On some maximal multipliers in L^p

Ciprian Demeter

Abstract

We extend an L^2 maximal multiplier result of Bourgain to all L^p spaces, $1 < p < \infty$.

1. Introduction

Consider a finite set $\Lambda = \{\lambda_1, \ldots, \lambda_N\} \subset \mathbb{R}$ and let $D > 0$ be the separation constant

$$D := \min_{1 \le i \ne j \le N} |\lambda_i - \lambda_j|.$$

For each $k \in \mathbb{Z}$ with $2^k < 10^{-2}D$ define R_k to be the collection[1] of all intervals of length 2^k centered at some element from Λ. The following result was proved in [1].

Theorem 1.1. *For each $f \in L^2(\mathbb{R})$,*

$$(1.1) \qquad \left\| \sup_k \left| \int_{R_k} \widehat{f}(\xi) e^{2\pi i \xi x} d\xi \right| \right\|_{L^2_x} \lesssim \left(\log N \right)^2 \|f\|_2.$$

In [4] it was proved that at least a factor of $(\log N)^{1/4}$ is needed on the right hand side. Inequality (1.1) found multiple applications in the work of Bourgain on various ergodic averages, see for example [1], [2] and [3]. Lacey [13] has used the inequality to prove bounds for the bilinear maximal function. More recently, it has become apparent that variants of (1.1) play a significant role in the analysis of maximal truncations associated with modulation invariant operators, see [6], [7] and [8].

---

[1]We will abuse notation and also denote by R_k the union of the intervals in the collection.

For each $1 \leq r < \infty$ and each sequence $(x_k)_{k \in \mathbb{Z}}$ in a Hilbert space $\mathcal{H}$, define the *r-variational norm* of $(x_k)_{k \in \mathbb{Z}}$ to be

$$\|x_k\|_{V_k^r(\mathcal{H})} := \sup_k \|x_k\|_{\mathcal{H}} + \|x_k\|_{\tilde{V}_k^r(\mathcal{H})}$$

where $\tilde{V}_k^r(\mathcal{H})$ is the homogeneous *r-variational seminorm*

$$\|x_k\|_{\tilde{V}_k^r(\mathcal{H})} := \sup_{M,\, k_0 < k_1 < \cdots < k_M} \left(\sum_{m=1}^{M} \|x_{k_m} - x_{k_{m-1}}\|_{\mathcal{H}}^r \right)^{1/r}.$$

If $\mathcal{H}$ is the set of complex numbers $\mathbb{C}$, then the dependence on $\mathcal{H}$ will be suppressed. The only other instance that will occur is when $\mathcal{H} = l^2(N)$ (i.e. $\mathbb{R}^N$ with the euclidean inner product). In that case we will use the notation $V_k^r(N)$ and $\tilde{V}_k^r(N)$. Also, $\|x_Q\|_{V_{v_1 \leq Q \leq v_2}^r}$ will refer to the variational norm of a sequence x_Q indexed by some quantity Q that takes dyadic values between v_1 and v_2.

For each interval $\omega \in R_k$, let m_ω be a complex valued Schwartz function adapted to ω, that is, supported on ω and satisfying

$$\|\partial^\alpha m_\omega\|_\infty \leq |\omega|^{-\alpha}, \quad \alpha \in \{0, 1\}.$$

Let $(w_\omega)_{\omega \in R_k}$ be a collection of complex weights. Define

$$\Delta_k f(x) := \sum_{\omega \in R_k} \int w_\omega m_\omega(\xi) \widehat{f}(\xi) e^{2\pi i \xi x} d\xi,$$

and also

$$\|w_\omega\|_{V^{r,*}} := \max_{1 \leq n \leq N} \left\| \{ w_{\omega_k} : \lambda_n \in \omega_k \in R_k \} \right\|_{V_k^r},$$

$$\|m_\omega\|_{V^{r,*}} := \max_{1 \leq n \leq N} \left\| \{ m_{\omega_k}(\lambda_n) : \lambda_n \in \omega_k \in R_k \} \right\|_{V_k^r}.$$

The following weighted version of (1.1) was proved in [8]

Theorem 1.2. *For each $r > 2$ and $f \in L^2(\mathbb{R})$ we have the inequality*

$$\left\| \sup_{2^k < 10^{-2} D} |\Delta_k f(x)| \right\|_{L_x^2(\mathbb{R})} \lesssim N^{1/2 - 1/r} \|w_\omega\|_{V^{r,*}} (1 + \|m_\omega\|_{V^{r,*}}) \|f\|_2,$$

with the implicit constant depending only on r.

Theorem 1.2 was one of the tools needed to prove the following extension of Bourgain's Return Times theorem:

Theorem 1.3. ([8]) *Let* $\mathbf{X} = (X, \Sigma, \mu, \tau)$ *be a dynamical system, and let* $1 < q \leq \infty$ *and* $p \geq 2$. *For each function* $g \in L^q(X)$ *there is a universal set* $X_0 \subseteq X$ *with* $\mu(X_0) = 1$, *such that for each second dynamical system* $\mathbf{Y} = (Y, \mathcal{F}, \nu, \sigma)$, *each* $f \in L^p(Y)$ *and each* $x \in X_0$, *the averages*

$$\frac{1}{N} \sum_{n=0}^{N} g(\tau^n x) f(\sigma^n y)$$

converge for ν-*almost every* y.

In [8], Theorem 1.3 was proved in the case $p = 2$, and the case $p > 2$ followed immediately due to the nestedness of $L^p(Y)$ spaces. If one wants to extend the approach from [8] to generalize Theorem 1.3 even further, to the case $p < 2$, then an L^p version of Theorem 1.2 needs to be proved. We achieve this here, see Theorem 1.5 bellow. This result gives hope that the following conjecture can be proved, and it is a first step in the direction of its resolution:

Conjecture 1.4. *Theorem* 1.3 *holds whenever* $1 < p, q \leq \infty$ *and*

$$\frac{1}{p} + \frac{1}{q} < \frac{3}{2}.$$

Theorem 1.5. *For each* $1 < p < 2$, *each* $\epsilon > 0$ *and each* $r > 2$ *we have the inequality*

$$\left\| \sup_{2^k < 10^{-2}D} |\Delta_k f(x)| \right\|_{L_x^p(\mathbb{R})} \lesssim N^{1/p - 1/r + \epsilon} \|w_\omega\|_{V^{r,*}} (1 + \|m_\omega\|_{V^{r,*}}) \|f\|_p,$$

with the implicit constant depending only on r, ϵ *and* p.

2. Where the difficulty lies

Theorem 1.5 would typically be applied to the Conjecture 1.4 with r very close to (but larger than[2]) 2, in which case the exponent of N approaches $1/p - 1/2$. On the other hand, this is also the best possible exponent one can hope for in Theorem 1.5. This was already observed in [8], and it suffices to consider a single scale. We reproduce the construction for completeness.

Proposition 2.1. *For each* $N \in \mathbb{N}$ *and* $p \in (1, 2)$ *there is a choice of signs* $(\varepsilon_n)_{1 \leq n \leq N}$ *such that if* $\widehat{f_N} = 1_{[0,N]}$ *then*

$$\left\| \int \widehat{f_N}(\xi) \sum_{l=0}^{N-1} \varepsilon_n 1_{[n,n+1]}(\xi) e^{2\pi i \xi x} d\xi \right\|_{L_x^p(\mathbb{R})} \gtrsim N^{1/p - 1/2} \|f_N\|_{L^p(\mathbb{R})}.$$

---

[2]The weights w_ω are typically averages at scale $|\omega|^{-1}$ of the weight function g; the r variation of these averages is only bounded when $r > 2$.

Proof. It immediately follows that

$$\|f_N\|_{L^p(\mathbb{R})} \sim N^{1-1/p},$$

$$\left\| \left(\sum_{n=0}^{N-1} | \int \widehat{f}_N(\xi) 1_{[n,n+1]}(\xi) e^{2\pi i \xi x} d\xi|^2 \right)^{1/2} \right\|_{L_x^p(\mathbb{R})} \sim N^{1/2}.$$

Khintchine's inequality ends the proof. ∎

It is worth mentioning that Theorem 1.5 also holds in the case $p > 2$, with the bound[3] $N^{\frac{1}{2} - \frac{2}{pr} + \epsilon}$ but the proof in this case is immediate. Indeed, due to interpolation with the L^2 result in Theorem 1.2, it suffices to achieve the crude bound $N^{1/2}$ for each $p > 2$. To get this, note that by the inequality of Cauchy-Schwartz we have

$$\sup_{2^k \le 10^{-2} D} \left| \Delta_k f(x) \right| \le N^{1/2} \|w_\omega\|_{V^{r,*}} \left(\sum_{n=1}^{N} (M(f_n)(x))^2 \right)^{1/2},$$

where

$$f_n(x) = \int_{|\xi - \lambda_n| < D/10} \widehat{f}(\xi) e^{2\pi i \xi} d\xi,$$

and $M(f)$ denotes the Hardy-Littlewood maximal function of f. The result then follows from a combination of the Fefferman-Stein inequality [10] and Rubio de Francia's result [17] on the L^p boundedness ($p > 2$) of the square function

$$SQ(f)(x) := \left(\sum_{n=1}^{N} |f_n(x)|^2 \right)^{1/2}.$$

The main difficulty in proving Theorem 1.5 is that the square function $SQ(f)$, while it has $L^p \to L^p$ norm independent of N in the case $p \ge 2$, it will have a norm of magnitude $N^{1/p-1/2}$ in the case $p < 2$. To overcome this deficit, what we will do instead is relate $\|\sup_k |\Delta_k f(x)|\|_{L_x^p(\mathbb{R})}$ to a more complicated square function, Vf, without the use of the Cauchy-Schwartz inequality (and thus without any further loss in powers of N). Most of the paper is then devoted to proving that V has essentially the same $L^p, 1 < p < 2$ operator norm as SQ (see Theorem 5.1). This itself is a result of independent interest, and we use time-frequency techniques to prove it. It would be interesting to know whether one could use a more standard approach, via interpolation with an $L^1 \to L^{1,\infty}$ result. In particular, it is not clear whether the operator norm $\|V\|_{L^1 \to L^{1,\infty}}$ is also of order $N^{1/p-1/2}$. Our approach seems to shed no light on this issue.

---

[3]Note that the argument above also shows that this bound is essentially best possible, when r is close to 2 and ϵ is close to 0.

We mention that both Theorem 1.1 and Theorem 1.2 have variants in [1] and [8] respectively, in which the separation restriction on λ_n is eliminated. In each case, the boundedness of the unrestricted operator $\sup_{k \in \mathbb{Z}} |\Delta_k f|$ follows from the boundedness of the restricted version and by using techniques related to the Rademacher-Menshov theorem. A similar approach can probably be pursued with Theorem 1.5 here, too. One would first have to bound locally the L^p norm by the L^2 norm and then apply Rademacher-Menshov type arguments to the localized operator. These local contributions are then summed over the whole real line, to produce the desired estimate. We refer the interested reader to the proof of Theorem 8.7 in [8] for details.

3. A variational inequality

The following simple inequalities will be useful throughout the rest of the paper.

Lemma 3.1. *For each k, let a_k, b_k be some complex numbers. Then, for each $r \geq 1$,*

$$\|a_k b_k\|_{V_k^r} \lesssim \|a_k\|_{V_k^r} \|b_k\|_{V_k^r}.$$

Lemma 3.2. *Let $\{c_k := (c_k^{(n)})_{1 \leq n \leq N} : k \in \mathbb{Z}\} \subseteq l^2(N)$. Then*

$$\|c_k\|_{V_k^r(N)} \leq \Big(\sum_{n=1}^N \|c_k^{(n)}\|_{V_k^r}^2 \Big)^{1/2}.$$

Lemma 3.3. *Let $(x_k)_{k_0 \leq k < k_L} \in \mathbb{C}$ and let $k_0 < k_1 < \cdots < k_L$. Then*

$$\|x_k\|_{V_{k_0 \leq k < k_L}^r} \leq \sum_{l=0}^{L-1} \|x_k\|_{V_{k_l \leq k < k_{l+1}}^r}.$$

There is an extensive literature on variational estimates in Harmonic Analysis and Ergodic Theory. We confine ourselves to mentioning only a few such papers: [5], [11] and [12].

For each[4] $x \in \mathbb{R}$ and $k \in \mathbb{Z}$ denote by $I(x, k)$ the unique dyadic interval of length 2^{-k} which contains x. Let $\phi^{(k)}$ be Schwartz functions with Fourier transform adapted to $[-\frac{1}{2}, \frac{1}{2}]$. For each interval I of length 2^{-k}, define

$$\phi_I(y) := \frac{1}{|I|} \phi^{(k)} \Big(\frac{y - c(I)}{|I|} \Big).$$

---

[4] We can ignore the dyadic points, since they have measure zero.

Lemma 3.4. *We have for each $1 < p < \infty$ and each $r > 2$*

$$\left\| \|\langle f, \phi_{I(k,x)} \rangle\|_{V_k^r} \right\|_{L_x^p} \lesssim \left(1 + \|\widehat{\phi^{(k)}}(0)\|_{V_k^r} \right) \|f\|_p,$$

with the implicit constant depending only on ϕ, r and p.

Proof. Consider the dyadic martingale

$$\mathbb{E}_k(f)(x) = \left\langle f, \frac{1}{|I(k,x)|} 1_{I(k,x)} \right\rangle.$$

A result of Lepingle [14] implies that

$$\left\| \|\mathbb{E}_k(f)(x)\|_{V_k^r} \right\|_{L_x^p} \lesssim \|f\|_p.$$

By using this and Lemma 3.1 we get that

$$\left\| \|\widehat{\phi^{(k)}}(0)\mathbb{E}_k(f)(x)\|_{V_k^r} \right\|_{L_x^p} \lesssim \|\widehat{\phi^{(k)}}(0)\|_{V_k^r} \|f\|_p.$$

It suffices now to note that the difference operator can be bounded as follows

$$\left\| \left\| \left\langle f, \frac{\widehat{\phi^{(k)}}(0)}{|I(k,x)|} 1_{I(k,x)} - \phi_{I(k,x)} \right\rangle \right\|_{V_k^r} \right\|_{L_x^p} \leq$$

$$\leq \left\| \left(\sum_k \left| \left\langle f, \frac{\widehat{\phi^{(k)}}(0)}{|I(k,x)|} 1_{I(k,x)} - \phi_{I(k,x)} \right\rangle \right|^2 \right)^{1/2} \right\|_{L_x^p} \lesssim \|f\|_p.$$

The first inequality holds since $r > 2$, while the second one is a consequence of the Littlewood-Paley theory, since $\phi_I - \frac{\widehat{\phi^{(k)}}(0)}{|I|} 1_I$ has mean zero (when $|I| = 2^{-k}$). ∎

4. Time-frequency interlude

The purpose of this section is to prepare the ground for the proof of Theorem 5.1 in the next section.

To prove Theorem 1.5, it will suffice by scaling invariance to assume that

$$(4.1) \qquad \min_{1 \leq i \neq j \leq N} |\lambda_i - \lambda_j| = 1.$$

We will do so throughout the rest of the paper.

Definition 4.1. A tile $s = s(I, n)$ will be a rectangle of the form

$$s(I, n) := I \times [\lambda_n - 2^{k-1}, \lambda_n + 2^{k-1}],$$

where $1 \leq n \leq N$, while I is a dyadic interval of length $2^{-k} > 100$. We refer to $I_s := I$ and $\omega_s := [\lambda_n - 2^{k-1}, \lambda_n + 2^{k-1}]$ as the time interval and frequency interval of the tile s.

The collection of all tiles will be denoted by $\mathbf{S}$. We define $\mathbf{S}_n$ to consist of the collection of all tiles $s \in \mathbf{S}$ such that $\lambda_n \in \omega_s$. Note that $(\mathbf{S}_n)_{n=1}^N$ forms a partition of $\mathbf{S}$.

Definition 4.2. A collection $\mathbf{S}' \subset \mathbf{S}$ will be referred to as convex, if whenever $s, s'' \in \mathbf{S}'$, $s' \in \mathbf{S}$, $\lambda_n \in \omega_s \cap \omega_{s'} \cap \omega_{s''}$ for some n and $I_s \subseteq I_{s'} \subseteq I_{s''}$, these also imply that $s' \in \mathbf{S}'$.

Definition 4.3. A tree $(\mathbf{T}, T)$ with top $T \in \mathbf{S}$ is a convex collection of tiles $\mathbf{T} \subset \mathbf{S}$ such that $I_s \subseteq I_T$ and $\omega_T \subseteq \omega_s$ for each $s \in \mathbf{T}$.

Note that each tree is entirely contained in some $\mathbf{S}_n$.

We will denote by T_m the Fourier projection associated with the multiplier m:

$$T_m f(x) := \int \widehat{f}(\xi) m(\xi) e^{2\pi i \xi x} d\xi.$$

We will use the notation

$$\tilde{\chi}_I(x) = \left(1 + \frac{|x - c(I)|}{|I|}\right)^{-1}.$$

Definition 4.4. Let f be a L^2 function and let $\mathbf{S}' \subset \mathbf{S}$. We define the size size($\mathbf{S}'$) of $\mathbf{S}'$ relative[5] to f as

$$\text{size}(\mathbf{S}') := \sup_{s \in \mathbf{S}'} \sup_{m_s} \frac{1}{|I_s|^{1/2}} \|\tilde{\chi}_{I_s}^{10}(x) T_{m_s} f(x)\|_{L_x^2},$$

where m_s ranges over all functions adapted to $10\omega_s$.

For a fixed tile s, the quantity

$$\text{size}(s) = \sup_{m_s} \frac{1}{|I_s|^{1/2}} \|\tilde{\chi}_{I_s}^{10}(x) T_{m_s} f(x)\|_{L_x^2},$$

is an approximate measure for the L^2 norm of the portion of f that is localized in time-frequency in s.

We need some notation and results from [16].

---

[5]The function with respect to which the size is computed will change throughout the paper; however, it will always be clear from the context.

Definition 4.5. Let $(\mathbf{T}, T)$ be a tree, and let $\mathbf{I}_T$ be the collection of all maximal dyadic intervals $I \subseteq I_T$ which have the property that $3I$ does not contain any of the intervals I_s with $s \in T$. For an integer j with $2^{-j} \leq |I_T|$ let $\tilde{E}_j$ be the union of all intervals I in $\mathbf{I}_T$ such that $|I| < 2^{-j}$

For an integer j with $2^{-j} > |I_T|$ we define $\tilde{E}_j = \emptyset$.

The sets $\tilde{E}_j$ obviously depend on the tree $\mathbf{T}$, but we suppress this dependence.

We recall the following Lemma 4.12 from [16].

Lemma 4.6. *Let Ω_j be the collection of connected components of $\tilde{E}_j$. Then Ω_j is a finite collection of dyadic intervals each of which has length equal to an integer multiple of 2^{-j}.*

For each $I \in \Omega_j$, let x_I^l and x_I^r denote the left and right endpoints of I, and let I_j^l and I_j^r denote the intervals

$$I_j^l := (x_I^l - 2^{-j-1}, x_I^l - 2^{-j-2})$$
$$I_j^r := (x_I^r + 2^{-j-2}, x_I^r + 2^{-j-1}).$$

Then the intervals I_j^l are disjoint as j varies in the integers with $2^{-j} \leq |I_T|$ and I varies in Ω_j.

Similar statements hold for the I_j^r.

Recall also the weight function from [16].

$$\mu_{j,\mathbf{T}}(x) := \sum_j 2^{-|j'-j|/100} \sum_{y \in \partial \tilde{E}_j} (1 + 2^{j'}|x - y|)^{-100}.$$

An important role in our future investigation is played by the function

$$W_\mathbf{T}(x) := \sum_{s \in \mathbf{T}} \frac{1_{I_s}(x)}{|I_s|} \int \mu_{j(s)}(y) \tilde{\chi}_{I_s}^2(y) dy,$$

where $|I_s| = 2^{-j(s)}$.

In the same Lemma 4.12 from [16] it was proved (and this is an immediate consequence of Lemma 4.6) that

$$\sum_j 2^{-j} \# \partial \tilde{E}_j \lesssim |I_T|.$$

This in turn automatically implies that

$$\int W_\mathbf{T}(x) dx \lesssim |I_T|.$$

We will need a stronger result, namely that

$$\int_J W_{\mathbf{T}}(x)dx \lesssim |J|$$

for each dyadic $J \subseteq I_T$. This however also turns out to be an easy consequence of Lemma 4.6. We leave the details to the reader. The following is then a standard corollary.

Corollary 4.7. *We have*

$$\|W_{\mathbf{T}}\|_{\mathrm{BMO}} \lesssim 1,$$

where $\| \cdot \|_{\mathrm{BMO}}$ stands for the dyadic BMO norm.

The following result is a simplified version[6] of Proposition 7.4 from [16]. In its current form, it appears essentially in [9].

Each tree $\mathbf{T}$ defines a region in the phase-space domain. The operator V from Theorem 5.1 will receive a contribution from all the trees. The lemma below will show that for each tree, the most important contribution comes from the part of the function f, denoted by $\Pi_{\mathbf{T}}(f)$, that essentially lives in the region defined by $\mathbf{T}$.

Lemma 4.8. (Phase-space projections) *Let $(\mathbf{T}, T)$ be a tree and $f \in L^2(\mathbb{R})$[7]. There is a function $\Pi_{\mathbf{T}}(f)$, called the phase-space projection of f on the tree $\mathbf{T}$, satisfying the properties:*

- *(i) (Control by size) $\Pi_{\mathbf{T}}(f)$ is supported in $2I_T$ and satisfies the L^∞ bound*

$$\|\Pi_{\mathbf{T}}(f)\|_\infty \lesssim \mathrm{size}(\mathbf{T}).$$

  As a consequence, for each $1 \le t \le \infty$,

  (4.2) $$\|\Pi_{\mathbf{T}}(f)\|_t \lesssim |I_T|^{1/t}\mathrm{size}(\mathbf{T})$$

  with the implicit constant independent of $\mathbf{T}$ or f.

- *(ii) ($\Pi_{\mathbf{T}}(f)$ approximates f on $\mathbf{T}$) For each $s \in \mathbf{T}$ with $|\omega_s| = 2^j$, each m adapted to ω_s we have*

  (4.3) $$\left\|\left(|I_s|\phi_{I_s}\right)^{1/2}T_m\left(\Pi_{\mathbf{T}}(f) - f\right)\right\|_2 \lesssim \mathrm{size}(\mathbf{T})|I_s|^{-1/2}\int \tilde{\chi}_{I_s}^2(x)\mu_j(x)dx.$$

---

[6]One of the simplifications arises from the simpler definition of size from Definition 4.4.
[7]Since we only want an L^2 estimate in (ii), there is no need for the extra assumption $\|f\|_\infty \le 1$ from [9] or [16].

Assume that for each $k \in \mathbb{Z}$ we have n functions $\phi_1^{(k)}, \ldots, \phi_n^{(k)}$ whose Fourier transforms are adapted to the interval $[-1/2, 1/2]$. Define also the modulated versions

(4.4)
$$\phi_{I,n}(x) := \frac{1}{|I|} \phi_n^{(k)}\left(\frac{x - c(I)}{|I|}\right) e^{2\pi i \lambda_n x},$$

whenever $|I| = 2^{-k}$. Also, if $s \in \mathbf{S}_n$, define $\phi_s := \phi_{I_s,n}$. We will assume throughout the rest of the section that

$$\max_n \|\widehat{\phi_n^{(k)}}(0)\|_{V_{k,2^{-k}>100}^r} \leq 1.$$

If $(\mathbf{T}, T)$ is a tree and $r > 2$, define

$$V_{\mathbf{T}} f(x) := \|1_{I_s}(x) \langle f, \phi_s \rangle\|_{V_{s \in \mathbf{T}}^r}.$$

As a consequence of Lemma 3.4, Corollary 4.7 and Lemma 4.8 we have

Proposition 4.9. *Let $(\mathbf{T}, T)$ be a tree and let $f \in L^2(\mathbb{R}) \cap L^t(\mathbb{R})$, for some $1 < t < \infty$. Then*

(4.5)
$$\|V_{\mathbf{T}} f\|_t \lesssim |I_T|^{1/t} \text{size}(\mathbf{T}).$$

Proof. Triangle's inequality implies

$$\|V_{\mathbf{T}} f\|_t \leq \|V_{\mathbf{T}} \Pi_{\mathbf{T}}(f)\|_t + \|V_{\mathbf{T}}(f - \Pi_{\mathbf{T}}(f))\|_t.$$

Since the first term above is well controlled by Lemma 3.4 and by (4.2), it suffices to control the second term. Note that by (4.3) and the fact that $\widehat{\phi}_s$ is supported in ω_s, we have for each $s \in \mathbf{T}$

$$
\begin{aligned}
|\langle f - \Pi_{\mathbf{T}}(f), \phi_s \rangle| &\leq \int |\phi_{I_s}(x) T_{m_s}(f - \Pi_{\mathbf{T}}(f))(x)| dx \\
&\lesssim \|(|I_s| \phi_{I_s}(x))^{1/2} T_{m_s}(f - \Pi_{\mathbf{T}}(f))(x)\|_2 \|(|I_s|^{-1} \phi_{I_s})^{1/2}\|_2 \\
&\lesssim \text{size}(\mathbf{T}) |I_s|^{-1} \int \tilde{\chi}_{I_s}^2(x) \mu_{j(s)}(x) dx,
\end{aligned}
$$

where m_s is an appropriate function, adapted to ω_s and equal to 1 on the support of $\widehat{\phi}_s$.

Next, by using the above and the trivial estimate $\|x_k\|_{V_k^r} \leq \|x_k\|_{l^1(\mathbb{Z})}$, we obtain the following pointwise estimate

$$|V_{\mathbf{T}}(f - \Pi_{\mathbf{T}}(f))(x)| \leq \text{size}(\mathbf{T}) W_{\mathbf{T}}(x).$$

The claim is now immediate from Corollary 4.7 and John-Nirenberg's inequality. ∎

The following lemma is immediate.

Lemma 4.10. *If $(\mathbf{T}, T)$ is a tree and let $f \in L^2(\mathbb{R}) \cap L^\infty(\mathbb{R})$ then*

$$\mathrm{size}(\mathbf{T}) \lesssim \sup_{s \in \mathbf{T}} \inf_{x \in I_s} Mf(x).$$

We now show how to split $\mathbf{S}$ into trees with special properties.

Lemma 4.11. *Let $\mathbf{S}' \subset \mathbf{S}$ be a finite convex collection of tiles and let $f \in L^2(\mathbb{R}) \cap L^\infty(\mathbb{R})$. Then we can split*

$$\mathbf{S}' = \bigcup_{2^{-m} \leq 2\mathrm{size}(\mathbf{S}')} \mathcal{S}_m,$$

where $\mathcal{S}_m$ is the union of disjoint trees

$$\mathcal{S}_m = \bigcup_{\mathbf{T} \in \mathcal{F}_m} \mathbf{T}$$

with $\mathrm{size}(\mathbf{T}) \leq 2^{-m}$ for each $\mathbf{T} \in \mathcal{F}_m$ and moreover

$$(4.6) \qquad \sum_{\mathbf{T} \in \mathcal{F}_m} |I_{\mathbf{T}}| \lesssim 2^{2m} \|f\|_2^2.$$

This lemma is obtained by iterating the following lemma

Lemma 4.12. *Let $\mathbf{S}' \subset \mathbf{S}$ be a finite convex collection of tiles, and assume*

$$\mathrm{size}(\mathbf{S}') \leq 2\lambda.$$

Then one can split

$$\mathbf{S}' = \mathbf{S}_1' \cup \mathbf{S}_2',$$

where $\mathbf{S}_1'$ is the union of disjoint trees

$$\mathbf{S}_1' = \bigcup_{\mathbf{T} \in \mathcal{F}} \mathbf{T}$$

with

$$\sum_{\mathbf{T} \in \mathcal{F}} |I_{\mathbf{T}}| \lesssim \lambda^{-2} \|f\|_2^2,$$

while

$$\mathrm{size}(\mathbf{S}_2') \leq \lambda.$$

Proof. The proof is entirely standard. We briefly sketch the argument.

- Step 0: Set $\mathcal{S} := \mathbf{S}'$ and $\mathcal{F} := \emptyset$.

- Step 1: Select a tile $T \in \mathcal{S}$ with maximal time interval I_T such that there exists m_T adapted to $10\omega_T$ satisfying

  (4.7) $$\frac{1}{|I_T|^{1/2}}\big\|\tilde{\chi}_{I_T}^{10}(x)T_{m_T}f(x)\big\|_{L_x^2} > \lambda.$$

  If there is more than one tile having a fixed (maximal) time interval that qualifies to be selected, just select any of them. Construct the maximal tree $\mathbf{T}$ in $\mathcal{S}$ having top T. Reset $\mathcal{F} := \mathcal{F} \cup \{\mathbf{T}\}$ and $\mathcal{S} := \mathcal{S} \setminus \mathbf{T}$.

  If there is no such tile to be selected, stop, the algorithm is over.

- Step 2: Go to Step 1.

After the algorithm stops, set

$$\mathbf{S}_1' := \bigcup_{\mathbf{T} \in \mathcal{F}} \mathbf{T} \qquad \text{and} \qquad \mathbf{S}_2' := \mathbf{S}' \setminus \mathbf{S}_1',$$

and note that $\text{size}(\mathbf{S}_2') \leq \lambda$, as desired.

Observe also that the rectangles $(I_T \times 10\omega_T)_{\mathbf{T} \in \mathcal{F}}$ are pairwise disjoint. Indeed, if $\mathbf{T}, \mathbf{T}'$ have the tops T and T' in the same $\mathbf{S}_n$ ($1 \leq n \leq N$) then I_T and $I_{T'}$ will be disjoint by maximality. If T and T' belong to distinct $\mathbf{S}_n$ and $\mathbf{S}_{n'}$, then, while their time intervals may intersect, the frequency components $10\omega_T, 10\omega_{T'}$ will be disjoint, due to (4.1) and the fact that $|\omega_T|, |\omega_{T'}| < \frac{1}{100}$.

Finally, inequality (4.7) combined with this pairwise disjointness implies via a (now) very standard TT^* argument (see for example Corollary 7.6 in [15]) that

$$\sum_{\mathbf{T} \in \mathcal{F}} |I_T| \lesssim \lambda^{-2}\|f\|_2^2.$$

∎

Remark 4.13. Note that if $\mathbf{S}'$ is convex, then each $\mathcal{S}_m$ is convex.

5. A square function estimate for the variational norm

Let as before $\phi_1^{(k)}, \ldots, \phi_n^{(k)}$ be Schwartz functions whose Fourier transforms are adapted to the interval $[-1/2, 1/2]$, and let $\phi_{I,n}$ be defined as in (4.4).

For $r > 2$ define

$$Vf(x) := \Big(\sum_{n=1}^N \|1_I(x)\langle f, \phi_{I,n}\rangle\|_{V_{I,|I|>100}^r}^2 \Big)^{1/2}.$$

Theorem 5.1. *We have for each $1 < p \le 2$ and each $\epsilon > 0$*

$$(5.1) \qquad \|Vf\|_p \lesssim N^{\frac{1}{p}-\frac{1}{2}+\epsilon}\Big(1 + \max_n \|\widehat{\phi_n^{(k)}}(0)\|_{V_{k,2^{-k}>100}^r}\Big)\|f\|_p,$$

with the implicit constant depending only on r, p and ϵ.

Proof. We can and will assume that

$$\max_n \|\widehat{\phi_n^{(k)}}(0)\|_{V_{k,2^{-k}>100}^r} \le 1.$$

The inequality when $p = 2$ (with $\epsilon = 0$) follows immediately from Lemma 3.4 and by orthogonality. By interpolating with the L^2 bound, it suffices to prove (5.1) with the bound $N^{1/2}$ rather than $N^{\frac{1}{p}-\frac{1}{2}+\epsilon}$

We rephrase the inequality to be proved in terms of tiles. We are to prove that

$$V_{\mathbf{S}'}f(x) := \Big(\sum_{n=1}^N \|1_{I_s}(x)\langle f, \phi_s\rangle\|_{V_{s\in\mathbf{S}'_n}^r}^2\Big)^{1/2}$$

satisfies

$$\|V_{\mathbf{S}'}f\|_p \lesssim N^{1/2}\|f\|_p,$$

uniformly over all finite $\mathbf{S}' \subseteq \mathbf{S}$ which in addition are also convex.

Also, by restricted type interpolation, it suffices to prove

$$(5.2) \qquad \big|\{x \in \mathbb{R} : V_{\mathbf{S}'}1_F(x) > \lambda^{1-\epsilon}N^{1/2}\}\big| \lesssim \frac{|F|}{\lambda^p},$$

for each $F \subset \mathbb{R}$ with finite measure and each $\epsilon > 0$, with an implicit constant depending just on p and ϵ. Moreover, due to the L^2 result, it suffices to prove (5.2) for $\lambda < 1$.

Consider the exceptional set

$$E := \{x \in \mathbb{R} : M1_F(x) \ge C\lambda\},$$

with C large enough, independent of F. Since

$$|E| \le 2\,C^{-p}\frac{|F|}{\lambda^p},$$

it further suffices to prove

$$\big|\{x \in \mathbb{R} \setminus E : V_{\mathbf{S}'}1_F(x) > \lambda^{1-\epsilon}N^{1/2}\}\big| \lesssim \frac{|F|}{\lambda^p}.$$

Obviously, on $\mathbb{R} \setminus E$ only the tiles $s \in \mathbf{S}'$ with $I_s \cap (\mathbb{R} \setminus E) \neq \emptyset$ will contribute to $V_{\mathbf{S}'} 1_F$. Thus we can assume $\mathbf{S}'$ consists only of such tiles. Note that $\mathbf{S}'$ remains convex. Lemma 4.10 implies now that $\text{size}(\mathbf{S}') \leq \lambda/2$, if C is sufficiently large. Here, the size is computed with respect to the function 1_F.

Recall now the decomposition in Lemma 4.11. For each $2^{-m} \leq \lambda$ and each $\mathbf{T} \in \mathcal{F}_m$ we define the exceptional set

$$E_{\mathbf{T}} := \{x \in I_T : V_{\mathbf{T}} 1_F(x) > \lambda^{\frac{1}{2}-\epsilon} 2^{-m/2}\},$$

and we note that due to Proposition 4.9,

$$|E_{\mathbf{T}}| \lesssim \lambda^{-t(\frac{1}{2}-\epsilon)} 2^{-mt/2} |I_{\mathbf{T}}|.$$

Furthermore, due to (4.6) we get that

$$\left| E' := \bigcup_{2^{-m} \leq \lambda} \bigcup_{\mathbf{T} \in \mathcal{F}_m} E_{\mathbf{T}} \right| \lesssim \sum_{2^{-m} \leq \lambda} \lambda^{-t(\frac{1}{2}-\epsilon)} 2^{-mt/2} 2^{2m} |F| \lesssim \frac{|F|}{\lambda^p},$$

assuming

$$t > \max\left\{4, \frac{2-p}{\epsilon}\right\}. \qquad \blacksquare$$

As a consequence, it further suffices to prove that for $x \in \mathbb{R} \setminus (E \cup E')$ we have

$$|V_{\mathbf{S}'} 1_F(x)| \leq \lambda^{1-\epsilon} N^{1/2}.$$

Observe that each $x \in \mathbb{R}$ belongs to at most N time intervals I_T of tops T of trees in $\mathcal{S}_m$, for each $2^{-m} \leq \lambda$. For each m, at most one of these trees is in each $\mathbf{S}_n$ ($1 \leq n \leq N$). Call these trees $\mathbf{T}_{1,m}, \ldots, \mathbf{T}_{n,m}$, where we allow some of them to be empty. Then

$$V_{\mathcal{S}_m} 1_F(x) = \left(\sum_{n=1}^{N} (V_{\mathbf{T}_{n,m}} 1_F(x))^2 \right)^{1/2}.$$

By the triangle inequality in Lemma 3.3 (possible since each $\mathcal{S}_m$ is convex) combined with the triangle inequality for the norm of $l^2(N)$, it follows that whenever $x \in \mathbb{R} \setminus (E \cup E')$ we have

$$|V_{\mathbf{S}'} 1_F(x)| \leq \sum_{2^{-m} \leq \lambda} |V_{\mathcal{S}_m} 1_F(x)| \leq \sum_{2^{-m} \leq \lambda} N^{1/2} \lambda^{\frac{1}{2}-\epsilon} 2^{-m/2} = \lambda^{1-\epsilon} N^{1/2}.$$

6. Proof of Theorem 1.5

We first use Fourier series to rewrite the maximal function $\sup_k |\Delta_k f|$ as a maximal expression involving sums of exponentials with variable coefficients. We then estimate the L^p norm by the L^2 norm, on each dyadic interval of unit length. The L^2 norm of these maximal exponential sums will then be estimated by using the metric entropy method, see Lemma 6.1 below. The result is that the local contribution, $\| \sup_k |\Delta_k f| \|_{L^p(J)}$, will be bounded by $\|Vf\|_{L^p(J)}$. These local contributions are then summed up and Theorem 5.1 finishes the argument.

The idea of reducing matters to maximal exponential sums goes back to [1]. The advantage of dealing with expressions of the form

$$\sup_k \Big| \sum_{n=1}^N c_{k,n} e^{2\pi i \lambda_n y} \Big|$$

resides both in the almost orthogonality of the exponentials, but also in the finite dimensionality of the coefficient space, $l^2(N)$. In [1], the reduction to exponential sums is performed via an averaging procedure, that relies crucially on Plancherel's theorem. However, in our case, this type of argument is not available, since we deal with L^p estimates. We will instead use windowed Fourier series expansions. The key is the fact that we can reduce matters to L^2 locally, without further loss in the constants.

The following lemma essentially appears in [1]. Its current formulation is taken from [8].

Lemma 6.1. ([8, Lemma 8.4]) *For each set* $C = \{c_k\} \subseteq l^2(N)$ *and each* $r > 2$ *we have*

$$\Big\| \sup_k \Big| \sum_{n=1}^N c_{k,n} e^{2\pi i \lambda_n y} \Big| \Big\|_{L_y^2([0,D^{-1}))} \lesssim N^{1/2 - 1/r} \|c_k\|_{V_k^r(N)},$$

with the implicit constant depending only on r.

Proof of Theorem 1.5. By rescaling, we can assume that the separation constant $D = 1$.

Let $\psi : \mathbb{R} \to \mathbb{C}$ be a smooth function such that $\widehat{\psi}$ is supported on $[-1, 1]$ and equals 1 on [-1/2,1/2]. For each dyadic interval $I := [l/2^k, (l+1)/2^k]$ denote by

$$\phi_{I,n}(x) := \frac{1}{|I|} \widehat{m}_{\omega,*} \Big(\frac{x - c(I)}{|I|} \Big) e^{2\pi i \lambda_n x},$$

where ω is the unique interval in R_k that contains λ_n, $m_{\omega,*}$ is the rescaled, shifted at the origin version of m_ω, that is

$$m_{\omega,*}\Big(\frac{\xi - \lambda_n}{|\omega|}\Big) = m_\omega(\xi).$$

Assume now ω is centered at some λ_n. By using windowed Fourier series on ω we can write

$$\widehat{f}(\xi)m_\omega(\xi) = \frac{1}{|\omega|}\sum_{l\in\mathbb{Z}}\Big[\int \widehat{f}(\eta)m_\omega(\eta)e^{2\pi i\frac{l}{|\omega|}\eta}d\eta\Big]e^{-2\pi i\frac{l}{|\omega|}\xi}\widehat{\psi}\Big(\frac{\xi - \lambda_n}{|\omega|}\Big),$$

so that

$$\int m_\omega(\xi)\widehat{f}(\xi)e^{2\pi i\xi x}d\xi = \sum_{\substack{|I|=|\omega|^{-1}\\ I\ dyadic}} e^{2\pi i\lambda_n x}\langle f,\phi_{I,n}\rangle\psi_I(x).$$

where

$$\psi_I(x) := \psi\Big(\frac{x - c(I)}{|I|}\Big).$$

Let now J be a dyadic interval of unit length. For each $2^k < \frac{1}{100}$ and each $l \in \mathbb{Z}$ denote by $I(J,k,l)$ the unique dyadic interval of length 2^{-k} such that $l2^{-k} + I(J,k,l)$ contains J. We have

$$\Big\|\sup_k |\Delta_k f(x)|\Big\|_{L_x^p(J)} \leq$$

$$\leq \sum_{l\in\mathbb{Z}}\Big\|\sup_k \Big|\Big(\sum_{n\leq N} e^{2\pi i\lambda_n x}\langle f,\phi_{I(J,k,l),n}\rangle w_{k,n}\Big)\psi_{I(J,k,l)}(x)\Big|\Big\|_{L_x^p(J)}$$

$$\lesssim \sum_{l\in\mathbb{Z}} 2^{-100|l|}\Big\|\sup_k \Big|\sum_{n\leq N} e^{2\pi i\lambda_n x}\langle f,\phi_{I(J,k,l),n}\rangle w_{k,n}\Big|\Big\|_{L_x^p(J)},$$

where $w_{k,n} := w_\omega$, given that ω is centered at λ_n and has scale 2^k. By invoking standard rescaling arguments it suffices to analyze the case $l = 0$, so any further dependence on l will be suppressed. We can now invoke Lemma 6.1, Lemma 3.2 and Lemma 3.1 to get

$$\Big\|\sup_k \Big|\sum_{n\leq N} e^{2\pi i\lambda_n x}\langle f,\phi_{I(J,k),n}\rangle w_{k,n}\Big|\Big\|_{L_x^p(J)} \leq$$

$$\leq \Big\|\sup_k \Big|\sum_{n\leq N} e^{2\pi i\lambda_n x}\langle f,\phi_{I(J,k),n}\rangle w_{k,n}\Big|\Big\|_{L_x^2(J)}$$

$$\lesssim N^{1/2-1/r}\max_n \|w_{k,n}\|_{V_k^r}\Big(\sum_{n=1}^N \|\langle f,\phi_{I(J,k),n}\rangle\|_{V_k^r}^2\Big)^{1/2}$$

$$= N^{1/2-1/r}\|w_\omega\|_{V^{r,*}}V_J f,$$

where $V_J f$ denotes the (constant) value of $Vf(x)$ on J.

By raising to the pth power and by summing up over all dyadic J of unit length, it suffices now to prove that

$$(6.1) \qquad \|Vf\|_{L^p(\mathbb{R})} \lesssim N^{\frac{1}{p}-\frac{1}{2}+\epsilon}(1 + \|m_\omega\|_{V^{r,*}})\|f\|_{L^p(\mathbb{R})}.$$

This however was proved in Theorem 5.1. ∎

References

[1] BOURGAIN, J.: Pointwise ergodic theorems for arithmetic sets. *Inst. Hautes Ètudes Sci. Publ. Math.* **69** (1989), 5–45.

[2] BOURGAIN, J.: Return times of dynamical systems. *C. R. Acad. Sci. Paris Sér. I Math.* **306** (1988), no. 12, 483–485.

[3] BOURGAIN, J.: Double recurrence and almost sure convergence. *J. Reine Angew. Math.* **404** (1990), 140–161.

[4] BOURGAIN, J., KOSTYUKOVSKY, S. AND OLEVSKII, A.: A remark on a maximal operator for Fourier multipliers. *Real Anal. Exchange* **26** (2000/01), no. 2, 901–904.

[5] CAMPBELL, J., JONES, R. L., REINHOLD, K. AND WIERDL, M.: Oscillation and variation for the Hilbert transform. *Duke Math. J.* **105** (2000), no. 1, 59–83.

[6] DEMETER, C.: Pointwise convergence of the ergodic bilinear Hilbert transform. *Illinois J. Math.* **51** (2007), no. 4, 1123–1158.

[7] DEMETER, C., TAO, T. AND THIELE, C.: Maximal multilinear operators. *Trans. Amer. Math. Soc.* **360** (2008), no. 9, 4989–5042.

[8] DEMETER, C., LACEY, M., TAO, T. AND THIELE, C.: Breaking the duality in the Return Times Theorem. *Duke Math. J.* **143** (2008), no. 2, 281–355.

[9] DEMETER, C. AND THIELE, C.: On the two dimensional bilinear Hilbert transform. *Amer J. Math.* **132** (2010), no. 1, 201–256.

[10] FEFFERMAN, C. AND STEIN, E. M.: Some maximal inequalities. *Amer. J. Math.* **93** (1971), 107–115.

[11] JONES, R. L., KAUFMAN, R., ROSENBLATT, J. M. AND WIERDL, M.: Oscillation in ergodic theory. *Ergodic Theory Dynam. Systems* **18** (1998), no. 4, 889–935.

[12] JONES, R. L., SEEGER, A. AND WRIGHT, J.: Strong variational and jump inequalities in harmonic analysis. *Trans. Amer. Math Soc.* **360** (2008), no. 12, 6711–6742.

[13] LACEY, M.: The bilinear maximal functions map into L^p for $2/3 < p \le 1$. *Ann. of Math. (2)* **151** (2000), no. 1, 35–57.

[14] LÉPINGLE, D.: La variation d'ordre p des semi-martingales. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **36** (1976), no. 4, 295–316.

[15] Muscalu, C., Tao, T. and Thiele, C.: Multi-linear operators given by singular multipliers. *J. Amer. Math. Soc.* **15** (2002), no. 2, 469–496.

[16] Muscalu, C., Tao, T. and Thiele, C.: Uniform estimates on multi-linear operators with modulation symmetry. *J. Anal. Math.* **88** (2002), 255–309.

[17] Rubio de Francia, J. L.: A Littlewood-Paley inequality for arbitrary intervals. *Rev. Mat. Iberoamericana* **1** (1985), no. 2, 1–14.

Ciprian Demeter
Department of Mathematics
Indiana University
831 East 3rd St.
Bloomington IN 47405, USA
demeterc@indiana.edu