PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 145, Number 5, May 2017, Pages 2139–2153
http://dx.doi.org/10.1090/proc/13396
Article electronically published on December 30, 2016

DIFFERENTIATION OF BESOV SPACES AND THE NIKODYM MAXIMAL OPERATOR

JASON MURCKO

(Communicated by Alexander Iosevich)

ABSTRACT. We study several questions related to differentiation of integrals for Besov spaces relative to the basis $\mathcal{R}$ of arbitrarily oriented rectangular parallelepipeds in $\mathbb{R}^d$, $d \geq 2$. We improve on positive and negative differentiation results of Aimar, Forzani, and Naibo and on capacitary and dimensional bounds for exceptional sets of Naibo. Our main tool in obtaining these improvements involves showing that bounds for the Nikodym maximal operator can be used to deduce boundedness properties of the local maximal operator associated to $\mathcal{R}$.

1. INTRODUCTION

A *differentiation basis* $\mathcal{B}$ in $\mathbb{R}^d$ consists of, for each $x \in \mathbb{R}^d$, a family $\mathcal{B}_x$ of bounded, Lebesgue measurable sets containing x, also satisfying:

(1) each set $F \in \mathcal{B}_x$ has positive Lebesgue measure, and
(2) each $\mathcal{B}_x$ contains sets of arbitrarily small positive diameter.

(We shall usually abbreviate "differentiation basis" simply to "basis".) We say that a basis $\mathcal{B}$ *differentiates the integrals* of an appropriate function f if, for Lebesgue–almost every x,

$$\lim_{\substack{\operatorname{diam}(F)\to 0 \\ F\in\mathcal{B}_x}} \frac{1}{|F|}\int_F f(y)\,dy = f(x).$$

And if $\mathcal{V}$ is a class or space of functions, we say that $\mathcal{B}$ differentiates $\mathcal{V}$ if $\mathcal{B}$ differentiates the integrals of every $f \in \mathcal{V}$.

In the present work we shall only concern ourselves with *translation-invariant* bases: those which satisfy $\mathcal{B}_x = \{x + F : F \in \mathcal{B}_0\}$ for all x. Moreover, we shall be working entirely with bases in $\mathbb{R}^d$ with $d \geq 2$.

Let us now recall two key differentiation results relating to concrete bases of interest. First, Nagel, Stein, and Wainger [7] showed that bases in $\mathbb{R}^2$ consisting of rectangles having "lacunary" directions differentiate L^p for $p > 1$. On the other hand, based on the construction of the Nikodym set [9], Zygmund observed that the basis of arbitrarily oriented rectangles in $\mathbb{R}^2$ does not even differentiate the class of characteristic functions of measurable sets. This final result also extends to $\mathbb{R}^d$ when we use the basis of arbitrarily oriented rectangular parallelepipeds, which we henceforth denote by $\mathcal{R}$.

Received by the editors July 11, 2016.
2010 *Mathematics Subject Classification.* Primary 42B25, 42B35.
The author would like to thank his advisor, Andreas Seeger, for his guidance and support.
The author was supported in part by the National Science Foundation.

For such bases, where there is no L^p theory available, one can still ask about regularity conditions on functions that are necessary or sufficient to recover a differentiation result. Stokolos [11, 12] seems to have been the first to pursue this line of inquiry, finding that the basis of convex sets in $\mathbb{R}^d$ differentiates the inhomogeneous Besov spaces $B^p_{(d-1)/p,1}$ for $p > d-1$. Note that this result implies that $\mathcal{R}$ differentiates these same spaces. Building on Stokolos's work, Aimar, Forzani, and Naibo [2] showed that $\mathcal{R}$ differentiates the spaces $B^p_{(d-1)/p,1}$ for $1 \leq p < \infty$ and that $\mathcal{R}$ does not differentiate $B^p_{\alpha,1}$ for $1 \leq p < d$ and $0 < \alpha < d/p - 1$. Note that the gap in these final results between the primary smoothness exponent for which there is a positive result ($\alpha = (d-1)/p$) and those for which there is a negative result ($\alpha < d/p-1$) leaves substantially open the question of the *minimal* regularity needed to ensure a differentiation result for $\mathcal{R}$.

A related question may be phrased generally in the following way. Suppose a basis $\mathcal{B}$ is known to differentiate a space of functions $\mathcal{V}$, for example a Lebesgue or Sobolev or Besov space. If $\mathcal{V}' \subseteq \mathcal{V}$ is a subspace consisting of functions possessing an additional degree of regularity beyond that in $\mathcal{V}$ as a whole, what can be said about the $\mathcal{B}$*–exceptional set*

$$E_{\mathcal{B}}(f) = \bigcap_{g:g=f\,\text{a.e.}} \left\{ x \in \mathbb{R}^d : \Gamma_{\mathcal{B}} g(x) = \limsup_{\substack{\operatorname{diam}(F)\to 0 \\ F\in\mathcal{B}_x}} \frac{1}{|F|} \int_F |g(y) - g(x)|\, dy > 0 \right\} \tag{1.1}$$

of functions $f \in \mathcal{V}'$? One typically knows that these exceptional sets have Lebesgue measure 0 as a consequence of the method for proving the underlying differentiation result, but it is natural to suspect that one can also control more refined measures of their size. For example, Bagby and Ziemer [3] found that with respect to the basis of centered balls in $\mathbb{R}^d$—for which Lebesgue's theorem holds—the exceptional set of a function in the inhomogeneous Sobolev space L^p_α, $1 < p < \infty$ and $0 < \alpha \leq d/p$, has Hausdorff dimension not exceeding $d - \alpha p$. Similarly but parallelling the differentiation results for $\mathcal{R}$ mentioned above, Naibo [8] found that functions in $B^p_{\alpha,q}$, with $p > 1$, $(d-1)/p < \alpha \leq d/p$ and $1 \leq q \leq \infty$, have $\mathcal{R}$–exceptional sets that are small in the sense of a certain Bessel capacity, implying that their Hausdorff dimension does not exceed $2d - 1 - \alpha p = d - [\alpha - (d-1)/p]p$.

Our main findings improve on the positive and negative differentiation results in [2] and the exceptional set results in [8], all in connection with the basis $\mathcal{R}$. Most of these improvements come from new estimates for the associated local maximal operator, given by

$$\mathcal{M}_{\mathcal{R}} f(x) = \sup_{\substack{\operatorname{diam}(R)\leq 1 \\ R\in\mathcal{R}_0}} \frac{1}{|R|} \int_{x+R} |f(y)|\, dy. \tag{1.2}$$

Here $\mathcal{R}_0$ is the family of arbitrarily oriented rectangular parallelepipeds containing the origin.

Perhaps the key insight is that we can use estimates for the Nikodym maximal operator to derive estimates for $\mathcal{M}_{\mathcal{R}}$. Recall that for a small parameter $\delta > 0$, this operator is defined by

$$\mathcal{N}_\delta f(x) = \sup_{T\in\mathcal{T}^\delta} \frac{1}{|T|} \int_{x+T} |f(y)|\, dy, \tag{1.3}$$

where $\mathcal{T}^\delta$ denotes the family of cylindrical tubes of length 1 and radius δ which are centered at the origin. The main conjecture for $\mathcal{N}_\delta$ is that for p and q satisfying $1 \le p \le d$ and $p \le q \le (d-1)p'$, there are $L^p \to L^q$ estimates for $\mathcal{N}_\delta$ as follows:

(1) If $p < d$, then $\|\mathcal{N}_\delta f\|_q \le C\delta^{-\alpha}\|f\|_p$ for $\alpha \ge d/p - 1$, with the constant C independent of δ.
(2) If $p = d$, then $\|\mathcal{N}_\delta f\|_d \le C\delta^{-\alpha}\|f\|_d$ for $\alpha > 0$, with the constant C again independent of δ.

The conjecture has been verified in dimension 2 (see, in part, [6]), but for $d \ge 3$ only partial results are known. For our purposes, the relevant facts are that the estimate holds for p in the range $1 \le p \le (d+1)/2$, with the full ranges of q and α indicated in the conjecture, and that it also holds for $(d+1)/2 < p < (d+2)/2$, with the full range of q but with $\alpha > d/p-1$, i.e. not including the endpoint $\alpha = d/p-1$. (For the case $p < (d+1)/2$ see [4]; the endpoint $p = (d+1)/2$ can be obtained by combining the X-ray transform estimates in [5] with arguments in [13], and for the remaining p, see [15].)

We now state our main results. Aside from the negative differentiation result in Theorem 1.3, we state them in a conditional form, dependent on estimates for $\mathcal{N}_\delta$.

Theorem 1.1. *Let $\mathcal{R}$ be the basis of arbitrarily oriented rectangular parallelepipeds in $\mathbb{R}^d$. An estimate $\|\mathcal{N}_\delta f\|_q \lesssim_{p,q,\alpha} \delta^{-\alpha}\|f\|_p$[1] for the Nikodym maximal operator on $\mathbb{R}^d$ with $1 \le p \le \infty$, $q \ge p$, and $\alpha > d(1/p - 1/q)$ implies that the maximal operator $\mathcal{M}_\mathcal{R}$ is bounded from the Besov space $B^p_{\alpha,1}$ to L^q.*

Theorem 1.2. *An estimate $\|\mathcal{N}_\delta f\|_p \lesssim_{p,\alpha} \delta^{-\alpha}\|f\|_p$ for the Nikodym maximal operator on $\mathbb{R}^d$ with $1 \le p < \infty$ and $\alpha > 0$ implies that the basis $\mathcal{R}$ in $\mathbb{R}^d$ differentiates the Besov space $B^p_{\alpha,1}$.*

Theorem 1.3. *Let $1 \le p < d$, and let $q > d$. Then the basis $\mathcal{R}$ does not differentiate $B^p_{d/p-1,q}$. More specifically, there exists $f \in B^p_{d/p-1,q}$ so that*

$$\limsup_{\substack{\operatorname{diam}(R)\to 0\\ R\in\mathcal{R}_0}} \frac{1}{|R|}\int_{x+R} f(y)\,dy = \infty$$

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

Suppose that for a given p in the range $1 \le p < d$, the optimal $L^p \to L^p$ Nikodym estimate with respect to α holds; that is, there is an estimate $\|\mathcal{N}_\delta f\|_p \lesssim \delta^{-(d/p-1)}\|f\|_p$. Then Theorems 1.2 and 1.3 close most of the gap from [2] mentioned above in terms of determining the minimal regularity needed to ensure a differentiation result for $\mathcal{R}$. More specifically, the only Besov spaces for which we do not then have a positive or negative result are $B^p_{d/p-1,q}$ with $1 < q \le d$.

Theorem 1.4. *For a fixed p in the range $1 < p \le d$, assume that the Nikodym estimate $\|\mathcal{N}_\delta f\|_d \lesssim_\varepsilon \delta^{-(d/p-1+\varepsilon)}\|f\|_p$ is known to hold for all $\varepsilon > 0$. Then if $g \in B^p_{\alpha,q}$, with $d/p - 1 < \alpha \le d/p$ and $1 \le q \le \infty$, the Hausdorff dimension of the $\mathcal{R}$–exceptional set $E_\mathcal{R}(g)$ is at most $d(d/p - \alpha)$.*

---

[1]We use the notation $A \lesssim B$ to represent the inequality $A \le CB$; we shall generally specify what parameters (for example, the Lebesgue exponent p, the smoothness exponent α, the frequency scale parameter j) the implicit constant C does or does not depend on, although we shall uniformly ignore any dependence on the dimension d. Most of the time, we shall indicate the dependence of the implicit constant on various parameters by using a subscript, as in $A \lesssim_p B$. We shall also write $A \approx B$ if $A \lesssim B$ and $B \lesssim A$.

The layout of this paper is as follows. In Section 2 we prove Theorems 1.1 and 1.2, with the former requiring most of the work. The idea is to reduce Theorem 1.1 to a result for frequency-localized functions, while also controlling $\mathcal{M}_{\mathcal{R}} f$ by averages over certain line segments. For frequency-localized functions, the averages over line segments can then be "smeared" to averages over tubes, which allows us to employ Nikodym estimates to our advantage. In Section 3 we prove Theorem 1.3. As is typically the case with differentiation counterexamples, showing that the maximal operator $\mathcal{M}_{\mathcal{R}}$ satisfies an unboundedness condition between appropriate spaces is key. Finally, in Section 4 we prove Theorem 1.4, using $B^p_{\alpha,1} \to L^q$ bounds for $\mathcal{M}_{\mathcal{R}}$—instead of the $B^p_{\alpha,1} \to L^p$ bounds employed in Theorem 1.2—to show that a certain Bessel capacity of the $\mathcal{R}$–exceptional set is 0.

2. Positive differentiation results

We begin this section by reviewing the fundamentals of (inhomogeneous) Besov spaces. First, let η be a C^∞, nondecreasing function on $\mathbb{R}$ such that $\eta(t) = 0$ for $t \le 1$ and $\eta(t) = 1$ for $t \ge 2$. Next define $\tilde{\eta}_j$ on $\mathbb{R}^d$ for $j \ge 0$ by $\tilde{\eta}_0(\xi) = 1 - \eta(|\xi|)$ and $\tilde{\eta}_j(\xi) = \eta(2^{-j+1}|\xi|) - \eta(2^{-j}|\xi|)$ for $j \ge 1$, so that $\sum_{j\ge 0} \tilde{\eta}_j \equiv 1$. Let Δ_j be the multiplier operator associated with $\tilde{\eta}_j$, i.e. $\widehat{\Delta_j f} = \tilde{\eta}_j \hat{f}$ for tempered distributions f. Finally, the Besov space $B^p_{\alpha,q}$ on $\mathbb{R}^d$, with $0 < p, q \le \infty$ and $\alpha \in \mathbb{R}$, consists of those tempered distributions f satisfying

$$\left(\sum_{j\ge 0} \left(2^{j\alpha} \|\Delta_j f\|_p \right)^q \right)^{1/q} < \infty,$$

with the obvious modification if $q = \infty$.

When $1 \le p, q \le \infty$, $B^p_{\alpha,q}$ is a Banach space with norm given by the quantity on the left above. Note that different choices of the function η give rise to equivalent norms and so do not affect the definition. We also mention several continuous embeddings which will be important to us at various times. The first two embeddings arise in a straightforward fashion from the form of the norm expression:

(1) $B^p_{\alpha_2,q_2} \subseteq B^p_{\alpha_1,q_1}$ if $\alpha_1 < \alpha_2$, for any $p, q_1, q_2 > 0$;
(2) $B^p_{\alpha,q_1} \subseteq B^p_{\alpha,q_2}$ if $q_1 < q_2$, for any $p > 0$ and $\alpha \in \mathbb{R}$.

A third embedding, analogous to the Sobolev embedding theorem, is

(3) $B^{p_1}_{\alpha_1,q} \subseteq B^{p_2}_{\alpha_2,q}$ if $p_1 < p_2$ and $\alpha_1 - d/p_1 = \alpha_2 - d/p_2$, for any $q > 0$.

Last, a basic embedding relationship with Lebesgue spaces is

(4) $B^p_{\alpha,q} \subseteq L^p$ if either $\alpha > 0$ or $\alpha = 0$ and $q = 1$, for any $p \ge 1$.

Moving more directly into our treatment of Theorems 1.1 and 1.2, we shall actually prove slightly stronger statements where the basis $\mathcal{R}$ is replaced by a larger basis. We define the translation-invariant basis $\mathcal{S}$ by specifying that $\mathcal{S}_0$ consists of all (bounded, Lebesgue measurable, nonnull) sets that are star-shaped with respect to the origin. That is, a set F belongs to $\mathcal{S}_0$ if, whenever $x \in F$, the line segment joining x to 0 lies within F.

Without fully rewriting the modifications of Theorems 1.1 and 1.2 that we now aim to prove (which require only the replacement of $\mathcal{R}$ by $\mathcal{S}$ in the theorem statements), let us agree to refer to the modified results as Theorems 1.1′ and 1.2′. The original theorems follow from these alternate versions since $\mathcal{R}$ is a subbasis of $\mathcal{S}$.

Let us also dispense with Theorem 1.2′ straightaway by seeing how it is implied by Theorem 1.1′. The argument is quite similar to the general schema for showing that a maximal function estimate implies a differentiation result.

Proof that Theorem 1.1′ ⇒ *Theorem* 1.2′. Fix $f \in B^p_{\alpha,1}$ and $\varepsilon > 0$. It suffices to show that the set $R_\varepsilon(f) = \{x : \Gamma_{\mathcal{S}} f(x) > \varepsilon\}$ has measure 0.

If g is a Schwartz function, then for $x \in R_\varepsilon(f)$ we have

$$\begin{aligned}\varepsilon &< \limsup_{\substack{\operatorname{diam}(F)\to 0\\ F\in\mathcal{S}_0}} \frac{1}{|F|}\int_{x+F} \big(\,|f(y)-g(y)| + |g(y)-g(x)| + |g(x)-f(x)|\,\big)\,dy \\ &\le \mathcal{M}_{\mathcal{S}}(f-g)(x) + |f(x)-g(x)|.\end{aligned}$$

Hence $R_\varepsilon(f) \subseteq \{\, x : \mathcal{M}_{\mathcal{S}}(f-g)(x) > \varepsilon/2 \,\} \cup \{\, x : |f(x)-g(x)| > \varepsilon/2 \,\}$, and so

$$|R_\varepsilon(f)| \lesssim_p \varepsilon^{-p}\|\mathcal{M}_{\mathcal{S}}(f-g)\|_p^p + \varepsilon^{-p}\|f-g\|_p^p.$$

The hypotheses of Theorem 1.2′, when combined with Theorem 1.1′, imply that $\mathcal{M}_{\mathcal{S}}$ is bounded from $B^p_{\alpha,1}$ to L^p. Using also the continuous embedding $B^p_{\alpha,1} \subseteq L^p$, we obtain $|R_\varepsilon(f)| \lesssim_{p,\alpha} \varepsilon^{-p}\|f-g\|^p_{B^p_{\alpha,1}}$. By choosing g arbitrarily close to f in $B^p_{\alpha,1}$, which is possible since $p < \infty$ (see [14, p. 48]), we now see that $|R_\varepsilon(f)| = 0$. □

2.1. **Frequency localization.** We now reduce Theorem 1.1′ to a maximal operator estimate for frequency-localized functions. Building up to the proof of the frequency-localized estimate will then occupy the remainder of Section 2.

Proposition 2.1.

(i) *An estimate* $\|\mathcal{N}_\delta f\|_q \lesssim_{p,q,\alpha} \delta^{-\alpha}\|f\|_p$ *for the Nikodym maximal operator on* $\mathbb{R}^d$, *where* $1 \le p \le q$ *and* $\alpha > d(1/p - 1/q)$, *implies an estimate*

$$\|\mathcal{M}_{\mathcal{S}} f\|_q \lesssim 2^{j\alpha}\|f\|_p,$$

valid for functions f *whose Fourier transform is supported in the annulus* $\mathcal{A}_j = \{\, 2^{j-1} \le |\xi| \le 2^{j+1} \,\}$, *with* $j \ge 1$. *The implicit constant may depend on* p, q, *and* α, *but is independent of* j.

(ii) *For functions* f *whose Fourier transform is supported in the ball* $\{\, |\xi| \le 2 \,\}$, *there is an estimate*

$$\|\mathcal{M}_{\mathcal{S}} f\|_q \lesssim \|f\|_p$$

when $1 \le p \le q$. *The implicit constant is independent of* p *and* q.

Proof that Proposition 2.1 ⇒ *Theorem* 1.1′. Assume that there is a Nikodym estimate as in the hypothesis of Theorem 1.1′, i.e. $\|\mathcal{N}_\delta f\|_q \lesssim_{p,q,\alpha} \delta^{-\alpha}\|f\|_p$. Let $f \in B^p_{\alpha,1}$, and expand f as $\sum_{j\ge 0}\Delta_j f$, where the Δ_j are the same Littlewood-Paley projections introduced in our definition of Besov spaces above. Since the $\Delta_j f$ are frequency-localized in accordance with Proposition 2.1, we obtain

$$\|\mathcal{M}_{\mathcal{S}} f\|_q \le \sum_{j\ge 0}\|\mathcal{M}_{\mathcal{S}}(\Delta_j f)\|_q \lesssim_{p,q,\alpha} \sum_{j\ge 0} 2^{j\alpha}\|\Delta_j f\|_p = \|f\|_{B^p_{\alpha,1}}.$$

□

2.2. **Averages over lines.** Here we prove the following elementary result:

Lemma 2.2. *Let f be any continuous function on $\mathbb{R}^d$. Then*

$$\mathcal{M}_{\mathcal{S}} f(x) \lesssim \sup_{\theta \in S^{d-1}} \sup_{0<L\leq 1} \frac{1}{L} \int_0^L |f(x+t\theta)|\, dt. \tag{2.1}$$

Proof. Consider a set $F \in \mathcal{S}_0$ with $\operatorname{diam}(F) \leq 1$. Set $r(\theta) = \sup\{\, t > 0 \mid t\theta \in F \,\}$ for $\theta \in S^{d-1}$, so that up to a null set $F = \{\, t\theta \mid \theta \in S^{d-1} \text{ and } 0 \leq t \leq r(\theta) \,\}$. Note that $r(\theta) \leq 1$ for all θ.

Now we write the integral of f over $x + F$ in polar coordinates and use Hölder's inequality:

$$\begin{aligned}
\frac{1}{|F|} \int_{x+F} |f(y)|\, dy &= \frac{1}{|F|} \int_{S^{d-1}} \int_0^{r(\theta)} |f(x+t\theta)|\, t^{d-1}\, dt\, d\sigma(\theta) \\
&\leq \left(\sup_{\theta \in S^{d-1}} \frac{1}{r(\theta)} \int_0^{r(\theta)} |f(x+t\theta)|\, dt \right) \cdot \frac{1}{|F|} \int_{S^{d-1}} r(\theta)^d\, d\sigma(\theta) \\
&\leq \left(\sup_{\theta \in S^{d-1}} \sup_{0<L\leq 1} \frac{1}{L} \int_0^L |f(x+t\theta)|\, dt \right) \cdot d.
\end{aligned}$$

Since the final expression is independent of $F \in \mathcal{S}_0$, we obtain (2.1). Note we have used that $|F| = \int_{S^{d-1}} \int_0^{r(\theta)} t^{d-1}\, dt\, d\sigma(\theta) = d^{-1} \int_{S^{d-1}} r(\theta)^d\, d\sigma(\theta)$. $\square$

2.3. **Averages over tubes.** We now assume that f is a function whose Fourier transform is supported either in the annulus $\mathcal{A}_j = \{\, 2^{j-1} \leq |\xi| \leq 2^{j+1} \,\}$, with $j \geq 1$, or in the ball $\{\, |\xi| \leq 2 \,\}$, which for convenience we denote by $\mathcal{A}_0$. Our goal is to show how an average of f over a line segment can be controlled by averages instead over cylindrical tubes of radius $\geq 2^{-j}$. To state our main result, let

$$\psi_j(x) = 2^{jd}(1 + 2^j|x|)^{-(d+1)}, \tag{2.2}$$

so that $\|\psi_j\|_1$ is finite and independent of j. (We note that the exponent $d+1$ could equally well be replaced, in all of what follows, by any $N > d$.)

Lemma 2.3. *Let f be a function with* $\operatorname{supp}(\hat{f}) \subseteq \mathcal{A}_j$, *$j \geq 0$, and let $0 < L \leq 1$.*

(i) *If $L \leq 2^{-j}$, then*

$$\frac{1}{2L} \int_{-L}^{L} |f(x_1, 0, \ldots, 0)|\, dx_1 \lesssim \psi_j * |f|(0).$$

(ii) *If $L > 2^{-j}$, with $2^{k-1} < L \leq 2^k$, then*

$$\frac{1}{2L} \int_{-L}^{L} |f(x_1, 0, \ldots, 0)|\, dx_1 \lesssim \psi_j * |f|(0) + \sum_{\ell=-j}^{k+1} \frac{2^{-(j+\ell)}}{|T_{k,\ell}|} \int_{T_{k,\ell}} |f(x)|\, dx;$$

here $T_{k,\ell}$ is the cylindrical tube, centered at the origin and with axis in the x_1 direction, having radius 2^ℓ and length 2^{k+2}.

In each case, the implicit constant is independent of j, L, and k.

The following auxiliary result is the key ingredient we need for Lemma 2.3.

Lemma 2.4. *For $0 < L \leq 1$, let μ_L be the measure on $\mathbb{R}^d$ given by*

$$\int f\, d\mu_L = \frac{1}{2L}\int_{-L}^{L} f(x_1, 0, \ldots, 0)\, dx_1$$

for continuous f.

(i) *If $L \leq 2^{-j}$, then there is a pointwise estimate*

$$\mu_L * \psi_j \lesssim \psi_j.$$

(ii) *If $L > 2^{-j}$ with $2^{k-1} < L \leq 2^k$, then there is a pointwise estimate*

$$\mu_L * \psi_j \lesssim \psi_j + \sum_{\ell=-j}^{k+1} \frac{2^{-(j+\ell)}}{|T_{k,\ell}|}\chi_{T_{k,\ell}};$$

here $T_{k,\ell}$ again denotes the cylindrical tube, centered at the origin and with axis in the x_1 direction, having radius 2^ℓ and length 2^{k+2}.

In each case, the implicit constant is independent of j, L, and k.

Proof. We first write out, for reference, that

$$\mu_L * \psi_j(x) = \int \psi_j(x-y)\, d\mu_L(y) = \frac{1}{2L}\int_{-L}^{L} 2^{jd}(1 + 2^j|x - se_1|)^{-(d+1)}\, ds,$$

where $e_1 = (1, 0, \ldots, 0)$.

Now suppose that $|x| \geq 2L$. (This includes, for (ii), all points not in any of the tubes $T_{k,\ell}$.) Then $|x - se_1| \geq |x|/2$ for $|s| \leq L$, and so

$$\mu_L * \psi_j(x) \leq \frac{1}{2L}\int_{-L}^{L} 2^{jd}(1 + 2^{j-1}|x|)^{-(d+1)}\, ds \lesssim \psi_j(x).$$

To complete (i), we additionally observe that $\mu_L * \psi_j(x) \leq 2^{jd} \approx \psi_j(x)$ for $|x| \leq 2L$.

For the remainder of (ii), suppose that $x = (x_1, x')$ belongs to one of the tubes $T_{k,\ell}$. It is not difficult to check that

$$\mu_L * \psi_j(x) \leq \mu_L * \psi_j(0, x') \approx \frac{1}{2L}\int_{-L}^{L} 2^{jd}(1 + 2^j|x'| + 2^j|s|)^{-(d+1)}\, ds.$$

We now consider two cases.

If $x \in T_{k,-j}$ so that $|x'| \leq 2^{-j}$, we estimate

$$\mu_L * \psi_j(x) \lesssim \frac{1}{2L}\int_{-\infty}^{\infty} 2^{jd}(1 + 2^j|s|)^{-(d+1)}\, ds \approx 2^{j(d-1)}2^{-k} \approx \frac{1}{|T_{k,-j}|}.$$

And if $x \in T_{k,\ell} \setminus T_{k,\ell-1}$ so that $2^{\ell-1} \leq |x'| \leq 2^\ell$, then

$$\mu_L * \psi_j(x) \lesssim \frac{1}{2L}\int_{-\infty}^{\infty} 2^{jd}(2^j|x'| + 2^j|s|)^{-(d+1)}\, ds \approx \frac{2^{-j}}{|x'|^d \cdot 2^k} \approx \frac{2^{-(j+\ell)}}{|T_{k,\ell}|}.$$

□

Proof of Lemma 2.3. Fix a Schwartz function φ whose Fourier transform is identically 1 on the annulus $\{\, 1/2 \leq |\xi| \leq 2 \,\}$. Set $\varphi_j(x) = 2^{jd}\varphi(2^j x)$ if $j \geq 1$, and let φ_0 be a fixed Schwartz function whose Fourier transform is identically 1 on the ball $\{\, |\xi| \leq 2 \,\}$. Then $\hat{\varphi}_j$ is identically 1 on $\mathcal{A}_j$ for all $j \geq 0$. We also have that $|\varphi_j| \lesssim \psi_j$ pointwise, with an implicit constant independent of j.

Now suppose $\hat{f}$ is supported in $\mathcal{A}_j$, so that $f = f * \varphi_j$. Then

$$\frac{1}{2L}\int_{-L}^{L} |f(x_1, 0, \ldots, 0)|\, dx_1 = \mu_L * |f * \varphi_j|(0) \lesssim (\mu_L * \psi_j) * |f|(0).$$

Substituting the estimates for $\mu_L * \psi_j$ from each part of Lemma 2.4 now immediately establishes the results. □

2.4. **Proof of Proposition 2.1.** As in the previous section, we assume here that $\hat{f}$ is supported in some $\mathcal{A}_j$, $j \geq 0$, and we take $\psi_j(x) = 2^{jd}(1 + 2^j|x|)^{-(d+1)}$. We also introduce the following notation: for integers k and ℓ with $\ell \leq k$, let $\mathcal{T}_{k,\ell}$ denote the family of cylindrical tubes centered at the origin having length 2^k and radius 2^ℓ. Let $\mathcal{N}_{k,\ell}$ denote the associated Nikodym-type maximal operator, i.e.

$$\mathcal{N}_{k,\ell} f(x) = \sup_{T \in \mathcal{T}_{k,\ell}} \frac{1}{|T|} \int_{x+T} |f(y)|\, dy. \tag{2.3}$$

By combining Lemmas 2.2 and 2.3, we obtain the following estimate of $\mathcal{M}_{\mathcal{S}} f$ using the $\mathcal{N}_{k,\ell}$.

Proposition 2.5. *Suppose that* $\operatorname{supp}(\hat{f}) \subseteq \mathcal{A}_j$. *Then there is a pointwise estimate*

$$\mathcal{M}_{\mathcal{S}} f \lesssim \psi_j * |f| + \sum_{k=-j+3}^{2} \sum_{\ell=-j}^{k-1} 2^{-(j+\ell)} \mathcal{N}_{k,\ell} f, \tag{2.4}$$

where the implicit constant is independent of j. (If $j = 0$, we treat the double sum as being empty.)

Proof. Because of the translation-invariance of the operators, we need only establish the estimate (2.4) at the origin.

By Lemma 2.2,

$$\mathcal{M}_{\mathcal{S}} f(0) \lesssim \sup_{\theta \in S^{d-1}} \sup_{0 < L \leq 1} \frac{1}{L} \int_0^L |f(t\theta)|\, dt \lesssim \sup_{\theta \in S^{d-1}} \sup_{0 < L \leq 1} \frac{1}{2L} \int_{-L}^{L} |f(t\theta)|\, dt. \tag{2.5}$$

We next show that the inner supremum in (2.5) can be estimated as follows:

$$\sup_{0 < L \leq 1} \frac{1}{2L} \int_{-L}^{L} |f(t\theta)|\, dt \lesssim \psi_j * |f|(0) + \sum_{k'=-j+1}^{0} \sum_{\ell=-j}^{k'+1} \frac{2^{-(j+\ell)}}{|T_{k',\ell}(\theta)|} \int_{T_{k',\ell}(\theta)} |f|\, dx. \tag{2.6}$$

Here $T_{k',\ell}(\theta)$ denotes the cylindrical tube, centered at the origin and having axis in the direction of θ, with length $2^{k'+2}$ and radius 2^ℓ. To establish the validity of (2.6), we show that the supremum in L over each of the intervals $0 < L \leq 2^{-j}$, $2^{-j} < L \leq 2^{-j+1}$, $\ldots$, $1/2 < L \leq 1$ is controlled by (a portion of) the right-hand side of (2.6).

By considering the function $f \circ R_\theta$, where R_θ is any rotation, i.e. matrix in $SO(d)$, taking $e_1 = (1, 0, \ldots, 0)$ to θ, we may reduce (2.6) to the case $\theta = e_1$. (This also makes use of the invariance of ψ_j under rotations.)

For the first supremum, we apply part (i) of Lemma 2.3:

$$\sup_{0 < L \leq 2^{-j}} \frac{1}{2L} \int_{-L}^{L} |f(te_1)|\, dt \lesssim \sup_{0 < L \leq 2^{-j}} \psi_j * |f|(0) = \psi_j * |f|(0),$$

and this final quantity is precisely the first term on the right-hand side of (2.6). (If $j = 0$, this estimate is all that is needed.)

The other suprema are dealt with analogously, but using instead part (ii) of Lemma 2.3: if $-j+1 \le k' \le 0$, then

$$\sup_{2^{k'-1}<L\le 2^{k'}} \frac{1}{2L}\int_{-L}^{L} |f(te_1)|\,dt \lesssim \psi_j * |f|(0) + \sum_{\ell=-j}^{k'+1} \frac{2^{-(j+\ell)}}{|T_{k',\ell}|}\int_{T_{k',\ell}} |f(x)|\,dx,$$

and the final quantity here corresponds to a portion of the double sum in (2.6). (We have also used the abbreviation $T_{k',\ell} \equiv T_{k',\ell}(e_1)$.)

Having established (2.6), the full estimate (2.4) at the origin follows by taking the supremum in θ of (2.6), moving the supremum past the double sum, and combining with (2.5). (We also relabel k' as $k-2$.) □

We conclude this section by proving Proposition 2.1 using the estimate in Proposition 2.5.

Proof of Proposition 2.1. We address the simpler part (ii) first. Since $\psi_0(x) = (1+|x|)^{-(d+1)}$, we have $\|\psi_0\|_r \lesssim 1$ independent of r. As $\hat{f}$ is supported in $\mathcal{A}_0$, we obtain from Proposition 2.5 and Young's inequality

$$\|\mathcal{M}_{\mathcal{S}} f\|_q \lesssim \big\|\psi_0 * |f|\big\|_q \le \|\psi_0\|_r \|f\|_p \lesssim \|f\|_p,$$

where r is now defined by $1/q + 1 = 1/p + 1/r$.

Turning to part (i), first observe that if r satisfies the equation just noted, then by a rescaling we have $\|\psi_j\|_r \lesssim 2^{jd(1-1/r)} = 2^{jd(1/p-1/q)} < 2^{j\alpha}$.

By another rescaling argument, the $L^p \to L^q$ operator norm of $\mathcal{N}_{k,\ell}$ is equal to $2^{kd(1/q-1/p)}$ times the $L^p \to L^q$ operator norm of $\mathcal{N}_\delta$ with $\delta = 2^{\ell-k}$. So by way of Proposition 2.5, an estimate $\|\mathcal{N}_\delta f\|_q \lesssim_{p,q,\alpha} \delta^{-\alpha}\|f\|_p$ implies, for f having Fourier transform supported in $\mathcal{A}_j$ with $j \ge 1$,

$$\begin{aligned}
\|\mathcal{M}_{\mathcal{S}} f\|_q &\lesssim \|\psi_j * |f|\|_q + \sum_{k=-j+3}^{2}\sum_{\ell=-j}^{k-1} 2^{-(j+\ell)}\|\mathcal{N}_{k,\ell} f\|_q \\
&\lesssim_{p,q,\alpha} \|\psi_j\|_r \|f\|_p + \sum_{k=-j+3}^{2}\sum_{\ell=-j}^{k-1} 2^{-(j+\ell)} \cdot 2^{kd(1/q-1/p)} \cdot 2^{-\alpha(\ell-k)}\|f\|_p \\
&\lesssim_{p,q,\alpha} 2^{j\alpha}\|f\|_p.
\end{aligned}$$

We have used $\alpha > d(1/p-1/q)$ here to estimate $\sum_k 2^{(\alpha+d/q-d/p)k} \lesssim_{p,q,\alpha} 1$. □

3. Negative differentiation results

We divide our work in this section into two pieces. First, we show generally how to obtain differentiation counterexamples for a Besov space given a certain unboundedness property for the local maximal operator associated with a translation-invariant basis. Our argument is essentially an adaptation of Stein's maximal principle, as in [10, Section 10.2], to the context of Besov spaces rather than Lebesgue spaces. Second, we specifically analyze the maximal operator $\mathcal{M}_{\mathcal{R}}$ and establish the necessary unboundedness property for the Besov spaces at play in Theorem 1.3. Throughout, we shall use $\|f\|^*_{L^{1,\infty}}$ to denote the Lorentz space quasinorm $\sup_{\lambda>0} \lambda\big|\{\, x : |f(x)| > \lambda \,\}\big|$.

Proposition 3.1. *Let $\mathcal{B}$ be a translation-invariant differentiation basis in $\mathbb{R}^d$. Fix a positive sequence $\{\varepsilon_\ell\}$ that decreases to 0, and let $\mathcal{B}^\ell$ be the translation-invariant basis defined by $\mathcal{B}_0^\ell = \{ F \in \mathcal{B}_0 : \operatorname{diam}(F) \le \varepsilon_\ell \}$. Let $1 \le p, q \le \infty$, let $\alpha > 0$, and let $\mathcal{X}$ be the cone in $B^p_{\alpha,q}$ consisting of nonnegative functions supported in the unit cube $Q_0 = [-1/2, 1/2]^d$. If each maximal operator $\mathcal{M}_{\mathcal{B}^\ell}$ is unbounded from $\mathcal{X}$ to $L^{1,\infty}$, that is,*

$$\sup_{f \in \mathcal{X} \setminus 0} \|\mathcal{M}_{\mathcal{B}^\ell} f\|^*_{L^{1,\infty}} / \|f\|_{B^p_{\alpha,q}} = \infty,$$

then there is a function $f \in B^p_{\alpha,q}$ such that, for almost every $x \in \mathbb{R}^d$,

$$\limsup_{\substack{\operatorname{diam}(F) \to 0 \\ F \in \mathcal{B}_0}} \frac{1}{|F|} \int_{x+F} f(y)\, dy = \infty.$$

Proof. We may assume without loss of generality that each $\varepsilon_\ell \le 1/2$, so that if f is supported in Q_0, then each $\mathcal{M}_{\mathcal{B}^\ell} f$ is supported in the cube $Q^* = [-1,1]^d$. Therefore, the assumed unboundedness means that for each ℓ we can find a $\lambda_\ell > 0$ and a nonnegative function $f_\ell \in B^p_{\alpha,q}$, supported in Q_0 and having norm 1, such that $\lambda_\ell \left| \{ x \in Q^* : \mathcal{M}_{\mathcal{B}^\ell} f_\ell(x) > \lambda_\ell \} \right| \ge 2^{2\ell}$.

Now set $g_\ell = 2^\ell f_\ell / \lambda_\ell$, so that $\left| \{ x \in Q^* : \mathcal{M}_{\mathcal{B}^\ell} g_\ell(x) > 2^\ell \} \right| \ge 2^\ell \|g_\ell\|_{B^p_{\alpha,q}}$. It follows that we can create a sequence $\{h_k\} = \{g_{\ell_k}\}$ (with ℓ_k nondecreasing), consisting of the entire sequence $\{g_\ell\}$ but with terms repeated as necessary, for which

$$\sum_k \left| \{ x \in Q^* : \mathcal{M}_{\mathcal{B}^{\ell_k}} h_k(x) > 2^{\ell_k} \} \right| = \infty \quad \text{and} \quad \sum_k \|h_k\|_{B^p_{\alpha,q}} < \infty. \tag{3.1}$$

Let E_k denote the set $\{ x \in Q^* : \mathcal{M}_{\mathcal{B}^{\ell_k}} h_k(x) > 2^{\ell_k} \}$. Since $\sum_k |E_k| = \infty$ while each E_k is contained in the compact set Q^*, we can find points $x_k \in \mathbb{R}^d$ so that almost every $x \in Q_0$ is contained in infinitely many of the translated sets $x_k + E_k$ (see [10, p. 442]).

We now consider the function $h = \sum_k h_k(\cdot - x_k)$, which lies in $B^p_{\alpha,q}$, in light of (3.1). If x is a point lying in infinitely many of the $x_k + E_k$, then

$$\limsup_{\substack{\operatorname{diam}(F) \to 0 \\ F \in \mathcal{B}_0}} \frac{1}{|F|} \int_{x+F} h(y)\, dy = \infty, \tag{3.2}$$

so that this condition holds almost everywhere in Q_0. To see this, fix an m and let k be such that (i) $x \in x_k + E_k$ and (ii) $\ell_k \ge m$. Then

$$\sup_{F \in \mathcal{B}_0^m} \frac{1}{|F|} \int_{x+F} h\, dy \ge \sup_{F \in \mathcal{B}_0^{\ell_k}} \frac{1}{|F|} \int_{x+F} h_k(y - x_k)\, dy = \mathcal{M}_{\mathcal{B}^{\ell_k}} h_k(x - x_k) > 2^{\ell_k}.$$

Since we can find k arbitrarily large meeting the requirements (i) and (ii), and since $\ell_k \to \infty$ as $k \to \infty$, the supremum on the left is infinite. This implies (3.2) because m was arbitrary.

So we now have a function h which locally meets the requirements of the proposition. The final step is to let $\{n_j\}_{j=1}^\infty$ be an enumeration of the lattice points $\mathbb{Z}^d$, and to let $f = \sum_j 2^{-j} h(\cdot - n_j)$. This function then fully satisfies the proposition. ☐

In light of Proposition 3.1, to establish Theorem 1.3 it suffices to show the following result for the basis $\mathcal{R}$.

Proposition 3.2. *Let $\mathcal{R}$ be the basis of arbitrarily oriented rectangular parallelepipeds in $\mathbb{R}^d$. Let $\ell \geq 1$, and let $\mathcal{R}^\ell$ be the translation-invariant basis defined by $\mathcal{R}_0^\ell = \{\, R \in \mathcal{R}_0 : \operatorname{diam}(R) \leq 2^{-\ell+1}\sqrt{d}\,\}$. Let $1 \leq p < d < q$, and let $\mathcal{X}$ be the cone in $B^p_{d/p-1,q}$ consisting of nonnegative functions supported in the unit cube $Q_0 = [-1/2, 1/2]^d$. Then the maximal operator $\mathcal{M}_{\mathcal{R}^\ell}$ is unbounded from $\mathcal{X}$ to $L^{1,\infty}$.*

Proof. Let Φ be a fixed nonnegative C^∞ function supported in the unit ball, and for $k \geq 0$ set $\Phi_k = 2^k\Phi(2^k\cdot)$. Let $\varepsilon < 2^{-\ell-1}$ be a small positive parameter, and take $y_1, y_2, \ldots, y_N$ to be a maximal ε-separated subset of the cube $[-1/4, 1/4]^d$, so that $N \approx \varepsilon^{-d}$. To demonstrate the unboundedness of $\mathcal{M}_{\mathcal{R}^\ell}$, we shall use the function(s)

$$f_\varepsilon = \sum_{m=1}^{N} \Phi_{\ell+m}(\cdot - y_m).$$

We first establish an upper bound for the $B^p_{d/p-1,q}$ norm of f_ε. To this end, we modify slightly the notation for Littlewood-Paley projections introduced in Section 2. We now use Δ_j for the multiplier operator associated with $\eta(2^{-j+1}|\xi|) - \eta(2^{-j}|\xi|)$ for *all* j (not just $j \geq 1$), and we use $\widetilde{\Delta}_0$ in place of Δ_0 for the multiplier operator associated with $1 - \eta(|\xi|)$. Since Φ is a Schwartz function, it follows fairly straightforwardly that $\|\Delta_j\Phi\|_p \lesssim \min(1, 2^{-jd})$.

Now $\|\Phi_k\|_p = 2^{-k(d/p-1)}\|\Phi\|_p$, so $\|\widetilde{\Delta}_0 f_\varepsilon\|_p \lesssim \|f_\varepsilon\|_p \lesssim_p 2^{-\ell(d/p-1)}$; and $\Delta_j\Phi_k = 2^k(\Delta_{j-k}\Phi)(2^k\cdot)$, so $\|\Delta_j\Phi_k\|_p \lesssim 2^{-k(d/p-1)}\min(1, 2^{-(j-k)d})$. By inserting these estimates at the end of

$$\begin{aligned}
\|f_\varepsilon\|_{B^p_{d/p-1,q}} &= \left(\|\widetilde{\Delta}_0 f_\varepsilon\|_p^q + \sum_{j\geq 0}\big(2^{j(d/p-1)}\|\Delta_j f_\varepsilon\|_p\big)^q\right)^{1/q} \\
&\leq \left(\|\widetilde{\Delta}_0 f_\varepsilon\|_p^q + \sum_{j\geq 0}\left(2^{j(d/p-1)}\sum_{m=1}^{N}\|\Delta_j\Phi_{\ell+m}\|_p\right)^q\right)^{1/q},
\end{aligned}$$

we obtain the estimate $\|f_\varepsilon\|_{B^p_{d/p-1,q}} \lesssim_{p,q} N^{1/q} \approx \varepsilon^{-d/q}$, with the primary contribution coming from the terms with $\ell + 1 \leq j \leq \ell + N$.

Now we proceed to estimating the $L^{1,\infty}$ quasinorm of $\mathcal{M}_{\mathcal{R}^\ell} f_\varepsilon$ from below. We begin here by noting that there is a constant C_0 such that for $m \geq 1$,

$$\mathcal{M}_{\mathcal{R}^\ell}\Phi_{\ell+m}(x) \geq \begin{cases} 2^{\ell+m}C_0, & \text{if } |x| \leq 2^{-(\ell+m)}; \\ C_0/|x|, & \text{if } 2^{-(\ell+m)} \leq |x| \leq 2^{-\ell}. \end{cases} \tag{3.3}$$

This can be seen by considering the average of $\Phi_{\ell+m}$ over $x + T_x$, where T_x is a minimal rectangular parallelepiped centered at the origin satisfying $x + T_x \supseteq B(0, 2^{-\ell-m}) \supseteq \operatorname{supp}\Phi_{\ell+m}$.

From (3.3) it follows that $\{\, x : \mathcal{M}_{\mathcal{R}^\ell}\Phi_{\ell+m}(x) > \lambda \,\} \supseteq B(0, \min(C_0/\lambda, 2^{-\ell}))$ if $\lambda < 2^{\ell+m}C_0$, and so

$$\{\, x : \mathcal{M}_{\mathcal{R}^\ell} f_\varepsilon(x) > \lambda \,\} \supseteq \bigcup_{\substack{1\leq m\leq N \\ 2^{\ell+m}C_0 > \lambda}} B(y_m, \min(C_0/\lambda, 2^{-\ell})).$$

Taking $\lambda = 2C_0/\varepsilon$, we have $\min(C_0/\lambda, 2^{-\ell}) = C_0/\lambda = \varepsilon/2$ in light of the condition $\varepsilon < 2^{-\ell-1}$, and this ensures that the balls in the above inclusion are disjoint. Thus

$$\|\mathcal{M}_{\mathcal{R}^\ell} f_\varepsilon\|^*_{L^{1,\infty}} \gtrsim \varepsilon^{-1} \cdot \#\big(\{\, m : 1 \le m \le N,\, 2^{\ell+m} > 2/\varepsilon \,\}\big) \cdot \varepsilon^d \approx \varepsilon^{-1},$$

as $N \approx \varepsilon^{-d}$ and the condition $2^{\ell+m} > 2/\varepsilon$ excludes a bounded fraction of the balls.

We have now shown that $\|\mathcal{M}_{\mathcal{R}^\ell} f_\varepsilon\|^*_{L^{1,\infty}} \gtrsim \varepsilon^{-1}$ and $\|f_\varepsilon\|_{B^p_{d/p-1,q}} \lesssim_{p,q} \varepsilon^{-d/q}$. It follows that $\mathcal{M}_{\mathcal{R}^\ell}$ is unbounded from $\mathcal{X} \subseteq B^p_{d/p-1,q}$ to $L^{1,\infty}$ when $q > d$. $\square$

We end this section with an observation regarding the possible optimality of Theorem 1.3 with respect to the Besov space exponent q. For a given p in the range $1 \le p < d$, let $q_{\mathcal{R}}(p)$ be the infimum of those q for which a differentiation counterexample of the type indicated in Theorem 1.3 exists in $B^p_{d/p-1,q}$. (So we have shown that $q_{\mathcal{R}}(p) \le d$.) Then $q_{\mathcal{R}}(p)$ must be a nonincreasing function of p, owing to the continuous Sobolev-type embeddings $B^{p_1}_{d/p_1-1,q} \subseteq B^{p_2}_{d/p_2-1,q}$ for $p_1 < p_2$. (We can make a similar argument involving Proposition 3.2, and indeed about positive differentiation results as well.) This property of $q_{\mathcal{R}}(p)$ is at least consistent with Theorem 1.3 being sharp, possibly modulo the endpoint $q = d$.

4. Capacity and dimension of exceptional sets

We start our discussion in this section by recalling the necessary background involving Bessel potentials and capacities. For $t \in \mathbb{R}$, the Bessel potential $\mathcal{G}_t$ is the operator defined (say on tempered distributions) by $\widehat{\mathcal{G}_t f} = (1+|\xi|^2)^{-t/2}\hat{f}$. If $t > 0$, then $\mathcal{G}_t$ corresponds to a positive, integrable convolution kernel G_t ([1, p. 10]). The Bessel potentials also evidently satisfy the group property $\mathcal{G}_{t_1}\mathcal{G}_{t_2} = \mathcal{G}_{t_1+t_2}$; in particular, for $t_1, t_2 > 0$ we have $G_{t_1} * G_{t_2} = G_{t_1+t_2}$. The Bessel potentials interact nicely with Besov spaces as well: $\mathcal{G}_t$ takes $B^p_{s,q}$ isomorphically onto $B^p_{s+t,q}$ for any $0 < p, q \le \infty$ and $s, t \in \mathbb{R}$ ([14, pp. 58–9]).

We now come to Bessel capacities, a two-parameter family of outer measures that are more refined than Lebesgue measure, in the sense that there are sets of Lebesgue measure 0 which have positive Bessel capacities. For $\beta > 0$ and $1 \le q < \infty$, define the (β, q) Bessel capacity of any set $E \subseteq \mathbb{R}^d$ by

$$C_{\beta,q}(E) = \inf\{\|f\|_q^q : f \in L^q,\ f \ge 0, \text{ and } G_\beta * f \ge \chi_E\}. \tag{4.1}$$

Or, if there are no such functions f, we set $C_{\beta,q}(E) = \infty$.

We now record three properties of Bessel capacities which we shall need later on.

(1) If h is a nonnegative function and $\lambda > 0$, then

$$C_{\beta,q}\big(\{x : G_\beta * h(x) \ge \lambda\}\big) \le \lambda^{-q}\|h\|_q^q. \tag{4.2}$$

This arises from the observation that $G_\beta * (h/\lambda) \ge \chi_{\{x : G_\beta * h(x) \ge \lambda\}}$.

(2) If we are given $\beta_1, \beta_2 > 0$ and $1 < q_1, q_2 < \infty$ satisfying either (i) $\beta_1 q_1 < \beta_2 q_2 \le d$ or (ii) $\beta_1 q_1 = \beta_2 q_2 \le d$ and $q_2 < q_1$, then ([1, p. 148])

$$C_{\beta_2,q_2}(E) = 0 \quad \Rightarrow \quad C_{\beta_1,q_1}(E) = 0. \tag{4.3}$$

(3) If $C_{\beta,q}(E) = 0$ with $q > 1$ and $\beta q \le d$, then the Hausdorff dimension of E is at most $d - \beta q$ ([1, pp. 137–8]).

With these preliminaries out of the way, we state this section's central result.

Proposition 4.1. *Let $\mathcal{B}$ be a translation-invariant basis in $\mathbb{R}^d$. Assume the local maximal operator $\mathcal{M}_{\mathcal{B}}$ is bounded from $B^p_{\alpha_0,1}$ to L^q, where $1 < p < \infty$, $0 \le \alpha_0 < d/p$, and $p \le q \le dp/(d-\alpha_0 p)$. If $f \in B^p_{\alpha,1}$ with $\alpha_0 < \alpha \le d/p$ and $(q-p)\alpha \le q\alpha_0$, then the $(\alpha-\alpha_0, q)$ Bessel capacity of the $\mathcal{B}$–exceptional set $E_{\mathcal{B}}(f)$ is 0.*

Before delving into the proof, let us see how Proposition 4.1 implies Theorem 1.4.

Proof that Proposition 4.1 $\Rightarrow$ *Theorem* 1.4. Let p, q, and α be as in Theorem 1.4, i.e. $1 < p \le d$, $1 \le q \le \infty$, and $d/p - 1 < \alpha \le d/p$. Using the hypothesized Nikodym estimate $\|\mathcal{N}_\delta f\|_d \lesssim_\varepsilon \delta^{-(d/p-1+\varepsilon)}\|f\|_p$ in Theorem 1.1, we see that $\mathcal{M}_{\mathcal{R}}$ is bounded from $B^p_{d/p-1+\varepsilon,1}$ to L^d for any $\varepsilon > 0$.

Now take $g \in B^p_{\alpha,q}$. Note that for all $\beta < \alpha$, we have $g \in B^p_{\beta,1}$. Apply Proposition 4.1, with its p having the same meaning as the present p, its q being d, its α_0 being $d/p - 1 + \varepsilon$, and its α being β. (For ε small and β close to α, the various inequalities included in the hypotheses of Proposition 4.1 are straightforward to verify.) We thereby conclude that $E_{\mathcal{R}}(g)$ has $(\beta - d/p + 1 - \varepsilon, d)$ capacity 0.

From the relationship between Bessel capacities and Hausdorff dimension indicated in property (3) above, it now follows that $E_{\mathcal{R}}(g)$ has Hausdorff dimension at most $d - (\beta - d/p + 1 - \varepsilon)d = d(d/p - \beta + \varepsilon)$. By letting $\varepsilon \to 0$ and $\beta \to \alpha$, we see that in fact the Hausdorff dimension of $E_{\mathcal{R}}(g)$ cannot exceed $d(d/p - \alpha)$. ☐

We now pause to remark on the definition of the exceptional set. The definition (1.1) given in Section 1 is a bit unwieldy owing to the intersection, but makes clear that $E_{\mathcal{B}}(f_1) = E_{\mathcal{B}}(f_2)$ if $f_1 = f_2$ a.e. To work with these sets, it will be convenient to define a related entity: the *naive* $\mathcal{B}$–exceptional set of a function f is

$$E^*_{\mathcal{B}}(f) = \{\, x \in \mathbb{R}^d : \Gamma_{\mathcal{B}} f(x) > 0 \,\}. \tag{4.4}$$

For this formulation, we need not have $E^*_{\mathcal{B}}(f_1) = E^*_{\mathcal{B}}(f_2)$ if $f_1 = f_2$ a.e. We also want to make an allowance in this definition for a.e.-defined functions: if f is such a function, then $E^*_{\mathcal{B}}(f)$ consists of the points in (4.4) together with those points where f is undefined. The utility of the naive exceptional set comes from its slightly simpler form relative to the exceptional set and the fact that if $f_1 = f_2$ a.e., then

$$E_{\mathcal{B}}(f_1) \subseteq E^*_{\mathcal{B}}(f_2). \tag{4.5}$$

This inclusion provides a bridge for establishing results about $E_{\mathcal{B}}(f_1)$.

In the spirit of this discussion, our next step is to prove an auxiliary lemma that gives us a way of picking a "good" f_2 relative to a given f_1.

Lemma 4.2. *Let $f \in B^p_{\alpha,1}$ with $1 \le p < \infty$ and $\alpha > 0$. Then for $\gamma > 0$, the integral $\int G_\gamma(x-y)f(y)\,dy$ is convergent for all x except those in a set having $(\alpha+\gamma, p)$ capacity 0.*

Proof. Since $\mathcal{G}_\alpha$ maps $B^p_{0,1}$ onto $B^p_{\alpha,1}$, we can find $g \in B^p_{0,1}$ satisfying $G_\alpha * g = f$. Note that g is also an L^p function. Define E to be the set of all x where

$$\iint G_\gamma(x-y)G_\alpha(y-z)|g(z)|\,dz\,dy = \infty. \tag{4.6}$$

If $x \notin E$, then $\int G_\gamma(x-y)f(y)\,dy$ is clearly finite. Thus it suffices to show that E has $(\alpha+\gamma, p)$ capacity 0. Now by reversing the order of integration in (4.6) and appealing to the identity $G_\alpha * G_\gamma = G_{\alpha+\gamma}$, we may write $E = \{x : G_{\alpha+\gamma} * |g|(x) = \infty\}$. Then evidently $E \subseteq \{x : G_{\alpha+\gamma} * |g|(x) \ge \lambda\}$ for any $\lambda > 0$, and so (4.2) implies $C_{\alpha+\gamma,p}(E) \le \lambda^{-p}\|g\|_p^p$. Letting $\lambda \to \infty$, we see that in fact $C_{\alpha+\gamma,p}(E) = 0$. ☐

We now conclude by establishing Proposition 4.1.

Proof of Proposition 4.1. Let $f \in B^p_{\alpha,1}$, and let $g \in B^p_{\alpha_0,1}$ satisfy $f = G_{\alpha'} * g$, where α' denotes $\alpha - \alpha_0$. More precisely, the convolution $G_{\alpha'} * g$ is an a.e.-defined function which is equal to f a.e. By (4.5), it thus suffices to show that the naive exceptional set $E^*_{\mathcal{B}}(G_{\alpha'} * g)$ has (α', q) capacity equal to 0.

Let us also subdivide $E^*_{\mathcal{B}}(G_{\alpha'} * g)$ into two sets: S_1, consisting of those points where $G_{\alpha'} * g$ is undefined, i.e. where the convolution integral is divergent, and $S_2 = \{x \notin S_1 : \Gamma_{\mathcal{B}}(G_{\alpha'} * g)(x) > 0\}$. We now show that $C_{\alpha',q}(S_1) = C_{\alpha',q}(S_2) = 0$.

From Lemma 4.2 we know that S_1 has (α, p) capacity 0. We now use the relationships between different capacities given in (4.3). It is an immediate consequence of the hypotheses of Proposition 4.1 that $(\alpha - \alpha_0)q \leq \alpha p \leq d$, and this is precisely as needed to deduce that $C_{\alpha',q}(S_1) = 0$.

Turning to S_2, for $t > 0$ define $S_{2,t} = \{x \in S_2 : \Gamma_{\mathcal{B}}(G_{\alpha'} * g)(x) > t\}$. By a standard subadditivity argument, we need only show that $C_{\alpha',q}(S_{2,t}) = 0$ for any $t > 0$; we henceforth think of t as being fixed.

Let ψ be a Schwartz function on $\mathbb{R}^d$. Using the triangulation

$$\begin{aligned}|G_{\alpha'} * g(y) - G_{\alpha'} * g(x)| \leq |G_{\alpha'} * (g - \psi)(y)| &+ |G_{\alpha'} * \psi(y) - G_{\alpha'} * \psi(x)| \\ &+ |G_{\alpha'} * (g - \psi)(x)|\end{aligned}$$

as well as $\Gamma_{\mathcal{B}}\big(G_{\alpha'} * \psi\big) \equiv 0$ owing to $G_{\alpha'} * \psi$ being continuous, we obtain

$$\begin{aligned}\Gamma_{\mathcal{B}}(G_{\alpha'} * g)(x) &\leq \mathcal{M}_{\mathcal{B}}\big(G_{\alpha'} * (g - \psi)\big)(x) + |G_{\alpha'} * (g - \psi)(x)| \\ &\leq G_{\alpha'} * \mathcal{M}_{\mathcal{B}}(g - \psi)(x) + G_{\alpha'} * |g - \psi|(x)\end{aligned}$$

for $x \notin S_1$. That $\mathcal{M}_{\mathcal{B}}\big(G_{\alpha'} * (g - \psi)\big) \leq G_{\alpha'} * \mathcal{M}_{\mathcal{B}}(g - \psi)$ can be seen by using Fubini's theorem and the nonnegativity of $G_{\alpha'}$.

In turn, $S_{2,t} \subseteq \{x : G_{\alpha'} * \mathcal{M}_{\mathcal{B}}(g - \psi)(x) > t/2\} \cup \{x : G_{\alpha'} * |g - \psi|(x) > t/2\}$, and therefore (4.2) yields

$$C_{\alpha',q}(S_{2,t}) \leq (t/2)^{-q}\big(\|\mathcal{M}_{\mathcal{B}}(g - \psi)\|_q^q + \|g - \psi\|_q^q\big).$$

We now invoke the assumed boundedness of $\mathcal{M}_{\mathcal{B}}$ and also that $B^p_{\alpha_0,1}$ embeds continuously into L^q. This latter fact can be deduced from the inequality $p \leq q \leq dp/(d - \alpha_0 p)$, for by the embedding properties for Besov spaces described in Section 2 we have that $B^p_{\alpha_0,1}$ embeds into both L^p and $L^{dp/(d-\alpha_0 p)}$. It follows that

$$C_{\alpha',q}(S_{2,t}) \lesssim_{p,q,\alpha_0} t^{-q}\|g - \psi\|^q_{B^p_{\alpha_0,1}}.$$

By picking ψ arbitrarily close to g in $B^p_{\alpha_0,1}$, we see that in fact $C_{\alpha',q}(S_{2,t}) = 0$. □

References

[1] David R. Adams and Lars Inge Hedberg, *Function spaces and potential theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 314, Springer-Verlag, Berlin, 1996. MR1411441

[2] Hugo Aimar, Liliana Forzani, and Virginia Naibo, *Rectangular differentiation of integrals of Besov functions*, Math. Res. Lett. **9** (2002), no. 2-3, 173–189, DOI 10.4310/MRL.2002.v9.n2.a4. MR1909636

[3] Thomas Bagby and William P. Ziemer, *Pointwise differentiability and absolute continuity*, Trans. Amer. Math. Soc. **191** (1974), 129–148. MR0344390

[4] J. Bourgain, *Besicovitch type maximal operators and applications to Fourier analysis*, Geom. Funct. Anal. **1** (1991), no. 2, 147–187, DOI 10.1007/BF01896376. MR1097257

[5] Michael Christ, *Estimates for the k-plane transform*, Indiana Univ. Math. J. **33** (1984), no. 6, 891–910, DOI 10.1512/iumj.1984.33.33048. MR763948

[6] Antonio Cordoba, *The Kakeya maximal function and the spherical summation multipliers*, Amer. J. Math. **99** (1977), no. 1, 1–22. MR0447949

[7] A. Nagel, E. M. Stein, and S. Wainger, *Differentiation in lacunary directions*, Proc. Nat. Acad. Sci. U.S.A. **75** (1978), no. 3, 1060–1062. MR0466470

[8] Virginia Naibo, *Bessel capacities and rectangular differentiation in Besov spaces*, J. Math. Anal. Appl. **324** (2006), no. 2, 834–840, DOI 10.1016/j.jmaa.2005.12.071. MR2265084

[9] Otton Nikodým, *Sur la mesure des ensembles plans dont tous les points sont rectilinéairement accessibles*, Fund. Math. **10** (1927), no. 1, 116–168.

[10] Elias M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals,* with the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993. MR1232192

[11] A. M. Stokolos, *On the differentiation of integrals by bases that do not possess the density property* (Russian, with Russian summary), Mat. Sb. **187** (1996), no. 7, 113–138, DOI 10.1070/SM1996v187n07ABEH000148; English transl., Sb. Math. **187** (1996), no. 7, 1061–1085. MR1404191

[12] A. M. Stokolos, *On differentiation of integrals with respect to bases of convex sets*, Studia Math. **119** (1996), no. 2, 99–108. MR1391470

[13] Terence Tao, *The Bochner-Riesz conjecture implies the restriction conjecture*, Duke Math. J. **96** (1999), no. 2, 363–375, DOI 10.1215/S0012-7094-99-09610-2. MR1666558

[14] Hans Triebel, *Theory of function spaces*, Monographs in Mathematics, vol. 78, Birkhäuser Verlag, Basel, 1983. MR781540

[15] Thomas Wolff, *An improved bound for Kakeya type maximal functions*, Rev. Mat. Iberoamericana **11** (1995), no. 3, 651–674, DOI 10.4171/RMI/188. MR1363209

Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, Wisconsin 53706

E-mail address: jmurcko@gmail.com