



Weighted norm inequalities for oscillatory integrals with finite type phases on the line [☆]

Jonathan Bennett ^{*}, Samuel Harrison

*School of Mathematics, The Watson Building, University of Birmingham, Edgbaston, Birmingham, B15 2TT, England,
United Kingdom*

Received 14 July 2011; accepted 14 December 2011

Available online 12 January 2012

Communicated by Charles Fefferman

Abstract

We obtain two-weighted L^2 norm inequalities for oscillatory integral operators of convolution type on the line whose phases are of finite type. The conditions imposed on the weights involve geometrically-defined maximal functions, and the inequalities are best-possible in the sense that they imply the full $L^p(\mathbb{R}) \rightarrow L^q(\mathbb{R})$ mapping properties of the oscillatory integrals. Our results build on work of Carbery, Soria, Vargas and the first author.

© 2011 Elsevier Inc. All rights reserved.

MSC: 44B20; 42B25

Keywords: Weighted inequalities; Oscillatory integrals; Convolution

1. Introduction

Weighted norm inequalities have been the subject of intense study in harmonic analysis in recent decades. On an informal level, given a suitable operator T (which maps functions on \mathbb{R}^m to functions on \mathbb{R}^n , say) and an exponent $p \in [1, \infty)$, such inequalities typically take the form

[☆] The first author was partially supported by EPSRC grant EP/E022340/1, and the second by an EPSRC Doctoral Training Grant.

^{*} Corresponding author. Fax: +44 121 414 3389.

E-mail addresses: J.Bennett@bham.ac.uk (J. Bennett), harriss@maths.bham.ac.uk (S. Harrison).

$$\int_{\mathbb{R}^n} |Tf|^p v \leq \int_{\mathbb{R}^m} |f|^p u, \tag{1}$$

where the weights u and v are certain nonnegative functions on \mathbb{R}^m and \mathbb{R}^n respectively. In such a context the general goal is to understand geometrically the pairs of weights u and v for which (1) holds for all admissible inputs f . As may be expected, sufficient conditions on u, v are often of the form $Mv \leq u$ for some appropriate maximal operator M capturing certain geometric characteristics of T . Under such circumstances a simple duality argument generally allows (1) to transfer bounds on M to bounds on T . For example, if $q, \tilde{q} \geq p$ then

$$\begin{aligned} \|Tf\|_{L^q(\mathbb{R}^n)} &= \sup_{\|v\|_{(q/p)'}=1} \left(\int_{\mathbb{R}^n} |Tf|^p v \right)^{1/p} \\ &\leq \sup_{\|v\|_{(q/p)'}=1} \left(\int_{\mathbb{R}^n} |f|^p \mathcal{M}v \right)^{1/p} \\ &\leq \sup_{\|v\|_{(q/p)'}=1} \|\mathcal{M}v\|_{L^{(\tilde{q}/p)'(\mathbb{R}^m)}}^{1/p} \|f\|_{L^{\tilde{q}}(\mathbb{R}^m)}, \end{aligned} \tag{2}$$

and so $\|T\|_{\tilde{q} \rightarrow q} \leq \|\mathcal{M}\|_{(q/p)' \rightarrow (\tilde{q}/p)'}$. Thus given such an operator T and an index p , it is of particular interest to identify a corresponding geometrically defined maximal operator \mathcal{M} which is *optimal* in the sense that all “interesting” $L^q \rightarrow L^{\tilde{q}}$ bounds for T may be obtained from those of \mathcal{M} in this way.¹

This endeavour has been enormously successful for broad classes of important operators T in euclidean harmonic analysis, such as maximal averaging operators, fractional integral operators, Calderón–Zygmund singular integral operators and square functions. For example if T is a standard Calderón–Zygmund singular integral operator, such as the Hilbert transform on the line, Córdoba and Fefferman [9] showed that for any $p, r > 1$, there is a constant $C_{p,r}$ for which

$$\int_{\mathbb{R}^n} |Tf|^p w \leq C_{p,r} \int_{\mathbb{R}^n} |f|^p M_r w.$$

Here $M_r w := (Mw^r)^{1/r}$, where M denotes the Hardy–Littlewood maximal function on \mathbb{R}^n .² The modern description of such inequalities belongs to the theory of A_p weights, and may be found in many sources; see for example [21,10,11].

It is pertinent at this stage to mention one further particular example of this perspective. Let I_α denote the fractional integral operator of order α on \mathbb{R} , defined via the Fourier transform by $\widehat{I_\alpha f}(\xi) = |\xi|^{-\alpha} \widehat{f}(\xi)$. In [17] Pérez showed that for any $0 < \alpha < 1$, there is a constant C_α for which

¹ One might interpret “interesting” here as those which generate the full mapping properties of T by duality considerations and interpolation with elementary inequalities.

² There are improvements of the above inequality due to Wilson [22] and Pérez [16]. We shall appeal to these in Section 4.

$$\int_{\mathbb{R}} |I_\alpha f|^2 w \leq C_\alpha \int_{\mathbb{R}} |f|^2 M_{2\alpha} M^2 w, \tag{3}$$

where now

$$M_\alpha w(x) = \sup_{r>0} \frac{1}{r^{1-\alpha}} \int_{|x-y|<r} w(y) dy$$

is a certain fractional Hardy–Littlewood maximal function and $M^2 = M \circ M$ denotes the composition of the Hardy–Littlewood maximal function M with itself.³ From the current perspective the factors of M in the maximal operator $M_\alpha M^2$ are of secondary importance since M_α and $M_\alpha M^2$ share the same $L^p \rightarrow L^q$ mapping properties provided $1 < p \leq \infty$. This follows from the $L^p \rightarrow L^p$ boundedness of M for $1 < p \leq \infty$.

It is of course natural to seek such weighted inequalities for other classes of operators which occupy a central place in harmonic analysis. Perhaps the most apparent context would be that of *oscillatory* integral operators, and this has indeed received some notable attention in the literature. In particular, in the Proceedings of the 1978 Williamstown Conference on Harmonic Analysis, Stein [20] raised the possibility that the disc multiplier operator or Bochner–Riesz multiplier operators may be controlled by Kakeya or Nikodym type maximal functions via weighted L^2 inequalities of the above general form. Although this question posed by Stein (which is sometimes referred to as Stein’s Conjecture) remains largely unsolved in all dimensions, it has generated a number of results of a similar nature (see for example [4,6,7,2,5,3,1,13]).

The aim of this paper is to establish a general result providing weighted norm inequalities of the form (1) for oscillatory integral operators of convolution type on the line with phases of finite type. Our approach builds on work of Carbery, Soria, Vargas and the first author [3].

2. Statement of results

Let $\ell \in \mathbb{N}$ satisfy $\ell \geq 2$ and $x_0 \in \mathbb{R}$ be given. Suppose that $\phi : \mathbb{R} \rightarrow \mathbb{R}$ is a smooth phase function satisfying the finite type condition

$$\phi^{(k)}(x_0) = 0 \text{ for } 2 \leq k < \ell, \quad \text{and} \quad \phi^{(\ell)}(x_0) \neq 0. \tag{4}$$

As we shall clarify in Section 3, this condition ensures that ϕ is close to the model phase $x \mapsto a + bx + c(x - x_0)^\ell$ in a neighbourhood of x_0 . Here a, b, c are real numbers.

For each $\lambda > 0$ define a convolution kernel $K_\lambda : \mathbb{R} \rightarrow \mathbb{C}$ by

$$K_\lambda(x) = e^{i\lambda\phi(x)} \psi(x), \tag{5}$$

where ψ is a smooth cutoff function supported in a small neighbourhood U of x_0 . Throughout we shall suppose that U is sufficiently small so that $\phi^{(\ell)}$ is bounded below by a positive constant on U . We define the oscillatory integral operator T_λ by

³ In [17] a similar statement is proved in all dimensions and for exponents $1 < p < \infty$.

$$T_\lambda f(x) = K_\lambda * f(x) = \int_{\mathbb{R}} e^{i\lambda\phi(x-y)} \psi(x-y) f(y) dy.$$

Our main result is the following:

Theorem 2.1. *There exists a constant $C > 0$ such that for all weight functions w and $\lambda \geq 1$,*

$$\int_{\mathbb{R}} |T_\lambda f(x)|^2 w(x) dx \leq C \int_{\mathbb{R}} |f(x)|^2 M^2 \mathcal{M}_{\ell,\lambda} M^4 w(x - x_0) dx \tag{6}$$

where M^k denotes the k -fold composition of the Hardy–Littlewood maximal function M , and $\mathcal{M}_{\ell,\lambda}$ is given by

$$\mathcal{M}_{\ell,\lambda} w(x) = \sup_{(y,r) \in \Gamma_{\ell,\lambda}(x)} \frac{1}{(\lambda r)^{\frac{1}{\ell-1}}} \int_{|y-y'| \leq r} w(y') dy'$$

where $\Gamma_{\ell,\lambda}(x)$ is the region

$$\{(y, r): \lambda^{-1} < r \leq \lambda^{-\frac{1}{\ell}}, \text{ and } |y - x| \leq (\lambda r)^{-\frac{1}{\ell-1}}\}.$$

The remainder of this section consists of a number of remarks on the context and implications of the above theorem, followed by an informal description of the ideas behind our proof.

First of all it should be noted that by translation-invariance it suffices to prove (6) with $x_0 = 0$.

We will actually prove a uniform version of the above theorem under the following quantified version of the hypothesis (4): Suppose that $\epsilon > 0$ and $(B_j)_{j \in \mathbb{N}}$ is a sequence of positive constants. In addition to (4), suppose that

$$\phi^{(k)}(x_0) = 0 \quad \text{for } 2 \leq k < \ell, \quad \text{and } \phi^{(\ell)}(x_0) \geq \epsilon, \tag{7}$$

and that

$$\|\phi^{(j)}\|_\infty \leq B_j$$

for all $j \in \mathbb{N}$. By the Mean Value Theorem, the neighbourhood U of x_0 may be chosen, depending only on ϵ and $B_{\ell+1}$, such that $\phi^{(\ell)} \geq \epsilon/2$ on U . As may be seen from the proof, the constant C in Theorem 2.1 depends only on ϵ , ℓ and finitely many of the B_j . As might be expected, such uniform versions allow one to handle multivariable phases which satisfy the hypotheses of the theorem in one scalar variable uniformly in the remaining variables. We do not elaborate on this.

By modulating the input f and output $T_\lambda f$ appropriately in (6), there is no loss of generality in strengthening the hypothesis (4) to

$$\phi^{(k)}(x_0) = 0 \quad \text{for } 0 \leq k < \ell, \quad \text{and } \phi^{(\ell)}(x_0) \neq 0. \tag{8}$$

Similarly, (7) may be replaced with

$$\phi^{(k)}(x_0) = 0 \quad \text{for } 0 \leq k < \ell, \quad \text{and } \phi^{(\ell)}(x_0) \geq \epsilon. \tag{9}$$

Notice that if ϕ satisfies the hypotheses (8) (or (9)) and χ is a local diffeomorphism in a neighbourhood of $y_0 \in \mathbb{R}$ with $\chi(y_0) = x_0$, then the phase function $\phi \circ \chi$ satisfies the hypotheses (8) (or (9)) at the point y_0 .

The maximal function $\mathcal{M}_{\ell,\lambda}$ resembles a fractional Hardy–Littlewood maximal operator corresponding to an “approach region”, somewhat reminiscent of (yet different from) the maximal operators studied by Nagel and Stein in [15]. It should be pointed out that $\mathcal{M}_{\ell,\lambda}$ actually behaves rather differently from a fractional maximal operator, exhibiting a homogeneity more suitable for $L^p \rightarrow L^p$ estimates; see the forthcoming Proposition 2.3 and the scaling identity (39) for further clarification. We note that $\mathcal{M}_{\ell,\lambda}$ is universal in the sense that it depends only on the parameters ℓ and λ , and is otherwise independent of the phase ϕ .

Since $\mathcal{M}_{\ell,\lambda}$ involves a fractional average, (6) bears some resemblance to the two-weighted inequality for the fractional integral (3) due to Pérez [17]. The root of this similarity lies in the fact that the Fourier multipliers for the operators T_λ and I_α both exhibit “power-like” decay. We remark that the factors of Hardy–Littlewood maximal function M are of secondary importance in Theorem 2.1 since for $1 < p, q \leq \infty$, $M^2 \mathcal{M}_{\ell,\lambda} M^4$ and $\mathcal{M}_{\ell,\lambda}$ share the same $L^p \rightarrow L^q$ bounds (up to absolute constants). Several of these factors of M do not appear to be essential and arise for technical reasons.

The model phase functions satisfying (4) are of course $\phi(x) = (x - x_0)^\ell$ for $\ell \geq 2$, although there are others of particular interest. For example the phase function $\phi(x) = \cos x$ (which clearly satisfies our hypotheses with $\ell = 2$ and $\ell = 3$, depending on the point x_0) arises naturally in the context of weighted inequalities for the Fourier extension operator associated with the circle \mathbb{S}^1 in \mathbb{R}^2 . The Fourier extension (or adjoint Fourier restriction) operator associated with \mathbb{S}^1 is the map $g \mapsto \widehat{g d\sigma}$, where

$$\widehat{g d\sigma}(\xi) = \int_{\mathbb{S}^1} g(x) e^{ix \cdot \xi} d\sigma(x).$$

Here σ denotes arclength measure on \mathbb{S}^1 and $g \in L^1(\mathbb{S}^1)$; thus $\widehat{g d\sigma}$ is simply the Fourier transform of the singular measure $g d\sigma$ on \mathbb{R}^2 . It should be noted that the adjoint of this map is the restriction $f \mapsto \widehat{f}|_{\mathbb{S}^1}$, where again $\widehat{}$ denotes the Fourier transform on \mathbb{R}^2 . On parametrising \mathbb{S}^1 , invoking a suitable partition of unity and applying Theorem 2.1 with $\phi(x) = \cos x$ for both $\ell = 2, 3$, one may deduce the following theorem of Carbery, Soria, Vargas and the first author.

Theorem 2.2. (See [3].) *There exists a constant $C > 0$ such that for all Borel measures μ supported on \mathbb{S}^1 and $R \geq 1$,*

$$\int_{\mathbb{S}^1} |\widehat{g d\sigma}(Rx)|^2 d\mu(x) \leq \frac{C}{R} \int_{\mathbb{S}^1} |g(\omega)|^2 M^2 \mathfrak{M}_R M^4(\mu)(\omega) d\sigma(\omega),$$

where

$$\mathfrak{M}_R \mu(\omega) = \sup_{R^{-1} \leq \alpha \leq R^{-2/3}} \frac{\mu(T(\alpha, \alpha^2 R))}{\alpha},$$

and M is the Hardy–Littlewood maximal function on \mathbb{S}^1 . Here $T(\alpha, \beta)$ denotes a rectangle in the plane, of short side α and long side β .

Our proof of Theorem 2.1 builds on that of Theorem 2.2 and is a testament to the robustness of the approach developed in [3].

The above application of Theorem 2.1 used the fact that the phase under consideration satisfied hypothesis (4) for different values of ℓ in different regions. A similar approach yields two-weighted inequalities associated with polynomial phases for example. The maximal operators that feature are linear combinations of translates of the operators $\mathcal{M}_{\ell,\lambda}$, determined by the local monomial structure of the polynomial. We do not pursue this matter further here, although it is pertinent to note that the weighted bounds on $T_\lambda f$ provided by Theorem 2.1 grow with the parameter ℓ . It is straightforward to verify that if $\ell \leq \ell'$ then $\Gamma_{\ell,\lambda}(x) \subseteq \Gamma_{\ell',\lambda}(x)$ for $x \in \mathbb{R}$. As a consequence

$$\mathcal{M}_{\ell,\lambda} w(x) \leq \mathcal{M}_{\ell',\lambda} w(x)$$

for all x .

Theorem 2.1 is sharp in the sense that the simple duality argument (2) when applied to (6) allows the interesting $L^p(\mathbb{R}) \rightarrow L^q(\mathbb{R})$ boundedness properties of the oscillatory integral operator T_λ to be deduced from those of the controlling maximal function $\mathcal{M}_{\ell,\lambda}$. For $\ell = 3$ the bounds on T_λ are already known and follow from work of Greenleaf and Seeger – [12]⁴. The central estimates are the following, from which all others may be obtained by interpolation with elementary estimates and duality.

Proposition 2.3. *For $\ell \geq 2$, there exists a constant $C > 0$ such that*

$$\|\mathcal{M}_{\ell,\lambda} f\|_{(\frac{\ell}{2})'} \leq C \lambda^{-\frac{2}{\ell}} \|f\|_{(\frac{\ell}{2})} \tag{10}$$

holds for all $f \in L^{(\frac{\ell}{2})'}(\mathbb{R})$ and $\lambda \geq 1$.

Theorem 2.1 combined with Proposition 2.3 yields the following:

Corollary 2.4. *For $\ell \geq 2$, there exists a constant $C > 0$ such that*

$$\|T_\lambda f\|_\ell \leq C \lambda^{-\frac{1}{\ell}} \|f\|_\ell, \tag{11}$$

holds for all $f \in L^\ell(\mathbb{R})$ and $\lambda \geq 1$.

As in the statement of Theorem 2.1, the constants C above depend only on ϵ , ℓ and finitely many of the B_j .

Theorem 2.1 (and thus Corollary 2.4) is only truly significant for $\ell > 2$. The case $\ell = 2$ of Theorem 2.1 may be handled by elementary methods using the local nature of the operator T_λ , Plancherel’s theorem and a simple stationary phase estimate on \widehat{K}_λ . Notice that if $\phi(x) = x^2$, then

⁴ This is only explicit for $\ell = 3$, although it is clear that their techniques extend to general ℓ in this setting.

$$e^{i\phi} * f(x) = \int_{\mathbb{R}} e^{i(x-y)^2} f(y) dy = e^{ix^2} \int_{\mathbb{R}} e^{-2ixy} (e^{iy^2} f(y)) dy,$$

and so an inequality of the form

$$\int_{\mathbb{R}} |e^{i\phi} * f(x)|^2 v(x) dx \leq C \int_{\mathbb{R}} |f(x)|^2 u(x) dx$$

is equivalent to

$$\int_{\mathbb{R}} |\widehat{f}(x)|^2 v(x) dx \leq C \int_{\mathbb{R}} |f(x)|^2 u(x) dx,$$

which is of course a weighted L^2 inequality for the Fourier transform. Such inequalities are known when u and v are certain power weights (Pitt’s inequality [18]), and generalisations thereof involving rearrangement-invariant conditions on the weights [14].

In the situation where the phase ϕ is homogeneous (that is, when $\phi(x) = x^\ell$ for some $\ell \geq 2$), a scaling and limiting argument allows one to pass from the local inequality (6) to the global inequality

$$\int_{\mathbb{R}} |e^{i(\cdot)^\ell} * f(x)|^2 dw(x) \leq C \int_{\mathbb{R}} |f(x)|^2 M^2 \widetilde{\mathcal{M}}_\ell M^4 w(x) dx, \tag{12}$$

where the maximal function $\widetilde{\mathcal{M}}_\ell$ is given by

$$\widetilde{\mathcal{M}}_\ell w(x) = \sup_{(y,r) \in \Gamma_\ell(x)} \frac{1}{r^{\frac{1}{\ell-1}}} \int_{|y-y'| \leq r} w(y') dy',$$

and

$$\Gamma_\ell(x) = \{(y, r): 0 < r \leq 1, \text{ and } |y - x| \leq r^{-\frac{1}{\ell-1}}\}.$$

Notice that (12) is only significant for $\ell > 2$, since $\widetilde{\mathcal{M}}_2 w \equiv 2\|w\|_\infty$.

The ideas behind the proof of Theorem 2.1. Our proof of Theorem 2.1 relies heavily on the convolution structure of the operator T_λ , with the Fourier transform playing a central role. Our strategy will involve decomposing the support of the Fourier transform of the input function f . It is therefore appropriate that we begin with the following elementary observation, which is a simple manifestation of the uncertainty principle.

Observation 2.5. Suppose $f \in L^1(\mathbb{R})$ is such that the support of \widehat{f} is contained in a bounded subset $I \subset \mathbb{R}$, and choose $\Psi \in \mathcal{S}(\mathbb{R})$ such that $\widehat{\Psi}(\xi) = 1$ for $\xi \in I$. Then

$$\int |T_\lambda f|^2 w \leq \|\Psi\|_1 \int |T_\lambda f|^2 |\Psi| * w \tag{13}$$

and

$$\int |T_\lambda f|^2 w \leq \|T_\lambda \Psi\|_1 \int |f|^2 |T_\lambda \Psi| * w. \quad (14)$$

Observation 2.5 is a simple consequence of the identities

$$T_\lambda f = \Psi * (T_\lambda f) = (T_\lambda \Psi) * f$$

combined with applications of the Cauchy–Schwarz inequality and Fubini’s theorem.

Since we wish to use (14) to prove (6), there are two questions that we must address:

- (i) How do we decompose frequency space so that the resulting functions $|T_\lambda \Psi|$ have a clear geometric interpretation?
- (ii) How do we then find sufficient (almost) orthogonality on $L^2(w)$ to allow us to put the pieces of the decomposition back together again?

The frequency decomposition that we employ comes in two stages. The first stage involves the use of fairly classical Littlewood–Paley theory to reduce to the situation where the support of \widehat{f} is contained in a dyadic interval. This is natural as the multiplier $\widehat{K}_\lambda(\xi)$ has power-like decay as $|\xi| \rightarrow \infty$. This Fourier support restriction allows us to mollify the weight function w via (13), and accounts for the integration in the definition of the maximal operator $\mathcal{M}_{\ell,\lambda}$. Since we wish to ultimately apply (14), it is necessary for us to decompose the support of \widehat{f} further, this time into intervals of equal length. A stationary phase argument then reveals that provided the scale of this finer frequency decomposition is sufficiently small, the corresponding objects $|T_\lambda \Psi|$ have a clear geometric interpretation (satisfying estimates similar to those of Ψ), allowing us to appeal to (14).

This analysis provides us with weighted norm inequalities for T_λ acting on functions f with very particular Fourier supports. However, a decomposition of a general function f (with unrestricted Fourier support) into such functions will not in general exhibit any (almost) orthogonality on $L^2(w)$. In order to overcome this obstacle we find an efficient way to dominate the weight w by a further weight w' which is sufficiently smooth for our decomposition to be almost orthogonal on $L^2(w')$. The process by which we pass to this larger weight involves a local supremum, and accounts for the presence of the approach region $\Gamma_{\ell,\lambda}$ in the definition of $\mathcal{M}_{\ell,\lambda}$.

This paper is organised as follows. In Section 3 we make some simple reductions and observations pertaining to the class of phase functions ϕ that we consider. In Section 4 we establish the Littlewood–Paley theory that we shall need in the proof of Theorem 2.1. The proof of Theorem 2.1 is provided in Sections 5, 6 and 7, and the proof of Proposition 2.3 in Section 8.

3. Properties of the phase

In this section we recall and further develop the properties of the phase ϕ introduced in Section 2.

Let $\epsilon > 0$ and let $(B_j)_{j=0}^\infty$ be a sequence of positive real numbers. An elementary calculation reveals that there is a constant C_ϵ , depending on at most ϵ such that the pointwise inequality

$$\mathcal{M}_{\ell,\epsilon\lambda} w \leq C_\epsilon \mathcal{M}_{\ell,\lambda} w$$

holds for all weight functions w . As a result we may assume without loss of generality that $\epsilon = 1$ in (9).

Assuming, as we may, that $x_0 = 0$, the hypothesis (9) on the phase ϕ becomes

$$\phi^{(k)}(0) = 0 \quad \text{for } 0 \leq k < \ell, \quad \text{and} \quad \phi^{(\ell)}(0) \geq 1. \tag{15}$$

For uniformity purposes we also assume that

$$\|\phi^{(j)}\|_\infty \leq B_j \tag{16}$$

for each $j \in \mathbb{N}_0$. By the mean value theorem we may of course choose a neighbourhood U of the origin, depending only on $B_{\ell+1}$, such that $\phi^{(\ell)} \geq 1/2$ on U , and insist that the cutoff function ψ in (5) is supported in U .

Our final observation clarifies the sense in which $x \mapsto x^\ell$ is a model for the phase function ϕ . If $0 \leq k \leq \ell - 1$, then by Taylor’s theorem, for each fixed x we have

$$\begin{aligned} \phi^{(k)}(x) &= \phi^{(k)}(0) + x\phi^{(k+1)}(0) + \dots + x^{\ell-k}\phi^{(\ell)}(y_{x,k}) \\ &= x^{\ell-k}\phi^{(\ell)}(y_{x,k}) \end{aligned}$$

for some $y_{x,k} \in (0, x)$. Since $1/2 \leq |\phi^{(\ell)}| \leq B_\ell$ on the support of ψ , we have

$$\frac{1}{2}|x|^{\ell-k} \leq |\phi^{(k)}(x)| \leq B_\ell|x|^{\ell-k} \tag{17}$$

for all x in the support of ψ and $0 \leq k \leq \ell - 1$.

Notation. Throughout this paper we shall write $A \lesssim B$ if there exists a constant c , possibly depending on finitely many of the B_j , such that $A \leq cB$. In particular, this constant will always be independent of λ , the function f and weight function w .

4. Some weighted Littlewood–Paley theory

In this section we collect together the weighted inequalities for Littlewood–Paley square functions that we will appeal to in our proof of Theorem 2.1. Some of these results, although very classical in nature, do not appear to be readily available in the literature, and hence may be of some independent interest.

4.1. Weighted inequalities for a dyadic square function

Let $Q : \mathbb{R} \rightarrow \mathbb{R}$ be a smooth function whose Fourier transform vanishes outside $[-3, -\frac{3}{4}] \cup [\frac{3}{4}, 3]$, and such that

$$\sum_{k \in \mathbb{Z}} \widehat{Q}(2^{-k}x) = 1$$

for all $x \neq 0$. For each $k \in \mathbb{Z}$ let the operator Δ_k be given by

$$\widehat{\Delta_k f}(\xi) = \widehat{Q}(2^{-k}\xi)\widehat{f}(\xi),$$

and define the square function S by

$$Sf(x) = \left(\sum_k |\Delta_k f(x)|^2 \right)^{1/2}.$$

Proposition 4.1. *For each weight w on \mathbb{R} ,*

$$\int_{\mathbb{R}} (Sf)^2 w \lesssim \int_{\mathbb{R}} |f|^2 Mw \tag{18}$$

and

$$\int_{\mathbb{R}} |f|^2 w \lesssim \int_{\mathbb{R}} (Sf)^2 M^3 w, \tag{19}$$

where M denotes the Hardy–Littlewood maximal function.

Proof. The forward inequality (18) is a straightforward consequence of a more general result of Wilson [23].

The reverse inequality (19) may be reduced to a well-known weighted inequality for Calderón–Zygmund singular integrals due to Wilson [22] and Pérez [16] as follows. For each $j = 0, 1, 2, 3$ let

$$T_j = \sum_{k \in 4\mathbb{Z} + \{j\}} \Delta_k,$$

so that $f = T_0 f + T_1 f + T_2 f + T_3 f$.

Let $\epsilon = (\epsilon_k)$ be a random sequence with $\epsilon_k \in \{-1, 1\}$ for each $j \in \mathbb{Z}$, and define

$$T_j^\epsilon = \sum_{k \in 4\mathbb{Z} + \{j\}} \epsilon_k \Delta_k$$

for each $j = 0, 1, 2, 3$.

Now let $Q' : \mathbb{R} \rightarrow \mathbb{R}$ be a smooth function such that \widehat{Q}' is equal to 1 on the support of \widehat{Q} and vanishing outside $[-4, -\frac{1}{2}] \cup [\frac{1}{2}, 4]$, and define the operator Δ'_k by

$$\widehat{\Delta'_k f}(\xi) = \widehat{Q}'(2^{-k}\xi) \widehat{f}(\xi).$$

Observe that if we set

$$S_j^\epsilon = \sum_{k \in 4\mathbb{Z} + \{j\}} \epsilon_k \Delta'_k,$$

then

$$S_j^\epsilon T_j^\epsilon = T_j,$$

since for each $j = 0, 1, 2, 3$, the supports of \widehat{Q} and \widehat{Q}' ensure that $\Delta'_{k'} \Delta_k = 0$ whenever k' and k are distinct in $4\mathbb{Z} + \{j\}$. Hence

$$\int_{\mathbb{R}} |f|^2 w \lesssim \sum_{j=0}^3 \int_{\mathbb{R}} |T_j f|^2 w = \sum_{j=0}^3 \int_{\mathbb{R}} |S_j^\epsilon T_j^\epsilon f|^2 w.$$

It is well known that the operator S_j^ϵ is a standard Calderón–Zygmund singular integral operator uniformly in the sequence ϵ , and so by [22] and [16] we have that for each $1 < p < \infty$,

$$\int_{\mathbb{R}} |S_j^\epsilon f|^p w \lesssim \int_{\mathbb{R}} |f|^p M^{[p]+1} w$$

uniformly in ϵ and $j = 0, 1, 2, 3$. Thus in particular

$$\int_{\mathbb{R}} |f|^2 w \lesssim \sum_{j=0}^3 \int_{\mathbb{R}} |T_j^\epsilon f|^2 M^3 w$$

uniformly in ϵ . Inequality (19) now follows on taking expectations and using Khinchine’s inequality. \square

We remark that the above proof also yields the weighted L^p inequality

$$\int_{\mathbb{R}} |f|^p w \lesssim \int_{\mathbb{R}} (Sf)^p M^{[p]+1} w \tag{20}$$

for all $1 < p < \infty$. It would be interesting to determine whether the power $[p] + 1$ of the Hardy–Littlewood maximal operator may be reduced here.

4.2. Weighted inequalities for an “equally-spaced” square function

The following result concerns a square function adapted to an equally-spaced frequency decomposition. Very similar results may be found in several places in the literature, including [8, 19] and [3]. We refer the reader to [3] for a variant on the circle, whose proof readily adapts to the current setting.

Proposition 4.2. *For $L > 0$, let W_L be a function on \mathbb{R} with $\text{supp } \widehat{W}_L \subset \{x \in \mathbb{R}: |x| \leq 2L\}$, such that*

$$\sum_{k \in \mathbb{Z}} \widehat{W}_L(x - kL) = 1$$

for all $x \in \mathbb{R}$. Suppose further that for each $N \in \mathbb{N}$,

$$|W_L(x)| \lesssim \frac{L}{(1 + L|x|)^N}$$

for all $x \in \mathbb{R}$. For a function f on \mathbb{R} and $k \in \mathbb{Z}$, define f_k by $\widehat{f}_k(\xi) = \widehat{f}(\xi)\widehat{W}_L(\xi - kL)$. Then for any weight function w on \mathbb{R} ,

$$\int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} |f_k|^2 w \lesssim \int_{\mathbb{R}} |f|^2 |W_L| * w.$$

5. An initial Littlewood–Paley reduction

In this section we use the dyadic Littlewood–Paley theory from the previous section to reduce the proof of Theorem 2.1 to a weighted inequality for functions f with restricted Fourier support. The nature of the frequency restrictions is motivated by the following estimates relating to the Fourier transform of K_λ .

Lemma 5.1.

$$\sup_{y \in \mathbb{R}} \left| \int_{-\infty}^y K_\lambda(x) e^{-ix\xi} dx \right| \lesssim \begin{cases} \lambda^{-\frac{1}{\ell}}, & |\xi| \leq \lambda^{\frac{1}{\ell}}, \\ \lambda^{-\frac{1}{2(\ell-1)}} |\xi|^{-\frac{\ell-2}{2(\ell-1)}}, & |\xi| > \lambda^{\frac{1}{\ell}}. \end{cases} \tag{21}$$

Moreover, for each $k, N \in \mathbb{N}$,

$$\left| \left(\frac{d}{d\xi} \right)^k \widehat{K}_\lambda(\xi) \right| \lesssim |\xi|^{-N} \tag{22}$$

for all $|\xi| \geq 2B_1\lambda$. The implicit constants above depend on k, N, ℓ and finitely many of the B_j .

Proof. The two estimates (21) follow from corresponding estimates on the integral

$$\int_I e^{i(\lambda\phi(x) - x\xi)} dx$$

that are uniform in I , where I is an interval contained in the support of ψ . This shall be achieved by routine applications of van der Corput’s lemma.

Writing $h(x) = \lambda\phi(x) - x\xi$ we see that $h^{(\ell)}(x) = \lambda\phi^{(\ell)}(x) \geq \lambda/2$ for all $x \in I$ and so

$$\left| \int_I e^{i(\lambda\phi(x) - x\xi)} dx \right| \lesssim \lambda^{-\frac{1}{\ell}}$$

for all $\xi \in \mathbb{R}$, with implicit constant depending only on ℓ .

For the second estimate, let $I' = \{x \in I : |x| \lesssim |\xi/\lambda|^{\frac{1}{\ell-1}}\}$, with suitably small implicit constant depending on B_ℓ . By (17), $|h'(x)| \gtrsim |\xi|$ for all $x \in I'$, and so

$$\left| \int_{I'} e^{i(\lambda\phi(x) - x\xi)} dx \right| \lesssim |\xi|^{-1}.$$

For $x \in I \setminus I'$ we have $|h''(x)| \gtrsim \lambda^{\frac{1}{\ell-1}} \xi^{\frac{\ell-2}{\ell-1}}$, so that

$$\left| \int_{I \setminus I'} e^{i(\lambda\phi(x)-x\xi)} dx \right| \lesssim \lambda^{-\frac{1}{2(\ell-1)}} |\xi|^{-\frac{\ell-2}{2(\ell-1)}}.$$

Now, if $\lambda^{1/\ell} \leq |\xi|$ we have $|\xi|^{-1} \leq \lambda^{-\frac{1}{2(\ell-1)}} |\xi|^{-\frac{\ell-2}{2(\ell-1)}}$, and so the second estimate in (21) is complete.

In order to establish (22) observe that if $|\xi| \geq 2B_1\lambda$ then the phase h has no stationary points and moreover $|h'(x)| \gtrsim |\xi|$ uniformly in x . Inequality (22) now follows by repeated integration by parts. \square

Given Lemma 5.1 it is natural to define the sets $(\mathcal{A}_p)_{p=0}^\infty$ by

$$\mathcal{A}_0 = \{ \xi \in \mathbb{R} : |\xi| \leq \lambda^{1/\ell} \}$$

and

$$\mathcal{A}_p = \{ \xi \in \mathbb{R} : 2^{p-3}\lambda^{1/\ell} < |\xi| \leq 2^{p+1}\lambda^{1/\ell} \}$$

for $p \geq 1$. By construction the sets \mathcal{A}_p cover \mathbb{R} with bounded multiplicity.

Proposition 5.2. *If the support of \widehat{f} is contained in \mathcal{A}_p then for all weight functions w and $\lambda \geq 1$,*

$$\int_{\mathbb{R}} |T_\lambda f(x)|^2 w(x) dx \lesssim \int_{\mathbb{R}} |f(x)|^2 M\mathcal{M}_{\ell,\lambda} M w(x) dx \tag{23}$$

uniformly in p , where $\mathcal{M}_{\ell,\lambda}$ is as in the statement of Theorem 2.1.

Before we come to the proof of Proposition 5.2 we show how it implies Theorem 2.1.

Suppose $f : \mathbb{R} \rightarrow \mathbb{C}$ has unrestricted Fourier support and observe that for each $k \in \mathbb{Z}$ the Fourier transform of $\Delta_k f$ is supported in \mathcal{A}_p for some p . Thus, by Proposition 4.1, Proposition 5.2, followed by Proposition 4.1 again,

$$\begin{aligned} \int_{\mathbb{R}} |T_\lambda f(x)|^2 w(x) dx &\lesssim \int_{\mathbb{R}} |ST_\lambda f(x)|^2 M^3 w(x) dx \\ &= \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} |\Delta_k T_\lambda f|^2 M^3 w(x) dx \\ &= \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} |T_\lambda \Delta_j f|^2 M^3 w(x) dx \\ &\lesssim \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} |\Delta_j f|^2 M\mathcal{M}_{\ell,\lambda} M^4 w(x) dx \end{aligned}$$

$$\begin{aligned}
 &= \int_{\mathbb{R}} (Sf)^2 M \mathcal{M}_{\ell, \lambda} M^4 w(x) dx \\
 &\lesssim \int_{\mathbb{R}} |f|^2 M^2 \mathcal{M}_{\ell, \lambda} M^4 w(x) dx,
 \end{aligned}$$

which is the conclusion of Theorem 2.1.

It thus remains to prove Proposition 5.2. By the definition of the sets \mathcal{A}_p and Lemma 5.1 we have that

$$|\widehat{K}_\lambda(\xi)| \lesssim \lambda^{-\frac{1}{\ell}} 2^{-\frac{p(\ell-2)}{2(\ell-1)}}$$

for all $\xi \in \mathcal{A}_p$ with $1 \leq 2^p < 4B_1\lambda^{(\ell-1)/\ell}$, and for each $N \in \mathbb{N}$,

$$|\widehat{K}_\lambda(\xi)| \lesssim |\xi|^{-N} \tag{24}$$

for all $\xi \in \mathcal{A}_p$ with $2^p \geq 4B_1\lambda^{(\ell-1)/\ell}$. We therefore divide the proof of Proposition 5.2 into two cases. Section 6 is devoted to the more interesting case $1 \leq 2^p < 4B_1\lambda^{(\ell-1)/\ell}$, while Section 7 handles the “error terms” corresponding to the case $2^p \geq 4B_1\lambda^{(\ell-1)/\ell}$.

6. The proof of Proposition 5.2 for $1 \leq 2^p < 4B_1\lambda^{(\ell-1)/\ell}$

Suppose that the function f has frequencies support in \mathcal{A}_p for some p with $1 < 2^p < 4B_1\lambda^{(\ell-1)/\ell}$; we shall deal with the case $p = 0$ separately. Now, if we choose a nonnegative $\Phi \in \mathcal{S}(\mathbb{R})$ such that $\widehat{\Phi}(\xi) = 1$ whenever $|\xi| \lesssim 1$, and define $\Phi_{2^p\lambda^{1/\ell}} \in \mathcal{S}(\mathbb{R})$ by $\widehat{\Phi}_{2^p\lambda^{1/\ell}}(\xi) = \widehat{\Phi}(2^{-p}\lambda^{-1/\ell}\xi)$, then $\widehat{\Phi}_{2^p\lambda^{1/\ell}}(\xi) = 1$ for all $\xi \in \mathcal{A}_p$. Hence by (13) of Observation 2.5,

$$\int_{\mathbb{R}} |T_\lambda f(x)|^2 w(x) dx \leq \int_{\mathbb{R}} |T_\lambda f(x)|^2 w_1(x) dx, \tag{25}$$

where $w_1 := \Phi_{2^p\lambda^{1/\ell}} * w$.

For general p we have no clear geometric control of the function $|T_\lambda \Phi_{2^p\lambda^{1/\ell}}|$, and so are not in a position to meaningfully apply (14) of Observation 2.5 at this stage. We proceed by performing a further frequency decomposition of the function f at a scale $0 < L \lesssim 2^p\lambda^{1/\ell}$ to be specified later.

Let $W \in \mathcal{S}(\mathbb{R})$ be such that \widehat{W} is supported in $[-2, 2]$ and

$$\sum_{k \in \mathbb{Z}} \widehat{W}(\xi - k) = 1$$

for all $\xi \in \mathbb{R}$. Define $W_L \in \mathcal{S}(\mathbb{R})$ by $\widehat{W}_L(\xi) = \widehat{W}(\xi/L)$, and $W_{L,k} \in \mathcal{S}(\mathbb{R})$ by $\widehat{W}_{L,k}(\xi) = \widehat{W}_L(\xi - kL)$, so that

$$\sum_{k \in \mathbb{Z}} \widehat{W}_{L,k}(\xi) = 1$$

for all $\xi \in \mathbb{R}$. Hence if $f_k = W_{L,k} * f$ then

$$f = \sum_{k \in \mathbb{Z}} f_k,$$

and so by the linearity of T_λ ,

$$T_\lambda f = \sum_{k \in \mathbb{Z}} T_\lambda f_k. \tag{26}$$

We note that since $\text{supp } \widehat{f} \subseteq \mathcal{A}_p$, the only nonzero contributions to the above sum occur when $|k| \sim 2^p \lambda^{1/\ell} / L$.

Unfortunately the decomposition (26) does not in general exhibit any (almost) orthogonality on $L^2(w_1)$ as the Fourier support of w_1 is too large. We thus seek an efficient way of dominating w_1 by a further weight which does have a sufficiently small Fourier support. This we achieve in two stages; the first involving a local supremum, and the second a carefully considered mollification. For each $x \in \mathbb{R}$ let

$$w_2(x) = \sup_{|x' - x| \leq (4B_1)^{-\frac{1}{\ell-1}} / L} w_1(x').$$

The factor of $4B_1$ appearing in the definition of w_2 is included for technical reasons that will become clear later.

Let $\Theta \in \mathcal{S}(\mathbb{R})$ be a nonnegative function whose Fourier transform is nonnegative and supported in $[-1, 1]$. Now let

$$w_3 = \Theta_L * w_2$$

where $\Theta_L \in \mathcal{S}(\mathbb{R})$ is defined by $\widehat{\Theta}_L(\xi) = \widehat{\Theta}(\xi/L)$. By construction w_3 has Fourier support in $[-L, L]$.

Lemma 6.1. $w_1 \leq w_2 \lesssim w_3$.

Proof. The first inequality is trivial and so we focus on the second. Observe that since $\widehat{\Theta}$ is nonnegative $\Theta(0) > 0$ and so by continuity there exists an absolute constant $0 < c \leq 1$ such that $\Theta(x) \gtrsim 1$ whenever $|x| \leq c$. Thus $\Theta_L(x) \gtrsim L$ whenever $|x| \leq c/L$. Consequently

$$\begin{aligned} w_3(x) = \Theta_L * w_2(x) &\gtrsim L \int_{|x'| \leq c/L} w_2(x - x') dx' \\ &\geq L \int_{|x'| \leq \tilde{c}/L} w_2(x - x') dx' \end{aligned}$$

where $\tilde{c} = \min\{c, (4B_1)^{-1/(\ell-1)}\}$. By the definition of w_2 and elementary considerations, either

$$w_2(x - x') \geq w_2(x) \quad \text{for all } -(4B_1)^{-\frac{1}{\ell-1}} / L \leq x' \leq 0,$$

or

$$w_2(x - x') \geq w_2(x) \quad \text{for all } 0 \leq x' \leq (4B_1)^{-\frac{1}{\ell-1}}/L,$$

and so $w_3(x) \gtrsim w_2(x)$ with implicit constant depending only on B_1 and ℓ . \square

We now see that the decomposition (26) is almost orthogonal in the smaller $L^2(w_3)$. By (25), Lemma 6.1, (26) and Parseval’s identity we have

$$\begin{aligned} \int_{\mathbb{R}} |T_\lambda f|^2 w &\leq \int_{\mathbb{R}} |T_\lambda f|^2 w_3 \\ &= \int_{\mathbb{R}} \left| \sum_{k \in \mathbb{Z}} T_\lambda f_k \right|^2 w_3 \\ &= \int_{\mathbb{R}} \sum_{k, k' \in \mathbb{R}} T_\lambda f_k \overline{T_\lambda f_{k'}} w_3 \\ &= \int_{\mathbb{R}} \sum_{k, k' \in \mathbb{R}} \widehat{T_\lambda f_k} * \widehat{T_\lambda f_{k'}} \widehat{w}_3. \end{aligned}$$

Since $\text{supp } \widehat{f_k} \subset [(k - 2)L, (k + 2)L]$ we have

$$\text{supp } \widehat{T_\lambda f_k} * \widehat{T_\lambda f_{k'}} \subseteq [(k - k' - 4)L, (k - k' + 4)L],$$

and so

$$\widehat{T_\lambda f_k} * \widehat{T_\lambda f_{k'}} \widehat{w}_3 \equiv 0$$

whenever $|k - k'| \geq 6$. Thus by the Cauchy–Schwarz inequality we have

$$\begin{aligned} \int_{\mathbb{R}} |T_\lambda f|^2 w &\lesssim \int_{\mathbb{R}} \sum_{k, k' \in \mathbb{R}: |k-k'| < 6} T_\lambda f_k \overline{T_\lambda f_{k'}} w_3 \\ &\lesssim \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} |T_\lambda f_k|^2 w_3. \end{aligned} \tag{27}$$

Now let $\psi \in \mathcal{S}(\mathbb{R})$ be such that

$$\widehat{\psi}(\xi) = \begin{cases} 1 & \text{if } |\xi| \leq 2, \\ 0 & \text{if } |\xi| \geq 4, \end{cases}$$

and define $\psi_L, \psi_{L,k} \in \mathcal{S}(\mathbb{R})$ as we did W_L and $W_{L,k}$ previously. Since $\widehat{\psi}_{L,k}(\xi) = 1$ when $\xi \in \text{supp } \widehat{f_k}$, by (27) followed by (14) of Observation 2.5 we have

$$\int_{\mathbb{R}} |T_\lambda f|^2 w \lesssim \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} |T_\lambda f_k|^2 w_3 \leq \sum_{k \in \mathbb{Z}} \|T_\lambda \Psi_{\ell, \lambda}\|_1 \int_{\mathbb{R}} |f|^2 |T_\lambda \Psi_{\ell, \lambda}| * w_3. \tag{28}$$

Our next lemma tells us that provided the scale L is small enough, the functions $|T_\lambda \Psi_{L, k}|$ satisfy estimates similar to those of Ψ_L uniformly in k .

Lemma 6.2. *If $L = 2^{-p/(\ell-1)} \lambda^{1/\ell}$ and $|k| \sim 2^p \lambda^{1/\ell} / L$ then*

$$|T_\lambda \Psi_{L, k}(x)| \lesssim \lambda^{-\frac{1}{\ell}} 2^{-\frac{p(\ell-2)}{2(\ell-1)}} H_L(x)$$

where for each $N \in \mathbb{N}$, the function H_L satisfies

$$H_L(x) \lesssim \frac{L}{(1 + L|x|)^N}$$

for all $x \in \mathbb{R}$. The implicit constants depend on at most ℓ , and finitely many of the B_j .

Proof. For $k \in \mathbb{Z}$ and $y \in \mathbb{R}$ let $h_k(y) = \phi(y) - \lambda^{-1} k L y$ and observe that

$$T_\lambda \Psi_{L, k}(x) = e^{ikLx} \int_{\mathbb{R}} e^{i\lambda h_k(y)} \psi(y) \Psi_L(x - y) dy.$$

Now, y_0 is a stationary point of the phase h_k precisely when $\phi'(y_0) = \lambda^{-1} k L$. However, since $|k| \sim 2^p \lambda^{1/\ell} / L$ and $L = 2^p \lambda^{1/\ell}$ we have

$$\lambda^{-1} |k| L \sim 1 / L^{\ell-1}, \tag{29}$$

and since $|\phi'(y)| \sim |y|^{\ell-1}$ (by (17)), we have $|y_0| \sim 1/L$. As usual the implicit constants depend only on ℓ and finitely many of the B_j . We note that since $\phi^{(\ell)} \neq 0$, there are at most $\ell - 1$ such stationary points y_0 in the support of ψ .

Let $(\eta_j)_{j=0}^\infty$ be a smooth partition of unity on \mathbb{R} with $\text{supp } \eta_j \subset \{x \in \mathbb{R}: 2^{j-1} \leq |x| \leq 2^{j+1}\}$ for $j \geq 1$ and $\text{supp } \eta_0 \subset \{x \in \mathbb{R}: |x| \leq 2\}$. For uniformity purposes we suppose that (η_j) is constructed in the standard way from a fixed smooth bump function and taking differences. Define $(\eta_{L, j})_{j=0}^\infty$ by $\eta_{L, j}(x) = \eta_j(Lx)$. Clearly $(\eta_{L, j})_{j=0}^\infty$ forms a partition of unity on \mathbb{R} with $\text{supp } \eta_{L, j} \subset \{x \in \mathbb{R}: |x| \sim 2^j / L\}$ for $j \geq 1$ and $\text{supp } \eta_{L, 0} \subset \{x \in \mathbb{R}: |x| \lesssim 1/L\}$. We may thus write

$$T_\lambda \Psi_{L, k} = \sum_{j=0}^\infty T_\lambda (\Psi_{L, k} \eta_{L, j}).$$

We now write

$$|T_\lambda \Psi_{L, k}(x)| \leq \sum_{j: 2^j \geq cL|x|} |T_\lambda (\Psi_{L, k} \eta_{L, j})(x)| + \sum_{j: 2^j < cL|x|} |T_\lambda (\Psi_{L, k} \eta_{L, j})(x)|, \tag{30}$$

where c is a positive constant depending on B_ℓ to be chosen later. We shall prove the required bounds for each of the sums in (30) separately.

Fix x and suppose that $2^j \geq cL|x|$, then integrating by parts we have

$$\begin{aligned} |T_\lambda(\Psi_{L,k}\eta_{L,j})(x)| &= \left| \int_{\mathbb{R}} \frac{d}{dy} \left(\int_{-\infty}^y e^{i(\lambda\phi(z)-kLz)} \psi(z) dz \right) \Psi_{L,k}(x-y)\eta_{L,j}(x-y) dy \right| \\ &\leq \int_{\mathbb{R}} \left| \int_{-\infty}^y e^{i(\lambda\phi(z)-kLz)} \psi(z) dz \right| \left| \frac{d}{dy} (\Psi_{L,k}(x-y)\eta_{L,j}(x-y)) \right| dy. \end{aligned}$$

By Lemma 5.1, we have the estimate

$$\left| \int_{-\infty}^y e^{i(\lambda\phi(z)-kLz)} \psi(z) dz \right| \lesssim \lambda^{-\frac{1}{\ell}} 2^{-\frac{p(\ell-2)}{2(\ell-1)}},$$

uniformly in $k \sim 2^p \lambda^{1/\ell} / L$ and y , and so

$$|T_\lambda(\Psi_{L,k}\eta_{L,j})(x)| \lesssim \lambda^{-\frac{1}{\ell}} 2^{-\frac{p(\ell-2)}{2(\ell-1)}} 2^{-Nj} L$$

for any $N \in \mathbb{N}$. Thus

$$\begin{aligned} \sum_{j: 2^j \geq cL|x|} |T_\lambda(\Psi_{L,k}\eta_{L,j})(x)| &\lesssim \lambda^{-\frac{1}{\ell}} 2^{-\frac{p(\ell-2)}{2(\ell-1)}} L \sum_{j: 2^j \geq cL|x|} 2^{-Nj} \\ &\sim \lambda^{-\frac{1}{\ell}} 2^{-\frac{p(\ell-2)}{2(\ell-1)}} \frac{L}{(1 + L|x|)^N} \end{aligned}$$

similarly uniformly.

We now suppose that $2^j < cL|x|$. Then

$$|T_\lambda(\Psi_{L,k}\eta_{L,j})(x)| \leq \lambda^{-N} \int_{\mathbb{R}} |(D_k^*)^N (\psi(y)\Psi_L(x-y)\eta_{L,j}(x-y))| dy \tag{31}$$

for any $N \in \mathbb{N}$, where the differential operator D_k^* is given by

$$D_k^* g(y) = \frac{d}{dy} \left(\frac{g(y)}{h'_k(y)} \right).$$

An elementary induction argument shows that $|(D_k^*)^N g|$ is bounded by a sum of terms (the number of which depending only on N) of the form

$$\frac{|g^{(r)}|}{|h'_k|^{ln}} \prod_{j=2}^{\ell-1} |\phi^{(j)}|^{mj}, \tag{32}$$

uniformly in k , where the indices m_j, n and r satisfy

$$(\ell - 1)n - \sum_{j=2}^{\ell-1} m_j(\ell - j) + r \leq \ell N,$$

$N \leq n \leq 2N$ and $0 \leq r \leq N$. Here only low derivatives of ϕ feature since we have used the bound $\|\phi^{(j)}\|_\infty \leq B_j$ for $j \geq \ell$.

Now, provided the constant $c > 0$ in (30) is chosen sufficiently small (depending on at most B_ℓ) then $|y| \sim |x|$ and by (29),

$$|h'_k(y)| = |\phi'(y) - \lambda^{-1}kL| \sim |y|^{\ell-1} \sim |x|^{\ell-1}$$

on the support of the integrand in (31). Recalling also that $|\phi^{(j)}(y)| \sim |y|^{\ell-j}$ for all $2 \leq j \leq \ell - 1$, the expression (32) is therefore comparable to

$$\frac{|g^{(r)}(y)||y|^{\sum_{j=2}^{\ell-1} m_j(\ell-j)}}{|y|^{(\ell-1)n}} \lesssim |g^{(r)}(y)||x|^{r-\ell N}$$

for all y in the support of the integrand in (31). Hence for $2^j < cL|x|$ and any $M \in \mathbb{N}$,

$$\begin{aligned} |T_\lambda(\Psi_{L,k\eta_{L,j}})(x)| &\lesssim \lambda^{-N}|x|^{r-\ell N} \sum_{r=0}^N \int_{\mathbb{R}} \left| \left(\frac{d}{dy} \right)^r (\psi(y)\Psi_L(x-y)\eta_{L,j}(x-y)) \right| dy \\ &\lesssim \lambda^{-N}|x|^{r-\ell N} \sum_{r=0}^N L^{r+1} 2^{-jM} 2^j / L \\ &= 2^{-(M-1)j} (\lambda^{-\frac{1}{\ell}} L)^{\ell(N-1/2)} \lambda^{-\frac{1}{\ell}} 2^{-\frac{p(\ell-2)}{2(\ell-1)}} \sum_{r=0}^N \frac{L}{(L|x|)^{\ell N-r}} \\ &\lesssim 2^{-(M-1)j} \lambda^{-\frac{1}{\ell}} 2^{-\frac{p(\ell-2)}{2(\ell-1)}} \frac{L}{(1+L|x|)^N}, \end{aligned}$$

since $\lambda^{-1/\ell}L = 2^{-p/(\ell-1)} \leq 1$. Thus

$$\sum_{j: 2^j < cL|x|} |T_\lambda(\Psi_{L,k\eta_{L,j}})(x)| \lesssim \lambda^{-\frac{1}{\ell}} 2^{-\frac{p(\ell-2)}{2(\ell-1)}} \frac{L}{(1+L|x|)^N},$$

completing the proof of the lemma. \square

If we let $w_4 = \lambda^{-\frac{2}{\ell}} 2^{-\frac{p(\ell-2)}{\ell-1}} H_L * w_3$, then by (28) and Lemma 6.2 we conclude that

$$\int_{\mathbb{R}} |T_\lambda f(x)|^2 w(x) dx \lesssim \int_{\mathbb{R}} \sum_k |f_k(x)|^2 w_4(x) dx.$$

Now on applying Lemma 4.2, our weighted estimate for T_λ becomes

$$\int_{\mathbb{R}} |T_\lambda f(x)|^2 w(x) dx \lesssim \int_{\mathbb{R}} |f(x)|^2 w_5(x) dx \tag{33}$$

where $w_5 = |W_L| * w_4$.

In order to complete the proof of Proposition 5.2 for $\text{supp } \widehat{f} \subseteq \mathcal{A}_p$ it remains to show that

$$w_5(x) \lesssim M\mathcal{M}_{\ell,\lambda} M w(x). \tag{34}$$

Since $w_5 = \lambda^{-\frac{2}{\ell}} 2^{-\frac{p(\ell-2)}{\ell-1}} |W_L| * H_L * \Theta_L * w_2$, we have

$$w_5 \lesssim \lambda^{-\frac{2}{\ell}} 2^{-\frac{p(\ell-2)}{\ell-1}} M w_2,$$

where M denotes the Hardy–Littlewood maximal function.

By translation-invariance it thus suffices to show that

$$\lambda^{-\frac{2}{\ell}} 2^{-\frac{p(\ell-2)}{\ell-1}} w_2(0) \lesssim \mathcal{M}_{\ell,\lambda} M w(0)$$

with implicit constant uniform in p . Now,

$$w_2(0) = \sup_{|y| \leq (4B_1)^{-\frac{1}{\ell-1}}/L} \Phi_{2^p \lambda^{1/\ell}} * w(y),$$

and so on setting $r = 2^{-p} \lambda^{-1/\ell}$ we obtain

$$\lambda^{-\frac{2}{\ell}} 2^{-\frac{p(\ell-2)}{\ell-1}} w_2(0) \lesssim r^{\frac{\ell-2}{\ell-1}} \lambda^{-\frac{1}{\ell-1}} \sup_{|y| \leq (4B_1 \lambda r)^{-\frac{1}{\ell-1}}} \Phi_{1/r} * w(y).$$

For each $N \in \mathbb{N}$ we may estimate

$$\begin{aligned} & r^{\frac{\ell-2}{\ell-1}} \lambda^{-\frac{1}{\ell-1}} \Phi_{1/r} * w(y) \\ & \lesssim (\lambda r)^{-\frac{1}{\ell-1}} \int_{\mathbb{R}} \frac{w(x)}{(1 + |x - y|/r)^N} dx \\ & \sim (\lambda r)^{-\frac{1}{\ell-1}} \int_{|x-y| \leq r} w(x) dx + \sum_{k=1}^{\infty} 2^{-kN} (\lambda r)^{-\frac{1}{\ell-1}} \int_{|x-y| \sim 2^k r} w(x) dx. \end{aligned} \tag{35}$$

If for $s > 0$ we define the averaging operator A_s by

$$A_s w(y) = \frac{1}{2s} \int_{|x-y| \leq s} w(x) dx,$$

then we may write

$$\begin{aligned}
 (\lambda r)^{-\frac{1}{\ell-1}} \int_{|x-y| \sim 2^k r} w(x) dx &= (\lambda r)^{-\frac{1}{\ell-1}} 2^k r A_{2^k r} w(y) \\
 &\lesssim 2^k (\lambda r)^{-\frac{1}{\ell-1}} \int_{|y-y'| \leq r} A_{2^{k+1} r} w(y') dy' \\
 &\leq 2^k (\lambda r)^{-\frac{1}{\ell-1}} \int_{|y-y'| \leq r} Mw(y') dy',
 \end{aligned}$$

since $A_{2^k r} w(y) \lesssim A_{2^{k+1} r} w(y')$ if $|y - y'| \leq r$. Hence by (35) we have

$$r^{\frac{\ell-2}{\ell-1}} \lambda^{-\frac{1}{\ell-1}} \Phi_{1/r} * w(y) \lesssim (\lambda r)^{-\frac{1}{\ell-1}} \int_{|x-y| \leq r} w(x) dx + (\lambda r)^{-\frac{1}{\ell-1}} \int_{|y-y'| \leq r} Mw(y') dy'. \tag{36}$$

Note that since $r = 2^{-p} \lambda^{-1/\ell}$ and $1 \leq 2^p \leq 4B_1 \lambda^{(\ell-1)/\ell}$ we have that $(4B_1 \lambda)^{-1} \leq r \leq \lambda^{-1/(\ell-1)}$. It will be convenient to consider separately the two cases $\lambda^{-1} \leq r \leq \lambda^{-1/(\ell-1)}$ and $(4B_1 \lambda)^{-1} \leq r \leq \lambda^{-1}$.⁵

If $\lambda^{-1} \leq r \leq \lambda^{-1/(\ell-1)}$, then $(y, r) \in \Gamma_{\ell, \lambda}(0)$ and so the first and second terms in the right hand side of (36) are dominated by $\mathcal{M}_{\ell, \lambda}(w)(0)$ and $\mathcal{M}_{\ell, \lambda}(Mw)(0)$ respectively.

If $(4B_1 \lambda)^{-1} \leq r \leq \lambda^{-1}$ then by taking $r' = \lambda^{-1}$ we have

$$(\lambda r)^{-\frac{1}{\ell-1}} \int_{|x-y| \leq r} w(x) dx \leq (4B_1)^{\frac{1}{\ell-1}} (\lambda r')^{-\frac{1}{\ell-1}} \int_{|x-y| \leq r'} w(x) dx \tag{37}$$

and

$$(\lambda r)^{-\frac{1}{\ell-1}} \int_{|y-y'| \leq r} Mw(y') dy' \leq (4B_1)^{\frac{1}{\ell-1}} (\lambda r')^{-\frac{1}{\ell-1}} \int_{|y-y'| \leq r'} Mw(y') dy'. \tag{38}$$

Furthermore, since $|y| \leq (4B_1 \lambda r)^{-1/(\ell-1)}$ an elementary calculation reveals that $(y, r') \in \Gamma_{\ell, \lambda}(0)$, and consequently (37) and (38) are dominated by constant multiples of $\mathcal{M}_{\ell, \lambda}(w)(0)$ and $\mathcal{M}_{\ell, \lambda}(Mw)(0)$ respectively. In each case we obtain the bound

$$r^{\frac{\ell-2}{\ell-1}} \lambda^{-\frac{1}{\ell-1}} \Phi_{1/r} * w(y) \lesssim \mathcal{M}_{\ell, \lambda}(Mw)(0),$$

as claimed, with implicit constant depending on (at most) B_1 and ℓ .

This completes the proof of Proposition 5.2 for $1 < 2^p < 4B_1 \lambda^{(\ell-1)/\ell}$, leaving only the case $p = 0$ to consider. To deal with this we observe that when $p = 1$, the second frequency decomposition (26) is effectively vacuous since $L = 2^{-p/(\ell-1)} \lambda^{1/\ell} \sim 2^p \lambda^{1/\ell}$. Given this, it is

⁵ By making B_1 larger if necessary, we may of course assume that $4B_1 \geq 1$.

straightforward to verify that our analysis in the case $p = 1$ is equally effective in the case $p = 0$. We leave this to the reader.

7. The proof of Proposition 5.2 for $2^p \geq 4B_1\lambda^{(\ell-1)/\ell}$

Suppose that the Fourier support of f is contained in \mathcal{A}_p for some p with $2^p \geq 4B_1\lambda^{(\ell-1)/\ell}$. As we shall see, the rapid decay in (24) means that such terms may be viewed as error terms in the sense that (23) holds with $\mathcal{M}_{\ell,\lambda}$ replaced with a much smaller operator.

Now, let $\Phi_{2^p\lambda^{1/\ell}}$ be constructed in the usual way by dilating a fixed Schwartz function such that $\widehat{\Phi}_{2^p\lambda^{1/\ell}}(\xi) = 1$ for $\xi \in \mathcal{A}_p$. By (14) of Observation 2.5, we have

$$\int_{\mathbb{R}} |T_\lambda f|^2 w \leq \|T_\lambda \Phi_{2^p\lambda^{1/\ell}}\|_1 \int_{\mathbb{R}} |f|^2 |T_\lambda \Phi_{2^p\lambda^{1/\ell}}| * w.$$

Now, for each $k \in \mathbb{N}$, repeated integration by parts gives

$$\begin{aligned} |T_\lambda \Phi_{2^p\lambda^{1/\ell}}(x)| &= \left| \int_{\mathbb{R}} e^{ix\xi} \widehat{K}_\lambda(\xi) \widehat{\Psi}_{2^p\lambda^{1/\ell}}(\xi) d\xi \right| \\ &\leq |x|^{-k} \int_{\mathbb{R}} \left| \left(\frac{d}{d\xi} \right)^k (\widehat{K}_\lambda(\xi) \widehat{\Psi}_{2^p\lambda^{1/\ell}}(\xi)) \right| d\xi. \end{aligned}$$

Using Lemma 5.1 and the fact that $|\xi| \sim 2^p\lambda^{1/\ell} \gtrsim \lambda$ for $\xi \in \text{supp } \widehat{\Phi}_{2^p\lambda^{1/\ell}}$ we conclude that for each $k, N \in \mathbb{N}$,

$$|T_\lambda \Phi_{2^p\lambda^{1/\ell}}(x)| \lesssim \lambda^{-N} \frac{\lambda}{(1 + \lambda|x|)^k}.$$

By repeating the analysis that leads to (34) it is now straightforward to verify that

$$\|T_\lambda \Phi_{2^p\lambda^{1/\ell}}\|_1 \|T_\lambda \Phi_{2^p\lambda^{1/\ell}}| * w \lesssim \mathcal{M}_{\ell,\lambda} M w(0),$$

uniformly in all parameters. This completes the proof of Proposition 5.2.

8. The proof of Proposition 2.3

Since our proof will use a Hardy space estimate and complex interpolation, we begin by bounding $\mathcal{M}_{\ell,\lambda}$ by a more “regular” maximal operator. For $x \in \mathbb{R}$ let

$$\widetilde{\Gamma}_{\ell,\lambda}(x) = \left\{ (y, r) : 0 < r \leq \lambda^{-\frac{1}{\ell}} \text{ and } |y - x| \leq (\lambda r)^{-\frac{1}{\ell-1}} \right\},$$

and observe that $\Gamma_{\ell,\lambda}(x) \subseteq \widetilde{\Gamma}_{\ell,\lambda}(x)$. Let P be a nonnegative smooth bump function supported in $[-2, 2]$ and positive on $[-1, 1]$, and for each $r > 0$, let $P_r(x) = \frac{1}{r} P(\frac{x}{r})$. Then for any weight function w ,

$$\mathcal{M}_{\ell,\lambda} w(x) \lesssim \widetilde{\mathcal{M}}_{\ell,\lambda} w(x) := \sup_{(y,r) \in \widetilde{\Gamma}_{\ell,\lambda}(x)} r(\lambda r)^{-\frac{1}{\ell-1}} |P_r * w(y)|.$$

Thus in order to prove Proposition 2.3 it suffices to show that

$$\|\widetilde{\mathcal{M}}_{\ell,\lambda}\|_{(\frac{\ell}{2})' \rightarrow (\frac{\ell}{2})'} \lesssim \lambda^{-\frac{2}{\ell}},$$

which, thanks to the scale-invariance property

$$\lambda^{2/\ell} \widetilde{\mathcal{M}}_{\ell,\lambda} w(\lambda^{-1/\ell}x) = \widetilde{\mathcal{M}}_{\ell,1}(w(\lambda^{-1/\ell}\cdot))(x), \tag{39}$$

is equivalent to

$$\|\widetilde{\mathcal{M}}_{\ell,1}\|_{(\frac{\ell}{2})' \rightarrow (\frac{\ell}{2})'} \lesssim 1. \tag{40}$$

In order to prove (40) we define

$$\mathcal{M}_{\ell}^{\beta} \phi(x) = \sup_{(y,r) \in \widetilde{\Gamma}_{\ell,1}(x)} r^{\frac{\ell\beta}{\ell-1}} |P_r * \phi(y)|$$

with P_r as above. By Stein’s method of analytic interpolation (see [21]), inequality (40) may be obtained from the estimates

$$\|\mathcal{M}_{\ell}^0 \phi\|_{\infty} \lesssim \|\phi\|_{\infty}$$

and

$$\|\mathcal{M}_{\ell}^1 \phi\|_{L^1} \lesssim \|\phi\|_{H^1}. \tag{41}$$

The first estimate is elementary, and the second may be verified by testing on atoms. Let a be an H^1 -atom with support interval I (by translation invariance we may suppose that I is centred at the origin). In order to establish (41) it suffices to prove that

$$\|\mathcal{M}_{\ell}^1 a\|_{L^1} \lesssim 1 \tag{42}$$

uniformly in a . For an atom a as described above, we have the standard pointwise bound

$$r^{\frac{\ell}{\ell-1}} |P_r * a(x)| \lesssim \begin{cases} r^{\frac{\ell}{\ell-1}}/|I|, & \text{if } r \leq |I| \text{ and } |x| \leq 5|I|/2, \\ |I|/r^{2-\frac{\ell}{\ell-1}}, & \text{if } r \geq |I| \text{ and } |x| \leq 5r/2, \\ 0, & \text{otherwise.} \end{cases} \tag{43}$$

In order to establish (42) we distinguish two cases.

Case 1. $|I| \geq 1$. From the definition of $\widetilde{\Gamma}_{\ell,1}(x)$ and (43) it is straightforward to verify that

$$\mathcal{M}_{\ell}^1 a(x) \lesssim \begin{cases} |I|^{-1}, & \text{if } |x| \leq 4|I|, \\ |I|^{-1}|x|^{-\ell}, & \text{if } |x| > 4|I| \end{cases}$$

uniformly in a . The desired bound (41) in the case $|I| \geq 1$ follows immediately.

Case 2. $|I| < 1$. Using (43) it is easily verified that $\mathcal{M}_\ell^1 a(x) \lesssim |I|^{\frac{1}{\ell-1}}$ uniformly in a and x , and so

$$\int_{|x| \leq 4|I|^{-\frac{1}{\ell-1}}} |\mathcal{M}_\ell^1 a(x)| dx \lesssim 1$$

uniformly in I (and thus a). Arguing as in Case 1 we have that $\mathcal{M}_\ell^1 a(x) \lesssim |I|^{-1} |x|^{-\ell}$ uniformly in a and $|x| \geq 4|I|^{-\frac{1}{\ell-1}}$, yielding

$$\int_{|x| \geq 4|I|^{-\frac{1}{\ell-1}}} |\mathcal{M}_\ell^1 a(x)| dx \lesssim |I|^{-1} \int_{|x| \geq 4|I|^{-\frac{1}{\ell-1}}} |x|^{-\ell} dx \sim 1$$

uniformly in I (and thus a). This completes the proof of (42), and thus Proposition 2.3.

Acknowledgments

It is a pleasure to acknowledge the influence that Tony Carbery has had on the content and perspective of this work. We would also like to thank the anonymous referee for a number of helpful suggestions.

References

- [1] J.A. Barceló, J. Bennett, A. Carbery, A note on localised weighted inequalities for the extension operator, *J. Aust. Math. Soc.* 84 (3) (2008) 289–299.
- [2] J.A. Barceló, A. Ruiz, L. Vega, Weighted estimates for the Helmholtz equation and some applications, *J. Funct. Anal.* 150 (2) (1997) 356–382.
- [3] J. Bennett, A. Carbery, F. Soria, A. Vargas, A Stein conjecture for the circle, *Math. Ann.* 336 (2006) 671–695.
- [4] A. Carbery, E. Romera, F. Soria, Radial weights and mixed norm estimates for the disc multiplier, *J. Funct. Anal.* 109 (1) (1992) 52–75.
- [5] A. Carbery, A. Seeger, Weighted inequalities for Bochner–Riesz means in the plane, *Q. J. Math.* 51 (2000) 155–167.
- [6] A. Carbery, F. Soria, Pointwise Fourier inversion and localisation in \mathbb{R}^n , *J. Fourier Anal. Appl.* 3 (1997) 847–858 (special issue).
- [7] A. Carbery, F. Soria, Sets of divergence for the localisation principle for Fourier integrals, *C. R. Acad. Sci. Paris Sér. I Math.* 325 (12) (1997) 1283–1286.
- [8] A. Córdoba, Geometric Fourier analysis, *Ann. Inst. Fourier (Grenoble)* 32 (3) (1982) 215–226.
- [9] A. Córdoba, C. Fefferman, A weighted norm inequality for singular integrals, *Studia Math.* 57 (1) (1976) 97–101.
- [10] J. Duoandikoetxea, *Fourier Analysis*, Grad. Stud. Math., vol. 29, American Mathematical Society, 2001.
- [11] J. Garcia-Cuerva, J.L. Rubio de Francia, *Weighted Norm Inequalities and Related Topics*, North-Holland Math. Stud., vol. 116, North-Holland, Amsterdam, 1985.
- [12] A. Greenleaf, A. Seeger, Fourier integral operators with fold singularities, *J. Reine Angew. Math.* 455 (1994) 35–56.
- [13] S. Lee, A. Seeger, K.M. Rogers, Improved bounds for Stein’s square functions, *Proc. London Math. Soc.*, in press.
- [14] B. Muckenhoupt, Weighted norm inequalities for the Fourier transform, *Trans. Amer. Math. Soc.* 276 (2) (1983) 729–742.
- [15] A. Nagel, E.M. Stein, On certain maximal functions and approach regions, *Adv. Math.* 54 (1984) 83–106.
- [16] C. Pérez, Weighted norm inequalities for singular integral operators, *J. London Math. Soc.* 49 (2) (1994) 296–308.
- [17] C. Pérez, Sharp L^p -weighted Sobolev inequalities, *Ann. Inst. Fourier* 45 (3) (1995) 809–824.
- [18] H.R. Pitt, Theorems on Fourier series and power series, *Duke Math. J.* 3 (4) (1937) 747–755.
- [19] J.L. Rubio de Francia, A Littlewood–Paley inequality for arbitrary intervals, *Rev. Mat. Iberoam.* 1 (2) (1985) 1–14.
- [20] E.M. Stein, Some problems in harmonic analysis, in: Williamstown, Mass., 1978, in: *Proc. Sympos. Pure Math.*, vol. 1, pp. 3–20.

- [21] E.M. Stein, Harmonic Analysis, Princeton University Press, 1993.
- [22] J.M. Wilson, Weighted norm inequalities for the continuous square functions, *Trans. Amer. Math. Soc.* 314 (1989) 661–692.
- [23] J.M. Wilson, The intrinsic square function, *Rev. Mat. Iberoam.* 23 (3) (2007) 771–791.