CARLESON CONDITIONS FOR ASYMPTOTIC WEIGHTS

MICHAEL BRIAN KOREY

ABSTRACT. The doubling and A_∞ conditions are characterized in terms of convolution with rapidly decreasing kernels. The Carleson-measure criterion for A_∞ of Fefferman, Kenig, and Pipher is extended to the case when all bounds become optimally small in the asymptotic limit.

1. INTRODUCTION

The doubling and A_∞ conditions that arise in many areas of mathematical analysis are usually defined in terms of mean values over simple compact sets, such as balls or cubes. The present paper formulates these conditions rather in terms of weighted averages over all of Euclidean space.

One reason to seek such formulations is briefly as follows. In their famous paper on Hardy spaces in $\mathbf{R}^n$, Fefferman and Stein [5] gave a square-function characterization of the space of functions of bounded mean oscillation: A function $f = \log w$ is in BMO($\mathbf{R}^n$) if and only if

$$(1) \qquad d\tilde{\mu}(x,t) = |\nabla_x(\log w * \varphi_t)(x)|^2 t\, dxdt$$

is a Carleson measure on $\mathbf{R}_+^{n+1}$. (See §2 below for the necessary definitions; here φ is a Schwartz function, such as the Gaussian, and $\varphi_t(x) = t^{-n}\varphi(t^{-1}x)$ gives its mass-preserving dilation.) In studying perturbation theory for divergence-form elliptic partial differential equations and, in particular, in seeking perturbation conditions that preserve the solvability of the Dirichlet problem with L^p boundary data, Fefferman, Kenig, and Pipher [6] obtained a related square-function criterion: A doubling weight w is in A_∞ on $\mathbf{R}^n$ if and only if

$$(2) \qquad d\mu(x,t) = |\nabla_x \log(w * \varphi_t)(x)|^2 t\, dxdt$$

is a Carleson measure. Note that the quantities (1) and (2) differ only in the order in which the logarithm is taken and the convolution formed. While commuting these two processes is generally not permissible, the result of [6] shows that for A_∞ weights it is. This is in accord with one of the standard formulations of A_∞: A weight is in this class exactly when the difference between the logarithm of its mean value and the mean value of its logarithm is uniformly bounded over all balls.

Unlike the earlier result in [5], the proof of the latter theorem in [6] does not immediately yield an estimate for the Carleson norm of $d\mu$ in terms of the BMO norm of $\log w$ (or the A_∞ bound of w).[1] That the characterization of A_∞ in [6] nevertheless carries over to the so-called "asymptotic case," in which all bounds approach their optimal values over small scales, is the main result here. While the statement of this result is but a modest refinement of the theorem in [6], its proof turns out to be surprisingly subtle; the challenge we repeatedly encounter is keeping the various constants of inequality that arise close to their smallest possible values. As a by-product of these efforts to obtain a quantitative Carleson-measure characterization of A_∞, we shall also make several new observations about the doubling condition.

For an excellent account of the links between such square functions as appear in (2), regularity theory for elliptic PDEs, quasiconformal mappings, and the classical theory of a.e. differentiability of real functions in several variables, see the introduction to [6]. Without further ado, we now present the definitions that will allow us to make precise the results sketched above.

2. Preliminaries

For reference, we record here the notation that is used throughout this work. The symbol $|E|$ denotes the Lebesgue measure of the set E in $\mathbf{R}^n$. The Lebesgue integral of the function f over E is written $\int_E f$ or $\int_E f\,dx$; if the region of integration is not shown, it is understood to be all of $\mathbf{R}^n$. When $0 < |E| < \infty$, the symbol f_E represents the mean value of f over E; that is, $f_E = (\int_E f)/|E|$. The Schwartz space of rapidly decreasing functions is denoted $\mathcal{S}$.

A *weight* is any non-negative, locally integrable function on $\mathbf{R}^n$. The symbol $w(E)$ denotes the mass of E under the weight w, i.e., $w(E) = \int_E w\,dx$. The r-fold concentric dilate of a ball B is written rB; the *double* of B is $2B$. The measure of the unit ball $B_1(0)$ in $\mathbf{R}^n$ is ω_n. The letter χ is reserved for the normalized characteristic function of the unit ball, i.e., $\chi(x)$ is $1/\omega_n$ when $x \in B_1(0)$ and is 0 otherwise. The (mass-preserving) dilation of an integrable function φ to scale t is defined by the rule $\varphi_t(x) = t^{-n}\varphi(t^{-1}x)$. To simplify notation, we shall consistently suppress the spatial variable x in estimates that hold uniformly over all of $\mathbf{R}^n$.

Let us now recall the definitions of four key concepts: doubling, the A_∞ condition, the space BMO, and Carleson measures. We shall first consider the basic form of each of these and then the respective asymptotically optimal variant.

Doubling. A weight w is doubling if its average over each ball is (uniformly) comparable to its average over the concentric double of that ball; that is, if there is some constant C such that

$$(3) \qquad\qquad C^{-1}w_B \leq w_{2B} \leq Cw_B$$

uniformly over all balls B. The smallest C for which (3) is valid is termed the *doubling constant* of w and is denoted $Db(w)$. Note that if $Db(w) = 1$, then w is a.e. constant. The collection of all doubling weights is written Db.

---

[1]The argument in [5] shows that the Carleson norm of $d\tilde\mu$ in (1) is equivalent to $\|\log w\|_*^2$, the square of the BMO norm of $\log w$. The local form of this says that $\log w \in$ VMO if and only if $d\tilde\mu$ is a Carleson measure, with vanishing trace. Once again, see the next section for the relevant definitions.

We call a weight *asymptotically doubling* (see [11] and the references cited there) if its doubling behavior becomes optimal over small scales. More precisely, we write $w \in Db_{as}$ if both $w \in Db$ and the ratio of its averages of over every pair of sufficiently small balls, each of which is contained in the double of the other, is arbitrarily close to 1.

There is a simple continuity criterion for doubling which will prove very useful in the sequel. While $w(rB) \to w(B)$ as $r \to 1$ for every weight w, this convergence of differences can be converted to a convergence of ratios precisely when $w \in Db$.

Lemma 1. *A weight w is doubling if and only if $w(rB)/w(B) \to 1$, as $r \to 1$, uniformly over all balls B. When this holds, the convergence occurs at a rate that depends only on the doubling constant $Db(w)$.*

See [2] or [11] for a proof; a related result in [3] applies in the general setting of doubling measures on homogeneous spaces.

The A_∞ condition. Next, a weight satisfies the A_∞ condition (see [4] and [12]) if its mass is "fairly" distributed not just over a ball and its central half (as in the case of doubling), but over all subsets of the ball of half its measure; more precisely, we write $w \in A_\infty$ if there are constants α and β, both less than 1, such that

$$w(E)/w(B) < \beta \quad \text{whenever} \quad |E|/|B| < \alpha.$$

(Here E is an arbitrary Borel subset of the ball B.) This scale-invariant form of absolute continuity has many equivalent formulations; the one of chief interest to us is the following: A weight is in A_∞ just in case its arithmetic and geometric means are uniformly comparable over all balls.[2] In other words, there is a constant C for which

$$
\text{(4)} \qquad \frac{1}{|B|} \int_B w \le C \exp\left(\frac{1}{|B|} \int_B \log w \right)
$$

over all balls B. If, in addition, the constant C can be taken arbitrarily close to 1 for all sufficiently small balls, then we say that w is *asymptotically absolutely continuous* and we write $w \in A_{\infty,as}$.[3]

The space BMO. Third, a real-valued, locally integrable function f has bounded mean oscillation if the quantity

$$
\text{(5)} \qquad \|f\|_* = \sup_B \frac{1}{|B|} \int_B |f - f_B|
$$

is finite, and the set of all such functions is written BMO. The corresponding asymptotic variant of BMO is termed VMO, for the space of functions of *vanishing mean oscillation*. This requires not only that the quantity $\|f\|_*$ be bounded, but also that the supremum in (5) be arbitrarily close to 0 when measured over all sufficiently small balls. BMO serves as a replacement for L^∞ in many contexts of harmonic analysis, and VMO plays within BMO a role analogous to that of the bounded, uniformly continuous functions in L^∞ (see [15]).

What relations exist among these conditions? It is well-known, for example, that every A_∞ weight is doubling and that its logarithm is in BMO (see [7]). Sharp

---

[2]The latter formulation is due to [7], [8], and [14]; see [4] or [7] for still other definitions of A_∞.
[3]As is shown in [11], all of the classical formulations of A_∞ in [4] remain equivalent when optimal bounds are required in the asymptotic limit.

asymptotic versions of these statements are derived in [11] and can be summarized qualitatively in the following form:

Lemma 2. *Every weight in $A_{\infty,as}$ is also in Db_{as} and its logarithm is in* VMO.

Combined with the fundamental inequality of John and Nirenberg [10], this result shows that *the logarithm of a weight is in* VMO *if and only if some positive power of the weight is in $A_{\infty,as}$.* Our aim here is to give another characterization of $A_{\infty,as}$. Before doing so, however, we must give one final definition.

Carleson measures. A non-negative Borel measure μ on $\mathbf{R}_+^{n+1} = \mathbf{R}^n \times (0,\infty)$ is a Carleson measure if the ratio

$$(6) \qquad \frac{1}{|B|} \int_{T(B)} d\mu$$

is bounded over all balls B, where $T(B)$ denotes the cylinder $B \times \big(0, \mathrm{radius}(B)\big)$ in $\mathbf{R}_+^{n+1}$; the smallest bound is likewise termed the *Carleson norm* of μ. If this bound can also be taken arbitrarily small for all small balls B, then we say that μ has *vanishing trace* (see [13]).

3. The main result

Our point of departure is a theorem of Fefferman, Kenig, and Pipher.

Basic Theorem ([6]). *Suppose that $\varphi \in \mathcal{S}$, $\varphi \geq 0$, and $\int \varphi > 0$. Let $\vec{\psi} = \nabla\varphi$. A weight w is in A_∞ if and only if it is doubling and*

$$(7) \qquad d\mu(x,t) = \frac{|w * \vec{\psi}_t(x)|^2}{|w * \varphi_t(x)|^2} \, dx \frac{dt}{t}$$

is a Carleson measure on $\mathbf{R}_+^{n+1}$.

That this Carleson-measure characterization of A_∞ in the presence of doubling carries over to the asymptotic setting is our main result.

Theorem 1. *Under the same assumptions as in the prior theorem, a weight w is in $A_{\infty,as}$ if and only if it is doubling and the measure $d\mu$ of (7) is a Carleson measure on $\mathbf{R}_+^{n+1}$, with vanishing trace.*

4. Outline of the proof

Following the argument in [6], we shall focus on the special case when φ is the Gaussian, i.e., $\varphi(x) = \pi^{-n/2} e^{-|x|^2}$, normalized so that $\int \varphi = 1$. In this setting, the *heat extension u* of a function w on $\mathbf{R}^n$ is the latter's convolution with this Gaussian φ, namely

$$u(x,t) = w * \varphi_{\sqrt{t}}(x), \quad \text{for } (x,t) \text{ in } \mathbf{R}_+^{n+1}.$$

The theorem then takes the following special form:

Theorem 2 (Gaussian special case). *Let u be the heat extension of a doubling weight w, as above. Set*

$$(8) \qquad h(s) = \sup_{y \in \mathbf{R}^n} \frac{1}{\omega_n s^n} \int_{t=0}^{s^2} \int_{|x-y|<s} \frac{|\nabla u(x,t)|^2}{u(x,t)^2} \, dx dt.$$

Then $w \in A_{\infty,as}$ if and only if

(9) $\qquad\qquad h \ \text{is bounded} \quad \text{and} \quad \lim_{s \to 0} h(s) = 0.$

By the translation invariance of the definitions of $A_{\infty,as}$, Db, and h, it suffices to consider only balls centered at the origin. Let $h(s)$ henceforth be the integral in (8) evaluated at $y = 0$, i.e.,

$$h(s) = \frac{1}{\omega_n s^n} \int_{t=0}^{s^2} \int_{|x|<s} \frac{|\nabla u(x,t)|^2}{u(x,t)^2} \, dxdt.$$

As in [6], we can use the heat operator $H = D_t - \triangle$ and the divergence theorem to split the integral defining h into two parts. Since $u(x,t) = w * \varphi_{\sqrt{t}}(x)$, then $Hu = 0$, so that $H(\log u) = u^{-1}D_t u - \sum_{i=1}^{n} D_i(u^{-1}D_i u) = |\nabla u|^2/u^2$. Hence

$$
\begin{aligned}
h(s) &= \frac{1}{\omega_n s^n} \int_{t=0}^{s^2} \int_{|x|<s} H(\log u)(x,t) \, dxdt \\
&= \frac{1}{\omega_n s^n} \int_{|x|<s} \int_{t=0}^{s^2} D_t(\log u)(x,t) \, dtdx - \frac{1}{\omega_n s^n} \int_{t=0}^{s^2} \int_{|x|<s} \operatorname{div}(\nabla \log u)(x,t) \, dxdt \\
&= \frac{1}{\omega_n s^n} \int_{|x|<s} \log \frac{u(x,s^2)}{w(x)} \, dx - \frac{1}{\omega_n s^n} \int_{t=0}^{s^2} \int_{|x|=s} \frac{\nabla u(x,t) \cdot \vec{n}_x}{u(x,t)} \, dxdt \\
&= h_1(s) - h_2(s).
\end{aligned}
$$

Here $\vec{n}_x = x/|x|$, the outer unit normal to the boundary at the point x.

Let us pause for a moment to compare the two terms, h_1 and h_2. The integrand of the first of these features the ratio of w to its Gaussian average at scale s; the integral $h_1(s)$ is then the mean value of the logarithm of this ratio over a ball of radius s. Modulo the reversal of averaging and taking the logarithm, the integral $h_1(s)$ is thus a comparison of the averages of w over two different positive kernels, the Gaussian and the characteristic function of the unit ball. By contrast, the integrand of $h_2(s)$ is the ratio of the average of w with respect to a kernel of integral zero (the normal derivative of the Gaussian) to its average with respect to a positive kernel (the Gaussian itself).

This difference will be crucial in the proof. Indeed, in the following pages, we shall show that the boundary integral $h_2(s)$ vanishes with s when w is merely asymptotically doubling; on the other hand, for the dominant term $h_1(s)$ to vanish, we shall need the much stronger condition that w is in the class $A_{\infty,as}$.

The argument breaks down into three steps.

- STEP I: $w \in A_{\infty,as} \implies \lim_{s \to 0} h_1(s) = 0$.
- STEP II: $w \in Db_{as} \implies \lim_{s \to 0} h_2(s) = 0$.
- STEP III: $[w \in Db, \lim_{s \to 0} h(s) = 0] \implies w \in Db_{as}$.

The first two of these combine to give the necessity of the condition (9) in the theorem, while the last will be used to show its sufficiency. We now address each of these claims in turn.

5. Necessity

As noted above, the integrand of the leading term h_1 of the Carleson norm is, in principle, the logarithm of the ratio of the Gaussian and standard averages of w. To show that the integral h_1 is small, we aim to show that this ratio is near 1.

Colloquially, we must first "cut off the tail" of the Gaussian and then "flatten out the bumps." That is, we must first pass from the full heat kernel to a truncated version and then, in turn, to the characteristic function of a ball. The challenge is to accomplish both tasks while insuring that the ratios of the interchanged quantities remain close to 1. Each task will be addressed in a separate lemma.

For later purposes (STEPS II and III), we choose to state the results for kernels somewhat more general than the Gaussian. The convolution kernels φ we shall consider are non-negative, non-trivial (i.e., $\int \varphi > 0$), continuous, and rapidly decreasing functions on $\mathbf{R}^n$; the last condition means that for each non-negative integer N there is a constant A_N such that

$$(10) \qquad |\varphi(x)| \le \frac{A_N}{(1 + |x|)^N}, \quad \text{for all } x \text{ in } \mathbf{R}^n.$$

For brevity, we denote by $\mathcal{K}$ the set of such kernels.

5.1. **Truncation of convolution kernels.** Our first lemma shows that "cutting off the tail" of a non-compactly-supported kernel has a negligible effect on the averages of a doubling weight. To state the result, let $\varphi^{(T)}$ denote the truncation of a kernel φ at radius T, i.e.,

$$\varphi^{(T)}(x) = \begin{cases} \varphi(x) & \text{if } |x| \le T, \\ 0 & \text{otherwise}, \end{cases}$$

and let $\varphi_t^{(T)}$ denote the usual, mass-preserving dilate of $\varphi^{(T)}$ at scale t (observe that truncation occurs before dilation).

Lemma 3. *Suppose that $\varphi \in \mathcal{K}$. If $w \in Db$, then*

$$\frac{w * \varphi_t^{(T)}}{w * \varphi_t} \to 1, \quad as \ T \to \infty,$$

uniformly over all positive t.

Proof. The ratio $(w * \varphi_t^{(T)})/(w * \varphi_t)$ is no larger than 1, since $w \ge 0$ and $\varphi^{(T)} \le \varphi$. The assertion here is rather that this ratio is not much smaller than 1 for large T. If we only use the doubling constant $Db(w)$ and the decay properties of φ, then it suffices to show this solely for convolutions formed about the origin at scale $t = 1$.

Divide $\mathbf{R}^n$ into dyadic rings: For T larger than 1 and $k = 0, 1, 2, \dots$, let $TR_k = B_{2^k T}(0) \setminus B_{2^{k-1} T}(0)$. Since φ decreases rapidly and w doubles over consecutive rings—that is, there is a constant C depending only on $Db(w)$ such that $w(TR_{k+1}) \le Cw(TR_k)$, independently of T and k—then

$$\begin{aligned}
w * \varphi(0) &= \int_{B_T(0)} w(x)\varphi(-x)\,dx + \sum_{k=1}^{\infty} \int_{TR_k} w(x)\varphi(-x)\,dx \\
&\le w * \varphi^{(T)}(0) + \sum_{k=1}^{\infty} \left(\frac{A_N}{(2^{k-1}T)^N} \int_{TR_k} w(x)\,dx \right) \\
&\le w * \varphi^{(T)}(0) + \left(\sum_{k=1}^{\infty} \frac{A_N C^k}{(2^{k-1}T)^N} \right) \int_{TR_0} w(x)\,dx.
\end{aligned}$$

The last series is summable for large N, hence

$$(11) \qquad w * \varphi(0) - w * \varphi^{(T)}(0) \le C_w T^{-N} \int_{B_T(0)} w(x)\,dx,$$

for some constant $C_w = C_w\big(Db(w), \varphi, N\big)$.

To complete the proof, insert the kernel φ back into the integral on the right-hand side of (11); this is possible with a further application of doubling. Indeed, suppose $\varphi(-x_0) > 0$ at some point $-x_0$. Continuity implies that φ is strictly positive in some neighborhood of this point: say, $\varphi(-x) > a > 0$ when $|x - x_0| < b < 1$. Then, by repeated doubling,

$$(12) \quad \int_{B_T(0)} w(x)\, dx \leq C'_w(T) \int_{B_b(x_0)} w(x)\, dx \leq a^{-1} C'_w(T) \int_{B_T(0)} w(x)\varphi(-x)\, dx,$$

with $C'_w(T) = \left(2^n Db(w)\right)^{\log(2T/b)+1}$. Now choose N so large that $T^{-N} C'_w(T) \to 0$ as $T \to \infty$. Then (11) and (12) combine to yield

$$w * \varphi(0) = \left(1 + o(1)\right) w * \varphi^{(T)}(0), \quad \text{as } T \to \infty,$$

which is the desired estimate. $\square$

5.2. **Doubling in terms of general kernels.** Let us now define the classes Db_φ and $Db_{\varphi,as}$ on the model of the doubling and asymptotically doubling conditions. We use the notation T^λ for the operator of translation by λ, i.e., $T^\lambda \varphi(x) = \varphi(x - \lambda)$.

If $\varphi \in \mathcal{K}$, then Db_φ is the set of all weights w for which there is some constant C such that[4]

$$C^{-1} w * \varphi_t \leq w * \varphi_{2t} \leq C w * \varphi_t$$

uniformly over all positive t; denote the smallest bound C by $Db_\varphi(w)$. If, in addition, both

$$\sup_{1/2 \leq \rho \leq 2} \frac{w * \varphi_{\rho t}}{w * \varphi_t} \quad \text{and} \quad \sup_{|\lambda| \leq 1} \frac{w * (T^\lambda \varphi)_t}{w * \varphi_t}$$

approach 1 as $t \to 0$, then we write $w \in Db_{\varphi,as}$.

Note that these are exactly the defining conditions for Db and Db_{as} when the kernel φ is replaced by χ, the normalized characteristic function of the unit ball. Since the truncation of a kernel in $\mathcal{K}$ looks much like χ, we might expect from the previous lemma that the classes Db and Db_φ coincide, and perhaps even that their asymptotic versions Db_{as} and $Db_{\varphi,as}$ agree. For kernels that resemble the Gaussian, we shall ultimately see that this is the case. The prior lemma provides the first step toward the proof of this result.

Corollary 4. *If $\varphi \in \mathcal{K}$, then $Db \subseteq Db_\varphi$. In fact, for each doubling weight w, there is a constant $C = C\left(Db(w), \varphi\right)$ such that $C^{-1} \leq (w * \varphi_t)/(w * \chi_t) \leq C$ uniformly over all positive t.*

Proof. Use the lemma to pick T such that $(w * \varphi_t)/(w * \varphi_t^{(T)}) \leq 2$. Since $\varphi^{(T)}$ is both strictly positive on some ball and bounded over the ball $B_T(0)$ containing its support, then the doubling property of w insures that $c^{-1} \leq (w * \varphi_t^{(T)})/(w * \chi_t) \leq c$, for some constant $c = c\left(Db(w), \varphi\right)$. Take $C = 2c$. $\square$

Under a mild additional assumption, the converse is also true.

Lemma 5. *If $\varphi \in \mathcal{K}$ and $\varphi(0) > 0$, then $Db_\varphi \subseteq Db$.*

---

[4]A functional inequality in which the spatial variable is not shown is always to be understood to hold uniformly over all of the underlying domain $\mathbf{R}^n$.

Proof. Note first that, when a weight $w \in Db_\varphi$, its mean values grow at most polynomially: There is a large number $M = M\big(Db_\varphi(w)\big)$ and a constant C such that

$$(13) \qquad w * \chi_{rt} \leq Cr^M \, w * \varphi_t, \quad \text{for } r \geq 1.$$

Indeed, suppose that $\chi \leq \varphi$; this represents no loss of generality, as φ can otherwise be replaced by a scalar multiple of a suitable dilate of itself. Iterating the Db_φ condition then shows that

$$w * \chi_{2^k t} \leq w * \varphi_{2^k t} \leq Db_\varphi(w)^k w * \varphi_t,$$

which gives (13), with $M = \log_2 Db_\varphi(w)$.

Now break up the region of integration for $w * \varphi_t$ into large rings of size $2^j r$ and use the rapid decay of φ and the bound (13). With $A = A_{2M+n}$ from (10), then

$$\begin{aligned}
w * \varphi_t &\leq& A\omega_n r^n w * \chi_{rt} + A\omega_n \sum_{j=1}^{\infty} (2^{j-1}r)^{-2M-n}(2^j r)^n w * \chi_{2^j rt} \\
&\leq& A\omega_n r^n w * \chi_{rt} + \left[CA\omega_n \sum_{j=1}^{\infty} (2^{j-1}r)^{-2M-n}(2^j r)^{n+M} \right] w * \varphi_t.
\end{aligned}$$

Choose $r = R$ to be a power of 2 so large that the bracketed term is less than $1/2$. Then

$$w * \varphi_t \leq C' w * \chi_{Rt},$$

for $C' = 2A\omega_n R^n$. Thus,

$$w * \chi_{2t} \leq w * \varphi_{2t} \leq C'' w * \varphi_{t/R} \leq C'C'' w * \chi_t,$$

for $C'' = Db_\varphi(w)^{\log_2 R + 1}$. This means that $(w * \chi_{2t})/(w * \chi_t)$ is uniformly bounded, i.e., $w \in Db$. $\qquad \square$

The last results guarantee the equivalence of the mean values and Gaussian averages of every doubling weight. We next wish to establish that this equivalence carries over to the asymptotic setting. To do so, we shall first prove a technical result that shows how modest dilations and translations of a positive kernel have only a negligible effect on the averages of a doubling weight (compare Lemma 1). This result will also prove essential in the final step of the main theorem.

Lemma 6. *Suppose that $\varphi \in \mathcal{K}$ and that $\varphi > 0$. If $w \in Db$, then*

$$(14) \qquad \frac{w * \varphi_{\rho t}}{w * \varphi_t} \to 1, \quad as \ \rho \to 1,$$

and

$$(15) \qquad \frac{w * (T^\lambda \varphi)_t}{w * \varphi_t} \to 1, \quad as \ |\lambda| \to 0,$$

uniformly for all positive t.

Proof. By Lemma 3, it suffices to show (14) for the truncated kernel $\varphi^{(T)}$ in place of φ. By translation and dilation invariance, it further suffices to work at scale $t = 1$ and position $x = 0$. That is, we wish to show, for fixed T, that the ratio

$$\frac{w * \varphi_\rho^{(T)}(0)}{w * \varphi^{(T)}(0)}$$

approaches 1 at a rate that depends only on φ and the doubling constant of w.

Since $\varphi^{(T)}$ has compact support, we can hope to relate this to the corresponding unweighted result, the "flat" case of averages over balls. In particular, as w is doubling, its mass over a ball differs negligibly from its mass over small dilates of the ball; that is, by Lemma 1,

$$(16) \qquad \frac{w\big(B_{\rho T}(0)\big)}{w\big(B_T(0)\big)} \to 1, \quad \text{as } \rho \to 1.$$

To exploit this, suppose ρ is larger than 1, expand $w * \varphi_\rho^{(T)}(0)$ as an integral, and split up the region of integration:

$$w * \varphi_\rho^{(T)}(0) = \int_{B_T(0)} w(x)\varphi_\rho^{(T)}(-x)\,dx + \int_{R_{\rho T}(0)} w(x)\varphi_\rho^{(T)}(-x)\,dx = I + II,$$

with $R_{\rho T}(0) = B_{\rho T}(0) \setminus B_T(0)$. The first integral approaches $w*\varphi^{(T)}(0)$ as $\rho \to 1^+$. Indeed,

$$I = \rho^{-n} \int_{B_T(0)} w(x)\varphi^{(T)}(-x/\rho)\,dx \leq \rho^{-n}\alpha^{(T)}(\rho) \int_{B_T(0)} w(x)\varphi^{(T)}(-x)\,dx,$$

with

$$(17) \qquad \alpha^{(T)}(\rho) = \sup\left\{ \frac{\varphi^{(T)}(x/\rho)}{\varphi^{(T)}(x)} : |x| < T \right\}.$$

Since the kernel $\varphi^{(T)}$ is uniformly continuous and strictly positive on its support (by the assumed positivity of φ), then $\alpha^{(T)}(\rho) \to 1$ as $\rho \to 1^+$. Hence,

$$(18) \qquad I \leq \big(1 + o(1)\big)\, w * \varphi^{(T)}(0), \quad \text{as } \rho \to 1^+.$$

The second integral II is small. In fact, as all the values of φ are comparable on $B_T(0)$ (again, by the assumptions on φ), the ratio

$$(19) \qquad \beta^{(T)} = \frac{\sup\{\varphi(x) : |x| < T\}}{\inf\{\varphi(x) : |x| < T\}}$$

is finite for each T. Pull the largest value of $\varphi^{(T)}$ out of the integral II, use (16), and re-insert $\varphi^{(T)}$ back into the integral by dividing by its smallest value. Then

$$
\begin{aligned}
II \;&\leq\; \beta^{(T)} \frac{w\big(R_{\rho T}(0)\big)}{w\big(B_T(0)\big)} \int_{B_T(0)} w(x)\varphi^{(T)}(-x)\,dx \\
&=\; o(1)\, w * \varphi^{(T)}(0), \quad \text{as } \rho \to 1^+.
\end{aligned}
$$
(20)

Combining the estimates for I and II yields

$$\limsup_{\rho \to 1^+} \frac{w * \varphi_\rho^{(T)}(0)}{w * \varphi^{(T)}(0)} = 1.$$

The corresponding bound from below and the translation result (15) follow similarly. $\qquad\square$

5.3. **Leveling of convolution kernels.** The next result gives the "flattening out the bumps" procedure alluded to earlier; it allows us to swap χ for the convolution kernel φ in the asymptotic setting.

Lemma 7. *Suppose that* $\varphi \in \mathcal{K}$, $\varphi > 0$, *and* $\int \varphi = 1$. *If* $w \in Db_{as}$, *then*

$$(21) \qquad \frac{w * \varphi_t}{w * \chi_t} \to 1, \quad as \ t \to 0.$$

The same conclusion holds when $w \in Db_{\varphi,as}$.

Proof. We shall often have occasion to state that the ratio between two quantities is close to 1. For notational convenience, let us write

$$X \overset{\varepsilon}{\sim} Y$$

for the assertion that

$$(1 + \varepsilon)^{-1} Y < X < (1 + \varepsilon)Y.$$

Quantities X and Y that satisfy this will be said to be ε-*comparable*.

Suppose that φ is given and that $w \in Db_{as}$. Fix a small, positive ε. Use Lemma 3 to find a truncation radius T such that

$$(22) \qquad w * \varphi_t(x) \overset{\varepsilon}{\sim} w * \varphi_t^{(T)}(x)$$

uniformly over all (x, t) in $\mathbf{R}_+^{n+1}$; suppose, in addition, that T is so large that $(1 + \varepsilon)^{-1} < \int \varphi^{(T)} \leq 1$. Next, use the argument of Lemma 6 to choose a number ρ just smaller than 1 so that

$$(23) \qquad w * \varphi_t^{(T)}(x) \overset{\varepsilon}{\sim} w * \varphi_{\rho t}^{(T)}(x).$$

Furthermore, suppose ρ is so close to 1 that there is a small, positive r such that

$$(24) \qquad \varphi_\rho^{(T)}(x) < (1 + \varepsilon)\varphi^{(T)} * \chi_r(x)$$

for all x in $\mathbf{R}^n$; this is possible because $\varphi^{(T)}$ is strictly positive and uniformly continuous over its support $B_T(0)$. Convert (24) to scale t and combine it with the previous estimates (22) and (23). Then

$$w * \varphi_t(x) < (1 + \varepsilon)^2 w * \varphi_{\rho t}^{(T)}(x) < (1 + \varepsilon)^3 w * (\varphi^{(T)} * \chi_r)_t(x)$$

for all (x, t) in $\mathbf{R}_+^{n+1}$. This last term may be written

$$(1 + \varepsilon)^3 w * \varphi_t^{(T)} * \chi_{rt}(x) = (1 + \varepsilon)^3 \int_{B_{Tt}(x)} \big(w * \chi_{rt}(y)\big)\varphi_t^{(T)}(x - y)\, dy,$$

which has the form of an average of an average.[5]

To this point, the argument has been scale-invariant. At small scales t, however, the first factor in the last integrand is nearly constant over the (bounded) region of integration. Indeed, since $w \in Db_{as}$, then

$$(25) \qquad w * \chi_{rt}(y) \overset{\varepsilon}{\sim} w * \chi_t(x)$$

---

[5]The idea for this method stems from [9]. There, truncations of the Poisson kernel (and other kernels with polynomial decay) are compared to their averages formed over small scales.

uniformly for all y in $B_{Tt}(x)$, when t is sufficiently small. Since $\int \varphi_t^{(T)} \leq 1$, inserting (25) into the integrand leads to the conclusion that

$$w * \varphi_t(x) < (1+\varepsilon)^4 \int_{B_{Tt}(x)} w * \chi_t(x) \varphi_t^{(T)}(x-y)\, dy \leq (1+\varepsilon)^4 w * \chi_t(x),$$

when t is small. This is one half of (21).

For the opposite inequality, find a ρ just larger than 1 and a small, positive r such that

(26) $$\varphi_\rho^{(T)}(x) > (1+\varepsilon)^{-1} \varphi^{(T)} * \chi_r(x).$$

Use this in place of (24) and invoke (22) and (23) once more. It follows that

$$w * \varphi_t(x) > (1+\varepsilon)^{-3} w * \varphi_t^{(T)} * \chi_{rt}(x),$$

and a similar analysis, coupled with the estimate $\int \varphi_t^{(T)} > (1+\varepsilon)^{-1}$, finishes the proof. To obtain (21) when $w \in Db_{\varphi,as}$, simply reverse the roles of $\varphi^{(T)}$ and χ in the above argument from (23) onward and use Lemma 1 in place of Lemma 6. $\quad\square$

We now summarize the results in this section.

Theorem 3. *If φ is positive, continuous, and rapidly decreasing, then*

$$Db = Db_\varphi \quad and \quad Db_{as} = Db_{\varphi,as}.$$

5.4. **Control of the leading term** (STEP I). Let us now return to considering the main term in the estimate of the Carleson norm, namely

$$h_1(s) = \frac{1}{\omega_n s^n} \int_{|x|<s} \log \frac{u(x, s^2)}{w(x)}\, dx.$$

Recall that $u(x, s^2) = w * \varphi_s(x)$, with φ the normalized Gaussian. Suppose that $w \in Db_{as}$. The Gaussian satisfies all the conditions of Lemma 7, so that

$$\left| \log \frac{u(x, s^2)}{w * \chi_s(x)} \right| = \left| \log \frac{w * \varphi_s(x)}{w * \chi_s(x)} \right| = o(1), \quad \text{as } s \to 0,$$

uniformly over all x in $\mathbf{R}^n$. Moreover, as a consequence of asymptotic doubling,

$$\sup_{|x|<s} \left| \log \frac{w * \chi_s(x)}{w * \chi_s(0)} \right| = o(1), \quad \text{as } s \to 0.$$

So

$$h_1(s) = \frac{1}{\omega_n s^n} \int_{|x|<s} \log \frac{w * \chi_s(0)}{w(x)}\, dx + o(1), \quad \text{as } s \to 0.$$

In sum,

(27) $$w \in Db_{as} \implies \lim_{s \to 0} h_1(s) = \lim_{s \to 0} \big(\log(w * \chi_s)(0) - (\log w) * \chi_s(0) \big).$$

Therefore, if $w \in A_{\infty,as}$, then $\lim_{s\to 0} h_1(s) = 0$, completing the proof of STEP I. $\quad\square$

5.5. **Control of the secondary term** (STEP II). The lateral boundary term in the estimate of the Carleson norm is given, after a change of variables, by

$$(28) \qquad h_2(s) = \frac{2}{\omega_n s^n} \int_{t=0}^s \int_{|x|=s} \frac{t \, \nabla u(x, t^2) \cdot \vec{n}_x}{u(x, t^2)} \, dx dt.$$

As noted, the kernel $(\nabla \varphi)_t \cdot \vec{n}_x$ implicit in the numerator of the integrand has integral zero over $\mathbf{R}^n$ for each fixed x. The normal vector $\vec{n}_x$ fixes a hyperplane through the origin that divides this kernel into its regions of positive and negative values. In order to obtain results analogous to those for the positive kernels discussed in the previous section, we shall treat each of these half-spaces separately. On account of the rotational invariance of the doubling condition, we may restrict our attention, without loss of generality, to the case when $\vec{n}_x$ is $\vec{e}_1$, the unit vector in the positive x_1-direction. Accordingly, let us set

$$(29) \qquad \tilde{\varphi}(x) = \tilde{\varphi}(x_1, \dots, x_n) = \begin{cases} kx_1 e^{-|x|^2} & \text{if } x_1 > 0, \\ 0 & \text{otherwise,} \end{cases}$$

with the constant $k = 2\pi^{(1-n)/2}$ chosen so that $\int \tilde{\varphi} = 1$.

We wish to compare the averages of an asymptotic doubling weight w formed with $\tilde{\varphi}$ to those formed with χ, that is, to the standard mean values of w over balls. How much of the previous analysis of Sections 5.1–5.3 applies to $\tilde{\varphi}$? Well, since $\tilde{\varphi}$ is continuous and rapidly decreasing, the truncation result (Lemma 3) holds directly as stated for $\tilde{\varphi}$. But since $\tilde{\varphi}$ is not positive, Lemmas 6 and 7 do not on the surface seem to apply. If we examine the proofs of these latter two lemmas, however, we find that positivity enters only at three points: the convergence of the quantity $\alpha^{(T)}(\rho)$ defined in (17) to 1 as ρ approaches 1 from above; the finiteness of the quantity $\beta^{(T)}$ defined in (19); and the possibility of finding a ρ close to 1 and a small, positive r such that (24) and (26) hold. We now show how to modify the arguments involved at each of these three points[6] to obtain versions of Lemmas 6 and 7 valid for $\tilde{\varphi}$.

First modification. The first point requires only a minor adjustment: Simply replace the references to "$|x| < T$" in the proof of Lemma 6 by "$|x| < T, x_1 > 0$." For with

$$\alpha^{(T)}(\rho) = \sup \left\{ \frac{\tilde{\varphi}^{(T)}(x/\rho)}{\tilde{\varphi}^{(T)}(x)} : |x| < T, x_1 > 0 \right\},$$

then $\alpha^{(T)}(\rho) = \sup\{\rho^{-1} \exp((1 - \rho^{-2})|x|^2) : |x| < T\}$, which converges to 1 when $\rho \to 1^+$. Setting

$$\tilde{I} = \int_{B_T(0)} w(-x) \tilde{\varphi}_\rho^{(T)}(x) \, dx,$$

then the argument of Lemma 6 yields an estimate analogous to (18), namely that $\tilde{I} \leq (1 + o(1)) \, w * \tilde{\varphi}^{(T)}(0)$, as $\rho \to 1^+$.

---

[6] The author wishes to thank the anonymous referee for pointing out several errors in a previous version of this section.

Second modification. Here, as well, only minor changes are needed to the argument of Lemma 6 to get around the difficulty that $\tilde{\varphi}^{(T)}(x)$ becomes arbitrarily small near $\{x_1 = 0\}$. Fix a large truncation radius T and set

$$\widetilde{II} = \int_{R_{\rho T}(0)} w(-x)\tilde{\varphi}_\rho^{(T)}(x)\, dx.$$

The task is then to show that $\widetilde{II} = o(1)\, w * \tilde{\varphi}^{(T)}(0)$, as $\rho \to 1^+$. Choose a ball B within $B_T(0)$ where the kernel $\tilde{\varphi}^{(T)}$ is strictly positive, say the ball with radius 1 and center $(2, 0, \dots, 0)$ on the x_1-axis. Then the argument leading to (20) yields

$$\begin{aligned}
\widetilde{II} &\leq \sup\{\tilde{\varphi}_\rho^{(T)}(x) : x \in R_{\rho T}(0)\} w\big(R_{\rho T}(0)\big) \\
&\leq \sup\{\tilde{\varphi}_\rho^{(T)}(x) : x \in R_{\rho T}(0)\} \frac{w\big(B_T(0)\big)}{w(-B)} \frac{w\big(R_{\rho T}(0)\big)}{w\big(B_T(0)\big)} \int_B w(-x)\, dx \\
&\leq \frac{\sup\{\tilde{\varphi}_\rho^{(T)}(x) : x \in R_{\rho T}(0)\}}{\inf\{\tilde{\varphi}^{(T)}(x) : x \in B\}} \frac{w\big(B_T(0)\big)}{w(-B)} \frac{w\big(R_{\rho T}(0)\big)}{w\big(B_T(0)\big)} \int_B w(-x)\tilde{\varphi}^{(T)}(x)\, dx.
\end{aligned}$$

In the last line, the first factor is bounded due to the choice of B and the second thanks to the doubling property of w; the third vanishes as $\rho \to 1^+$, and the fourth can be increased by expanding the region of integration from B to $B_T(0)$. Altogether, the right-hand side is then $o(1)\, w * \tilde{\varphi}^{(T)}(0)$, which is the desired result.

Third modification. The proof of Lemma 7 carries over to $\tilde{\varphi}^{(T)}$ save in one place: The bounds (24) and (26) are not valid for points x near the plane $\{x_1 = 0\}$. A substitute for (24), however, is provided by the upper bound

$$(30) \qquad \tilde{\varphi}_\rho^{(T)}(x) < (1 + \varepsilon)\tilde{\varphi}^{(T)} * \chi_r(x) + c_1 \xi^{(r_0, T)}(x),$$

where $\xi^{(r_0, T)}$ denotes the characteristic function of the thin n-dimensional interval $[-2r_0, 2r_0] \times [-2T, 2T]^{n-1}$ and χ retains its previous meaning as the normalized characteristic function of the unit ball. By choosing the parameters r and r_0 carefully, we shall see that the extra term on the right-hand side of (30) introduces only a neglible error. The details are as follows.

Fix ε and let the truncation radius T and the dilation factor ρ be chosen so large, respectively, so close to (and just smaller than) 1, that the estimates (22) and (23) are valid for $\tilde{\varphi}$ in place of φ; we have just argued that this is possible. Doubling insures that a suitably thin slice of the cube $[-2T, 2T]^n$ near the plane $\{x_1 = 0\}$ has small mass relative to that of the whole cube: There exists a small width r_0, depending only on T and $Db(w)$, such that

$$w([-2r_0, 2r_0] \times [-2T, 2T]^{n-1}) < \varepsilon w([-2T, 2T] \times [-2T, 2T]^{n-1}).$$

Since the unit ball comprises a substantial share (for fixed T) of $[-2T, 2T]^n$, then another application of doubling allows us to re-write the last estimate as

$$w * \xi_t^{(r_0, T)}(x) < c_2 \varepsilon w * \chi_t(x).$$

Now, on the one hand, as $\tilde{\varphi}^{(T)}$ is strictly positive and uniformly continuous on $B_T(0) \cap \{x_1 > r_0\}$, it is possible to find r smaller than r_0 such that the original upper bound (24) holds for $\tilde{\varphi}^{(T)}$ away from the critical plane $\{x_1 = 0\}$:

$$\tilde{\varphi}_\rho^{(T)}(x) < (1 + \varepsilon)\tilde{\varphi}^{(T)} * \chi_r(x), \quad \text{when } |x_1| > 2r_0 \text{ or } |x| > \rho T.$$

Near this plane (say, when $|x_1| \leq 2r_0$ and $|x| \leq \rho T$), on the other hand, the kernel $\tilde{\varphi}_\rho^{(T)}$ is dominated by $\xi^{(r_0, T)}$, since $\tilde{\varphi}$ is bounded and ρ is near 1. These last two

observations combine to yield (30), for a suitable constant c_1. Proceeding as in the proof of Lemma 7, we then find, for small t, that

$$
\begin{aligned}
w * \tilde{\varphi}_t^{(T)}(x) &< (1+\varepsilon)^3 (w * \chi_{rt}) * \tilde{\varphi}_t^{(T)}(x) + c_1 w * \xi_t^{(r_0,T)}(x) \\
&< (1+\varepsilon)^4 w * \chi_t(x) + c_1 c_2 \varepsilon w * \chi_t(x) \\
&= \big(1 + O(\varepsilon)\big) w * \chi_t(x).
\end{aligned}
$$

This is the desired result. The converse estimate follows similarly, with the lower bound

$$
(31) \qquad \tilde{\varphi}_\rho^{(T)}(x) > (1+\varepsilon)^{-1} \tilde{\varphi}^{(T)} * \chi_r(x) - c_1 \xi^{(r_0,T)}(x)
$$

(for ρ just larger than 1) serving in place of (26).

As a result, the conclusions of all the lemmas in question hold for $\tilde{\varphi}$.[7] In particular, we can re-state Lemma 7 in this context as follows:

Lemma 8. *Let $\tilde{\varphi}$ be as in (29). If $w \in Db_{as}$, then*

$$
(32) \qquad \frac{w * \tilde{\varphi}_t}{w * \chi_t} \to 1, \quad as\ t \to 0.
$$

We now apply this to the analysis of $h_2(s)$. In terms of $\tilde{\varphi}$, the numerator in the integrand of h_2 can be written as a difference,

$$
(33) \qquad |t\, \nabla u(x, t^2) \cdot \vec{n}_x| = \pi^{-1/2} |w * (\tau^{(x)} \tilde{\varphi})_t(x) - w * (\tau^{(-x)} \tilde{\varphi})_t(x)|,
$$

for $\tau^{(x)}$ the rotation that maps $\vec{e}_1$ into $\vec{n}_x$. Since the class Db_{as} is rotation-invariant, then the preceding lemma implies that the two terms in the difference are ε-comparable at all sufficiently small scales, when $w \in Db_{as}$. That is, given any small ε, we can find δ such that if $0 < t < s < \delta$, then

$$
w * (\tau^{(x)} \tilde{\varphi})_t(x) \overset{\varepsilon}{\sim} w * \chi_t(x) \overset{\varepsilon}{\sim} w * (\tau^{(-x)} \tilde{\varphi})_t(x).
$$

Hence the numerator in (28) is small:

$$
|t\, \nabla u(x, t^2) \cdot \vec{n}_x| \leq C\varepsilon w * \chi_t(x)
$$

whenever $|x| = s$ and $0 < t < s < \delta$. As for the denominator in (28), Lemma 7 shows that

$$
\frac{u(x, t^2)}{w * \chi_t(x)} \to 1, \quad as\ t \to 0,
$$

uniformly over all x in $\mathbf{R}^n$. Thus, the integrand of $h_2(s)$ is pointwise no larger than $C'\varepsilon$, so that $h_2(s) \leq C''\varepsilon$, provided only that s is sufficiently small. Therefore,

$$
(34) \qquad w \in Db_{as} \implies \lim_{s \to 0} h_2(s) = 0,
$$

which completes the proof of STEP II and of the necessity of condition (9) in Theorem 2. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ☐

---

[7]We have given above only a proof that the dilation result (14) in Lemma 6 continues to be valid for $\tilde{\varphi}$; the corresponding translation result (15) can be obtained by a variant of the argument in the paragraph headed "Third modification." In fact, the method of adding an extra term to cover the small region where the kernel $\tilde{\varphi}^{(T)}$ vanishes could also have been used to obtain the dilation result for $\tilde{\varphi}$ directly.

6. SUFFICIENCY

Our aim is now to prove the implication

$$[w \in Db, \lim_{s \to 0} h(s) = 0] \implies w \in Db_{as}.$$

This gives the sufficiency of (9) in the theorem. For when $w \in Db_{as}$, then (27) and (34) combine to give the conclusion that

$$\lim_{s \to 0} h(s) = \lim_{s \to 0} \big(\log(w * \chi_s)(0) - (\log w) * \chi_s(0)\big) = 0,$$

i.e. (using translation invariance), that $w \in A_{\infty,as}$.

Proving this implication is rather complex, however; we begin with a few preliminary observations. The Carleson-measure condition that we here assume, namely $\lim_{s \to 0} h(s) = 0$, is an averaged form of the statement that

$$(35) \qquad \frac{w * (\nabla \varphi)_t}{w * \varphi_t} \to 0, \quad \text{as } t \to 0.$$

Note that the key to our analysis so far has been a careful study of terms of precisely this sort: the ratio of the average of a weight w with respect to a kernel of integral zero to its average with respect to a non-negative kernel. Under the assumption that w is a doubling weight and that the Carleson measure based on (35) has vanishing trace, our goal is now to show that w is asymptotically doubling. Fortunately, this goal can also be formulated by means of a similar ratio of averages. Indeed, since $Db_{as} = Db_{\varphi,as}$ by Theorem 3, the statement $w \in Db_{as}$ means exactly that

$$\frac{w * (T^\lambda \varphi)_{\rho t}}{w * \varphi_t} \to 1, \quad \text{as } t \to 0,$$

uniformly for all dilation factors ρ in $[1/2, 2]$ and all translations by a vector λ within the unit ball. Set $\psi^{(\lambda,\rho)} = \varphi - (T^\lambda \varphi)_\rho$. Then $\int \psi^{(\lambda,\rho)} = 0$, and the last condition becomes the requirement that

$$(36) \qquad \frac{w * \psi_t^{(\lambda,\rho)}}{w * \varphi_t} \to 0, \quad \text{as } t \to 0,$$

uniformly over the indicated range of λ and ρ.

The crux of the remainder of this section is the passage from (35) to (36), that is, from $\nabla \varphi$ to $\psi^{(\lambda,\rho)}$. As in much of Littlewood-Paley theory, the underlying principle we shall invoke is that one sufficiently smooth kernel with integral zero is as good as any other.

6.1. **A decomposition lemma.** The first lemma begins this "change-of-kernel" procedure. It decomposes the Fourier transform of a Schwartz function into its smooth components of various dyadic frequencies.[8] To shorten the statement of our results, we introduce some notation for certain special classes of Schwartz functions. Let $\mathcal{S}_0 = \{\psi \in \mathcal{S} : \int \psi = 0, \psi \not\equiv 0\}$ and $\mathcal{S}_+ = \{\varphi \in \mathcal{S} : \varphi \geq 0, \int \varphi > 0\}$. For any finite collection $\mathcal{F} = \{\|\cdot\|_{\alpha,\beta}\}$ of seminorms on $\mathcal{S}$, let

$$\|f\|_{\mathcal{F}} = \sup\{\|f\|_{\alpha,\beta} : \|\cdot\|_{\alpha,\beta} \in \mathcal{F}\} \quad \text{and} \quad \mathcal{S}_{\mathcal{F}} = \{f \in \mathcal{S} : \|f\|_{\mathcal{F}} \leq 1\}.$$

---

[8]This is essentially the technique used by Fefferman and Stein [5] to characterize Hardy spaces via non-tangential maximal functions formed with respect to various averaging kernels.

Lemma 9. *Fix a function φ in $\mathcal{S}_+$ and a finite family $\mathcal{F}$ of seminorms on $\mathcal{S}$. Then there exists a larger (finite) family $\mathcal{F}_0$ so that for each $M > 0$ there is a constant $c = c(\varphi, \mathcal{F}, M)$ with the following property: For each Φ in $\mathcal{S}_{\mathcal{F}_0}$, there are functions $\{\eta^{(k)}\}_{k=0}^{\infty}$ in $\mathcal{S}$ such that*

$$\Phi = \sum_{k=0}^{\infty} \eta^{(k)} * \varphi_{2^{-k}} \quad and \quad \|\eta^{(k)}\|_{\mathcal{F}} \leq c2^{-Mk}.$$

For the proof, see Stein [17, pp. 93-94].

6.2. **Primitives of Schwartz functions.** To be able to apply this result fruitfully, we shall need to decompose not the zero-integral kernels themselves (say, $\nabla\varphi$ or $\varphi_1 - \varphi_2$) in terms of φ, but rather their "primitives." The next lemma shows how this is possible; it states that the divergence operator has a right inverse defined on $\mathcal{S}_0$, the set of all Schwartz functions with integral zero.

Lemma 10. *There exists a continuous, linear map $\mathcal{I} : \mathcal{S}_0 \to \mathcal{S} \times \cdots \times \mathcal{S}$ such that*

$$(37) \qquad\qquad \Psi = \mathrm{div}\, \mathcal{I}(\Psi), \quad for\ all\ \Psi\ in\ \mathcal{S}_0.$$

Proof. The one-variable case is easy. Given Ψ in $\mathcal{S}_0$, let $\Phi(x) = \int_{-\infty}^{x} \Psi(s)\,ds$. Then $\Phi' = \Psi$ and $\Phi \in C^\infty$. Since $\Phi(x) \to 0$ as $|x| \to \infty$, it remains only to show that Φ satisfies the decay estimates

$$(38) \qquad\qquad |x^N \Phi(x)| \leq C_N,$$

for constants $\{C_N\}$ that depend only on the Schwartz-seminorms of Ψ. But, by assumption, there are constants $\{A_N\}$ such that $|\Psi(x)| \leq A_{N+1}(1 + |x|)^{-(N+1)}$; when $x \leq 1$, the estimate (38) follows immediately from this by integration. We can obtain the same estimate when $x > 1$ by integrating from the right, noting that $\int \Psi = 0$; for then $\Phi(x) = -\int_x^\infty \Psi(s)\,ds$.

The construction is slightly trickier in higher dimensions, since the analogues of the one-variable "antiderivatives" of a function in $\mathcal{S}_0$ need not lie in $\mathcal{S}$. To counteract this, we introduce a C^∞ cut-off function ζ on $\mathbf{R}$; we require that $\zeta(x) = 0$ when $x \leq -1$ and that $\zeta(x) = 1$ when $x \geq 1$.

The construction in $\mathbf{R}^2$ is already typical.[9] Suppose that Ψ is in $\mathcal{S}(\mathbf{R}^2)$ and that $\int_{\mathbf{R}^2} \Psi = 0$. Set

$$\Phi^{(1)}(x, y) = \int_{-\infty}^{x} \Psi(s, y)\,ds - \zeta(x) \int_{-\infty}^{\infty} \Psi(s, y)\,ds.$$

Then

$$(\Psi - D_1 \Phi^{(1)})(x, y) = \zeta'(x) \int_{-\infty}^{\infty} \Psi(s, y)\,ds,$$

so that $(\Psi - D_1 \Phi^{(1)})$ is a C^∞ function supported in the strip $[-1, 1] \times \mathbf{R} \subset \mathbf{R}^2$. In fact, $(\Psi - D_1 \Phi^{(1)}) \in \mathcal{S}$, whence $D_1 \Phi^{(1)} \in \mathcal{S}$. A further calculation (as in the one-variable case, separately for $x \to -\infty$ and $x \to \infty$) shows that $\Phi^{(1)}$ itself is a Schwartz function.

---

[9]A similar construction is used by Bott and Tu [1, pp. 37-40] to prove the Poincaré lemma for compactly-supported cohomology. The author wishes to thank Sebastian Behm for this tip.

Now apply this construction to the remainder $(\Psi - D_1\Phi^{(1)})$ in the y-direction: Set

$$\Phi^{(2)}(x, y) = \int_{-\infty}^{y} (\Psi - D_1\Phi^{(1)})(x, t)\, dt - \zeta(y) \int_{-\infty}^{\infty} (\Psi - D_1\Phi^{(1)})(x, t)\, dt.$$

Then $\Phi^{(2)} \in \mathcal{S}(\mathbf{R}^2)$ and

$$(\Psi - D_1\Phi^{(1)} - D_2\Phi^{(2)})(x, y) = \zeta'(y)\zeta'(x) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Psi(s, t)\, ds\, dt = 0.$$

Define $\mathcal{I}(\Psi) = (\Phi^{(1)}, \Phi^{(2)})$; continuity follows as in the one-variable case. $\qquad\square$

6.3. **Change-of-kernel estimates.** We can use the last results to change from one kernel with integral zero to another in the numerator of (35). We show this first in the uniform setting, then for Carleson measures.

Lemma 11. *Fix a doubling weight w and a function φ in $\mathcal{S}_+$. There is a family $\mathcal{F}_1$ of seminorms on $\mathcal{S}$ and a constant $C = C(Db(w), \varphi)$ such that if $\Psi \in \mathcal{S}_{\mathcal{F}_1}$ and $\int \Psi = 0$, then*

$$(39) \qquad \sup_{s \le t} \frac{|w * \Psi_s|}{w * \varphi_s} \le C \sup_{s \le t} \frac{|w * (\nabla\varphi)_s|}{w * \varphi_s}.$$

Proof. By translation and dilation invariance, it suffices to consider only averages centered at the origin and at scale $t = 1$. For brevity, let us abbreviate the supremum on the right-hand side of (39) by $\nu(t)$. Let $\mathcal{F}$ be the family of seminorms $\{\|x^\alpha f(x)\|_\infty : |\alpha| \le N\}$, for some $N = N(w, \varphi)$ to be specified below. Let $\mathcal{F}_0$ be the larger family of seminorms whose existence is guaranteed by Lemma 9. Finally, let $\mathcal{F}_1$ be a still larger family with the following property: There is a constant K such that if $\Psi \in \mathcal{S}_{\mathcal{F}_1}$ and $\int \Psi = 0$, then Ψ is the divergence of an n-tuple $\vec{\Phi}$, each of the components of which is K times an element in $\mathcal{S}_{\mathcal{F}_0}$. (The existence of such a family $\mathcal{F}_1$ follows from the continuity of $\mathcal{I}$ in Lemma 10.)

It thus suffices to show (39) in the special case when $\Psi = D_j\Phi$, where Φ is a function in $\mathcal{S}_{\mathcal{F}_0}$ and j is a fixed index between 1 and n. By Lemma 9, each such Φ can be decomposed as a sum $\Phi = \sum_{k=0}^{\infty} \eta^{(k)} * \varphi_{2^{-k}}$, with $\|\eta^{(k)}\|_{\mathcal{F}} \le c_M 2^{-Mk}$. Then

$$
\begin{aligned}
|w * \Psi(0)| &= |w * D_j\Phi(0)| \\
&= \left| \sum_{k=0}^{\infty} 2^k \eta^{(k)} * w * (D_j\varphi)_{2^{-k}}(0) \right| \\
&\le \sum_{k=0}^{\infty} 2^k \int |\eta^{(k)}(-x)| \, |w * (D_j\varphi)_{2^{-k}}(x)|\, dx \\
&\le \sum_{k=0}^{\infty} 2^k \nu(2^{-k}) \int |\eta^{(k)}(-x)| \, |w * \varphi_{2^{-k}}(x)|\, dx.
\end{aligned}
$$

As the last integrand features the convolution of w with a non-negative kernel, it is now possible to exploit the doubling behavior of w. To do so, we can once again split up the region of integration dyadically: Let $R_l = B_{2^l}(0) \setminus B_{2^{l-1}}(0)$ and $R_0 = B_1(0)$. Since $w \in Db$ and $\varphi \in \mathcal{K}$, then $w \in Db_\varphi$. So there is a constant $C = C(Db(w), \varphi)$ such that

$$w * \varphi_{2^{-k}}(x) \le C^{k+2l+1} w * \varphi(0), \quad \text{when } x \in R_l;$$

in fact, $w * \varphi_{2^{-k}}(x) \leq C^{k+l} w * \varphi_{2^l}(x) \leq C^{k+l+1} w * \varphi_{2^l}(0) \leq C^{k+2l+1} w * \varphi(0)$. Inserting this into the integral gives

$$
\begin{aligned}
|w * \Psi(0)| &\leq \sum_{k=0}^{\infty} 2^k \left[\sum_{l=0}^{\infty} \int_{R_l} |\eta^{(k)}(-x)| \, |w * \varphi_{2^{-k}}(x)| \, dx \right] \nu(1) \\
&\leq \sum_{k=0}^{\infty} C(2C)^{k+1} \left[\sum_{l=0}^{\infty} C^{2l} \int_{R_l} |\eta^{(k)}(x)| \, dx \right] \nu(1) w * \varphi(0) \\
&\leq \sum_{k=0}^{\infty} C(2C)^{k+1} c_M 2^{-kM} \sum_{l=0}^{\infty} C^{2l} 2^{(1-l)N} \omega_n 2^{ln} \nu(1) w * \varphi(0);
\end{aligned}
$$

the last line follows from the bound $|x|^N |\eta^{(k)}(x)| \leq c_M 2^{-kM}$. Now, choose M and N so large that both $(2C)2^{-M}$ and $C^2 2^{n-N}$ are less than 1. Then the double series is convergent, and $|w * \Psi(0)| \leq C' \nu(1) w * \varphi(0)$. $\square$

We also need a variant of this result in an averaged form.

Corollary 12. *Fix a doubling weight w and a function φ in $\mathcal{S}_+$. If*

$$
d\mu(x, t) = \frac{|w * (\nabla\varphi)_t(x)|^2}{|w * \varphi_t(x)|^2} \, dx \frac{dt}{t}
$$

is a Carleson measure, with vanishing trace, then the same is true of the measure

$$
d\mu^{\Psi}(x, t) = \frac{|w * \Psi_t(x)|^2}{|w * \varphi_t(x)|^2} \, dx \frac{dt}{t},
$$

for every Ψ in $\mathcal{S}_0$. A quantitative statement on the model of that in the previous lemma also holds: There is a family $\mathcal{F}_1$ of seminorms on $\mathcal{S}$ and a constant $C = C\big(Db(w), \varphi\big)$ such that if $\Psi \in \mathcal{S}_{\mathcal{F}_1}$ and $\int \Psi = 0$, then

$$
(40) \qquad \sup_{B: |B| \leq s} \frac{\mu^{\Psi}\big(T(B)\big)}{|B|} \leq C \left(\sup_{B: |B| \leq 2^N s} \frac{\mu\big(T(B)\big)}{|B|} + 2^{-N} \sup_{B} \frac{\mu\big(T(B)\big)}{|B|} \right),
$$

for all sufficiently large numbers N.

Proof. The proof is in many respects similar to that just given; we thus only sketch the argument. Let $\mathcal{F}$, $\mathcal{F}_0$, and $\mathcal{F}_1$ be as above. Once again, for (40) it suffices to consider $\Psi = D_j \Phi$ for some fixed function Φ in $\mathcal{S}_{\mathcal{F}_0}$ and some fixed j, to take $s = 1$, and to set $B = B_1$, the unit ball. Let $\psi = D_j \varphi$. Then, via the decomposition lemma,

$$
\begin{aligned}
\mu^{\Psi}\big(T(B_1)\big) &\leq \sum_{k=0}^{\infty} 2^{2k} \int_{T(B_1)} \frac{|\eta_t^{(k)} * w * \psi_{2^{-k}t}(x)|^2}{|w * \varphi_t(x)|^2} \, dx \frac{dt}{t} \\
&\leq c \sum_{k=0}^{\infty} 2^{2k} \int_{T(B_1)} \frac{(|\eta_t^{(k)}| * |w * \psi_{2^{-k}t}|^2)(x)}{|w * \varphi_t(x)|^2} \, dx \frac{dt}{t} \\
&\leq c \sum_{k=0}^{\infty} 2^{2k} \int_{T(B_1)} \int_{\mathbf{R}^n} \frac{|\eta_t^{(k)}(x-z)| |w * \psi_{2^{-k}t}(z)|^2}{|w * \varphi_t(x)|^2} \, dz \, dx \frac{dt}{t} \\
&\leq c \sum_{k=0}^{\infty} 2^{2k} \int_{T(B_1)} \sum_{l=0}^{\infty} \int_{R_l(x)} |\eta_t^{(k)}(x-z)| \frac{|w * \psi_{2^{-k}t}(z)|^2}{|w * \varphi_{2^{-k}t}(z)|^2} \\
&\qquad \times \frac{|w * \varphi_{2^{-k}t}(z)|^2}{|w * \varphi_t(x)|^2} \, dz \, dx \frac{dt}{t}.
\end{aligned}
$$

Here $R_l(x) = B_{2^l}(x) \backslash B_{2^{l-1}}(x)$ when $l \geq 1$ and $R_0(x) = B_1(x)$. The second estimate holds by an application of the Cauchy-Schwarz inequality, since $\eta^{(k)} \in \mathcal{S}_{\mathcal{F}}$ implies that $\|\eta^{(k)}\|_{L^1}$ is bounded independently of k. Doubling comes into play once again: There is a constant C so that $|w * \varphi_{2^{-k}t}(z)| \leq C^{k+2l+1}|w * \varphi_t(x)|$ when $x \in B_1$ and $z \in R_l(x)$. The estimate (40) follows from inserting this, enlarging the region of integration, and applying Fubini's theorem. $\qquad\square$

6.4. **Completion of the proof** (STEP III). Let $u(x,t) = w * \varphi_{\sqrt{t}}(x)$ be the heat extension of a doubling weight w. Suppose that

$$\frac{|w * (\nabla \varphi)_t(x)|^2}{|w * \varphi_t(x)|^2} \, dx \frac{dt}{t}$$

is a Carleson measure, with vanishing trace. Then we claim that $w \in Db_{as}$, and thus by STEPS I and II, that $w \in A_{\infty,as}$.

Were it the case that instead of the averaged (i.e., Carleson measure) condition, we knew the uniform pointwise condition

$$\frac{|w * (\nabla \varphi)_t|}{w * \varphi_t} \to 0, \quad \text{as } t \to 0,$$

then the matter would be settled. For since the Gaussian φ is in $\mathcal{S}_+$, we could use Lemma 11 to replace $\nabla\varphi$ in the numerator by other smooth kernels of integral zero, and the result would be

$$(41) \qquad \frac{|w * (T^\lambda \varphi)_{\rho t} - w * \varphi_t|}{w * \varphi_t} \to 0, \quad \text{as } t \to 0,$$

uniformly over all λ in the unit ball and all ρ in $[1/2, 2]$.[10] This would mean that $w \in Db_{\varphi,as} = Db_{as}$, exactly as claimed.

But instead of the estimate (41), Corollary 12 allows us only to conclude a priori that the measures

$$(42) \qquad d\mu^{(\lambda,\rho)}(x,t) = \frac{|w * (T^\lambda \varphi)_{\rho t}(x) - w * \varphi_t(x)|^2}{|w * \varphi_t(x)|^2} \, dx \frac{dt}{t}$$

are Carleson measures, with uniformly vanishing trace over the same range of λ and ρ. How can we close the gap?[11] The answer lies in the doubling behavior of w.

For suppose that $w \in Db$, that the measures in (42) have uniformly vanishing trace, but that, arbitrarily close to the boundary $\mathbf{R}^n$ of the half-space $\mathbf{R}^{n+1}_+$, the density of (at least) one of these is somewhere "large"; specifically, suppose that

$$(43) \qquad \frac{|w * (T^{\lambda_0}\varphi)_{\rho_0 t_0}(x_0) - w * \varphi_{t_0}(x_0)|^2}{|w * \varphi_{t_0}(x_0)|^2} > \varepsilon,$$

for some ε less than 1, some particular point (x_0, t_0) in the upper half-space, and some ρ_0 and λ_0 in the indicated range. Since $w \in Db$, then a similar inequality holds for all (x,t) in an appropriately-scaled neighborhood of (x_0, t_0). Indeed, our prior result on the negligible effect of modest dilations and translations of the averaging

---

[10]The uniformity is due to the boundedness of the set $\{\varphi - (T^\lambda\varphi)_\rho : 1/2 \leq \rho \leq 2, |\lambda| \leq 1\}$ in $\mathcal{S}_0$.

[11]Note that $F(x,t) \, dxdt/t$ being a Carleson measure with vanishing trace does not imply that $F(x,t) \to 0$ uniformly in x as $t \to 0$; take F, for example, to be the characteristic function of the cuspidal set $\{(x,t) : |x| < t^2, t < 1\}$.

kernel φ (Lemma 6) insures that there is a ρ close to 1 and a λ close to 0, both depending only on $Db(w)$, φ, and ε (not on ρ_0, λ_0, x_0, or t_0), such that

$$\frac{|w * (T^{\lambda_0}\varphi)_{\rho_0 t}(x) - w * \varphi_t(x)|^2}{|w * \varphi_t(x)|^2} > \frac{\varepsilon}{2}$$

for all (x, t) in the set

$$E_{x_0, t_0} = \{(x, t) : \rho^{-1} t_0 < t < t_0, |x - x_0| < \lambda t_0\}$$

near the point (x_0, t_0). Thus, if $w \in Db$ and (43) holds, then

$$\mu^{(\lambda_0, \rho_0)}\Big(T\big(B_{t_0}(x_0)\big)\Big) > \frac{\varepsilon}{2} \iint_{E_{x_0, t_0}} dx \frac{dt}{t}.$$

Since E_{x_0, t_0} fills a substantial share of the Carleson cylinder, then the right-hand side exceeds $(\varepsilon/2)(\lambda^n \log \rho)|B_{t_0}(x_0)|$; this means that the Carleson norm of $\mu^{(\lambda_0, \rho_0)}$ is large (i.e., it exceeds a fixed share of ε) at scale t_0. Hence, if the Carleson measures in (42) have uniformly vanishing trace, then their densities must vanish uniformly, as well. This proves the desired estimate (41), and thus $w \in Db_{as}$.

Step III is therefore complete and with it the proof of Theorem 2, the Gaussian special case of the main result. The general case (Theorem 1) follows from this by Corollary 12. $\qquad\square$

References

[1] R. Bott and L. W. Tu, *Differential Forms in Algebraic Topology*, Graduate Texts in Math., vol. 82, Springer-Verlag, Berlin, Heidelberg, and New York, 1982. MR **83i:**57016
[2] S. M. Buckley, *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc. **340** (1993), 253–272. MR **94a:**42011
[3] M. Christ, *A T(b) theorem with remarks on analytic capacity and the Cauchy integral*, Colloq. Math. **61** (1990), 601–628. MR **92k:**42020
[4] R. R. Coifman and C. Fefferman, *Weighted norm inequalities for maximal functions and singular integrals*, Studia Math. **51** (1974), 241–250. MR **50:**10670
[5] C. Fefferman and E. M. Stein, *H^p spaces of several variables*, Acta Math. **129** (1972), 137–193. MR **56:**6263
[6] R. Fefferman, C. E. Kenig, and J. Pipher, *The theory of weights and the Dirichlet problem for elliptic equations*, Ann. of Math. (2) **134** (1991), 65–124. MR **93h:**31010
[7] J. García-Cuerva and J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, North-Holland, Amsterdam, New York, and Oxford, 1985. MR **87d:**42023
[8] S. V. Hruščev, *A description of weights satisfying the A_∞ condition of Muckenhoupt*, Proc. Amer. Math. Soc. **90** (1984), 253–257. MR **85k:**42049
[9] D. Jerison and C. E. Kenig, *The logarithm of the Poisson kernel for a C^1 domain has vanishing mean oscillation*, Trans. Amer. Math. Soc. **273** (1982), 781–794. MR **83k:**31004
[10] F. John and L. Nirenberg, *On functions of bounded mean oscillation*, Comm. Pure Appl. Math. **14** (1961), 415–426. MR **24:**A1348
[11] M. B. Korey, *Ideal weights: doubling and absolute continuity with asymptotically optimal bounds*, Ph.D. Thesis, University of Chicago, 1995.
[12] B. Muckenhoupt, *The equivalence of two conditions for weight functions*, Studia Math. **49** (1974), 101–106. MR **50:**2790
[13] S. C. Power, *Vanishing Carleson measures*, Bull. London Math. Soc. **12** (1980), 207–210. MR **82c:**30057
[14] H. M. Reimann and T. Rychener, *Funktionen beschränkter mittlerer Oszillation*, Lecture Notes in Math. **487**, Springer-Verlag, Berlin, Heidelberg, and New York, 1975. MR **58:**23564
[15] D. Sarason, *Functions of vanishing mean oscillation*, Trans. Amer. Math. Soc. **207** (1975), 391–405. MR **51:**13690

[16] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, Princeton, NJ, 1970. MR **44:**7280

[17] E. M. Stein, *Harmonic analysis: Real-variable methods, orthogonality, and oscillatory integrals*, Princeton Univ. Press, Princeton, NJ, 1993. MR **95c:**42002

MAX-PLANCK-ARBEITSGRUPPE "PARTIELLE DIFFERENTIALGLEICHUNGEN UND KOMPLEXE ANA-LYSIS", UNIVERSITÄT POTSDAM, 14415 POTSDAM, GERMANY

Current address: Institut für Mathematik, Universität Potsdam, 14415 Potsdam, Germany

E-mail address: `mike@mpg-ana.uni-potsdam.de`