xi, \mathbf{e})) = \frac{\sigma'(r)}{\sigma(r)} \fint_{\mathcal{B}} \nabla f(\eta)(-v(\eta, \zeta)) \frac{\sigma(d(\zeta, \eta))}{\sigma'(d(\zeta, \eta))} \, \mathrm{d}\sigma(\eta).$$

In the case $\zeta = \mathbf{e}$, estimate (3.16) follows directly from (3.17) and Lemma 11. From now on we assume that $\zeta \ne \mathbf{e}$. From Lemma 5 we also know that

$$(3.18) \qquad \nabla \widetilde{\mathcal{M}} f(\xi)(-v(\xi, \mathbf{e})) = \fint_{\mathcal{B}} \nabla f(\eta) \, S(\eta) \, \mathrm{d}\sigma(\eta),$$

with $S(\eta) = (\eta \cdot v(\xi, \mathbf{e}))\xi - (\eta \cdot \xi)v(\xi, \mathbf{e})$. The idea is to compare the identities (3.17) and (3.18) in order to bound $\left|\nabla\widetilde{\mathcal{M}}f(\xi)\right| = \left|\nabla\widetilde{\mathcal{M}}f(\xi)(-v(\xi, \mathbf{e}))\right|$. To do so, we write the right-hand side of (3.18) as a sum of three terms, one being comparable to $\left|\nabla\widetilde{\mathcal{M}}f(\xi)\right|$, the second one being small, and the third one being close to the right-hand side of (3.17) in a suitable sense. We start by writing

$$1 = \frac{\theta(\xi) - \theta(\zeta)}{r} = \frac{d(\mathbf{e}, \xi) - d(\mathbf{e}, \zeta)}{r}.$$

Let us define $v_1(\eta) = S(\eta)/|S(\eta)|$. We then have

$$
\begin{aligned}
(3.19)\quad & \fint_{\mathcal{B}} \nabla f(\eta)\, S(\eta)\, \mathrm{d}\sigma(\eta) \\
&= \fint_{\mathcal{B}} \nabla f(\eta)\, |S(\eta)| \left(\frac{\theta(\xi) - \theta(\zeta)}{r}\right) v_1(\eta)\, \mathrm{d}\sigma(\eta) \\
&= \fint_{\mathcal{B}} \nabla f(\eta)\, |S(\eta)| \frac{\theta(\xi)}{r} v_1(\eta)\, \mathrm{d}\sigma(\eta) \\
&\quad - \fint_{\mathcal{B}} \nabla f(\eta)\, (|S(\eta)| - 1) \frac{\theta(\zeta)}{r} v_1(\eta)\, \mathrm{d}\sigma(\eta) - \fint_{\mathcal{B}} \nabla f(\eta)\frac{\theta(\zeta)}{r} v_1(\eta)\, \mathrm{d}\sigma(\eta).
\end{aligned}
$$

Step 1. Let us start by bounding the quantity

$$\frac{\sigma'(r)}{\sigma(r)} \fint_{\mathcal{B}} \nabla f(\eta)(-v(\eta, \zeta))\, \frac{\sigma(d(\zeta, \eta))}{\sigma'(d(\zeta, \eta))}\, \mathrm{d}\sigma(\eta) + \fint_{\mathcal{B}} \nabla f(\eta)\frac{\theta(\zeta)}{r} v_1(\eta)\, \mathrm{d}\sigma(\eta).$$

This last expression is equal to (recall the definition of u in Lemma 11)

$$(3.20)\qquad \frac{\sigma'(r)}{\sigma(r)} \fint_{\mathcal{B}} \nabla f(\eta) \big[d(\zeta, \eta)\, u(d(\zeta, \eta))\, (-v(\eta, \zeta)) + d(\mathbf{e}, \zeta)\, u(r)\, v_1(\eta)\big]\, \mathrm{d}\sigma(\eta).$$

Note now that

$$
\begin{aligned}
& d(\zeta, \eta)\, u(d(\zeta, \eta))\, (-v(\eta, \zeta)) + d(\mathbf{e}, \zeta)\, u(r)\, v_1(\eta) \\
&= u(d(\zeta, \eta)) \big[d(\zeta, \eta)(-v(\eta, \zeta)) + d(\mathbf{e}, \zeta)v_1(\eta)\big] \\
&\quad - d(\mathbf{e}, \zeta) \big[u(d(\zeta, \eta)) - u(r)\big] v_1(\eta).
\end{aligned}
$$

From Lemma 11 we know that $u(t)$ is Lipschitz for $0 \le t \le 1/4$. We then have $|u(d(\zeta, \eta)) - u(r)| \lesssim_d r$ and another application of Lemma 11 yields

$$
\begin{aligned}
(3.21)\quad & \frac{\sigma'(r)}{\sigma(r)} \left|\fint_{\mathcal{B}} \nabla f(\eta)\, d(\mathbf{e}, \zeta) \big[u(d(\zeta, \eta)) - u(r)\big] v_1(\eta)\, \mathrm{d}\sigma(\eta)\right| \\
& \lesssim_d \fint_{\mathcal{B}} |\nabla f(\eta)|\, d(\mathbf{e}, \zeta)\, \mathrm{d}\sigma(\eta).
\end{aligned}
$$

Let us now deal with the remaining piece. Observe that

$$
\begin{aligned}
& d(\zeta, \eta)\, (-v(\eta, \zeta)) + d(\mathbf{e}, \zeta)\, v_1(\eta) \\
&= d(\zeta, \eta)\big(v_1(\eta) \cos\alpha + v_1(\eta)^* \sin\alpha\big) + d(\mathbf{e}, \zeta)\, v_1(\eta) \\
(3.22)\quad &= \big[d(\zeta, \eta)v_1(\eta) \cos\beta + d(\mathbf{e}, \zeta)\, v_1(\eta)\big] + \big[d(\zeta, \eta)v_1(\eta)^* \sin\alpha\big] \\
&\quad + \big[d(\zeta, \eta)v_1(\eta)(\cos\alpha - \cos\beta)\big] \\
&= [I] + [II] + [III],
\end{aligned}
$$

where $\cos\alpha = -v(\eta, \zeta) \cdot v_1(\eta)$ $(0 \le \alpha \le \pi)$, $v_1(\eta)^*$ is unitary and orthogonal to $v_1(\eta)$ (in the plane determined by $v_1(\eta)$ and $v(\eta, \zeta)$), and $\cos\beta = v(\zeta, \eta) \cdot (-v(\zeta, \mathbf{e}))$

$(0 \le \beta \le \pi)$. Naturally, we may assume without loss of generality that $\eta \ne \zeta$. We now proceed with the analysis of the three terms in (3.22).

Analysis of $[I]$. Observe that

$$|d(\zeta, \eta) v_1(\eta) \cos \beta + d(\mathbf{e}, \zeta) \, v_1(\eta)| = |d(\zeta, \eta) \cos \beta + d(\mathbf{e}, \zeta)|.$$

Consider the geodesic triangle with vertices $\mathbf{e}, \zeta, \eta$ (that has angle $\angle \mathbf{e} \zeta \eta = \pi - \beta$). Assuming ρ small, if $\beta > \pi/2$ we may use Lemma 10 (ii) to find

$$|d(\zeta, \eta) \cos \beta + d(\mathbf{e}, \zeta)| \le d(\mathbf{e}, \eta).$$

In case $0 \le \beta \le \pi/2$ we have

$$0 \le \operatorname{sgn}(\cos \beta) = \operatorname{sgn}\big[(\eta - (\zeta \cdot \eta)\zeta) \cdot (-\mathbf{e} + (\zeta \cdot \mathbf{e})\zeta)\big] = \operatorname{sgn}\big[-(\eta \cdot \mathbf{e}) + (\zeta \cdot \mathbf{e})(\zeta \cdot \eta)\big],$$

which implies that

$$\cos(\theta(\zeta)) = (\zeta \cdot \mathbf{e}) \ge (\zeta \cdot \mathbf{e})(\zeta \cdot \eta) \ge (\eta \cdot \mathbf{e}) = \cos(\theta(\eta)).$$

From this we conclude that $d(\mathbf{e}, \zeta) = \theta(\zeta) \le \theta(\eta) = d(\mathbf{e}, \eta)$ and hence

$$|d(\zeta, \eta) \cos \beta + d(\mathbf{e}, \zeta)| \le d(\zeta, \eta) + d(\mathbf{e}, \zeta) \le (d(\mathbf{e}, \zeta) + d(\mathbf{e}, \eta)) + d(\mathbf{e}, \zeta) \le 3d(\mathbf{e}, \eta).$$

Analysis of $[II]$ *and* $[III]$. We note that the angles α and β are close, and it is important for our purposes to actually quantify this discrepancy. In order to do this, let us parametrize the points as follows. We write $\zeta = (\cos \theta, \sin \theta, \mathbf{0})$, with $\mathbf{0} \in \mathbf{R}^{d-1}$, and $\eta = (\cos \theta_1, \sin \theta_1 \cos \varphi, \sin \theta_1 \sin \varphi \, \omega)$ with $\omega \in \mathbf{S}^{d-2} \subset \mathbf{R}^{d-1}$. Here we set $0 \le \theta, \theta_1, \varphi \le \pi$. Recall that in this notation we have $\mathbf{e} = (1, 0, \mathbf{0})$. We then have $-v(\zeta, \mathbf{e}) = (-\sin \theta, \cos \theta, \mathbf{0})$. Recall also that the vector $v_1(\eta)$ is the unitary vector tangent to η in the direction of the derivative of the curve that takes the point η along the rotation in the first two coordinates (in the direction from $\mathbf{e}$ to ζ). A direct computation yields

(3.23) $$S(\eta) = (-\sin \theta_1 \cos \varphi, \cos \theta_1, \mathbf{0})$$

and

$$v_1(\eta) = \frac{1}{\sqrt{1 - \sin^2 \theta_1 \sin^2 \varphi}} (-\sin \theta_1 \cos \varphi, \cos \theta_1, \mathbf{0}).$$

Using that $v(\zeta, \mathbf{e}) \perp \zeta$ and $v_1(\eta) \perp \eta$ we then find

$$\cos \beta = v(\zeta, \eta) \cdot (-v(\zeta, \mathbf{e})) = \frac{\eta - (\eta \cdot \zeta)\zeta}{|\eta - (\eta \cdot \zeta)\zeta|} \cdot (-v(\zeta, \mathbf{e}))$$

$$= \frac{-\sin \theta \cos \theta_1 + \cos \theta \sin \theta_1 \cos \varphi}{|\eta - (\eta \cdot \zeta)\zeta|}$$

and

$$\cos \alpha = -v(\eta, \zeta) \cdot v_1(\eta) = \frac{-\zeta + (\eta \cdot \zeta)\eta}{|-\zeta + (\eta \cdot \zeta)\eta|} \cdot v_1(\eta)$$

$$= \frac{-\sin \theta \cos \theta_1 + \cos \theta \sin \theta_1 \cos \varphi}{\sqrt{1 - \sin^2 \theta_1 \sin^2 \varphi} \, |-\zeta + (\eta \cdot \zeta)\eta|}.$$

Since $|\eta - (\eta \cdot \zeta)\zeta| = |-\zeta + (\eta \cdot \zeta)\eta| = \sqrt{1 - (\eta \cdot \zeta)^2}$, we plainly obtain that $|\cos \beta| \le |\cos \alpha|$ and hence $\sin \alpha \le \sin \beta$. Using Lemma 10 (i) we then find

$$|d(\zeta, \eta) v_1(\eta)^* \sin \alpha| \le d(\zeta, \eta) \sin \beta \lesssim d(\mathbf{e}, \eta).$$

This takes care of the term $[II]$ in (3.22). Finally, we recall that all the action takes place inside a small ball $\mathcal{B}_\rho(\mathbf{e})$, which means that the angles θ and θ_1 are small. This yields an estimate for the term $[III]$ of the form

$$|d(\zeta, \eta) v_1(\eta) (\cos\alpha - \cos\beta)| \lesssim |\zeta - \eta| |\cos\alpha - \cos\beta|$$

$$= \frac{\sqrt{2(1 - (\eta \cdot \zeta))}}{\sqrt{1 - (\eta \cdot \zeta)^2}} |-\sin\theta\cos\theta_1 + \cos\theta\sin\theta_1\cos\varphi| \left(\frac{1}{\sqrt{1 - \sin^2\theta_1 \sin^2\varphi}} - 1 \right)$$

$$\lesssim \sin^2\theta_1 \lesssim \theta_1 = d(\mathbf{e}, \eta).$$

Combining (3.20), (3.21) and the bounds for the terms $[I], [II], [III]$ in (3.22), and using Lemma 11, we arrive at

$$(3.24) \quad \left| \frac{\sigma'(r)}{\sigma(r)} \fint_\mathcal{B} \nabla f(\eta)(-v(\eta, \zeta)) \frac{\sigma(d(\zeta, \eta))}{\sigma'(d(\zeta, \eta))} \, d\sigma(\eta) + \fint_\mathcal{B} \nabla f(\eta) \frac{\theta(\zeta)}{r} v_1(\eta) \, d\sigma(\eta) \right|$$

$$\lesssim_d \fint_\mathcal{B} |\nabla f(\eta)| \, \theta(\zeta) \, d\sigma(\eta) + \frac{1}{r} \fint_\mathcal{B} |\nabla f(\eta)| \, \theta(\eta) \, d\sigma(\eta).$$

Step 2. We continue our analysis with the term

$$\fint_\mathcal{B} \nabla f(\eta) \left(|S(\eta)| - 1 \right) \frac{\theta(\zeta)}{r} v_1(\eta) \, d\sigma(\eta).$$

From (3.23) we know that $|S(\eta)|^2 = \eta \cdot p(\eta)$, where $p(\eta)$ is the projection of η over the plane generated by ζ and $\mathbf{e}$. Therefore

$$(3.25) \quad \left| \fint_\mathcal{B} \nabla f(\eta) \left(|S(\eta)| - 1 \right) \frac{\theta(\zeta)}{r} v_1(\eta) \, d\sigma(\eta) \right|$$

$$\le \fint_\mathcal{B} |\nabla f(\eta)| \left(1 - |S(\eta)|^2 \right) \frac{\theta(\zeta)}{r} \, d\sigma(\eta)$$

$$\le \fint_\mathcal{B} |\nabla f(\eta)| \, |\eta \cdot (\eta - p(\eta))| \frac{\theta(\zeta)}{r} \, d\sigma(\eta)$$

$$\le \fint_\mathcal{B} |\nabla f(\eta)| \, |\eta - p(\eta)| \frac{\theta(\zeta)}{r} \, d\sigma(\eta)$$

$$\le \fint_\mathcal{B} |\nabla f(\eta)| \, \theta(\zeta) \, d\sigma(\eta).$$

Step 3. Combining (3.17), (3.18), (3.19), (3.24) and (3.25) we find that

$$\left| \fint_\mathcal{B} \nabla f(\eta) \, |S(\eta)| \frac{\theta(\xi)}{r} v_1(\eta) \, d\sigma(\eta) \right|$$

$$\lesssim_d \fint_\mathcal{B} |\nabla f(\eta)| \, \theta(\zeta) \, d\sigma(\eta) + \frac{1}{r} \fint_\mathcal{B} |\nabla f(\eta)| \, \theta(\eta) \, d\sigma(\eta),$$

and therefore

$$\left| \nabla \widetilde{\mathcal{M}} f(\xi) \right| = \left| \fint_\mathcal{B} \nabla f(\eta) \, S(\eta) \, d\sigma(\eta) \right|$$

$$\lesssim_d \frac{1}{\theta(\xi)} \fint_\mathcal{B} |\nabla f(\eta)| \, \theta(\eta) \, d\sigma(\eta) + \frac{r \, \theta(\zeta)}{\theta(\xi)} \fint_\mathcal{B} |\nabla f(\eta)| \, d\sigma(\eta).$$

This concludes the proof of the lemma. $\qquad\square$

3.2.4. Proof of Theorem 2 – Lipschitz case. We are now in position to move on to the proof of Theorem 2 when our initial datum f is a Lipschitz function. In this

case we also have $\widetilde{\mathcal{M}}f$ Lipschitz. Consider the set $\mathcal{H}_d = \{\xi \in \mathbf{S}^d \colon \widetilde{\mathcal{M}}f(\xi) > \widetilde{\mathcal{M}}^I f(\xi)\}$. In light of Proposition 8 it suffices to show that

$$\int_{\mathcal{H}_d} \left|\nabla\widetilde{\mathcal{M}}f(\xi)\right| \mathrm{d}\sigma(\xi) \lesssim_d \int_{\mathbf{S}^d} |\nabla f(\xi)| \, \mathrm{d}\sigma(\xi).$$

For each $\xi \in \mathbf{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ let us choose a ball $\overline{\mathcal{B}_{r_\xi}(\zeta_\xi)} \in \mathbf{B}_\xi$ with r_ξ minimal and, subject to this condition, with ζ_ξ in the half great circle connecting $\mathbf{e}, \xi, -\mathbf{e}$ in a way that $w(\zeta_\xi) = \min\{d(\mathbf{e}, \zeta_\xi), d(-\mathbf{e}, \zeta_\xi)\}$ is minimal. If there are two potential choices for ζ_ξ we choose the one with $0 \leq \theta(\zeta_\xi) \leq \theta(\xi)$.

First let us observe that we can restrict our attention to small balls. For $c > 0$, define the set $\mathcal{R}_c = \{\xi \in \mathcal{H}_d \setminus \{\mathbf{e}, -\mathbf{e}\} \colon r_\xi \geq c\}$. By Lemma 5 we find

$$\int_{\mathcal{R}_c} \left|\nabla\widetilde{\mathcal{M}}f(\xi)\right| \mathrm{d}\sigma(\xi) \leq \int_{\mathcal{R}_c} \frac{1}{\sigma(\mathcal{B}_{r_\xi}(\zeta_\xi))} \int_{\mathcal{B}_{r_\xi}(\zeta_\xi)} |\nabla f(\eta)| \, \mathrm{d}\sigma(\eta) \, \mathrm{d}\sigma(\xi)$$

$$\lesssim_{c,d} \int_{\mathbf{S}^d} |\nabla f(\eta)| \, \mathrm{d}\sigma(\eta).$$

If $\xi \in \mathcal{H}_d \setminus \{\mathbf{e}, -\mathbf{e}\}$ and r_ξ is small we must have $w(\zeta_\xi) < 4r_\xi$ (otherwise we would fall in the regime of the operator $\widetilde{\mathcal{M}}^I$). Assuming that $\xi \in \mathcal{H}_d \setminus \{\mathbf{e}, -\mathbf{e}\}$, that $\widetilde{\mathcal{M}}f$ is differentiable at ξ, and that $\nabla\widetilde{\mathcal{M}}f(\xi) \neq 0$ (which implies that $\xi \in \partial\mathcal{B}_{r_\xi}(\zeta_\xi)$), we may restrict ourselves to the situation where $d(\mathbf{e}, \xi) \leq \rho$ or $d(-\mathbf{e}, \xi) \leq \rho$ (where ρ is given by Lemma 12). By symmetry let us assume that $\theta(\xi) = d(\mathbf{e}, \xi) \leq \rho$. We call such set $\mathcal{G}_d$ and further decompose it in $\mathcal{G}_d^- = \{\xi \in \mathcal{G}_d \colon 0 \leq \theta(\zeta_\xi) < \theta(\xi)\}$ and $\mathcal{G}_d^+ = \{\xi \in \mathcal{G}_d \colon 0 < \theta(\xi) < \theta(\zeta_\xi)\}$. We bound the integrals over these two sets separately.

Step 1. For $\mathcal{G}_d^+$ we use Lemma 5 and proceed as follows:

(3.26)
$$\int_{\mathcal{G}_d^+} \left|\nabla\widetilde{\mathcal{M}}f(\xi)\right| \mathrm{d}\sigma(\xi) \leq \int_{\mathcal{G}_d^+} \fint_{\mathcal{B}_{r_\xi}(\zeta_\xi)} |\nabla f(\eta)| \, \mathrm{d}\sigma(\eta) \, \mathrm{d}\sigma(\xi)$$
$$= \int_{\mathbf{S}^d} |\nabla f(\eta)| \int_{\mathcal{G}_d^+} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta)}{\sigma(\mathcal{B}_{r_\xi}(\zeta_\xi))} \, \mathrm{d}\sigma(\xi) \, \mathrm{d}\sigma(\eta).$$

Note that $\theta(\eta) \geq \theta(\xi)$ in this case. Observe that

(3.27)
$$r_\xi > \frac{w(\zeta_\xi)}{4} = \frac{\theta(\zeta_\xi)}{4} \geq \frac{\theta(\xi)}{4},$$

and also, by triangle inequality,

(3.28)
$$r_\xi \geq \frac{d(\eta, \xi)}{2} \geq \frac{\theta(\eta)}{2} - \frac{\theta(\xi)}{2}.$$

Dividing (3.28) by 2 and adding up to (3.27) we get

$$r_\xi \geq \frac{\theta(\eta)}{6}.$$

Returning to the computation (3.26) we have, for a fixed η,

$$\int_{\mathcal{G}_d^+} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta)}{\sigma(\mathcal{B}_{r_\xi}(\zeta_\xi))} \, \mathrm{d}\sigma(\xi) \leq \int_{\mathcal{B}_{\theta(\eta)}(\mathbf{e})} \frac{1}{\sigma\left(\frac{\theta(\eta)}{6}\right)} \, \mathrm{d}\sigma(\xi) \simeq_d 1,$$

from which the required bound follows.

Step 2. We now bound the integral over $\mathcal{G}_d^-$ using Lemma 12. If $\xi \in \mathcal{G}_d^-$ then

$$(3.29) \qquad\qquad r_\xi \leq \theta(\xi) < 5r_\xi.$$

We then have

$$
\int_{\mathcal{G}_d^-} |\nabla \widetilde{\mathcal{M}} f(\xi)| \, \mathrm{d}\sigma(\xi) \lesssim_d \int_{\mathcal{G}_d^-} \left(\frac{1}{\theta(\xi)} \fint_{\mathcal{B}_{r_\xi}(\zeta_\xi)} |\nabla f(\eta)| \, \theta(\eta) \, \mathrm{d}\sigma(\eta) \right.
$$

$$
\left. + \frac{r_\xi \, \theta(\zeta_\xi)}{\theta(\xi)} \fint_{\mathcal{B}_{r_\xi}(\zeta_\xi)} |\nabla f(\eta)| \, \mathrm{d}\sigma(\eta) \right) \mathrm{d}\sigma(\xi)
$$

$$(3.30)$$

$$
\lesssim \int_{\mathbf{S}^d} |\nabla f(\eta)| \int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta) \, \theta(\eta)}{r_\xi \, \sigma(r_\xi)} \, \mathrm{d}\sigma(\xi) \, \mathrm{d}\sigma(\eta)
$$

$$
+ \int_{\mathbf{S}^d} |\nabla f(\eta)| \int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta) \, \theta(\zeta_\xi)}{\sigma(r_\xi)} \, \mathrm{d}\sigma(\xi) \, \mathrm{d}\sigma(\eta).
$$

Using (3.29) and the fact that $\theta(\zeta_\xi) \leq \theta(\xi)$ in this case, we have, for a fixed η,

$$
\int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta) \, \theta(\zeta_\xi)}{\sigma(r_\xi)} \, \mathrm{d}\sigma(\xi) \leq \int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta) \, \theta(\xi)}{\sigma(r_\xi)} \, \mathrm{d}\sigma(\xi)
$$

$$(3.31)$$

$$
\lesssim_d \int_0^\rho \frac{\theta \, (\sin\theta)^{d-1}}{\sigma(\theta)} \, \mathrm{d}\theta \lesssim_d 1,
$$

where we used Lemma 11 in the last inequality. For the other integral, we use (3.29), the fact that $\theta(\eta) \leq \theta(\xi)$ in this case, and Lemma 11 again to get

$$(3.32) \quad \int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta) \, \theta(\eta)}{r_\xi \, \sigma(r_\xi)} \, \mathrm{d}\sigma(\xi) \leq \theta(\eta) \int_{\theta(\eta)}^\rho \frac{(\sin\theta)^{d-1}}{r_\xi \, \sigma(r_\xi)} \, \mathrm{d}\theta \lesssim_d \theta(\eta) \int_{\theta(\eta)}^\rho \frac{1}{\theta^2} \, \mathrm{d}\theta \lesssim 1.$$

Our desired inequality plainly follows from inserting the bounds given by (3.31) and (3.32) into (3.30). This completes the proof of Theorem 2 in the Lipschitz case.

3.3. Passage to the general case. We will be brief here since the outline is the same as in §2.2. The following lemma is the analogue of Lemma 4 in the case of the sphere and we omit its proof.

Lemma 13. (i) *A polar function $f(\xi)$ is weakly differentiable in $\mathbf{S}^d \backslash \{\mathbf{e}, -\mathbf{e}\}$ if and only if its polar restriction $f(\theta)$ is weakly differentiable in $(0, \pi)$. In this case, the weak gradient ∇f of $f(\xi)$ and the weak derivative f' of $f(\theta)$ are related by*

$$
\nabla f(\xi) = f'(\theta(\xi))(-v(\xi, \mathbf{e})) = f'(\theta(\xi)) \frac{-\mathbf{e} + (\xi \cdot \mathbf{e})\xi}{|-\mathbf{e} + (\xi \cdot \mathbf{e})\xi|}.
$$

(ii) *In the situation above, if $f(\xi)$ and $\nabla f(\xi)$ are locally integrable in neighborhoods of $\mathbf{e}$ and $-\mathbf{e}$, then f is weakly differentiable in $\mathbf{S}^d$.*

Consider now a (nonnegative) polar function $f(\xi)$ in $W^{1,1}(\mathbf{S}^d)$. Then, by Lemma 13, its polar version $f(\theta)$ is weakly differentiable in $(0, \pi)$ and verifies

$$
\int_0^\pi |f'(\theta)| \, (\sin\theta)^{d-1} \, \mathrm{d}\theta < \infty.
$$

In particular, after a possible redefinition on a set of measure zero, one can take $f(\theta)$ continuous in $(0, \pi)$ (in fact, absolutely continuous in each compact interval of $(0, \pi)$). This is equivalent to assuming that $f(\xi)$ is continuous in $\mathbf{S}^d \backslash \{\mathbf{e}, -\mathbf{e}\}$.

In this case the detachment set

$$\mathcal{D}_d := \{\xi \in \mathbf{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\} \colon \widetilde{\mathcal{M}}f(\xi) > f(\xi)\}$$

is an open set. One can also show that $\widetilde{\mathcal{M}}f$ is continuous in $\mathbf{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ (see the ideas in Step 3 of §2.2), being indeed locally Lipschitz in $\mathcal{D}_d$ (see the ideas in the proof of Lemma 7, passage (3.5) and the remark thereafter). In particular, $\widetilde{\mathcal{M}}f$ is differentiable almost everywhere in $\mathcal{D}_d$.

Let $\{f_n\} \subset C^\infty(\mathbf{S}^d)$ be a sequence of nonnegative smooth functions such that $f_n \to f$ in $W^{1,1}(\mathbf{S}^d)$. We may simply assume that f_n is given by the spherical convolution of f with a smooth polar kernel φ_n (say, non-increasing in the polar angle) of integral 1 supported in the geodesic ball of radius $1/n$ centered at the north pole; see [10, Chapter 2, Sections 2.1 and 2.3, and Proposition 2.6.4] for details on the spherical convolution. We may also assume that $f_n \to f$ and $\nabla f_n \to \nabla f$ pointwise almost everywhere in $\mathbf{S}^d$ (say, outside a set $\mathcal{X} \subset \mathbf{S}^d$ of measure zero). Let $\xi \in \mathcal{D}_d \setminus \mathcal{X}$ be a point at which $\widetilde{\mathcal{M}}f$ is differentiable and all $\widetilde{\mathcal{M}}f_n$ are differentiable (this is still almost everywhere in $\mathcal{D}_d$). Note that for n large we shall have $\xi \in \{\widetilde{\mathcal{M}}f_n(\xi) > f_n(\xi)\}$. We now observe that if $\mathcal{B}_n = \mathcal{B}_{r_n}(\zeta_n)$ is a ball that realizes the maximal function $\widetilde{\mathcal{M}}f_n(\xi)$ with $r_n \to r$ and $\zeta_n \to \zeta$, then we must have $r > 0$ and the limiting ball $\mathcal{B}_r(\zeta)$ realizing the maximal function $\widetilde{\mathcal{M}}f(\xi)$. This plainly implies that

$$\widetilde{\mathcal{M}}f_n(\xi) \to \widetilde{\mathcal{M}}f(\xi)$$

as $n \to \infty$, and also, by Lemma 5,

$$\nabla \widetilde{\mathcal{M}}f_n(\xi) \to \nabla \widetilde{\mathcal{M}}f(\xi)$$

as $n \to \infty$.

Since we have proved Theorem 2 for Lipschitz functions, using Fatou's lemma we have

$$(3.33) \qquad \int_{\mathcal{D}_d} \left|\nabla \widetilde{\mathcal{M}}f(\xi)\right| \mathrm{d}\sigma(\xi) \leq \liminf_{n\to\infty} \int_{\mathcal{D}_d} \left|\nabla \widetilde{\mathcal{M}}f_n(\xi)\right| \mathrm{d}\sigma(\xi)$$
$$\lesssim_d \liminf_{n\to\infty} \|\nabla f_n\|_{L^1(\mathbf{S}^d)} = \|\nabla f\|_{L^1(\mathbf{S}^d)}.$$

This places us in position to adapt the one-dimensional argument of [9, Section 5.4] to show that $\widetilde{\mathcal{M}}f(\theta)$ is weakly differentiable in $(0, \pi)$, with weak derivative given by

$$(3.34) \qquad \chi_{\mathcal{D}_1^c} f'(\theta) + \chi_{\mathcal{D}_1}(\widetilde{\mathcal{M}}f)'(\theta),$$

where $\mathcal{D}_1 = \{\theta(\xi) \colon \xi \in \mathcal{D}_d\}$ is the polar version of $\mathcal{D}_d$. In fact, if $\theta \in \mathcal{D}_1^c$ is a point of differentiability of f (which are almost all points of $\mathcal{D}_1^c$) one can verify that $f'(\theta) = 0$, otherwise θ would belong to $\mathcal{D}_1$ instead. The weak derivative of $\widetilde{\mathcal{M}}f(\theta)$ is then simply $\chi_{\mathcal{D}_1}(\widetilde{\mathcal{M}}f)'(\theta)$. This in turn implies that $\widetilde{\mathcal{M}}f$ is weakly differentiable in $\mathbf{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ by Lemma 13. From (3.33) and (3.34) we have

$$(3.35) \qquad \left\|\nabla \widetilde{\mathcal{M}}f\right\|_{L^1(\mathbf{S}^d)} \lesssim_d \|\nabla f\|_{L^1(\mathbf{S}^d)},$$

which is our desired bound. From the Sobolev embedding we know that $f \in L^{d/(d-1)}(\mathbf{S}^d)$, and hence so does $\widetilde{\mathcal{M}}f$. In particular, $\widetilde{\mathcal{M}}f$ is locally integrable in $\mathbf{S}^d$. From (3.35) we already know that $\nabla \widetilde{\mathcal{M}}f$ is locally integrable in $\mathbf{S}^d$, and a further application of Lemma 13 shows that $\widetilde{\mathcal{M}}f$ is in fact weakly differentiable in $\mathbf{S}^d$, which completes our proof.

4. Proof of Theorem 3

We now turn our attention to the proof of Theorem 3. As presented in the introduction, the notation here is slightly different, as we denote our initial datum by u_0 and our maximal function by u^*. As usual, throughout this section, we assume that u_0 is real-valued and nonnegative (or $+\infty$).

4.1. Lipschitz case. As in the proofs of the previous two theorems in this paper, we address first the case when our polar $u_0 \in W^{1,1}(\mathbf{S}^d)$ is a Lipschitz function. In this case we have that u^* is a polar function that is also Lipschitz (see [5, Lemma 16 (ii)]).

4.1.1. A preliminary lemma. The following result will be important for our purposes.

Lemma 14. *Let $u_0 \colon \mathbf{S}^d \to \mathbf{R}^+$ be a polar and Lipschitz function. Then, in polar coordinates,*

$$\widetilde{\mathcal{M}}u_0\left(\tfrac{\pi}{2}\right) - u_0\left(\tfrac{\pi}{2}\right) \lesssim_d \|\nabla u_0\|_{L^1(\mathbf{S}^d)}.$$

Proof. Let us assume that $\widetilde{\mathcal{M}}u_0\left(\tfrac{\pi}{2}\right) > u_0\left(\tfrac{\pi}{2}\right)$. First observe that

$$\widetilde{\mathcal{M}}u_0\left(\tfrac{\pi}{2}\right) - u_0\left(\tfrac{\pi}{2}\right) = \left(\widetilde{\mathcal{M}}u_0\left(\tfrac{\pi}{2}\right) - \sup_{\theta \in [\frac{\pi}{4}, \frac{3\pi}{4}]} u_0(\theta)\right) + \left(\sup_{\theta \in [\frac{\pi}{4}, \frac{3\pi}{4}]} u_0(\theta) - u_0\left(\tfrac{\pi}{2}\right)\right),$$

and

$$\sup_{\theta \in [\frac{\pi}{4}, \frac{3\pi}{4}]} u_0(\theta) - u_0\left(\tfrac{\pi}{2}\right) \leq \int_{\frac{\pi}{4}}^{\frac{3\pi}{4}} |u_0'(\theta)| \, \mathrm{d}\theta \lesssim_d \int_{\frac{\pi}{4}}^{\frac{3\pi}{4}} |u_0'(\theta)| \, (\sin\theta)^{d-1} \, \mathrm{d}\theta \lesssim_d \|\nabla u_0\|_{L^1(\mathbf{S}^d)}.$$

Therefore it suffices to bound $\widetilde{\mathcal{M}}u_0\left(\tfrac{\pi}{2}\right) - \sup_{\theta \in [\frac{\pi}{4}, \frac{3\pi}{4}]} u_0(\theta)$. Bringing things back to the notation of § 3.1, let $\xi \in \mathbf{S}^d$ be such that $\theta(\xi) = \tfrac{\pi}{2}$ and let $\overline{\mathcal{B}} = \overline{\mathcal{B}_r(\zeta)} \in \mathbf{B}_\xi$. Let $\mathcal{Z} = \{\eta \in \mathbf{S}^d \colon \tfrac{\pi}{4} \leq \theta(\eta) \leq \tfrac{3\pi}{4}\}$. If $\mathcal{B} \subset \mathcal{Z}$, then $\widetilde{\mathcal{M}}u_0\left(\tfrac{\pi}{2}\right) - \sup_{\theta \in [\frac{\pi}{4}, \frac{3\pi}{4}]} u_0(\theta) \leq 0$ and we are done. Assume henceforth that $\mathcal{B} \not\subset \mathcal{Z}$ and that $\widetilde{\mathcal{M}}u_0\left(\tfrac{\pi}{2}\right) - \sup_{\theta \in [\frac{\pi}{4}, \frac{3\pi}{4}]} u_0(\theta) \geq 0$. Writing $\eta = (\cos\theta, (\sin\theta)\,\omega)$, with $\omega \in \mathbf{S}^{d-1}$, we define

$$\ell(\theta) = \int_{\mathbf{S}^{d-1}} \chi_{\mathcal{B}}(\eta) \, (\sin\theta)^{d-1} \, \mathrm{d}\sigma_{d-1}(\omega)$$

(that is, the $(d-1)$-dimensional measure of the intersection of $\mathcal{B}$ with the level set $d(\mathbf{e}, \eta) = \theta$). We then have

$$
\begin{aligned}
\widetilde{\mathcal{M}}u_0\left(\tfrac{\pi}{2}\right) &= \fint_{\mathcal{B}} u_0(\eta) \, \mathrm{d}\sigma(\eta) = \frac{1}{\sigma(\mathcal{B})} \int_0^\pi u_0(\theta) \, \ell(\theta) \, \mathrm{d}\theta \\
&= \frac{1}{\sigma(\mathcal{B})} \left(\int_0^{\frac{\pi}{4}} u_0(\theta) \, \ell(\theta) \, \mathrm{d}\theta + \int_{\frac{\pi}{4}}^{\frac{3\pi}{4}} u_0(\theta) \, \ell(\theta) \, \mathrm{d}\theta + \int_{\frac{3\pi}{4}}^\pi u_0(\theta) \, \ell(\theta) \, \mathrm{d}\theta \right) \\
&\leq \left(\sup_{\theta \in [\frac{\pi}{4}, \frac{3\pi}{4}]} u_0(\theta) \right) \frac{1}{\sigma(\mathcal{B})} \int_{\frac{\pi}{4}}^{\frac{3\pi}{4}} \ell(\theta) \, \mathrm{d}\theta \\
&\quad + \frac{1}{\sigma(\mathcal{B})} \left(\int_0^{\frac{\pi}{4}} u_0(\theta) \, \ell(\theta) \, \mathrm{d}\theta + \int_{\frac{3\pi}{4}}^\pi u_0(\theta) \, \ell(\theta) \, \mathrm{d}\theta \right).
\end{aligned}
$$

(4.1)

Now observe that

$$
\begin{aligned}
(4.2) \qquad \int_{\frac{3\pi}{4}}^{\pi} u_0(\theta)\,\ell(\theta)\,\mathrm{d}\theta &= \int_{\frac{3\pi}{4}}^{\pi} \left(\int_{\frac{3\pi}{4}}^{\theta} u_0'(\tau)\,\mathrm{d}\tau + u_0(\tfrac{3\pi}{4}) \right) \ell(\theta)\,\mathrm{d}\theta \\
&= u_0(\tfrac{3\pi}{4}) \int_{\frac{3\pi}{4}}^{\pi} \ell(\theta)\,\mathrm{d}\theta + \int_{\frac{3\pi}{4}}^{\pi} u_0'(\tau) \left(\int_{\tau}^{\pi} \ell(\theta)\,\mathrm{d}\theta \right) \mathrm{d}\tau.
\end{aligned}
$$

Plugging the bound

$$
\int_{\tau}^{\pi} \ell(\theta)\,\mathrm{d}\theta \lesssim_d \int_{\tau}^{\pi} (\sin\theta)^{d-1}\,\mathrm{d}\theta \lesssim (\sin\tau)^{d-1}
$$

into (4.2) we get

$$
(4.3) \qquad \int_{\frac{3\pi}{4}}^{\pi} u_0(\theta)\,\ell(\theta)\,\mathrm{d}\theta \leq \left(\sup_{\theta\in[\frac{\pi}{4},\frac{3\pi}{4}]} u_0(\theta) \right) \int_{\frac{3\pi}{4}}^{\pi} \ell(\theta)\,\mathrm{d}\theta + C_d \,\|\nabla u_0\|_{L^1(\mathbf{S}^d)},
$$

where C_d is a universal constant. In an analogous way we obtain

$$
(4.4) \qquad \int_{0}^{\frac{\pi}{4}} u_0(\theta)\,\ell(\theta)\,\mathrm{d}\theta \leq \left(\sup_{\theta\in[\frac{\pi}{4},\frac{3\pi}{4}]} u_0(\theta) \right) \int_{0}^{\frac{\pi}{4}} \ell(\theta)\,\mathrm{d}\theta + C_d \,\|\nabla u_0\|_{L^1(\mathbf{S}^d)}.
$$

Combining (4.1), (4.3) and (4.4) we get

$$
\widetilde{\mathcal{M}} u_0\big(\tfrac{\pi}{2}\big) \leq \sup_{\theta\in[\frac{\pi}{4},\frac{3\pi}{4}]} u_0(\theta) + C_d \,\|\nabla u_0\|_{L^1(\mathbf{S}^d)},
$$

from where our result follows. $\square$

4.1.2. Proof of Theorem 3 – Lipschitz case. Assume $d \geq 2$ since the case $d = 1$ has already been treated in [5, Theorem 3]. Define the detachment set (excluding the poles)

$$
\mathcal{A}_d = \{\xi \in \mathbf{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\} \colon u^*(\xi) > u_0(\xi)\}
$$

and its one-dimensional polar version

$$
\mathcal{A}_1 = \{\theta(\xi) \colon \xi \in \mathcal{A}_d\} \subset (0, \pi).
$$

These sets are open and from [5, Lemma 17] we know that u^* is subharmonic on $\mathcal{A}_d$. We write

$$
\mathcal{A}_1 = \bigcup_{i=0}^{\infty} (a_i, b_i)
$$

as a countable union of disjoint open intervals. If $\frac{\pi}{2} \in \mathcal{A}_1$ we let $\frac{\pi}{2} \in (a_0, b_0)$ and let

$$
\mathcal{A}_1^- = \bigcup_{(a_i,b_i)\subset\left(0,\frac{\pi}{2}\right)} (a_i, b_i) \quad \text{and} \quad \mathcal{A}_1^+ = \bigcup_{(a_i,b_i)\subset\left(\frac{\pi}{2},\pi\right)} (a_i, b_i).
$$

If $\frac{\pi}{2} \notin \mathcal{A}_1$ we just regard (a_0, b_0) as empty, and keep $\mathcal{A}_1^{\pm}$ as above.

Let (a, b) denote a generic interval (a_i, b_i) of this union. As in the proof of Theorem 1, the subharmonicity implies that u^* has no strict local maximum in (a, b) and then there exists τ with $a \leq \tau \leq b$ such that u^* is non-increasing in $[a, \tau]$ and non-decreasing in $[\tau, b]$. We then have $(u^*)'(\theta) \leq 0$ a.e. in $a < \theta < \tau$, and $(u^*)'(\theta) \geq 0$ a.e. in $\tau < \theta < b$.

An important idea of this proof is to proceed via the comparison (1.8) to the uncentered Hardy–Littlewood maximal function when appropriate, and make use

of the gradient bound established in Theorem 2. We consider first the case when $(a, b) \subset \mathcal{A}_1^-$. Using integration by parts we get

$$
\int_a^b \left|(u^*)'(\theta)\right| (\sin\theta)^{d-1}\, \mathrm{d}\theta
$$

$$
= -\int_a^\tau (u^*)'(\theta)\,(\sin\theta)^{d-1}\, \mathrm{d}\theta + \int_\tau^b (u^*)'(\theta)\,(\sin\theta)^{d-1}\, \mathrm{d}\theta
$$

$$
= u^*(a)\,(\sin a)^{d-1} + u^*(b)\,(\sin b)^{d-1} - 2\,u^*(\tau)\,(\sin\tau)^{d-1}
$$

(4.5)
$$
\quad + \int_a^\tau u^*(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\, \mathrm{d}\theta - \int_\tau^b u^*(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\, \mathrm{d}\theta
$$

$$
\leq u_0(a)\,(\sin a)^{d-1} + u_0(b)\,(\sin b)^{d-1} - 2\,u_0(\tau)\,(\sin\tau)^{d-1}
$$

$$
\quad + \int_a^\tau \widetilde{\mathcal{M}}u_0(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\, \mathrm{d}\theta - \int_\tau^b u_0(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\, \mathrm{d}\theta
$$

$$
\leq \int_a^b \left|u_0'(\theta)\right| (\sin\theta)^{d-1}\, \mathrm{d}\theta + \int_a^\tau \left(\widetilde{\mathcal{M}}u_0(\theta) - u_0(\theta)\right) \frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\, \mathrm{d}\theta.
$$

In the computation above we have taken advantage of the fact that $\frac{\partial}{\partial\theta}(\sin\theta)^{d-1} \geq 0$. Note also that we have no problem if $a = 0$ since $\lim_{a\to 0} u^*(a)\,(\sin a)^{d-1} = 0$ as $d \geq 2$. If we sum (4.5) over all the intervals $(a, b) \subset \mathcal{A}_1^-$ we find

$$
\int_{\mathcal{A}_1^-} \left|(u^*)'(\theta)\right| (\sin\theta)^{d-1}\, \mathrm{d}\theta
$$

$$
\leq \int_0^{\frac{\pi}{2}} \left|u_0'(\theta)\right| (\sin\theta)^{d-1}\, \mathrm{d}\theta + \int_0^{\frac{\pi}{2}} \left(\widetilde{\mathcal{M}}u_0(\theta) - u_0(\theta)\right) \frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\, \mathrm{d}\theta
$$

(4.6)
$$
= \int_0^{\frac{\pi}{2}} \left|u_0'(\theta)\right| (\sin\theta)^{d-1}\, \mathrm{d}\theta - \int_0^{\frac{\pi}{2}} \left((\widetilde{\mathcal{M}}u_0)'(\theta) - u_0'(\theta)\right)(\sin\theta)^{d-1}\, \mathrm{d}\theta
$$

$$
\quad + \left(\widetilde{\mathcal{M}}u_0\left(\tfrac{\pi}{2}\right) - u_0\left(\tfrac{\pi}{2}\right)\right) \lesssim_d \int_0^\pi \left|u_0'(\theta)\right| (\sin\theta)^{d-1}\, \mathrm{d}\theta,
$$

where we have used Theorem 2 and Lemma 14.

Finally we have to consider the case when $\frac{\pi}{2} \in \mathcal{A}_1$ and bound the integral $\int_{a_0}^{\pi/2} \left|(u^*)'(\theta)\right| (\sin\theta)^{d-1}\, \mathrm{d}\theta$. Let τ_0 be the corresponding local minimum over the interval (a_0, b_0). Let $c_0 = \min\{\tau_0, \frac{\pi}{2}\}$. Proceeding as in (4.5) and (4.6) we obtain

$$
-\int_{a_0}^{c_0} (u^*)'(\theta)(\sin\theta)^{d-1}\mathrm{d}\theta
$$

$$
= u^*(a_0)(\sin a_0)^{d-1} - u^*(c_0)(\sin c_0)^{d-1} + \int_{a_0}^{c_0} u^*(\theta)\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\mathrm{d}\theta
$$

(4.7)
$$
\leq u_0(a_0)\,(\sin a_0)^{d-1} - u_0(c_0)\,(\sin c_0)^{d-1} + \int_{a_0}^{c_0} \widetilde{\mathcal{M}}u_0(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\, \mathrm{d}\theta
$$

$$
= -\int_{a_0}^{c_0} u_0'(\theta)\,(\sin\theta)^{d-1}\, \mathrm{d}\theta + \int_{a_0}^{c_0} \left(\widetilde{\mathcal{M}}u_0(\theta) - u_0(\theta)\right) \frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\, \mathrm{d}\theta
$$

$$
\lesssim_d \int_0^\pi \left|u_0'(\theta)\right| (\sin\theta)^{d-1}\, \mathrm{d}\theta.
$$

The last estimate we need is the following

$$\int_{c_0}^{\frac{\pi}{2}} (u^*)'(\theta)\,(\sin\theta)^{d-1}\,\mathrm{d}\theta = u^*\big(\tfrac{\pi}{2}\big) - u^*(c_0)(\sin c_0)^{d-1} - \int_{c_0}^{\frac{\pi}{2}} u^*(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\,\mathrm{d}\theta$$

$$\leq \widetilde{\mathcal{M}}u_0\big(\tfrac{\pi}{2}\big) - u_0(c_0)(\sin c_0)^{d-1} - \int_{c_0}^{\frac{\pi}{2}} u_0(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\,\mathrm{d}\theta$$

(4.8)
$$= \Big(\widetilde{\mathcal{M}}u_0\big(\tfrac{\pi}{2}\big) - u_0\big(\tfrac{\pi}{2}\big)\Big) + \int_{c_0}^{\frac{\pi}{2}} u_0'(\theta)\,(\sin\theta)^{d-1}\,\mathrm{d}\theta$$

$$\lesssim_d \int_0^{\pi} \big|u_0'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta.$$

By combining (4.6), (4.7) and (4.8), and adding the integral over the set $\{u^* = u_0\}$ we find

$$\int_0^{\frac{\pi}{2}} \big|(u^*)'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta \lesssim_d \int_0^{\pi} \big|u_0'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta.$$

By symmetry we then have

$$\int_{\frac{\pi}{2}}^{\pi} \big|(u^*)'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta \lesssim_d \int_0^{\pi} \big|u_0'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta,$$

and the proof is complete by adding these two estimates.

4.2. Passage to the general case. The passage to the general case of a polar $f \in W^{1,1}(\mathbf{S}^d)$ follows closely the outline of §2.2, with Lemma 4 replaced by Lemma 13 when appropriate. We omit the details.

References

[1] Aldaz, J. M., L. Colzani, and J. Pérez Lázaro: Optimal bounds on the modulus of continuity of the uncentered Hardy–Littlewood maximal function. - J. Geom. Anal. 22, 2012, 132–167.

[2] Aldaz, J. M., and J. Pérez Lázaro: Functions of bounded variation, the derivative of the one dimensional maximal function, and applications to inequalities. - Trans. Amer. Math. Soc. 359:5, 2007, 2443–2461.

[3] Beltran, D., and J. Madrid: Endpoint Sobolev continuity of the fractional maximal function in higher dimensions. - Int. Math. Res. Not. IMRN (to appear).

[4] Beltran, D., J. P. Ramos, and O. Saari: Regularity of fractional maximal functions through Fourier multipliers. - J. Funct. Anal. 276:6, 2019, 1875–1892.

[5] Carneiro, E., R. Finder, and M. Sousa: On the variation of maximal operators of convolution type II. - Rev. Mat. Iberoam. 34, 2018, 739–766.

[6] Carneiro, E., and K. Hughes: On the endpoint regularity of discrete maximal operators. - Math. Res. Lett. 19:6, 2012, 1245–1262.

[7] Carneiro, E., and J. Madrid: Derivative bounds for fractional maximal functions. - Trans. Amer. Math. Soc. 369:6, 2017, 4063–4092.

[8] Carneiro, E., J. Madrid, and L. B. Pierce: Endpoint Sobolev and BV continuity for maximal operators. - J. Funct. Anal. 273, 2017, 3262–3294.

[9] Carneiro, E., and B. F. Svaiter: On the variation of maximal operators of convolution type. - J. Funct. Anal. 265, 2013, 837–865.

[10] Dai, F., and Y. Xu: Approximation theory and harmonic analysis on spheres and balls. - Springer, 2013.

[11] Hajłasz, P., and J. Malý: On approximate differentiability of the maximal function. - Proc. Amer. Math. Soc. 138, 2010, 165–174.

[12] Hajłasz, P., and J. Onninen: On boundedness of maximal functions in Sobolev spaces. - Ann. Acad. Sci. Fenn. Math. 29:1, 2004, 167–176.

[13] Hart, J., F. Liu, and Q. Xue: Regularity and continuity of local multilinear maximal type operator. - J. Geom. Anal. (to appear).

[14] Kinnunen, J.: The Hardy–Littlewood maximal function of a Sobolev function. - Israel J. Math. 100, 1997, 117–124.

[15] Kinnunen, J., and P. Lindqvist: The derivative of the maximal function. - J. Reine Angew. Math. 503, 1998, 161–167.

[16] Kinnunen, J., and E. Saksman: Regularity of the fractional maximal function. - Bull. London Math. Soc. 35:4, 2003, 529–535.

[17] Kurka, O.: On the variation of the Hardy–Littlewood maximal function. - Ann. Acad. Sci. Fenn. Math. 40, 2015, 109–133.

[18] Liu, F., and H. Wu: A note on the endpoint regularity of the discrete maximal operator. - Proc. Amer. Math. Soc. 147:2, 2019, 583–596.

[19] Luiro, H.: Continuity of the maximal operator in Sobolev spaces. - Proc. Amer. Math. Soc. 135:1, 2007, 243–251.

[20] Luiro, H.: The variation of the maximal function of a radial function. - Ark. Mat. 56:1, 2018, 147–161.

[21] Luiro, H., and J. Madrid: The variation of the fractional maximal function of a radial function. - Int. Math. Res. Not. IMRN 17, 2019, 5284–5298.

[22] Madrid, J.: Sharp inequalities for the variation of the discrete maximal function. - Bull. Aust. Math. Soc. 95:1, 2017, 94–107.

[23] Madrid, J.: Endpoint Sobolev and BV continuity for maximal operators II. - Rev. Mat. Iberoam. 35:7, 2019, 2151–2168.

[24] Pérez, C., T. Picon, O. Saari, and M. Sousa: Regularity of maximal functions on Hardy–Sobolev spaces. - Bull. Lond. Math. Soc. 50:6, 2018, 1007–1015.

[25] Ramos, J. P.: Sharp total variation results for maximal functions. - Ann. Acad. Sci. Fenn. Math. 44, 2019, 41–64.

[26] Saari, O.: Poincaré inequalities for the maximal function. - Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 19, 2019, 1065–1083.

[27] Schoen, R., and S.-T. Yau: Lectures in differential geometry. - International Press, 2010.

[28] Stein, E. M.: Singular integrals and differentiability properties of functions. - Princeton Univ. Press, 1970.

[29] Tanaka, H.: A remark on the derivative of the one-dimensional Hardy–Littlewood maximal function. - Bull. Austral. Math. Soc. 65:2, 2002, 253–258.