Harmonic analysis, Carleson theorems, and multilinear analysis

Mathematical Research Community and Summer School *

Snowbird, Utah, June 27th - July 3rd 2009

Organizers:

Ciprian Demeter, University of Indiana, Bloomington

Michael Lacey, Georgia Institute of Technology

Christoph Thiele, University of California, Los Angeles

*supported by the AMS as part of the MRC program and by NSF grant DMS 0701302

Contents

1 Breaking the duality in the return times theorem (part II)

after Demeter, Lacey, Tao, and Thiele, [3]

A summary written by Michael Bateman

Abstract

We outline a method for controlling a model for the return times operator. For more information about how this model operator relates to the return times theorem, see the summary by Patrick LaVictoire in these conference proceedings.

1.1 Notation and prior results

For a sequence $(x_j)_j \subseteq \mathbb{C}$, define

$$||x_j||_{\tilde{V}_k^r} = \sup_{M,k_0,\ldots,k_M} \left(\sum_{m=1}^{M} |x_{k_m} - x_{k_{m-1}}|^r \right)^{\frac{1}{r}} \tag{1}$$

$$||x_j||_{V_k^r} = \sup |x_j| + ||x_j||_{\tilde{V}_k^r}. \tag{2}$$

These are called variational norms. We will use the fact that they satisfy the triangle inequality.

For a sequence of multipliers, f_k, and $1 \leq q \leq \infty$, define

$$||(f_k)_{k\in\mathbb{Z}}||_{M_q^*} = \sup_{||g||_q=1} \left|\left| \sup_k \int f_k(\xi)\hat{g}(\xi)e^{2\pi i z\xi} d\xi \right|\right|_{L_z^q}. \tag{3}$$

These are called maximal multiplier norms. We will use the fact that they satisfy the triangle inequality.

Given a set of points $\Lambda = \{\lambda_1, ..., \lambda_N\} \subseteq \mathbb{R}$, let R_k be the set of dyadic intervals of length 2^k such that $R_k \cap \Lambda \neq \emptyset$. For each $\omega \in R_k$, let m_ω be such that $\text{supp}(m_\omega) \subseteq \omega$ and such that for $n \in \{0, 1\}$ we have $||\partial^n m_\omega||_\infty \leq C|w|^{-n}$. We say that such an m_ω is C-adapted to ω. Let $\omega_k(\lambda)$ be the unique interval ω in R_k containing λ. Define

$$M_{\Lambda,k}(\xi) = \sum_{\omega\in R_k} m_w(\xi), \tag{4}$$

and

$$||m_\omega||_{V^{r,*}} = \sup_{\lambda\in\Lambda} || \left(m_{\omega_k(\lambda)}(\lambda)\right)_k ||_{V_k^r} \tag{5}$$

Theorem 1 (Demeter, [2])**.** *For $1 < q < 2$, $\epsilon > 0$, and $r > 2$, we have*

$$||(M_k)_{k\in\mathbb{Z}}||_{M_q^*} \lesssim |\Lambda|^{\frac{1}{q}-\frac{1}{r}+\epsilon}(C + . \tag{6}$$

This theorem is an extension of a similar result in [3], which is itself an extension of a result of Bourgain. We will use it below in Subsection 1.3. Here are some other results that we need:

Proposition 2. *If* $\mathbf{S}$ *is a collection of tiles, then*

$$\mathit{size}(\mathbf{S}, f) \lesssim \sup_{s\in\mathbf{S}} \inf_{x\in I_s} Mf(x) \tag{7}$$

We have used the notation

$$\text{size}(\mathbf{S}, f) = \sup_{s\in\mathbf{S}} \sup_{m_s} \frac{1}{|I_s|^{\frac{1}{2}}} \left|\left| \left(1 + \frac{|x - c(I_s)|}{|I_s|}\right)^{-10} \int \hat{f}(\xi)m(\xi)e^{2\pi i x\xi}d\xi \right|\right|_{L_x^2} \tag{8}$$

Later, we will consider a set $E = \{x\colon Mf(x) > \alpha\}$, and this lemma allows us to conclude that if $I_s \not\subseteq E$ for all $s \in \mathbf{S}$, then $\text{size}(\mathbf{S}, f) \lesssim \alpha$. The next proposition will help us sort a collection of tiles into subcollections each of which has uniform size.

Proposition 3. *Let* $\mathbf{S}$ *be a convex collection of tiles, and* $\delta = -\log_2(\mathit{size}(\mathbf{S}, f))$*, Then we may partition* $\mathbf{S} = \bigcup_{n\geq\delta} \mathcal{P}_n$*, so that* $\mathit{size}(\mathcal{P}_n, f) \leq 2^{-n}$*, and so that each* $\mathcal{P}_n$ *may be partitioned as* $\mathcal{P}_n = \bigcup_{\mathbf{T}\in\mathcal{F}_n} \mathbf{T}$*, where each* $\mathbf{T}$ *is a tree and* $\sum_{\mathbf{T}\in\mathcal{F}_n} |I_\mathbf{T}| \lesssim 2^{2n}||f||_2^2$*. We have used* $I_\mathbf{T}$ *to denote the space interval of the top of* $\mathbf{T}$*.*

For what follows, we assume ϕ_s and φ_s denote Schwartz functions associated to a tile s such that

$$\varphi_s\colon \mathbb{R} \to \mathbb{R} \tag{9}$$

$$\phi_s\colon \mathbb{R}^2 \to \mathbb{R} \text{ and } \text{supp}_\theta\phi_s(x, \theta) \subseteq \omega_s, \tag{10}$$

These functions will also satisfy some standard technical decay conditions. Also assume that for each $l \in \mathbb{N}$, for any tree $\mathbf{T} = (I_\mathbf{T}, \xi_\mathbf{T})$, and for each $s \in \mathbf{T}$, we have a decomposition $\phi_{s,\mathbf{T}} = \phi_{s,\mathbf{T}}^{(l)} + \tilde{\phi}_{s,\mathbf{T}}^{(l)}$ such that $\text{supp}_x\tilde{\phi}_s^{(l)}(x, \theta) \subseteq 2^{l-1}I_s$. Further assume that φ_s and $\phi_s^{(l)}$ are such that the following proposition holds:

Proposition 4. *For $1 < t < \infty$, and for any tree* $\mathbf{T}$*,*

$$\left|\left|\left|\left|\left(\sum_{s\in\mathbf{T},|I_s|=2^k} < f,\varphi_s > \phi^{(l)}_{s,\mathbf{T}}(x,\theta)\right)_k\right|\right|_{V^r_k}\right|\right|_{L^t_x} \lesssim ||f||^2_2 size(\mathbf{T},f)|I_{\mathbf{T}}|^{\frac{1}{t}}. \quad (11)$$

The properties of the functions ϕ_s and φ_s that arise from modeling the return times operator have the properties assumed above. The need for splitting ϕ_s into two parts is so that one can use the proposition above. This should not be clear from what has been said here, but for reasons of space, we explain no further.

1.2 Main result

For a set of tiles $\mathbf{S}$, define

$$V_{\mathbf{S}}f(x) = \left|\left|\left(\sum_{s\in\mathbf{S},|I_s|=2^k} < f,\varphi_s > \phi_s(s,\theta)\right)_k\right|\right|_{M^*_{q,\theta}} \quad (12)$$

Later we will use the fact that if $\mathbf{S}$ is partitioned into $\mathbf{S}_1$ and $\mathbf{S}_2$, then $V_{\mathbf{S}}f(x) \leq V_{\mathbf{S}_1}f(x) + V_{\mathbf{S}_2}f(x)$. We can now state the main theorem. This is proved in [3] for $p = 2$ and in [1] for the full range stated here.

Theorem 5. *If φ_s, ϕ_s have the properties above, if $\mathbf{S}$ is a convex collection of tiles, if $0 < \lambda \leq 1$,if $\delta > 0$, if $F \subseteq \mathbb{R}$, if $1 < p < 2$ and $\frac{1}{p} + \frac{1}{q} < \frac{3}{2}$, then*

$$|\{x\colon V_{\mathbf{S}}1_F(x) > \lambda\}| \lesssim \frac{|F|}{\lambda^{p+\delta}}. \quad (13)$$

To prove this theorem, we define several exceptional sets, show that they are small, and show that $V_{\mathbf{S}}\mathbf{1}_F(x)$ is small outside these sets. We sketch this in slightly more detail below, ignoring numerological issues. Fix λ. Define $E = \{x\colon M1_f(x) > \lambda\}$. Assume that for all $s \in \mathbf{S}$, $I_s \cap E^c \neq \emptyset$. (Of course this will not always be true, but a similar argument handles that case as well. Then one can use the sublinearity of $V_{\mathbf{S}}$ in the set of tiles $\mathbf{S}$, as noted above.) By Proposition 2, we know that $\text{size}(\mathbf{S}, 1_F) \lesssim \lambda$. By Proposition 3, if $\delta = -\log \text{size}(\mathbf{S}, 1_F)$, we may write $\mathbf{S} = \bigcup_{n\geq\delta}\mathcal{P}_n$ where

$$\sum_{\mathbf{T}\in\mathcal{F}_n} |I_{\mathbf{T}}| \lesssim 2^{2n}||f||^2_2. \quad (14)$$

Now for each n, we define

$$E_n^{(1)} = \bigcup_{l\geq 0} \{x\colon \sum_{\mathbf{T}\in\mathcal{F}_n} 1_{2^l I_T}(x) > \beta_n 2^{2l}\} \tag{15}$$

$$E_n^{(2)} = \bigcup_{l,m\geq 0} \bigcup_{\mathbf{T}} \{x\colon || \sum_{s\in\mathbf{T},|I_s|=2^k} < f, \varphi_s > \phi_{s,\mathbf{T}}^{\alpha(l,m)}(x,\xi_{\mathbf{T}})||_{V_k^r} > \gamma_n 2^{-10l} \frac{1}{(|m|+1)^2}\}, \tag{16}$$

for some appropriate β_n and γ_n, and where the second union in the definition of $E_n^{(2)}$ is over an appropriate collection of trees. We use the notation $\alpha(l,m) = 0$ if $m \in \{-1,0,1\}$ and $\alpha(l,m) = l + \log_2 m$ otherwise. Notice that we can control $|E_n^{(1)}|$ just by using Chebyshev's inequality and the estimate in 14. Further, we can control $|E_n^{(2)}|$ by using Proposition 4. These estimates are important, since they now allow us to focus on $V_{\mathbf{S}}1_F(x)$ for $x \notin \bigcup_n (E_n^{(1)} \cup E_n^{(2)})$. This is the topic of the next section.

1.3 Pointwise estimates outside exceptional sets

For the following theorem, we need to introduce some notation. Let $\mathbf{S} = \bigcup_{\mathbf{T}\in\mathcal{F}} \mathbf{T}$ be a collection of tiles partitioned into families of trees. For $\mathbf{T} \in \mathcal{F}$, let $G(\mathbf{T}) = \{s \in \mathbf{S}\colon \omega_s \supseteq \omega_T\}$. Now fix $l \in \mathbb{N}$. For a tree $\mathbf{T}$, define

$$\mathbf{T}_{l,0} = \{s \in G(\mathbf{T})\colon I_s \cap 2^l I_T\} \tag{17}$$

$$\mathbf{T}_{l,m} = \{s \in G(\mathbf{T})\colon I_s \cap (2^l I_T + m2^l I_T \text{ but } s \notin \bigcup_{|j|<m} \mathbf{T}_{l,m}\}. \tag{18}$$

Also define $\mathcal{F}_{l,m} = \{\mathbf{T}_{l,m}\colon \mathbf{T} \in \mathcal{F}\}$.

Theorem 6. *Use assumptions and notation from above. Assume for each tile s we have a number a_s and that $\frac{a_s}{|I_s|^{\frac{1}{2}}} \leq \sigma$. Define*

$$E^{(1)} = \bigcup_{l\geq 0} \{x\colon \sum_{\mathbf{T}\in\mathcal{F}} 1_{2^l I_T}(x) \geq \beta 2^{2l}\} \tag{19}$$

$$E^{(2)} = \bigcup_{l,m\geq 0} \bigcup_{\mathbf{T}\in\mathcal{F}_{l,m}} \{x\colon \left|\left| \sum_{s\in\mathbf{T},|I_s|=2^k} a_s \phi_{s,\mathbf{T}}^{\alpha(l,m)}(x,\xi_{\mathbf{T}}) \right|\right|_{V_k^r} > \gamma 2^{-10l} \frac{1}{(|m|+1)^2}\}. \tag{20}$$

Then for $\epsilon > 0$, if $x \notin E^{(1)} \cup E^{(2)}$, then

$$\left|\left|\left(\sum_{s\in\mathbf{S},|I_s|=2^k} a_s\phi_s(x,\theta)\right)_k\right|\right|_{M^*_{q,\theta}} \lesssim \beta^{\frac{1}{q}-\frac{1}{r}+\epsilon}(\gamma+\sigma). \tag{21}$$

Recall that proving this estimate for x outside $E^{(1)}$ and $E^{(1)}$ is exactly what we need to finish proving Theorem 5. So fix an $x \notin E^{(1)} \cup E^{(2)}$. We will partition the tiles $\mathbf{S} = \bigcup_l \mathcal{P}_{l,x}$, where $\mathcal{P}_{l,x}$ is defined below:

$$\mathcal{F}_{0,x} = \{\mathbf{T}\colon x \in I_T\} \tag{22}$$

$$\mathcal{F}_{l,x} = \{\mathbf{T}\colon x \in 2^l I_T \setminus 2^{l-1} I_T\} \tag{23}$$

$$\mathcal{P}_{l,x} = \left(\bigcup_{\mathbf{T}\in\mathcal{F}_{l,x}} G(\mathbf{T}\right) \setminus \bigcup_{0\le j<l} \mathcal{P}_{l,x}. \tag{24}$$

Let $m_{\omega,l}(\theta) = \sum_{s\in\mathcal{P}_{l,x},\omega_s=\omega} a_s\phi_s(x,\theta)$. Note that it is supported in ω and is adapted with constant $2^{-10l}\sigma$, by assumption on a_s and by decay properties of ϕ_s. Let

$$M_{k,l}f(x) = \sum_{\omega\colon |\omega|=2^{-k}} m_{\omega,l}(\theta) \tag{25}$$

$$M_k f(x) = \sum_l M_{k,l}f(x). \tag{26}$$

Note $||M_k f(x)||_{M^*_{q,\theta}} \le \sum_l ||M_{k,l}f(x)||_{M^*_{q,\theta}}$, so we will focus on $||M_{k,l}f(x)||_{M^*_{q,\theta}}$ for a fixed l. Note that if $s \in \mathcal{P}_{l,x}$, then there is a tree $\mathbf{T} \in \mathcal{F}_{l,x}$ such that $\omega_s \ni c(\omega_T)$. So let $\Lambda_{x,l} = \{c(\omega_T)\colon \mathbf{T} \in \mathcal{F}_{l,x}\}$. Then

$$|\Lambda_{x,l}| \le \sum_{\mathbf{T}\in\mathcal{F}_{l,x}} 1_{2^l I_T(x)} \le \beta 2^{2l} \tag{27}$$

by our assumption on x. For a given tree $\mathbf{T}$, denote

$$W_{\mathbf{T}} = \left|\left|\sum_{s\in\mathcal{P}_{l,x}\cap G(\mathbf{T})} a_s\phi_s(x,\theta)\right|\right|_{V^r_k}. \tag{28}$$

We now apply the maximal multiplier theorem (Theorem 1) with the set $\Lambda_{x,l}$ and the functions $(M_{k,l})_k$ to see that

$$||M_{k,l}f(x)||_{M^*_{q,\theta}} \leq \beta^{\frac{1}{q}-\frac{1}{r}+\epsilon} 2^{2l}(2^{-10l}\sigma + \sup_{\mathbf{T}\in\mathcal{F}_{l,x}} W_{\mathbf{T}}). \tag{29}$$

Finally, we control the numbers $W_{\mathbf{T}}$:

$$\begin{aligned} W_{\mathbf{T}} &\leq \sum_{m=0}^{\infty} \left|\left| \sum_{s\in\mathcal{P}_{l,x}\cap\mathbf{T}_{l,m}} a_s\phi_s(x,\theta) \right|\right|_{V_k^r} \\ &= \sum_{m=0}^{\infty} \left|\left| \sum_{s\in\mathcal{P}_{l,x}\cap\mathbf{T}_{l,m}} a_s\phi_{s,\mathbf{T}}^{\alpha(l,m)}(x,\theta) \right|\right|_{V_k^r} \\ &\leq \sum_{m=0}^{\infty} \left|\left| \sum_{s\in\mathbf{T}_{l,m}} a_s\phi_{s,\mathbf{T}}^{\alpha(l,m)}(x,\theta) \right|\right|_{V_k^r}. \end{aligned} \tag{30}$$

This final quantity is well-controlled because $x \notin E^{(2)}$. We are able to replace $\phi_s(x,\theta)$ with $\phi_{s,\mathbf{T}}^{\alpha(l,m)}(x,\theta)$ because of the support of $\phi_s - \phi_{s,\mathbf{T}}^{\alpha(l,m)}$. The last inequality holds because if

$$\{s \in \mathcal{P}_{l,x} \cap \mathbf{T}_{l,m} \colon |I_s| = 2^k\} \neq \{s \in \mathbf{T}_{l,m} \colon |I_s| = 2^k\}, \tag{31}$$

then in fact the left side is empty. This is helpful because it allows us to eliminate the dependence on x of the set of tiles in the final sum of 30. This is the reason for organizing the tiles into the collections $\mathcal{P}_{l,x}$ as we did.

References

[1] Demeter, C., *Improved range in the Return Times Theorem* , available at http://arxiv.org/abs/0904.1254

[2] Demeter, C., *On some maximal multipliers in L^p,* available at http://arxiv.org/abs/0901.4084

[3] Demeter C., Lacey M., Tao T., Thiele C., *Breaking the duality in the return times theorem.* Duke Math. J. **143** (2008), no. 2, 281-355.

Michael Bateman, Indiana University / UCLA
email: **mdbatema@indiana.edu**

2 On the multilinear restriction and Kakeya conjectures

after J. Bennett, A. Carbery, and T. Tao [1]
A summary written by Matthew Bond

Abstract

Here we will summarize the main ideas of Bennett, Carbery, and Tao's paper, **On the multilinear restriction and Kakeya conjectures** [1].

2.1 Introduction - The linear and multilinear Kakeya conjectures

An unsolved problem, the (linear) **Kakeya conjecture**, states that any compact set in $\mathbb{R}^d$([1]) containing line segments in every direction must have full Hausdorff dimension. Suppose we have a proposed counterexample E, and choose from it a collection of line segments in δ-separated directions. To each such line segment corresponds a δ tube (i.e., a δ-neighborhood of that line segment). Then one can investigate the L^p norm of the sum of the corresponding characteristic functions. If it were possible to cleverly construct the set E so that the tubes overlap so much that the L^p norm is large, then one might expect the Kakeya conjecture to be false. On the other hand, if for some ranges of p the L^p norm always behaves on the same order as if the tubes all went through the origin, then it seems that a uniformly high overlap is impossible, and so the Kakeya conjecture is probably true. Indeed, the Kakeya conjecture has an equivalent formulation in this spirit:

The linear Kakeya conjecture: Let $\delta << 1$. Let $\mathbf{T}$ be a collection of δ-tubes T, pointing in directions δ-separated on S^{d-1}. Then for each $p > \frac{d}{d-1}$,([2]) we have

$$|| \sum_{T \in \mathbf{T}} \chi_T||_p^p \leq C_p \delta^{d-1} \cdot (\#\mathbf{T})^{p-\frac{1}{d-1}}.$$

The object of study in [1] is a related **multilinear Kakeya conjecture**. It is similar in spirit, but substantially different in that the linear Kakeya

[1] Here, and in the entire paper, we will be working in $\mathbb{R}^d$ with $d \geq 2$.

[2] One can show that $p < \frac{d}{d-1}$ is impossible just by letting the tubes be disjoint.

conjecture is not a special case, for example. In the multilinear case, we choose families $\mathbf{T}_j$ of tubes, $j = 1, ..., d$, with each tube $T_{j,l} \in \mathbf{T}_j$ pointing in directions a small distance ($\leq \varepsilon_0$, a small enough absolute constant maybe depending on d) away from the j-th standard basis element e_j. Unlike the linear case, we allow tubes to be repeated. As before, we sum the characteristic functions in each family, but now we also take the product over the families, and what one gets is

The multilinear Kakeya conjecture: In the setting of the previous paragraph, one has

$$|| \prod_{j=1}^{d} \sum_{l=1}^{\#\mathbf{T}_j} \chi_{T_{j,l}} ||_p^p \leq C_p \delta^d (\prod_{j=1}^{d} \#\mathbf{T}_j)^p$$

whenever $p \geq \frac{1}{d-1}$.

Again, the upper bound is true if, up to a constant, the largest L^p norm is attained by letting all tubes pass through the origin, so that it is not possible to have lots of overlap on a large set. [1] proves the conjecture when $p > \frac{1}{d-1}$. One may also replace the e_j by another basis, changing only C_p and ε_0, by linear algebra considerations. In contrast to the linear case, taking the product allows us to simplify by assuming the tubes to have infinite length. We will make this assumption from now on.

2.2 The joints problem

The boundary exponent of the multilinear Kakeya conjecture easily be shown to be necessary by letting the function take the value 1 on a large set. In this case, one can allow the tubes to align along the integer grid $[1, (\frac{N}{d})^{\frac{1}{d-1}}]^d$. This integer grid is the set of **joints** of the family $\bigcup_{j \leq d} \mathbf{T}_j$, i.e., the set of points x for which d distinct tubes have their axes pass through x in linearly independent directions. Conversely, for larger p, the multilinear Kakeya conjecture puts a bound on the number of joints, subject to transversality conditions. A weak form of this can be stated and easily proved by comparing the multilinear Kakeya estimate with the lower bound $\#joints \cdot \delta^d$, since from above we have $\#\mathbf{T}_j \leq N$, the total number of tubes. For simplicity, we limit our attention to the case $d = 3$ (so $p = \frac{1}{2} + \varepsilon$), where we can state it as:

Watered-down joints theorem: Let the lines $\ell_{j,l}$ be the axes of tubes as in the multilinear Kakeya conjecture, and let $\sum_j \#\mathbf{T}_j = N$. Then the number of joints obtained by taking a single tube $T_{j,l}$ from each $\mathbf{T}_j$ is no more than $C_\varepsilon N^{3/2+\varepsilon}$.

In order to bootstrap this result into something more interesting, we need to define transversality. We say that a basis $v_1, ..., v_d$ has transversality $\theta = |det(v_1...v_d)|$, and let this also be the definition of the transversality of the joint formed by d lines in the directions v_j.

Joints theorem: Let $\mathbf{T}$ be an arbitrary set of δ tubes, let $\theta_0 > 0$, and let $N = \#\mathbf{T}$. Let $J(\theta_0)$ be the set of (distinct) joints with transversality $\theta \geq \theta_0$. Then for all $\varepsilon > 0$,

$$J(\theta_0) \leq C_\varepsilon N^{3/2+\varepsilon} \theta_0^{-1/2-\varepsilon}.$$

The presence of ε corresponds to the fact that the $p = \frac{1}{d-1}$ case of the multilinear Kakeya conjecture was unknown at the time of [1].

2.3 The Loomis-Whitney inequality

If we limit ourselves to the case $\varepsilon_0 = 0$, the multilinear Kakeya conjecture is already known to be true, and it can be seen to imply the **Loomis-Whitney** inequality [2] (up to a constant). Let $\pi_j : \mathbb{R}^d \to \mathbb{R}^{d-1}$ act by deleting the j-th coordinate. Then if $f_j \in L^{d-1}(\mathbb{R}^{d-1})$ and we define $\tilde{f}_j = f_j \circ \pi_j$, the Loomis-Whitney inequality states:

$$|| \prod_{j=1}^{d} \tilde{f}_j ||_{L^1(\mathbb{R}^d)} \leq \prod_{j=1}^{d} ||f_j||_{L^{d-1}(\mathbb{R}^{d-1})}.$$

Indeed, one can let $\tilde{f}_j = (\sum_{T_{j,l} \in \mathbf{T}_j} \chi_{T_{j,l}})^p$ (on the support of the integrand) in the multilinear Kakeya conjecture inequality at the endpoint $p = \frac{1}{d-1}$. Letting δ be small, one can use approximations by characteristic functions to get the Loomis-Whitney inequality up to a constant for arbitrary $f_j \in L^P(\mathbb{R}^{d-1})$.

Since $\varepsilon_0 = 0$ is not, in general, required, the multilinear Kakeya conjecture is also called a **perturbed Loomis-Whitney inequality**.

2.4 Proof of the (unperturbed) Loomis-Whitney inequality (and more)

Most of the main ideas of [1] are on display even when we require $\varepsilon_0 = 0$, i.e., that the tubes $T_{j,l} \in \mathbf{T}_j$ be parallel. So let us consider this case now, and leave the reader to consult [1] for the perturbed case. The main idea is monotonicity - one compares an arbitrary configuration of tubes in the given directions to another in which all tubes are centered at the origin, by moving between the two cases linearly with respect to a time variable t. Namely, as the tubes slide away from the origin, one might hope for the L^p norm to decrease. [1] shows that this happens, if not literally, then certainly after one estimates the characteristic functions of tubes from above by smoother functions which exhibit Gaussian decay as one moves in any direction orthogonal to the tube's axis. The main convenience of Gaussians comes from working with a product formula and adding exponents. The characteristic functions $\chi_{T_{j,l}}$, then, can be thought of as

$$\chi_{T_{j,l}}(x) = e^{-\pi < \pi_j(x - v_{j,l}), \pi_j(x - v_{j,l}) >},$$

where the $v_{j,l}$ give the location of the tube. It is also natural to think of π_j as being given by the corresponding diagonal matrix A_j^0 having 1's on the diagonal, except for at the j-th diagonal entry, which is 0, and to only write A_j^0 in the first slot of the inner product.

Let us single out the properties of A_j^0 which make the desired inequality possible by the proof of [1], and refer to any such set of matrices as **good**. The upshot is that one can prove a much more general multilinear inequality, with the level of multilinearity n independent of the dimension d, and the positive real n-vector of exponents p allowed to vary its entries independently subject to linear algebraic conditions.

First, a definition: For $d \times d$ matrices A and B, we will say that $A \geq B$ if $A - B$ is positive semi-definite, and $A > B$ if $A - B$ is positive definite.([3])

Let $p \in (0, \infty)^n$, for some integer $n > 0$. A p-good collection of matrices, then, is defined to be a collection $A_1, ..., A_n$ of positive semi-definite real symmetric $d \times d$ matrices such that $A_k \leq \sum_j p_j A_j =: A_*$ for all $k = 1, ..., n$ and $\bigcap_j ker(A_j) = \{0\}$([4]). We can see that $\{A_j^0\}_{j=1}^d$ is p-good for $p_1 = ... = p_d = \frac{1}{d-1}$, where $d = n$.

Finally, let us introduce the time dependence to the formula, and generalize once more by replacing finite sums by integration with respect to finite, compactly supported Borel measures μ_j. Our "sum of characteristic

[3] We will never have occasion to employ any other partial ordering on matrices.

[4] This shows that $\sum p_j A_j$ is nonsingular

functions", then, becomes

$$f_j(t,x) = \int_{\mathbb{R}^d} e^{-\pi < A_j(x_j - v_j t),(x_j - v_j t)>} d\mu_j(v_j).(^5) \tag{1}$$

This is a sum in the usual sense if the measure μ_j is a sum of delta measures.

Theorem 1. *Let $n > 0$ be an integer, and let $p \in (0,\infty)^n$. Let A_j, $j = 1,...,n$, be a p-good collection of matrices, and let μ_j be finite, compactly supported Borel measures on $\mathbb{R}^d$. Define functions f_j as in 1. Then the quantity*

$$Q_p(t) := \int_{\mathbb{R}^d} \prod_{j=1}^{n} f_j(t,x)^{p_j} dx \tag{2}$$

is nonincreasing in time t when $t > 0$.

Proof. One shows that $Q'_p(t) < 0$. To do this, we first assume that p is a vector of integers, and then show that a certain representation formula for $Q'_p(t)$ must in fact be valid for all $p \in (0,\infty)^n$ by a polynomial density lemma (8.2 of [1]). When each p_j is an integer, we may find each f_j^p by integrating p_j independent copies of each integrand, i.e.,

$$Q_p(t) = \int_{\mathbb{R}^d} \int_{(\mathbb{R}^d)^{p_1}} \cdots \int_{(\mathbb{R}^d)^{p_n}} e^{-\pi \sum_{j=1}^{n} \sum_{k=1}^{p_j} < A_j(x - v_{j,k}t),(x - v_{j,k}t)>} \prod_{j=1}^{n} \prod_{k=1}^{p_j} d\mu(v_{j,k}) dx. \tag{3}$$

Complete the square to write

$$\sum_{j=1}^{n} \sum_{k=1}^{p_j} < A_j(x - v_{j,k}t),(x - v_{j,k}t) > = < A_*(x - \bar{v}t),(x - \bar{v}t) > + \delta t^2, \tag{4}$$

where

$$\bar{v} := A_*^{-1} \sum_{j=1}^{n} A_j \sum_{k=1}^{p_j} v_{j,k}$$

is an average weighted velocity,

$$\delta := \sum_{j=1}^{n} \sum_{k=1}^{p_j} < A_j v_{j,k}, v_{j,k} > - < A_* \bar{v}, \bar{v} >,$$

[5] The subscript l is no longer needed since v_j is a function with respect to the (possibly discrete) measure μ_j.

and recalling $A_* := \sum_{j=1}^n p_j A_j$. Putting 4 into 3 and changing the x variable, one sees that the time dependence is captured entirely by the δt^2 term, i.e., the integrand takes the form $\Phi(x)e^{-\pi\delta t^2}$, where δ depends on the variables $v_{j,k}$. One gets

$$Q_p'(t) =$$

$$2\pi t \int_{\mathbb{R}^d} \int_{(\mathbb{R}^d)^{p_1}} \cdots \int_{(\mathbb{R}^d)^{p_n}} \delta e^{-\pi \sum_{j=1}^n \sum_{k=1}^{p_j} <A_j(x-v_{j,k}t),(x-v_{j,k}t)>} \prod_{j=1}^n \prod_{k=1}^{p_j} d\mu(v_{j,k}) dx.$$

One can clean this up with probabilistic language by treating the $v_{j,k}$ as random variables with the measures

$$\frac{e^{-\pi<A_j(x-v_{j,k}),(x-v_{j,k})>} d\mu_j(v_{j,k})}{f_j(t,x)}.$$

Then using expectation with respect to the product, we can write

$$Q_p'(t) = -2\pi t \int_{\mathbb{R}^d} \mathbf{E}(\delta) \prod_{j=1}^n f_j(t,x)^{p_j} dx. \tag{5}$$

Some linear algebra shows that

$$\mathbf{E}(\delta) = \sum_{j=1}^n p_j \mathbf{E}(< (A_j - A_j A_*^{-1} A_j)(v_j - \mathbf{E}(v_j), (v_j - \mathbf{E}(v_j))) >) \tag{6}$$

$$+ \sum_{j=1}^n p_j < A_j(\mathbf{E}(v_j - \mathbf{E}(\bar{v}))), (\mathbf{E}(v_j - \mathbf{E}(\bar{v}))) > .$$

Additional considerations show that if we multiply $\mathbf{E}(\delta)$ by $det(A_*)$, the dependence of this expression on p is polynomial **for general** p, and lemma 8.2 from [1] shows that formula 5 holds under the interpretation 6 for **all** $p \in (0,\infty)^n$. Multiplying $A_* \geq A_j$ on the right by $A_*^{-1}A_j$, one gets positivity of $\mathbf{E}(\delta)$, and the result follows. □

References

[1] J. Bennett, A. Carbery, and T. Tao, *On the multilinear restriction and Kakeya conjectures*, *Acta Math.* **196** (2006), 261–302.

[2] L.H. Loomis and H. Whitney, *An inequality related to the isoperimetric inequality*, *Bull. Amer. Math. Soc* **55** (1949), 961-962.

Matthew Bond, MSU
email: **bondmatt@msu.edu**

3 Boundedness of the Carleson operator

after M. Lacey and C. Thiele [1]
A summary written by Alberto A. Condori

Abstract

We outline a proof of the weak type L^2 estimate for the maximal Carleson operator based on time-frequency analysis due to M. Lacey and C. Thiele.

3.1 Introduction

For any function $f \in L^1$ on $\mathbb{R}$, we define the *Fourier transform* $\hat{f}$ *of* f by $\hat{f}(\xi) = \int_{\mathbb{R}} f(x)e^{-2\pi i\xi x}dx$ and denote the operator $f \mapsto \hat{f}$ by $\mathcal{F}$.

It is well known that the Fourier transform (operator) $\mathcal{F}$ is isometric on $L^1 \cap L^2$ under the L^2 norm and $L^1 \cap L^2$ is dense in L^2. Hence, $\mathcal{F}$ admits a unique extension to a unitary operator on L^2 and

$$f(x) = \lim_{N\to\infty} \int_{-N}^{N} \hat{f}(\xi)e^{2\pi ix\xi}d\xi \tag{1}$$

holds for each $f \in L^2$ in the sense of L^2 convergence. In particular,

$$f(x) = \int_{\mathbb{R}} \hat{f}(\xi)e^{2\pi i\xi x}d\xi \ \text{ for a.e. } x \in \mathbb{R} \tag{2}$$

holds for $f \in L^2$ such that $\hat{f} \in L^1$ by (1) and the dominated convergence theorem. Thus, it seems natural to ask whether (1) holds a.e. for any $f \in L^2$.

In view of the density of the Schwartz space $\mathcal{S}$ (of rapidly decreasing functions on $\mathbb{R}$) in L^2, it seems natural to use the result (2) for $\mathcal{S}$ to establish the a.e. convergence in (1) for arbitrary $f \in L^2$. This reasoning leads one to

Theorem 1. *The a.e. convergence in* (1) *holds for arbitrary* $f \in L^2$ *if there is a constant* $C > 0$ *such that*

$$\|\mathcal{C}f\|_{2,\infty}^2 := \sup_{\lambda>0} \lambda^2 |\{x \in \mathbb{R} : \mathcal{C}f(x) > \lambda\}| \leq C\|f\|_2^2, \quad \textit{for all } f \in L^2, \tag{3}$$

where $\mathcal{C}$ *denotes the double-sided* Carleson operator $\mathcal{C}$*,*

$$\mathcal{C}f(x) = \sup_{N>0} \left| \int_{-N}^{N} \hat{f}(\xi)e^{2\pi ix\xi}d\xi \right| \quad \textit{for } f \in L^2.$$

In this note, we provide an outline of the time-frequency analysis technique used in [1] to prove the validity of the weak-type estimate in (3).

3.2 Notation and preliminary reductions

Let us (formally) define translation, modulation, and dilation operators by

$$\operatorname{Tr}_\lambda f(x) = f(x-\lambda),\ \operatorname{Mod}_\lambda f(x) = f(x)e^{2\pi i x\lambda} \text{ and } \operatorname{Dil}_\lambda f(x) = \lambda^{-1/2} f(x/\lambda).$$

Clearly, $\mathcal{F}\operatorname{Tr}_\lambda = \operatorname{Mod}_{-\lambda}\mathcal{F}$, $\mathcal{F}\operatorname{Mod}_\lambda = \operatorname{Tr}_\lambda\mathcal{F}$ and $\mathcal{F}\operatorname{Dil}_\lambda = \operatorname{Dil}_{1/\lambda}\mathcal{F}$.

We write $|E|$ for the Lebesgue measure a set $E \subset \mathbb{R}$. Given an interval I (i.e. a subset of $\mathbb{R}$ of the form $[x, y)$ with $x < y$) and $\alpha > 0$, we write $c(I)$ for the center of I, and αI for the interval with center $c(I)$ and length $\alpha|I|$.

An interval is called *dyadic* if it is of the form $[n2^k, (n+1)2^k)$ for some integers k and n. Dyadic intervals have three useful properties: Any dyadic interval is the disjoint union of two dyadic intervals of half length; if two dyadic intervals intersect, then one is contained in the other; and, for each $x \in \mathbb{R}$ and $k \in \mathbb{Z}$, there is only one dyadic interval I with $|I| = 2^k$ and $x \in I$.

Throughout, we think of the first coordinate x as the time interval and the second coordinate ξ as the frequency coordinate. Thus, we refer to (x, ξ) coordinate plane as the *time-frequency plane*.

Let $\mathcal{T}$ denote the collection of rectangles $I \times w$ with I and w dyadic and $|I||w| = 1$. Any $s = I_s \times w_s \in \mathcal{T}$ is called a *tile*.[6]

Let $\varphi \in \mathcal{S}$ such that $\chi_{[-1/10,1/10]} \le \hat{\varphi} \le \chi_{[-1/9,1/9]}$. For each tile $s = I_s \times w_s$, we define the "lower and upper halves" $w_{1s} := w \cap (-\infty, c(w_s))$ and $w_{2s} := w \cap (c(w_s), \infty)$ of w_s and the function $\varphi_s = \operatorname{Mod}_{c(w_{2s})}\operatorname{Tr}_{c(I_s)}\operatorname{Dil}_{|I_s|}\varphi$. We observe that $\hat{\varphi}_s$ is supported in $\frac{1}{2}w_{1s}$ and

$$|\varphi_s(x)| \le C_\nu |I_s|^{-1/2}(1 + |x - c(I_s)|/|I_s|)^{-\nu}, \nu \ge 0,$$

holds[7], because $(1+|x|)^\nu \varphi(x)$ is bounded for each $\nu \ge 0$ as $\varphi \in \mathcal{S}$.

For technical reasons, we consider (instead of $\mathcal{C}$) the *one-sided Carleson operator* $\mathcal{C}_*$ defined by $\mathcal{C}_* f(x) = \sup_N \left| \int_{-\infty}^N \hat{f}(\xi) e^{2\pi i x\xi} d\xi \right|$ for $f \in L^2$. Moreover, to prove the weak-L^2 bound for $\mathcal{C}_*$, we consider instead the dyadic model[8]:

$$A_\xi f := \sum_{s\in\mathcal{T}} \chi_{w_{2s}}(\xi) \langle f, \varphi_s\rangle \varphi_s \tag{4}$$

[6]Notice that our rectangles $I_s \times w_s$ have area 1 so that $\mathcal{F}\operatorname{Dil}_{|I_s|} = \operatorname{Dil}_{|w_s|}\mathcal{F}$.

[7]Note that the constant C_ν depends only on the choice of φ and ν.

[8]A_ξ is well-defined in $\mathcal{S}$: In fact, the series in (4) converges absolutely for $f \in \mathcal{S}$.

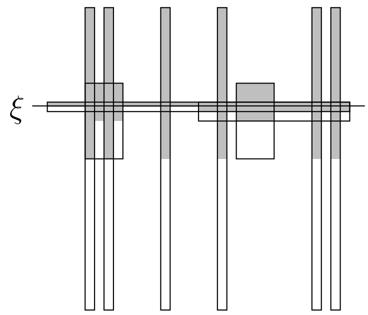

Figure 1: Some of the tiles s that contribute to the sum for A_ξ in the time-frquency plane. The shaded areas are the tiles $I_s \times \omega_{2s}$.

(See Figure 3.2.) After all, the weak-L^2 bound for $\mathcal{C}_*$ follows from

$$\|\sup_\xi |A_\xi f|\|_{2,\infty} \le C\|f\|_2 \text{ for all } f \in L^2. \tag{5}$$

By a duality argument, it can be seen that the estimate (5) follows from

$$\sum_{s\in\mathcal{P}} |\langle f, \varphi_s\rangle \langle \varphi_s(\chi_{w_{2s}} \circ N), \chi_E\rangle| \le C\|f\|_2|E|^{1/2} \tag{6}$$

for all $f \in \mathcal{S}$, measurable functions N, measurable sets E, and finite subsets $\mathcal{P}$ of $\mathcal{T}$. Hence, to prove (1) holds a.e., it suffices to prove (6).

3.3 The main argument

Let us fix an $f \in \mathcal{S}$, a measurable function N, a measurable set E, and finite subset $\mathcal{P}$ of $\mathcal{T}$. In view of the possible overlapping of tiles in $\mathcal{P}$, it seems natural to order tiles as follows: We say $s < s'$ if $I_s \subset I_{s'}$ and $w_{s'} \subset w_s$.

A set of tiles T is called a *tree* if there is a tile[9] $s_T = I_T \times w_T$, the *top of the tree* T, such that $s < s_T$ for all $s \in T$. A tree T is called a j-tree if $w_{js_T} \subset w_{js}$ for all $s \in T$.

In order to prove (6), we first make some definitions and establish some needed results.

Definition 2. *Let N be a measurable function on $\mathbb{R}$, E a set of finite measure and $\mathcal{P} \subset \mathcal{T}$. The* mass of E with respect to $\mathcal{P}$ *is defined by*

$$\text{mass}(E;\mathcal{P}) = \frac{1}{|E|}\sup_{s\in\mathcal{P}}\sup_{t\in\mathcal{T},s<t}\int_{E\cap N^{-1}[\omega_t]} \frac{1}{\left(1+\frac{|x-c(I_t)|}{|I_t|}\right)^{10}}\frac{dt}{|I_t|}.$$

[9]Note that s_T may not belong to T.

Theorem 3. *Let N be a measurable function on $\mathbb{R}$ and E a set of finite measure. If $\mathcal{P}$ is a finite set of tiles, then $\mathcal{P}$ can be decomposed as the union of sets $\mathcal{P}_{light}$ and $\mathcal{P}_{heavy}$ with*

$$\text{mass}(E; \mathcal{P}_{light}) \le 2^{-2} \, \text{mass}(E; \mathcal{P})$$

and $\mathcal{P}_{heavy}$ is the union of trees T_j such that

$$\sum_j |I_{T_j}| \le C_1 \, \text{mass}(E; \mathcal{P})^{-1}.$$

Definition 4. *Given a finite subset $\mathcal{P} \subset \mathcal{T}$ and a function $f \in L^2$, we define the* energy of f with respect to $\mathcal{P}$ *by*

$$\text{energy}(f; \mathcal{P}) = \frac{1}{\|f\|_2} \sup_{T \in \mathcal{P}:\, T \, is \, a \, 2-tree} \left(\frac{1}{|I_T|} \sum_{s \in T} |\langle f, \varphi_s \rangle|^2 \right)^{1/2}.$$

Theorem 5. *If $\mathcal{P}$ is a finite set of tiles, then $\mathcal{P}$ can be decomposed as the union of sets $\mathcal{P}_{low}$ and $\mathcal{P}_{high}$ with*

$$\text{energy}(f; \mathcal{P}_{low}) \le 2^{-1} \, \text{energy}(f; \mathcal{P})$$

and $\mathcal{P}_{high}$ is the union of trees T_j such that

$$\sum_j |I_{T_j}| \le C_2 \, \text{energy}(f; \mathcal{P})^{-2}.$$

Let us now show how Theorems 3 and 5 allow us to decompose $\mathcal{P}$ into sets $\mathcal{P}_n$, where n runs over a finite set of integers, so that

$$\text{mass}(E; \mathcal{P}_n) \le 2^{2n} \quad \text{and} \quad \text{energy}(E; \mathcal{P}_n) \le 2^n \tag{7}$$

and $\mathcal{P}_n$ is a union of trees $T_{n,k}$, $k \ge 1$, with

$$\sum_k |I_{T_{n,k}}| \le C 2^{-2n}.$$

To begin, we observe that $\mathcal{P}$ must satisfy the estimates in (7) for sufficiently large n. For such n, let us set $\mathcal{P}_n = \emptyset$. Next, if the mass of E relative to $\mathcal{P}$ is greater than $2^{2(n-1)}$, we split $\mathcal{P}$ into $\mathcal{P}_{light}$ and $\mathcal{P}_{heavy}$, replace $\mathcal{P}$ with $\mathcal{P}_{light}$ and add $\mathcal{P}_{heavy}$ to $\mathcal{P}_n$ so that $\text{mass}(E; \mathcal{P}) \le 2^{-2} 2^{2n} = 2^{2(n-1)}$ and $\mathcal{P}_n$ is a union of trees $T_{n,j}$ such that $\sum_j |I_{T_{n,j}}| \leq C_1 \operatorname{mass}(E;\mathcal{P}_{heavy})^{-1} \leq C_1/2^{2(n-1)}$. Likewise, if the energy of E relative to $\mathcal{P}$ is greater than 2^{n-1}, we split $\mathcal{P}$ into $\mathcal{P}_{low}$ and $\mathcal{P}_{high}$, replace $\mathcal{P}$ with $\mathcal{P}_{low}$ and add $\mathcal{P}_{high}$ to $\mathcal{P}_n$ so that $\operatorname{energy}(f;\mathcal{P}) \leq 2^{-1}2^n = 2^{n-1}$ and the last collection of trees $T_{n,j}$ added to $\mathcal{P}_n$ satisfy $\sum_j |I_{T_{n,j}}| \leq C_1 \operatorname{energy}(E;\mathcal{P}_{heavy})^{-2} \leq C_2/2^{2(n-1)}$. Then $\mathcal{P}$ satisfies the estimates in (7) with n replaced by $n-1$ and we continue this procedure in a similar fashion until there are no tiles left in $\mathcal{P}$.

Finally, to obtain the desired estimate (6), we need

Theorem 6 (Main estimate)**.** *There is a constant $C_3 > 0$ such that for any tree T, function $f \in L^2$, measurable function N on $\mathbb{R}$ and measurable set E,*

$$\sum_{s\in T} \left| \langle f, \varphi_s \rangle \left\langle \chi_{E\cap N^{-1}(\omega_{s(2)})}, \varphi_s \right\rangle \right| \leq C_3 |I_T| \operatorname{energy}(f;T) \operatorname{mass}(E;T) \|f\|_2 |E|.$$

In particular, Theorem 6 and our previous construction yield

$$\begin{aligned}
&\sum_{s\in P} | \langle f, \varphi_s \rangle \left\langle \chi_{E\cap N^{-1}(\omega_{s(2)})}, \varphi_s \right\rangle | \\
&\qquad \leq \sum_n \sum_j C_3 |I_{T_{n,j}}| \operatorname{energy}(f;T_{n,j}) \operatorname{mass}(E;T_{n,j}) \|f\|_2 |E| \\
&\qquad \leq C_3 \sum_n \sum_j |I_{T_{n,j}}| 2^{n+1} \min\{|E|^{-1}, 2^{2n+2}\} \|f\|_2 |E| \\
&\qquad \leq C_4 \|f\|_2 |E|^{1/2}.
\end{aligned}$$

This completes the proof of (6) and so the weak-type estimate for C_* and the a.e. convergence of (1) for arbitrary $f \in L^2$ hold, as desired.

We refer the reader to [1] and [2] for more material on Carleson's Theorem.

References

[1] M. Lacey and C. Thiele, *A proof of boundedness of the Carleson operator.* Math. Res. Lett. **7** (2000), no. 4, 361–370;

[2] M. Lacey, *Carleson's Theorem: proof, complements, variations.* Publ. Mat. **48** (2004), 251–307.

Alberto A. Condori, Michigan State University
email: **condoria@msu.edu**

4 A Carleson Type Theorem For A Cantor Group Model of the Scattering Transform

after C. Muscalu, T. Tao, and C. Thiele [1]
A summary written by Michael Dabkowski

Abstract

We prove the weak L^2 bounds for a Carleson type maximal operator for the d-adic model of the scattering transform on the line.

4.1 Introduction

Scattering transforms are seen as non-linear variants of the Fourier transform. We therefore ask if the classical a priori estimates such as Riemann-Lebesgue, Plancherel, or Hausdorff-Young hold for a given scattering transform. We will look to find a variant of the AKNS-ZS matrix system: $G' = WG$, where

$$G = \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \quad W = \begin{pmatrix} 0 & F(x)\exp(2ikx) \\ \overline{F(x)}\exp(-2ikx) & 0 \end{pmatrix},$$

where F is compactly supported (later we will extend to $L^2(R^+)$). Under the compactly supported assumption we can solve the differential equation under the initial condition $G(-\infty) = Id$ to obtain an absolutely continuous solution that is constant near ∞. The *scattering transform* of the potential F at spectral value k is $G(\infty)$. We would like to prove a Carleson-Hunt type theorem for this transform, namely,

$$\| \sup_x \sqrt{\log(a(x,\cdot))} \|_{L^2} \leq C \|F\|_2^2.$$

We choose to study the d-adic version of the problem. Starting with the base d expansions of two positive numbers x and k we define a character

$$w(k,x) = w\Big(\sum_{n\in\mathbb{Z}} k_n d^n, \sum_{n\in\mathbb{Z}} x_n d^n\Big) = \gamma^{\sum_{n\in\mathbb{Z}} k_n x_{-1-n}}, \tag{1}$$

where γ is a primitive d-th root of unity. We will now replace the exponential factor in the above matrix W with the character w. Doing so we arrive at the differential system

$$\partial_x G(k,x) = W(k,x)G(k,x) \tag{2}$$

where

$$G(k,0) = Id \quad \text{and} \quad W(k,x) = \begin{pmatrix} 0 & F(x)w(k,x) \\ \overline{F(x)w(k,x)} & 0 \end{pmatrix}. \tag{3}$$

For this problem we have the following theorem

Theorem 1. *For $d > 1$ and $F \in L^2(\mathbb{R}^+)$, for almost all $k \in \mathbb{R}^+$ the limit*

$$\lim_{x\to\infty} G(k,x) = G(k,\infty)$$

exists and satisfies the Plancherel inequality

$$\int_0^\infty \log|a(k,\infty)|\,dk \le C\int_0^\infty |F(x)|^2\,dx. \tag{4}$$

Moreover, we have the weak-type inequality with a constant $C(poly(d))$

$$\left|\left(k : \sup_x \log|a(k,x)| > \lambda\right)\right| \le C(poly(d))\lambda^{-1}\|F\|_2^2. \tag{5}$$

4.2 Localization by Tiles

We localize (2) and (3) in space about a d-adic interval $\omega = [d^k n, d^{k(n+1)})$. Doing so we arrive at the localized system

$$\partial_x G_\omega(k,x) = W_\omega(k,x)G_\omega(k,x) \tag{6}$$

where

$$G_\omega(k,0) = Id \quad \text{and} \quad W_\omega(k,x) = \begin{pmatrix} 0 & F(x)w(k,x)\chi_\omega(x) \\ \overline{F(x)w(k,x)\chi_\omega(x)} & 0 \end{pmatrix}. \tag{7}$$

Now we say that a **tile** is a rectangle of the form $p = I\times\omega$, where I and ω are d-adic intervals of the form $I = [d^\kappa n, d^\kappa(n+1))$ and $\omega = [d^{-\kappa}l, d^{-\kappa}(l+1))$ for $\kappa \in \mathbb{Z}$ and $n, l \in \mathbb{Z}^+$. If we suppose that (k,x) is a point in some tile p and k_0 is the leftmost point in I_p, then we can split (1) into parts

$$w(k,x) = \gamma^{\sum_{n<\kappa} k_n x_{-1-n}}\gamma^{\sum_{n\ge\kappa} k_n x_{-1-n}}.$$

Since we are in a tile the first factor doesn't change as x varies in ω_p and similarly the second factor doesn't change as k varies in I_p. It follows that $w(k,x) = \gamma^{j(k)} w(k_0,x)\chi_\omega(x)$ and hence the solution to (6) at the point (k,x) is conjugate to the solution to (6) at the point (k_0,x) by the diagonal matrix with eigenvalues $\gamma^j(k)$ and 1. This reasoning says that G_{ω_p} doesn't change too much in the tile p. Bearing this in mind we set $G_p = G_\omega(k_0,\infty)$.

To understand what it means for tiles to be "nearby" we say a **multitile** is be a rectangle of the form $P = I \times \omega$ with

$$P = [d^\kappa n, d^\kappa(n+1)) \times [d^{1-\kappa} l, d^{1-\kappa}(l+1)), \quad \kappa \in \mathbb{Z}, \quad l,n \in \mathbb{Z}^+.$$

An important but elementary consequence of this is that given a multitile P we can write it as a union of d horizontal tiles: $P = p_0 \cup p_2 \cup \cdots \cup p_{d-1}$ or as a union of d vertical tiles: $P = q_0 \cup q_2 \cup \cdots \cup q_{d-1}$. Also if $P = I \times \omega$ is a multitile then the horizontal and vertical decompositions are related by

$$G_{p_j} = \begin{pmatrix} a_j & b_j \\ \overline{b_j} & \overline{a_j} \end{pmatrix} \quad \Rightarrow \quad G_{q_m} = \begin{pmatrix} a_{d-1} & \gamma^{m(d-1)} b_{d-1} \\ \overline{\gamma^{m(d-1)} b_{d-1}} & \overline{a_{d-1}} \end{pmatrix} \cdots \begin{pmatrix} a_1 & \gamma^m b_1 \\ \overline{\gamma^m b_1} & \overline{a_1} \end{pmatrix} \begin{pmatrix} a_0 & b_0 \\ \overline{b_0} & \overline{a_0} \end{pmatrix}$$

4.3 Swapping Functions and Plancherel Inequality

Thinking of the Plancherel Inequality as a type of d-adic imbedding theorem we look for a function $\beta : SU(1,1) \to \mathbb{R}^+$ (using the Bellman function technique, see [2]) such that $\beta(G) \asymp \log|a|$ and has the swapping property

$$\sum_{j=0}^{d-1} \beta(G_{q_j}) \leq d \sum_{j=0}^{d-1} \beta(G_{p_j}). \tag{8}$$

The swapping inequality says that for a multitile P, up to a factor of d, the horizontal multitiles dominate the vertical multitiles with respect to β. Suppose that the support of F is contained in the interval $[0, d^K)$ for some large K and consider the rectangle $[0, d^K) \times [0, d^K)$. If we let $\mathbf{p}_k$ be the set of all tiles with $|I_p| = d^k$. Applying the swapping inequality iteratively we have

$$d^{-K} \sum_{p \in \mathbf{p}_{-K}} \log|a_p| \leq C d^K \sum_{p \in \mathbf{p}_K} \log|a_p|. \tag{9}$$

If we note that the set $\mathbf{p}_{-K}$ contains only the tile $[0, d^{-K}) \times [0, d^K)$, then by the results in section 2 we see that the left hand side of the above inequality is exactly $\| \log |a(\cdot, \infty)| \|_{L^1(dk)}$. The right hand of the inequality can be shown to be less than a constant times $\|F\|_2^2$ by using the operator Gronwall inequality.

Finding a function $\beta : SU(-1,1) \to \mathbb{R}^+$ such that $\beta(G) \asymp \log |a|$ is not difficult; rather, the difficultly is finding a function that satisfies the swapping inequality. In the case when $d = 2$, we can take $\beta(G) = \log \|G\|_{HS}$ (where $\|G\|_{HS}$ is the Hilbert-Schmidt norm of G). Verifying that the function β satisfies the swapping inequality is a trivial computation which boils down to the quasi-triangle inequality. Unfortunately the logarithm of the Hilbert-Schmidt norm doesn't work when $d \geq 3$, as a laborious computation shows. In the case when $d \geq 3$ we choose the smallest possible number r such that the function

$$\beta(z) = \begin{cases} |z|^2 - |z|^3 & |z| \leq r \\ \epsilon^{10} + \epsilon^{20} \operatorname{arcsinh}(|z|) & |z| > r \end{cases}$$

is continuous and ϵ is some small parameter that depends on d only. We then set $\beta(G) = \beta(b)$. This function β is the desired function. It is clear that up to a factor of d that $\beta(G) \asymp \log |a|$, as the $\operatorname{arcsinh}(|z|) = \log(|z| + \sqrt{|z|^2 + 1})$. Verification of the swapping inequalities is a tedious task which involves carefully looking at the products G_{q_m} using polynomial inequalities and Taylor approximations.

4.4 John-Nirenberg Type and Carleson Weak Type Inequalities

We define a partial order on the set of all tiles (similarly for multitiles) by saying $p < q$ if and only if $I_p \subseteq I_q$ and $\omega_q \subseteq \omega_p$. The swapping inequality (8) and several inductions implies the following

Lemma 2. *Let $\boldsymbol{q}$ be a finite set of pairwise disjoint tiles and $\boldsymbol{p}$ a collection of tiles such that*

$$q \subseteq \bigcup_{p \in \boldsymbol{p}} p, \quad \forall q \in \boldsymbol{q}$$

and whenever p and q have non-empty intersection we have $q < p$. Then

$$\sum_{q \in \boldsymbol{q}} |I_q| \beta(G_q) \leq \sum_{p \in \boldsymbol{p}} |I_p| \beta(G_p) \tag{10}$$

4.4.1 Some Terminology

We can similarly define what it means for two multitiles P and Q to be ordered. A collection of multitiles is *convex* if whenever $P < P' < P''$ with P and P'' in the collection implies P' is also in the collection. Suppose that a collection $\mathbf{P}$ of multitiles can be written as a union

$$\mathbf{P} = \bigcup_{n \in N} \mathbf{P}_n$$

with the property that if $P \in \mathbf{P}_n$ and $Q \in \mathbf{P}_m$ with $P < Q$, then $n \leq m$. We call such a union and *ordered splitting* of $\mathbf{P}$. A *tree* is a set T of multitiles which has a maximal element with respect to the ordering "$<$". Now each element of a tree P has a unique horizontal tile $p_{j(P)}$ that intersects the maximal element of the tree, which is called the *top*. Using this horizontal tile as a level for an element of a tree we can define a notion of size for a collection of multitiles: Given a collection of multitiles $\mathbf{P}$, we define

$$\text{size}(\mathbf{P}) = \sup_{T \subseteq \mathbf{P}} |I_T|^{-1} \sum_{P \in T} \sum_{j < j(P)} |I_P| \beta(G_{p_j}).$$

4.4.2 Calderón-Zygmund and John-Nirenberg Type Results

Given any collection of multitiles $\mathbf{P}$ we run the standard stopping time argument with respect to the size of $\mathbf{P}$ to obtain an ordered splitting

$$\mathbf{P} = \mathbf{P}_\infty \cup \bigcup_{n \in \mathbb{Z}} \mathbf{P}_n$$

such that size$(\mathbf{P}_k) \leq 2^{-4k}$ and $\mathbf{P}_k$ is a union of trees $\mathbf{T}_k$ such that

$$\sum_{T \in \mathbf{T}_k} |I_T| \leq C(poly(d)) 2^{4k} \|F\|_2^2.$$

Additionally, for each tree $T \in \mathbf{T}_k$ with top P contains all elements $P' \in \mathbf{P}_k$ with $P' < P$.

Assume we have a convex tree T and F vanishes above the lower endpoint of ω_{top}. If P is a multitile in T and q is a vertical tile of P, then for $k \in I_q$, we have

$$G_q(k) = \prod_{P' \in T : I_P \subseteq I'_P \subsetneq I_{P_T}} \prod_{j < j(P')} G_{p_j}(k) \tag{11}$$

With this in mind we define the following

$$M_T(k) = \sup_{k\in I, J, I\subseteq J\subsetneq I_T} \log\Big|\prod_{P\in T: I\subseteq I_P\subsetneq J}\prod_{j<j(P')} G_{p_j}(k)\Big|. \tag{12}$$

For this function we have a weaker John-Nirenberg inequality, namely,

$$\Big|\big(k\in I_T : M_T(k)\geq 4d\Gamma 2^{2\lambda}\mathrm{size}(T)\big)\Big| \leq 2^{-c\lambda^2}|I_T|, \tag{13}$$

where Γ is the constant that compares $\log|a|$ and $\beta(G)$ up to a polynomial in d.

4.4.3 Carleson Inequality

The previous subsection provides us all we need to prove the Carleson inequality. As in the classical Carleson proof, we define an exceptional set and show that it has measure controlled by $\Gamma^2\lambda\|F\|_2^2$ and off the exceptional set we have a good bound on the supremum. The components of the exceptional set are defined by

$$E_k = \begin{cases} \bigcup_{T\in\mathbf{T}_k} I_T & k<K \\ \bigcup_{T\in\mathbf{T}_k}\{k\in I_T : M_T(k)\geq 4d2^{2(k-K)}2^{-4k}\} & k\geq K \end{cases}$$

where K is a large parameter and the trees $\mathbf{T}_k$ are defined in the stopping time argument. Setting $E=\bigcup E_k$ and using the above argument we have that $|E|\leq\Gamma^2\lambda\|F\|_2^2$. Finally we show that off of the exceptional set $\log|G(k,x)|$ is smaller than $Cd\lambda^{-1}$.

References

[1] Muscalu, C., Tao, T., and Thiele, C., *A Carleson Type Theorem for a Cantor Group Model of the Scattering Transform* Nonlinearity 16 (2003), no. 1, 219–246 ;

[2] Nazarov, F. and Treil, S. , *The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems in harmonic analysis.* (Russian) Algebra i Analiz **8** [1996] no. 5 pp. 32-162. translation in St. Petersburg Math J. **8** [1997] pp. 721-824 ;

Michael Dabkowski, University of Wisconsin-Madison
email: **dabkowsk@math.wisc.edu**

5 The endpoint case of the Bennett-Carbery-Tao multilinear Kakeya conjecture

after Larry Guth [5]
A summary written by Yen Do

Abstract

We prove the Bennett-Carbery-Tao multilinear Kakeya conjecture using a variant of the Ham Sandwich theorem and a geometrical analogue of Dvir's argument.

5.1 Introduction

The Kakeya set conjecture [6] asserts that if a set $E \subset \mathbb{R}^n$ contains a unit line in every direction, it must have Hausdorff and Minkowski dimension equal to n. Wolff [7] introduced a finite field model of the conjecture, which was recently proved by Dvir [3] using a dimension argument. In the Euclidean setting, it is possible to formulate a maximal-but-discretised version, which implies the set conjecture (using standard arguments, see for instance [2]).

To state the maximal conjecture, let T be a collection of $\delta \times 1$ cylindrical tubes (i.e. δ-long tube whose cross section is an $(n-1)$-dimensional unit disk) pointing in a set of sufficiently separated directions: any two directions are at least δ-away from each other.

Conjecture 1 (Kakeya maximal)**.** *The set E_μ of points with intersection multiplicity $\gtrsim \mu$ (i.e. contained in $\gtrsim \mu$ tubes) has volume not less than*

$$C(\epsilon, n)\delta^{-\epsilon}\mu^{-\frac{n}{n-1}}$$

for any $\epsilon > 0$.

A related problem is to study overlap properties of a set of tubes, for instance a multilinear Kakeya conjecture was formulated by Bennett-Carbery-Tao [1]. A particular consequence of this conjecture says that: in a counter example to conjecture 1, a typical point in E_μ will have intersection multiplicity contributed largely by tubes whose directions almost belong to a common hyperplane.

To formulate this multilinear conjecture, we divide T into n subsets $T_1, \ldots, T_n$ containing respectively $A(1), \ldots, A(n)$ tubes, so that:

(i) in each subset the directions of the tubes are close;

(ii) any family $v_1, \dots, v_n$ of n unit directions with exactly one representative from every T_k will not be degenerate: the solid angle formed by these directions has nontrivial degree:

$$\det(v_1, \dots, v_n) \gtrsim \theta \tag{1}$$

For $x \in \mathbb{R}^n$, let $N(x)$ be the number of families (as in (ii)) such that every member of the family possesses x, i.e.

$$N(x) = \prod_{k=1}^{n} \sum_{U \in T_k} 1_U(x)$$

The multilinear Kakeyea conjecture, which is now proved by Guth [5], says:

Theorem 2 (Multilinear Kakeya estimate, Guth)**.**

$$\int N(x)^{\frac{1}{n-1}} dx \lesssim_n \theta^{-\frac{1}{n-1}} \prod_{j=1}^{n} A_j^{\frac{1}{n-1}}$$

Note that the bound on the RHS doesn't depends on δ, thus we can (and will) take the tubes in T to be of infinite length (in our proof).

Below we sketch why theorem 2 implies certain structural properties of a counter example T to conjecture 1. Assume that T can be divided it into the above configuration. Let

$$E^*_\mu := \{x \in \mathbb{R}^n : \text{for every } 1 \le k \le n,\ x \text{ belongs to} \gtrsim \mu \text{ tubes in } T_k\}$$

be a subset of E_μ where the global intersection multiplicity of x is contributed almost equally by every T_k. Since $N(x) \gtrsim \mu^n$ for $x \in E^*_\mu$, theorem 2 implies that

$$E^*_\mu \lesssim \mu^{\frac{-n}{n-1}}.$$

This means that for a typical point of E_μ (i.e. in $E_\mu \setminus E^*_\mu$), its multiplicity is contributed mostly by tubes comming from the same subcollection, i.e. with close directions.

5.2 Proof of theorem 2

To prove theorem 2, we will need a generalized version of the following Ham Sandwich theorem (will be referred to as HS theorem):

Ham Sandwich theorem. *Let $U_1, \dots, U_n$ be finite volume open sets in $\mathbb{R}^n$. Then there is a hyperplane H that bisect every U_i.*

This theorem was generalized by Gromov [4] to $N = \binom{n+d}{d} - 1$ sets, where the hyperplane should be replaced by an algebraic hypersurface of degree $\leq d$. Guth's generalization of the HS theorem is formulated using the notion of visibility of a surface.

For any vector $v \in \mathbb{R}$ we denote by π_v the projection along direction v (into the hyperplane $v^\perp \subset \mathbb{R}^n$).

Definition 3 (Directed volume)**.** *For any surface $S \subset \mathbb{R}^n$ and any unit vector $v \in \mathbb{R}^n$, the volume of S in direction v is defined by*

$$V_S(v) = \int_{v^\perp} |S \cap \pi_v^{-1}(y)| dy.$$

If v is not a unit vector, define $V_S(v) := |v| V_S(\frac{v}{|v|})$.

Essensially, $V_S(v)$ is the volume of the projection of S along direction v, counting multiplicity. Equivalently, we can define:

$$V_S(v) = \int_S |v \cdot N_x| dS(x)$$

here the integration is over S using the surface area measure, and N_x denotes the normal vector at x.

Definition 4 (Visibility)**.** *The visibility of S is defined by:*

$$Vis[S] := \frac{1}{Vol\big(Hull\{|v| \leq 1 : V_S(v) \leq 1\}\big)}$$

For an unit vector v, S is geometrically less visible in direction v when $V_S(v)$ is small. Essentially, $\frac{1}{\mathrm{Vis}[S]}$ is large if there are only a small number of directions where S is not visible. This definition depends on the relative size of S, however in our applications S will have size comparable to 1.

When S is a variety of degree $\leq n$ (which is easily parametrized by $\mathbb{R}P^N$ through the tuple of its coefficients - we'll simply say $S \in \mathbb{R}P^N$), it is not hard to see that, as function of S, the directed volume $V_S(v)$ (hence the visibility $\text{Vis}[S]$) is *not* continuous on $\mathbb{R}P^N$. Since continuity will be important for our generalization of the Ham Sandwich theorem, we'll need to modify our directed volume (and hence the visibility) by averaging over small open neighborhood of $S \in \mathbb{R}P^N$ (using the angle metric). Under this operation, all the needed properties of V_S and $\text{Vis}[S]$ are preserved up to some small $c > 0$, which will be harmless in our proof. We'll denote the new directed volume and visibility by $\overline{V}_S(v)$ and $\overline{\text{Vis}}[S]$.

The following basic estimate is useful: For any nonempty surface S and any family of directions $v_1, \ldots, v_n$, if $\overline{V}_S(v_j) \geq 1$ for every j then

$$\overline{\text{Vis}}[S] \lesssim_n \frac{\prod_{j=1}^n \overline{V}_S(v_j)}{\det(v_1, \ldots, v_n)} \tag{2}$$

This estimate can be easily proved using the definition of visibility. Below is the needed generalization of the Ham Sandwich theorem:

HS theorem and visibility. *If M is a function that assigns an integer value to each cube in the standard unit lattice of $\mathbb{R}^n$, then there exists an algebraic hypersurface Z such that $\overline{Vis}[Z \cap Q_k] \geq M(Q_k)$ for every cube Q_k, while the degree d of Z is controlled by $[\sum_k M(Q_k)]^{\frac{1}{n}}$.*

Proof of theorem 2. For each cube Q_k in the standard unit lattice, let $M_j(Q_k)$ be the number of tubes in subcollection T_j that intersect Q_k, and $F(Q_k) = \prod_{1 \leq j \leq n} M_j(Q_k)$ the number of families (with one from each subcollection) that intersect Q_k. Then $\int_{Q_k} N(x)^{\frac{1}{n-1}} dx \lesssim F(Q_k)^{\frac{1}{n-1}}$, so it sufficies to show

$$\begin{aligned} \sum_k F(Q_k)^{\frac{1}{n-1}} &\lesssim_n \theta^{-\frac{1}{n-1}} \prod_{j=1}^n A(j)^{\frac{1}{n-1}} \\ \Leftrightarrow \theta \Big[\sum_k F(Q_k)^{\frac{1}{n-1}} \Big]^{n-1} &\lesssim_n \prod_{j=1}^n A(j) \end{aligned} \tag{3}$$

Since T is finite, the tranversality condition (1) implies that only a finite number of Q_k contributes to the sum $\sum F(Q_k)$. Let Q be the set of these cubes.

The next argument essentially resembles Dvir's argument in the finite field setting. Using the above variant of the HS theorem, for any function m that assigns a nonnegative integer value to each Q_k, we can find a variety Z of degree d such that it intersect lots of cubes in a highly visible surface,

$$\overline{\text{Vis}}[Z \cap Q_k] \gtrsim m(Q_k) \quad \forall Q_k \tag{4}$$

while having a controlled degree

$$d \ \lesssim_n \ [\sum_k m(Q_k)]^{\frac{1}{n}} \tag{5}$$

We'll choose m such that the right hand side of (5) is very large compared to $|Q|$. At the end of this argument, it is the relative ratio between $m(Q_k)$'s that matters, so this assumption will be harmless.

Now, by adding a number of hyperplane to Z if necessary, we can assume that $\overline{V}_{Z\cap Q_k}(v) \geq 1$ for every unit direction v and for every $Q_k \in Q$ (basically for each Q_k we need about $O_n(1)$ hyperplanes, so totally need about $O_n(|Q|)$ hyperplanes). Adding hyperplane to Z is (at worst) equivalent to multiplying a polynomial of degree 1 to the polynomial associated with Z. Thus, these additions won't significantly increase the (bound (5) on the) degree of Z, due to our choice of m. Therefore, we can safely assume that (5) remains valid.

For each tube $T_{j,a}$ in the subcollection T_j we always have:

$$\sum_{Q_k \in Q \text{ that intersects } T_{j,a}} \overline{V}_{Z\cap Q_k}(\text{direction of } T_{j,a}) \lesssim d \tag{6}$$

This is because of the following reasons:

(i) For any Q_k that intersects $T_{j,a}$, the projection of every $Z \cap Q_k$ along the direction of of $T_{j,a}$ will be inside a slightly thickened version of the section of the cylinder $T_{j,a}$ (so will have volume $\lesssim_n 1$); and

(ii) The projection multiplicity of any point on Z, along any given direction, is controlled by the degree of the polynomial, which is d.

So essentially (6) is analogous to the finite field statement that a nontrivial homogeneous polynomial of degree d cannot vanish on more than d points. Now, by *choosing* a tube $T_{j,a}$ that *nicely* intersects *lots* of cubes Q_k's, we can then get a lower bound on the LHS of (6) and get a lower estimate for d. Then, by combining with (5) and an optimization argument, we can get an estimate for $\prod A(j)$ in terms of $F(Q_k)$ and deduce the desired estimate (3).

To *choose* such $T_{j,a}$, we'll use an averaging argument. First, consider a family of tubes $T(1), \ldots, T(n)$, one representative from each subcollection T_k, with respective unit directions $v_1, \ldots, v_n$. If for every $1 \leq j \leq n$, $Q_k \cap T(j) \neq \emptyset$, then using (2), (1), and (4), we have

$$m(Q_k) \lesssim \overline{\text{Vis}}[Z \cap Q_k] \lesssim_n \frac{\prod_{j=1}^n \overline{V}_{Z \cap Q_k}(v_j)}{\det(v_1, \ldots, v_n)} \lesssim \theta^{-1} \prod_{j=1}^n \overline{V}_{Z \cap Q_k}(v_j)$$

so $m(Q_k) \lesssim \theta^{-1} \prod_{j=1}^n \overline{V}_{Z \cap Q_k}(v_j)$. Fixing Q_k and summing over all possible such combinations of $T(1), \ldots, T(n)$, we get

$$F(Q_k) m(Q_k) \quad \lesssim_n \quad \theta^{-1} \prod_{j=1}^n \Big(\sum_{T_{j,a} \in T_j \text{ that intersects } Q_k} \overline{V}_{Z \cap Q_k}(T_{j,a}\text{'s direction}) \Big)$$

So by Holder's inequality, we have

$$\begin{aligned}
\sum_{Q_k} F(Q_k)^{\frac{1}{n}} m(Q_k)^{\frac{1}{n}} \quad &\lesssim_n \quad \theta^{-\frac{1}{n}} \sum_{Q_k} \prod_{j=1}^n \Big(\sum_{T_{j,a} \text{ intersects } Q_k} \overline{V}_{Z \cap Q_k}(T_{j,a}\text{'s direction}) \Big)^{\frac{1}{n}} \\
&\leq \quad \theta^{-\frac{1}{n}} \prod_{j=1}^n \Big(\sum_{Q_k} \sum_{T_{j,a} \text{ intersects } Q_k} \overline{V}_{Z \cap Q_k}(T_{j,a}\text{'s direction}) \Big)^{\frac{1}{n}} \\
&\leq \quad \theta^{-\frac{1}{n}} \prod_{j=1}^n \Big(dA(j) \Big)^{\frac{1}{n}} \quad \text{by (6)}
\end{aligned}$$

Consequently, we get a lower bound for d:

$$d \gtrsim_n \frac{\theta^{\frac{1}{n}} \sum_{Q_k} F(Q_k)^{\frac{1}{n}} m(Q_k)^{\frac{1}{n}}}{\prod_{j=1}^n A(j)^{\frac{1}{n}}}$$

Combining with (5), we get

$$\begin{aligned}
\big[\sum_k m(Q_k) \big]^{\frac{1}{n}} \quad &\gtrsim_n \quad \frac{\theta^{\frac{1}{n}} \sum_{Q_k} F(Q_k)^{\frac{1}{n}} m(Q_k)^{\frac{1}{n}}}{\prod_{j=1}^n A(j)^{\frac{1}{n}}} \\
\Longrightarrow \prod_{j=1}^n A(j) \quad &\gtrsim_n \quad \theta \big[\sum_{Q_k} F(Q_k)^{\frac{1}{n}} \big(\frac{m(Q_k)}{\sum_k m(Q_k)} \big)^{\frac{1}{n}} \big]^n
\end{aligned}$$

Optimizing the RHS over $\alpha_k := \frac{m(Q_k)}{\sum_k m(Q_k)} \geq 0$ under the constraint $\{\sum \alpha_k = 1\}$ (this can be done using Holder's inequality), we see that the best choice

$$\alpha_k = \frac{F(Q_k)^{\frac{1}{n-1}}}{\sum_{Q_k} F(Q_k)^{\frac{1}{n-1}}}$$

will give us the desired estimate (3). □

References

[1] Bennett,J., Carbery, A., and Tao, T., *On the multilinear restriction and Kakeya conjectures.* Acta Math. **196** (2006), 261–302.

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

[3] Dvir, *On the size of Kakeya sets in finite fields.* arXiv:0803.2336

[4] Gromov, M., *Isoperimetry of waists and concentration of maps.* Geom. Funct. Anal. **13** (2003), no. 1, 178–215

[5] Guth, L., *The endpoint case of the Bennett-Carbery-Tao multilinear Kakeya conjecture.* arXiv:0811.225v1

[6] Tao, T., *From rotating needles to stability of waves: emerging connections between combinatorics, analysis, and PDE.* Notices Amer. Math. Soc. **48** (2001), no. 3, 294–303.

[7] Wolff, T., *Lectures on Harmonic Analysis.* , AMS, University Lecture Series, vol. **29**, Oct. 2003

Yen Do, UCLA
email: **qdo@math.ucla.edu**

6 The (weak-L^2) boundedness of the quadratic Carleson operator

after Victor Lie
A summary written by S. Zubin Gautam

Abstract

We summarize the proof in [Lie08a] of the weak-type $(2,2)$ boundedness of the generalized Carleson maximal operator with arbitrary quadratic phase functions.

6.1 Introduction

(Note to the reader: This summary is admittedly quite long. The first seven pages are basically crucial, the next three are fairly important, and the remainder should be read only if stamina permits.)

By the Stein maximal principle ([Ste61]), Carleson's celebrated theorem on the almost-everywhere convergence of Fourier series of L^2 functions ([Car]) is equivalent to the boundedness of the Carleson maximal operator C from $L^2(\mathbb{T})$ to $L^{2,\infty}(\mathbb{T})$, where

$$\mathrm{C}\, f(x) := \sup_{a>0} \left| \int_{\mathbb{T}} f(x-y)\, e^{iay}\, \frac{1}{y}\, \mathrm{d}y \right| .$$

As a generalization of Carleson's theorem, Stein conjectured that the degree-d polynomial Carleson operator C_d given by

$$\mathrm{C}_d\, f(x) := \sup_{\substack{p\in\mathbb{R}[y] \\ \deg(p)\le d}} \left| \int_{\mathbb{T}} f(x-y) e^{ip(y)} \frac{1}{y}\, \mathrm{d}y \right|$$

enjoys this same boundedness (in fact, he conjectured that C_d should be strong-type (p,p) for all $1 < p < \infty$); of course, the operator C_1 coincides with the Carleson operator C.

A weaker form of the conjecture was verified by Stein ([Ste93]) for $d = 2$ and by Stein and Wainger ([SW]) for general d; their results treat the modified operators $\widetilde{\mathrm{C}}_d$ in which the supremum is taken only over polynomials $p \in \mathbb{R}[y]$ with *no linear term*, so of course they cannot be viewed as actual generalizations of Carleson's theorem.[10] Our goal is to prove the following theorem, which removes this restriction in the case $d = 2$:

Theorem 1. *Let* $\mathrm{T} = \mathrm{C}_2$*, so that*

$$\mathrm{T}f(x) = \sup_{a,b\in\mathbb{R}} \left| \int_{\mathbb{T}} f(x-y) e^{i(ay+by^2)} \frac{1}{y}\, \mathrm{d}y \right|.$$

Then

$$\|\mathrm{T}f\|_{L^p(\mathbb{T})} \lesssim_p \|f\|_{L^2(\mathbb{T})}$$

for all $1 \leq p < 2$*.*

Combined with the techniques of [Ste61], this theorem implies that T is weak-type $(2, 2)$; in fact, a little modification of the arguments given below, together with interpolation, can show that T is in fact of strong type (p, p) for all $1 < p < 2$. The analogous result for the more general polynomial Carleson operators C_d is proven in [Lie08b].[11]

From a macroscopic point of view, the proof of Theorem 1 in [Lie08a] follows the steps of C. Fefferman's proof of Carleson's theorem in [Fef] quite faithfully; as such, and as with most time-frequency proofs, the coarsest decomposition of the argument is into two main steps: a decomposition or "discretization" of the operator T as a sum of operators T_P whose outputs are well-localized in both time and frequency, and a selection algorithm by which one reassembles the pieces T_P in a suitable manner to provide good bounds on T.

The discretization procedure used in [Fef] splits the Carleson operator C as a sum of operators C_P, where each P is a "Heisenberg tile" in $\mathbb{R}^2$ (or $\mathbb{T}\times\mathbb{R}$): a rectangle of the form $P = I \times \omega$, with I and ω dyadic intervals such that $|I|\,|\omega| = 1$ (so that P has area 1).[12] Here, loosely speaking, I and ω are the intervals on which $\mathrm{C}_P\, f$ and $\widehat{\mathrm{C}_P\, f}$, respectively, are large; in other words, I

[10]The absence of modulation-invariance afforded by this restriction moves one out of the province of time-frequency analysis; the proofs of [Ste93] and [SW] rely mainly on oscillatory integral techniques.

[11]The full-force version of Stein's conjecture (*viz.*, with L^p-boundedness for all $1 < p < \infty$) remains open. However, the techniques of [Lie08a] can be used to show $\mathrm{T} : L^p \to L^{p-\varepsilon}$; removal of the ε seems to be more of a technical obstacle than a conceptual one.

[12]Discretization via Heisenberg tiles has by now become more or less a template for decomposing operators that commute with translations and dilations and are modulation-invariant.

and ω reflect the "L^∞ distribution" of $\mathrm{C}_P f$ and its Fourier transform, and the tile $P = I \times \omega$ should be viewed as a representation of the L^∞ time-frequency localization of $\mathrm{C}_P f$. By basic properties of the Fourier transform, one sees that if $\varphi \in L^2$ is time-frequency localized to a tile P, then any composition of dyadic translations, modulations, and dilations applied to φ will move its localization to another tile P'; in other words, the group generated by such transformations may be viewed as acting on the space of tiles. The quadratic Carleson operator T, however, enjoys an additional symmetry which should be respected in its decomposition: it is quadratic-modulation–invariant; *i.e.*, $\mathrm{T}\,\mathrm{Q}_b = \mathrm{T}$ for all $b \in \mathbb{R}$, where $\mathrm{Q}_b f(x) := e^{ibx^2} f(x)$. The group $\{\mathrm{Q}_b \mid b \in \mathbb{R}\}$ no longer acts on the space of tiles we have defined thus far; if φ is L^∞ localized to $P = I \times \omega$, $\mathrm{Q}_b\,\varphi$ will be L^∞ time-frequency localized to $I \times \Omega$ for some interval Ω which is generically much larger than ω, so Q_b destroys optimal time-frequency localization.[13] In light of this state of affairs, the method of decomposition via standard Heisenberg tiles appears to be unsuitable in the presence of quadratic-modulation–invariance, and moreover the very concept of localization in the sense of L^∞ distribution seems inapt. The key insight toward overcoming this difficulty is that one can approach time-frequency localization from a *relative* perspective, by considering when $|\langle \mathrm{Q}_b\,\varphi, \mathrm{Q}_{b'}\,\varphi\rangle|$ is large, as opposed to simply studying the size of $|\,\mathrm{Q}_b\,\varphi|$ and $|\widehat{\mathrm{Q}_b\,\varphi}|$. The upshot of this approach is that one generalizes the notion of a "tile" from rectangles to suitable area-one parallelograms in the phase plane $\mathbb{T} \times \mathbb{R}$, and to each such tile P one associates a "piece" T_P of the operator T, as we discuss in Sections 6.2 and 6.3.

Once this decomposition $\mathrm{T} = \sum_P \mathrm{T}_P$ has been accomplished, the second stage of the proof, the reassembly of the pieces T_P, essentially follows Fefferman's algorithm for the Carleson operator step by step; as we shall see, the main idea is carefully to isolate situations in which our pieces do not enjoy as much cancellation as Fefferman's and to apply "positive" or maximal methods in these situations. That being said, significant effort is required to effect this adaptation to the quadratic-phase setting.

6.2 Tiles: The relative perspective

We now elaborate on the "relative" or "relational" approach to time-frequency localization alluded to above. The idea is to interpret the localization of, say,

[13] The optimality here is in view of the uncertainty principle.

$\mathrm{Q}_b\,\varphi$ by studying $|\langle \mathrm{Q}_b\,\varphi, \mathrm{Q}_{b'}\,\varphi\rangle|$ for $b, b' \in \mathbb{R}$. Let φ be a nice L^2 function supported in some dyadic interval I with $\hat{\varphi}$ adapted to an interval of length $|I|^{-1}$ near the origin and $\|\varphi\|_2 \lesssim 1$; appealing to stationary/non-stationary phase considerations, we have

$$|\langle \mathrm{Q}_b\,\varphi, \mathrm{Q}_{b'}\,\varphi\rangle| \lesssim (\sup_{x\in I} |2bx - 2b'x|\,|I|)^{-1/2} := \left(\frac{\mathrm{dist}^I(l_b, l_{b'})}{|I|^{-1}}\right)^{-1/2}.$$

Here l_s is the line $\xi = 2sx$ in the x-ξ plane; note that we define the "distance relative to I" $\mathrm{dist}^I(l_b, l_{b'})$ with a supremum rather than an infimum. Heuristically, by varying b', we can think of $\mathrm{Q_b}\,\varphi$ as being localized near a parallelogram P_b in the time-frequency plane with the following properties:

- P_b has two sides of length $|I|^{-1}$ parallel to the ξ-axis and centered on the line $\xi = 2bx$.
- The projection of P_b to the x-axis is the interval $I = \mathrm{supp}(\varphi)$.

The reader is urged here and throughout this summary to sketch pictures. Additionally, he or she is invited to carry out a similar analysis of $|\langle \mathrm{M}_c\,\varphi, \mathrm{M}_{c'}\,\varphi\rangle|$, where $\mathrm{M}_c\,f := e^{ic\,\cdot} f$ is a modulation operator; in this case our relative perspective will yield exactly the same rectangular tile given by the classical approach via L^∞ distribution. Combining these reasonings, we view $\mathrm{M}_c\,\mathrm{Q}_b\,\varphi$ as localized to a parallelogram obtained by translating the above parallelogram P_b by c units in the ξ-direction (*i.e.*, vertically). We can now precisely define our tiles.

Definition 2. *A* tile *is a triple $P = [\alpha, \omega, I]$, where α, $\omega \subset \mathbb{R}$ and $I \subset \mathbb{T}$ are half-open dyadic intervals with $|\alpha| = |\omega| = |I|^{-1}$. The collection of all such tiles P is denoted $\mathbb{P}$.*

In the sequel, we will glibly identify a tile P with the corresponding parallelogram in the x-ξ time-frequency plane obtained as above. We will occasionally abuse terminology by referring to other parallelograms (*e.g.* P_ℓ^* and P_r^* below) as "tiles," provided they arise from genuine tiles in a suitably natural manner. Finally, we define the "central line" of a tile P to be the unique line bisecting both sides of P parallel to the frequency axis.

6.3 Decomposition

We first note that $\mathrm{T}f$ can be rewritten as

$$\mathrm{T}f(x) = \sup_{b,c\in\mathbb{R}} |\,\mathrm{M}_c\,\mathrm{Q}_b\,\mathrm{H}\,\mathrm{Q}_b^*\,\mathrm{M}_c^*\,f(x)| = \sup_l |\mathrm{T}_l f(x)|,$$

where H is the Hilbert transform, the latter supremum is taken over lines $l(x) = c + 2bx$, and the operator T_l is given by

$$\mathrm{T}_l f(x) = \int_{\mathbb{T}} f(x-y)e^{i\left(l(x)y - by^2\right)}\frac{1}{y}\,\mathrm{d}y.$$

As in Fefferman's proof of Carleson's theorem, we begin by linearizing the supremum in our operator by defining a measurable function $x \mapsto l_x$, which should be thought of as picking out the l attaining the supremum over all lines. Then, provided that the bounds we obtain be uniform over all choices of $x \mapsto l_x$, we can replace T by

$$\mathrm{T}f(x) = \mathrm{T}_{l_x} f(x) = \int_{\mathbb{T}} f(x-y)e^{i\left(l_x(x)y - b(x)y^2\right)}\frac{1}{y}\,\mathrm{d}y,$$

where $l_x(\cdot) = c(x) + 2b(x)\,\cdot\,$.

As a preliminary decomposition, we break T up according to time scales; namely, we split the portion $\frac{1}{y}$ of the kernel as $\frac{1}{y} = \sum_{k\geq 0}\psi_k$, with $\psi_k = 2^k\psi(2^k\,\cdot\,)$ for some $\psi \in C_c^\infty$ supported away from 0, and we obtain

$$\mathrm{T}f(x) = \sum_{k\geq 0}\mathrm{T}_k f(x) := \sum_{k\geq 0}\int_{\mathbb{T}} f(x-y)e^{i\left(l_x(x)y - b(x)y^2\right)}\psi_k(y)\,\mathrm{d}y.$$

Now our goal is further to refine the T_k into pieces T_P associated to tiles P at scale k; the motivation for the following definition is that a tile P should contribute only when it captures some of both the time and the frequency content of $\mathrm{T}f$. To this end, let $P = [\alpha, \omega, I] \in \mathbb{P}$, and set

$$E(P) := \{x \in I \mid l_x \in P\};$$

here we say "$l \in P$" when the line l crosses both sides of P that are parallel to the frequency axis. At last, for $P = [\alpha, \omega, I]$ with $|I| = 2^{-k}$, we define

$$\mathrm{T}_P f(x) := \mathrm{T}_k f(x)\chi_{E(P)}(x).$$

Now for a fixed scale k the collection $\{E([\alpha,\omega,I]) \mid |I| = 2^{-k}\}$ is a partition of $\mathbb{T}$, whence

$$\mathrm{T}_k f(x) = \sum_{\substack{P=[\alpha,\omega,I]\\ |I|=2^{-k}}} \mathrm{T}_P f(x).$$

This in turn gives $\mathrm{T} = \sum_{P\in\mathbb{P}} \mathrm{T}_P$, which is our final decomposition.

For posterity, we record the explicit forms of our "building blocks" T_P and their adjoints:

$$\mathrm{T}_P f(x) = \left(\int_{\mathbb{T}} f(x-y) e^{i\left(l_x(x)y - b(x)y^2\right)} \psi_k(y)\, \mathrm{d}y \right) \chi_{E(P)}(x), \tag{1}$$

and

$$\mathrm{T}_P^* f(x) = -\int_{\mathbb{T}} \left(\chi_{E(P)} f\right)(x-y) e^{i\left(l_{x-y}(x-y)y + b(x-y)y^2\right)} \psi_k(y)\, \mathrm{d}y. \tag{2}$$

In accordance with the heuristics we have set forth, we think of T_P as time-frequency localized to to the tile P as before, and we think of T_P^* as localized to a "bi-tile" $P^* = P_\ell^* \sqcup P_r^*$. Here $P_r^* = [\alpha_r^*, \omega_r^*, I_r^* := 2I + \frac{9}{2}|I|]$ is a parallelogram obtained by time-dilating P and sliding it to the right along its central line; of course $|\alpha_r^*| = |\omega_r^*| = |\alpha| = |\omega|$. $P_\ell^* = [\alpha_\ell^*, \omega_\ell^*, I_\ell^*]$ is obtained similarly by sliding to the left; again, the reader is urged to draw a picture.

From (1), we obtain the pointwise estimate $|\mathrm{T}_P f| \lesssim \left(\frac{1}{|I^*|}\int_{I^*} |f|\right) \chi_{E(P)}$ and the L^2 operator norm estimate $\|\mathrm{T}_P\| \sim \left(\frac{|E(P)|}{|I|}\right)^{1/2}$.

6.4 Basic tools and philosophy of estimation

In this section we will give a heuristic overview (borrowed from [Fef]) of the proof of Theorem 1, which will hopefully make the productivity of our decomposition apparent. Here and in the sequel, for a collection of tiles $\mathcal{S} \subset \mathbb{P}$, we write

$$\mathrm{T}^{\mathcal{S}} := \sum_{P\in\mathcal{S}} \mathrm{T}_P.$$

Given the estimates on T_P noted just above, the following definition presents itself naturally:

Definition 3. *For* $P = [\alpha, \omega, I] \in \mathbb{P}$*, the* density factor *of* P *is*

$$A_0(P) := \frac{|E(P)|}{|I|}.$$

Partition the space of tiles as $\mathbb{P} = \bigcup_n \mathcal{S}_n$, with $\mathcal{S}_n := \{P \mid A_0(P) \sim 2^{-n}\}$; for each tile $P \in \mathcal{S}_n$, we have $\|\mathrm{T}_P\| \sim 2^{-n/2}$. If we actually had

$$\|\mathrm{T}^{\mathcal{S}_n}\| \sim 2^{-n/2},$$

we could sum in n and obtain L^2-boundedness of T by our decomposition. It turns out, perhaps surprisingly, that this latter estimate is almost true. The idea is to consider increasingly rich families of tiles $\mathcal{P} \subset \mathbb{P}$ such that

$$\|\mathrm{T}^{\mathcal{P}}\| \sim \max_{P \in \mathcal{P}} \|\mathrm{T}_P\|, \tag{3}$$

with the hope of eventually filling up the $\mathcal{S}_n$. The first step is to prove estimate (3) for basic clustered families called "trees" (cf. Definition 8 and Lemma 12). From this point, we hope to combine multiple trees and still preserve (3). This can be guaranteed for certain well-arranged collections of trees, morally by arranging for almost-orthogonality between the constituent trees; these families are called "forests" (cf. Definition 9 and Proposition 11). Finally, we will show roughly that $\mathcal{S}_n$ can be partitioned into n forests, from which we obtain a summable estimate in n.

Now of course the density factor $A_0(P)$ yields the operator norm $\|\mathrm{T}_P\|$ for a single tile, and it turns out to suffice for controlling operators associated to trees. However, A_0 cannot capture the relative oscillation between $\mathrm{T}_{P_1} f$ and $\mathrm{T}_{P_2} g$; given the outlook of Section 6.2, we would naturally hope to exploit this oscillation to obtain some cancellation when studying $|\langle \mathrm{T}^*_{P_1} f, \mathrm{T}^*_{P_2} g \rangle|$.[14] The following definition will help us accomplish this:

Definition 4. *Let* P_1*,* $P_2 \in \mathbb{P}$ *with* $|I_1| \geq |I_2|$*. The* geometric factor of the pair (P_1, P_2) *is*

$$\big(1 + \Delta(P_1, P_2)\big)^{-1},$$

where

$$\Delta(P_1, P_2) := \frac{\inf_{\substack{l_1 \in P_1 \\ l_2 \in P_2}} \mathrm{dist}^{I_2}(l_1, l_2)}{|\omega_2|}.$$

[14] This will be crucial in particular for the aforementioned almost-orthogonality. We examine $\mathrm{T}^*_P f$ rather than $\mathrm{T}_P f$ here due to smoothness properties; cf. (1) and (2).

To see how the geometric factor actually encodes oscillation, we appeal to stationary phase considerations similar to those of Section 6.2, as follows. Draw two tiles P_1 and P_2 for which $I_1^* \cap I_2^* \neq \emptyset$ (here I_j^* is the projection of the bi-tile P_j^* to the time axis), so that $\langle \mathrm{T}_{P_1}^* f, \mathrm{T}_{P_2}^* g \rangle$ is not trivial by support considerations. A good heuristic at this point is simply to suppose that for all $x \in E(P_j)$, we actually have $l_x = l_j$, the central line of the tile P_j. Now consider the point $x_{1,2}^i$ on the time axis over which these central lines intersect, *i.e.* $l_1(x_{1,2}^i) = l_2(x_{1,2}^i)$. The crucial observation, given the form (2) of T_P^*, is the following:

- For x near the intersection point $x_{1,2}^i$ the phases of $\mathrm{T}_{P_1}^* f(x)$ and $\mathrm{T}_{P_2}^* g(x)$ are roughly the same (since $l_{P_1}(x_{1,2}^i) = l_{P_2}(x_{1,2}^i)$), so we expect little cancellation.

- For x far from $x_{1,2}^i$, we have significant phase difference between the two terms, and we can hope to exploit cancellation.

To be a bit more precise, one fixes some small ε_0, and defines the ε_0-*critical intersection interval* $I_{1,2}$ centered at $x_{1,2}^i$, of length

$$\min\{|I_1|, |I_2|\}\big(1 + \Delta(P_1, P_2)\big)^{\varepsilon_0 - \frac{1}{2}}.$$

Applying L^∞ estimates inside $I_{1,2}$ and the method of nonstationary phase (*i.e.* integration by parts) outside, we obtain the following estimates:

Lemma 5. *For P_1, $P_2 \in \mathbb{P}$, let $\tilde{\chi}_{I_{1,2}^c}$ be a smoothed-out version of $\chi_{I_{1,2}^c}$. Then*

$$\left| \int \tilde{\chi}_{I_{1,2}^c} \mathrm{T}_{P_1}^* f \, \overline{\mathrm{T}_{P_2}^* g} \right| \lesssim_{n,\varepsilon_0} \big(1 + \Delta(P_1, P_2)\big)^{-n} \frac{\int_{E(P_1)} |f| \int_{E(P_2)} |g|}{\max\{|I_1|, |I_2|\}},$$

$$\left| \int_{I_{1,2}} \mathrm{T}_{P_1}^* f \, \overline{\mathrm{T}_{P_2}^* g} \right| \lesssim \big(1 + \Delta(P_1, P_2)\big)^{\varepsilon_0 - \frac{1}{2}} \frac{\int_{E(P_1)} |f| \int_{E(P_2)} |g|}{\max\{|I_1|, |I_2|\}},$$

and

$$\|\mathrm{T}_{P_1} \mathrm{T}_{P_2}^*\|^2 \lesssim \min\left\{\frac{|I_1|}{|I_2|}, \frac{|I_2|}{|I_1|}\right\} \big(1 + \Delta(P_1, P_2)\big)^{-1} A_0(P_1) A_0(P_2).$$

The moral one should take from this whole discussion is that far away from intersections we will morally be at liberty to apply Fefferman's methods, while close to intersections we will need to argue our problems away with "positive" methods (*i.e.* without recourse to cancellation).

Finally, we can introduce the correct substitute for A_0 by which we will actually partition $\mathbb{P}$; in view of Lemma 5, it measures how strongly P interacts with other tiles:

Definition 6. *The* mass *of $P \in \mathbb{P}$ is (roughly), for $N \gg 1$ fixed,*

$$A(P) := \sup_{\substack{P' \in \mathbb{P} \\ I \subset I'}} A_0(P')\big(1 + \Delta(P, P')\big)^{-N}.$$

6.5 Ordering, trees, and forests

Armed with these basic tools, we will now start on our way to reassembling our building blocks T_P. The crude initial outlook is that, by our previous discussions, families $\mathcal{P}$ consisting of (geometrically) well-separated tiles P should give rise to pieces $\mathrm{T}_P f$ which are essentially pairwise orthogonal; on the other hand, if the tiles comprising $\mathcal{P}$ overlap, more consideration will be required. Since we will suppress a good number of necessary technical considerations, many of our definitions and arguments will appear identical to those of [Fef]. Suffice it to say that the actual treatment of [Lie08a] is more technically involved, and the remainder of our discussion should be considered as only morally accurate.

Indeed, as in [Fef], we begin by defining a partial order on $\mathbb{P}$:

Definition 7. *Let P_1, $P_2 \in \mathbb{P}$. We declare $P_1 \trianglelefteq P_2$ iff $I_1 \subseteq I_2$ and $l \in P_1$ for every $l \in P_2$. Furthermore, we say $P_1 \triangleleft P_2$ iff $P_1 \trianglelefteq P_2$ and $I_1 \subsetneq I_2$.*

This order relation is a qualitative counterpart to the quantitative geometric factor, and it too encodes the "degree of overlap" of two tiles. We can now define our most basic "clustered" sets of tiles; they have the structure of set-theoretic trees under $\trianglelefteq$.

Definition 8. *A* tree $\mathcal{P}$ with top $P_0 \in \mathbb{P}$ *is a collection $\mathcal{P} \subset \mathbb{P}$ such that:*

- *If $P \in \mathcal{P}$, then $P \trianglelefteq P_0$.*
- *If $P_1, P_2 \in \mathcal{P}$ and $P_1 \trianglelefteq P \trianglelefteq P_2$, then $P \in \mathcal{P}$.*

In general we do not require that $P_0 \in \mathcal{P}$.

Finally, we will group together reasonably well-separated trees of uniformly bounded mass.

Definition 9. *A* (δ-)forest *is a collection* $\{\mathcal{P}_j\}_j$ *of trees with respective tops* $P_j = [\alpha_j, \omega_j, I_j]$ *such that:*

1. *For all* j *and all* $P \in \mathcal{P}_j$, $A(P) < \delta$.
2. *For all* $k \neq j$, $P \not\trianglelefteq P_k$ *for all* $P \in \mathcal{P}_j$.
3. *No point of* $[0,1]$ *belongs to more than* $K\delta^{-2}$ *of the* I_j.

The following two propositions, which are true up to our aforementioned technical deficiencies, are the key ingredients used to prove Theorem 1. They treat the "extremal" geometric configurations of sparse and clustered sets of tiles.

Proposition 10. *Let* $\mathcal{P}$ *be a family of pairwise incomparable (re* $\trianglelefteq$*) tiles with* $A(P) < \delta$ *for all* $P \in \mathcal{P}$. *Then there is an absolute constant* $0 < \eta < 1/2$ *such that*

$$\|\mathrm{T}^{\mathcal{P}}\| \lesssim \delta^{\eta}.$$

Proposition 11. *Let* $\{\mathcal{P}_j\}_j$ *be a* δ*-forest. Then there is an absolute constant* $0 < \eta < 1/2$ *and a set* $F \subset \mathbb{T}$ *with* $|F| \lesssim \delta^{50} K^{-1}$ *such that*

$$\Big\| \sum_j \mathrm{T}^{\mathcal{P}_j} f \Big\|_{L^2(\mathbb{T}\setminus F)} \lesssim \delta^{\eta} \log K \|f\|_2$$

for all $f \in L^2(\mathbb{T})$.

Proposition 10 will be used extensively for "garbage collection" purposes, to allow us to reduce families of tiles to nicer subcollections. Its proof is almost identical to that of its counterpart in [Fef], using a combination of TT^* and maximal methods. The proof of Proposition 11 is more involved and will be described later.

6.6 Proof of Theorem 1

Here we will selectively describe a few aspects of the proof of Theorem 1, just to give a flavor of the arguments involved.

We begin by dividing $\mathbb{P}$ into dyadic mass blocks as $\mathbb{P} = \bigcup_0^\infty \mathcal{P}_n$, where

$$\mathcal{P}_n := \{P \in \mathbb{P} \mid 2^{-n-1} < A(P) \leq 2^{-n}\}.$$

The reader will note the use of A rather than A_0, as alluded to in Section 6.4.

Our first goal is to reduce $\mathcal{P}_n$ to a collection of tiles clustered around certain maximal tiles whose time intervals do not overlap excessively. To this end, we take $\{\bar{P}_k = [\bar{\alpha}_k, \bar{\omega}_k, \bar{I}_K]\}$ to be the set of maximal (re $\trianglelefteq$) tiles in $\mathbb{P}$ subject to the condition $A_0(P) \geq 2^{-n-1}$. Let $\mathcal{C}_n \subset \mathcal{P}_n$ consist of those tiles $P \in \mathcal{P}_n$ with no ascending chains $P \lhd P_1 \lhd \ldots \lhd P_n$ of length $n+1$ contained in $\mathcal{P}_n$, and set $\mathcal{P}_n^0 := \mathcal{P}_n \setminus \mathcal{C}_n$.

It turns out that every tile P in $\mathcal{P}_n^0$ (roughly) satisfies $P \lhd \bar{P}_k$ for some $k \in \mathbb{N}$. But by the ascending chain condition $\mathcal{C}_n$ can be decomposed as a disjoint union of at most n families, each comprised of pairwise incomparable tiles. Applying Proposition 10 gives $\|\mathrm{T}^{\mathcal{C}_n}\| \lesssim 2^{-n\eta'}$ for some $\eta' > 0$; thus we are free to throw out $\mathcal{C}_n$ and replace $\mathcal{P}_n$ with $\mathcal{P}_n^0$.

Furthermore, by excising a suitable subset of $\mathbb{T}$ that is small enough to estimate away trivially, and by throwing out all tiles whose time intervals are contained in this set, we can replace $\mathcal{P}_n^0$ by a subcollection $\mathcal{P}_n^G$ (roughly) satisfying:

- $A(P) \leq 2^{-n}$ for all $P \in \mathcal{P}_n^G$.
- $P \in \mathcal{P}_n^G \Rightarrow P \trianglelefteq \bar{P}_k$ for some $k \in \mathbb{N}$.
- No $x \in \mathbb{T}$ belongs to more than $K2^{2n}$ of the intervals $\bar{I}_k$.

Note that conditions 1 and 3 of Definition 9 are now trivially guaranteed by the structure of $\mathcal{P}_n^G$; all of the work to obtain actual forests goes toward satisfying condition 2. For $P \in \mathcal{P}_n^G$, we define

$$B(P) := \mathrm{card}\{j \mid P \trianglelefteq \bar{P}_j\}.$$

Note that, by the final property of $\mathcal{P}_n^G$ listed above, we have $B(P) \leq 2^M$ with $M = 2n \log_2 K$. We split $\mathcal{P}_n^G$ dyadically with respect to B as $\mathcal{P}_n^G = \bigcup_{j=0}^M \mathcal{P}_{nj}$, with $\mathcal{P}_{nj} := \{P \in \mathcal{P}_n^G \mid 2^j \leq B(P) < 2^{j+1}\}$. In Fefferman's setting of rectangular tiles, the sets $\mathcal{P}_{nj}$ themselves turn out to be forests whose constituent trees are of the form $\{P \mid P \trianglelefteq \bar{P}_k\}$. For us, more trimming and technical effort is required to obtain the forest decomposition.

Once this decomposition has been effected, one simply adds up the $M \sim 2n \log K$ forest estimates of Proposition 11 for each n and then sums in n, keeping track of the small sets excised in the proposition and in the reduction to $\mathcal{P}_n^G$. One eventually obtains the estimate $|\{|\mathrm{T}f| > \lambda\}| \lesssim_\varepsilon (\|f\|_2/\lambda)^{2-\varepsilon}$, and the desired $L^2 \to L^p$ estimate follows for all $1 \le p < 2$.

6.7 Proof of the forest estimate

To conclude, we give a brief sketch of the ideas used to prove Proposition 11. Since a forest is by definition a collection of trees, the obvious plan of attack is to control $\mathrm{T}^{\mathcal{P}}$ for trees $\mathcal{P}$ and then to study interactions of such operators.

Lemma 12. *Let $\mathcal{P}$ be a tree with $A(P) < \delta$ for all $P \in \mathcal{P}$. Then $\|\mathrm{T}^{\mathcal{P}}\| \le \delta^{1/2}$.*

The idea here is to conjugate $\mathrm{T}^{\mathcal{P}}$ by a suitable operator $\mathrm{M}_c\,\mathrm{Q}_b$ to move the top of the tree $\mathcal{P}$ to the real axis; from this point, one proceeds just as in [Fef] with the heuristic "$\mathrm{T}^{\mathcal{P}}$ behaves like a maximal truncation of the Hilbert transform." By this we mean that, up to small error, we have

$$\mathrm{T}^{\mathcal{P}} f(x) \approx \sum_{k_0(x) \le k \le k_1(x)} \int f(x-y)\, \psi_k(y)\, \mathrm{d}y$$

for $x \in \operatorname{supp}(\mathrm{T}^{\mathcal{P}} f)$; the second part of Definition 8 is the crux of the matter here. To obtain the decay in mass, we basically note that the mass bound guarantees that $\operatorname{supp}(\mathrm{T}^{\mathcal{P}})$ is a "thin" set.[15]

Next, we declare two trees $\mathcal{P}_1$ and $\mathcal{P}_2$ to be *δ-separated* if every tile from $\mathcal{P}_1$ has geometric factor less than δ with respect to the top of $\mathcal{P}_2$ and vice versa. Consider two such δ-separated trees (with no *a priori* mass bounds) whose tops have the same time interval. Then the previous lemma (with $\delta = 1$) and the philosophy behind Lemma 5[16] give the estimate

$$\left|\left\langle \mathrm{T}^{\mathcal{P}_1\,*} f, \mathrm{T}^{\mathcal{P}_2\,*} g\right\rangle\right| \lesssim_n \delta^n \|f\|_{L^2(13I_0)} \|g\|_{L^2(13I_0)} + \left\|\chi_{I_c} \mathrm{T}^{\mathcal{P}_1\,*} f\right\|_2 \left\|\chi_{I_c} \mathrm{T}^{\mathcal{P}_2\,*} g\right\|_2.$$

Here I_c should be thought of as similar to the critical intersection interval of the tops of our trees defined above, modified to take δ into account.

[15] Due to the clustered nature of trees, we shouldn't expect the geometric factor to play for us significantly; thus, the accent of the mass bound falls on its density factor component.

[16] Lemma 5 treats the simplest non-empty trees, namely those consisting of a single tile.

To finish the proof, one first trims the trees in the forest down to a certain nice form, which accounts for the excised set in the statement of Proposition 11. After throwing away some tiles via Proposition 10 to give δ-separation between our trees, we arrange our forest into a controlled number of "rows" $\mathcal{R}_k$, each of which consists of trees whose tops have mutually disjoint time intervals. These rows are arranged so that the above interaction estimate together with maximal methods yields bounds for the $\mathrm{T}^{\mathcal{R}_k}$ as well as almost-orthogonality between these operators. An application of the Cotlar–Stein Lemma completes the proof.

References

[Car] Lennart Carleson. On convergence and growth of partial sums of Fourier series. *Acta Math.* **116** (1966), 135–157.

[Fef] Charles Fefferman. Pointwise convergence of Fourier series. *Ann. of Math.* **98** (1973), 551–571.

[Lie08a] Victor Lie. The (weak-L^2) boundedness of the quadratic Carleson operator. To appear in *Geom. Funct. Anal.*

[Lie08b] Victor Lie. On Stein's conjecture on the polynomial Carleson operator. Preprint, arXiv:0805.1580v1 [math.CA].

[Ste61] Elias M. Stein. On limits of sequences of operators. *Ann. of Math.* **74** (1961), 140–170.

[Ste93] Elias M. Stein. Oscillatory integrals related to Radon-like transforms. Proceedings of the Conference in Honor of Jean-Pierre Kahane (Orsay, 1993). *J. Fourier Anal. Appl.*, Special Issue (1995), 535–551.

[SW] Elias M. Stein and Stephen Wainger. Oscillatory integrals related to Carleson's theorem. *Math. Res. Lett.* **8** (2001), 789–800.

S. Zubin Gautam, UCLA
email: **sgautam@ucla.edu**

7 On a Conjecture of E. M. Stein on the Hilbert Transform on Vector Fields – Part II

after Michael T. Lacey and Xiaochun Li [1]
A summary written by Vjekoslav Kovač

Abstract

We prove the L^2 estimate for the Hilbert transform on Lipschitz vector fields in the plane, assuming the conjectured estimate for the Lipschitz Kakeya maximal function, presented in the first part of this exposition.

7.1 Statement of the main result

In the first part of this presentation the following version of the *Lipschitz Kakeya maximal function* is introduced. Let v be a Lipschitz vector field on the plane, i.e. a Lipschitz continuous map $v\colon \mathbb{R}^2 \to \mathrm{S}^1$. For $0 < \delta < 1$ and $0 < w < \frac{1}{100}\|v\|_{\mathrm{Lip}}$, we define:

$$\mathrm{M}_{v,\delta,w} f(x) := \sup_{\substack{R \text{ rectangle} \\ |V(R)| \geq \delta |R| \\ w \leq W(R) \leq 2w}} \frac{\mathbf{1}_R(x)}{|R|} \int_R |f(y)|\, dy,$$

where $W(R)$ is the width of R and $\frac{|V(R)|}{|R|}$ is the portion of R that "respects" the direction of v.

The authors state the following conjecture in [1]:

Conjecture 1. *For some* $1 < p < 2$*, some* $N \in \mathbb{N}$*, all* $0 < \delta < 1$*, all Lipschitz vector fields* v*, and* $0 < w < \frac{1}{100}\|v\|_{\mathrm{Lip}}$*, the maximal function* $\mathrm{M}_{\delta,v,w}$ *is bounded from* $L^p(\mathbb{R}^2)$ *to* $L^{p,\infty}(\mathbb{R}^2)$ *with norm* $\lesssim \delta^{-N}$*.*

The main result we present is a conditional one: assuming that a given vector field satisfies the above conjecture, we derive boundedness of the Hilbert transform on that vector field.

For a Schwartz function $f\colon \mathbb{R}^2 \to \mathbb{C}$ we define:

$$\mathrm{H}_{v,\varepsilon} f(x) := \mathrm{p.v.} \int_{-\varepsilon}^{\varepsilon} f(x - yv(x))\, \frac{dy}{y}\,.$$

In words, $\mathrm{H}_{v,\varepsilon}$ is a truncated Hilbert transform along the line through x with direction $v(x)$. Locally it is one-dimensional, but is performed on functions of two variables.

To state the result, let us denote by S_t the operator that restricts the frequency to an annulus $1/t \leq |\xi| \leq 2/t$, or explicitly:

$$\mathrm{S}_t\, f(x) = \int_{1/t \leq |\xi| \leq 2/t} \widehat{f}(\xi) e^{i\xi\cdot x}\, d\xi\,.$$

In their earlier paper [2] the authors show:

Theorem 2. *For any* measurable *vector field v we have the $L^2 \to L^{2,\infty}$ estimate*

$$\sup_{\lambda>0} \lambda |\{|\,\mathrm{H}_{v,\infty} \circ \mathrm{S}_t\, f| > \lambda\}|^{1/2} \lesssim \|f\|_2\,.$$

This weak estimate is sharp for measurable vector fields. In this exposition we prove the best known result for Lipschitz vector fields:

Theorem 3. *Assume that conjecture 1 holds for some vector field v. There exists a universal constant $K > 0$ such that for $\varepsilon = K/\|v\|_{\mathrm{Lip}}$ and $0 < t < \|v\|_{\mathrm{Lip}}$ we have*

$$\|\,\mathrm{H}_{v,\varepsilon} \circ \mathrm{S}_t\,\|_2 \lesssim 1.$$

Moreover, if $v \in C^{1+\eta}$, $\eta > 0$, then for $\varepsilon = K/\|v\|_{C^{1+\eta}}$ we have

$$\|\,\mathrm{H}_{v,\varepsilon}\,\|_2 \lesssim (1 + \log \|v\|_{C^{1+\eta}})^2.$$

The authors also state the following conjecture, which we do not tackle in this exposition.

Conjecture 4. *Assume that conjecture 1 holds for some vector field $v \in C^{1+\eta}$, $\eta > 0$. There exists a universal constant $K > 0$ such that for $\varepsilon = K/\|v\|_{C^{1+\eta}}$ and $2 < p < \infty$ we have*

$$\|\,\mathrm{H}_{v,\varepsilon}\,\|_p \lesssim (1 + \log \|v\|_{C^{1+\eta}})^2.$$

7.2 Notation and terminology

Throughout the note κ will denote a fixed small positive constant.

A *grid* is a collection of intervals $\mathcal{G}$ so that for $I, J \in \mathcal{G}$ we have $I \cap J = \emptyset$ or $I \subset J$ or $J \subset I$. A grid is called *central* if for all $I, J \in \mathcal{G}$ with $I \subsetneqq J$ we have $500\kappa^{-20} I \subset J$.

Let ρ denote the rotation by $\pi/2$ and $\mathrm{e}_\perp := \rho(\mathrm{e})$. A *rectangle* is $\omega \subset \mathbb{R}^2$ that is a product of intervals with respect to the coordinate system $(\mathrm{e}, \mathrm{e}_\perp)$. We say that ω is an *annular rectangle* if $\omega = (-2^{l-1}, 2^{l-1}) \times (a, 2a)$ for an integer l with $2^l < \kappa\, a$. In that case the *scale* of ω is $\mathrm{scl}(\omega) := 2^l$ and the *annular parameter* of ω is $\mathrm{ann}(\omega) := a$.

Annular rectangles will decompose functions in the frequency variables. The uncertainty principle motivates the following definition.

Two rectangles R and R' are said to be *dual* if they are rectangles with respect to the same coordinate system $(\mathrm{e}, \mathrm{e}_\perp)$, and if we write $R = r_1 \times r_2$ and $R' = r'_1 \times r'_2$, then $1 \le |r_1|\,|r'_1| \le 4$, $1 \le |r_2|\,|r'_2| \le 4$. The product of two dual rectangles could be called a *phase rectangle*, the first component being the frequency one, the second being the spatial one.

We consider collections of phase rectangles $\mathcal{AT}$ that satisfy the following conditions.

(1) For each $s \in \mathcal{AT}$, $s = \omega_s \times R_s$, the frequency component ω_s is an annular rectangle.

(2) ω_s and R_s are dual.

(3) The spatial components R_s are from the product of central grids.

(4) For each $s \in \mathcal{AT}$ the family $\{1000\kappa^{-100} R : \omega_s \times R \in \mathcal{AT}\}$ covers $\mathbb{R}^2$.

(5) $\mathrm{ann}(\omega_s)$ is of the form 2^j for some $j \in \mathbb{Z}$.

(6) $\sharp\{\omega_s : \mathrm{scl}(s) = \mathrm{scl},\ \mathrm{ann}(s) = \mathrm{ann}\} \gtrsim \frac{\mathrm{ann}}{\mathrm{scl}}$

(7) $\mathrm{scl}(s) \le \kappa\, \mathrm{ann}(s)$

(8) There are auxiliary sets $\omega_s, \omega_{s1}, \omega_{s2} \subset \mathbb{T}$ associated to s such that $\Omega := \{\omega_s, \omega_{s1}, \omega_{s2} : s \in \mathcal{AT}\}$ is a grid in $\mathbb{T}$.

(9) $\omega_{s1} \cap \omega_{s2} = \emptyset \quad$ and $\quad |\omega_s| \ge 32(|\omega_{s1}| + |\omega_{s2}| + \mathrm{dist}(\omega_{s1}, \omega_{s2}))$

(10) ω_{s1} lies clockwise from ω_{s2} on $\mathbb{T}$.

(11) $|\omega_s| \le K \frac{\mathrm{scl}(\omega_s)}{\mathrm{ann}(\omega_s)}$

(12) $\{\frac{\xi}{|\xi|} \,:\, \xi \in \omega_s\} \subset \rho(\omega_{s1})$

Collections $\mathcal{AT}$ satisfying the above conditions will be called *annular tiles.*

* * *

We are going to associate three different functions to each phase rectangle.

Let φ be a fixed Schwartz function for which $\hat{\varphi}$ is nonnegative, supported on a small ball $\mathrm{B}(\mathbf{0}, \kappa)$, and equal to 1 on $\mathrm{B}(\mathbf{0}, \kappa/2)$. We associate the following wave packet to a tile $s \in \mathcal{AT}$:

$$\varphi_s := \mathrm{Mod}_{\mathrm{center}(\omega_s)} \mathrm{Tran}_{\mathrm{center}(R_s)} \mathrm{Dil}^{(2)}_{R_s} \varphi.$$

Here $\mathrm{Dil}^{(2)}_{R_s}$ denotes the L^2 normalized dilation operator, i.e. $\mathrm{D}^{(2)}_{I\times J} f(x_1, x_2) := (|I|\,|J|)^{-1/2} f(\frac{x_1}{|I|}, \frac{x_2}{|J|})$.

Suppose that $(\psi_t)_{t>0}$ are such that $\hat{\psi}_t$ is supported in $[-\theta-\kappa, -\theta+\kappa]$, $\theta > 0$, and $|\psi_t(x)| \lesssim_N (1+|x|)^{-N}$ for $N \in \mathbb{N}$. Define

$$\psi_s(y) := \mathrm{scl}(s)\psi_{\mathrm{scl}(s)}(\mathrm{scl}(s)y)$$

and

$$\phi_s(x) := \int_{\mathbb{R}} \varphi_s(x - yv(x))\psi_s(y)\, dy = \mathbf{1}_{\omega_{s2}}(v(x)) \int_{\mathbb{R}} \varphi_s(x - yv(x))\psi_s(y)\, dy$$

for every $s \in \mathcal{AT}$.

7.3 Main ingredients of the proof

The model operator we consider acts on Schwartz functions and is defined by

$$\mathcal{C}_{\mathrm{ann}} f := \sum_{\substack{s \in \mathcal{AT} \\ \mathrm{ann}(s)=\mathrm{ann} \\ \mathrm{scl}(s) \geq \|v\|_{\mathrm{Lip}}}} \langle f, \varphi_s \rangle \phi_s\,.$$

Remember that $\mathrm{ann}(s) = 2^j$ for some $j \in \mathbb{Z}$, so let us also add up over *ann* to define:

$$\mathcal{C} := \sum_{j=1}^{\infty} \mathcal{C}_{2^j}\,.$$

The proof of theorem 3 follows from the following two lemmas by averaging over over all translations, dilations and rotations of grids.

Lemma 5. *Assume that the vector field v is Lipschitz and satisfies conjecture 1. For all* ann $\geq \|v\|_{\mathrm{Lip}}^{-1}$ *the operator* $\mathcal{C}_{\mathrm{ann}}$ *extends to a bounded linear map on* $L^2(\mathbb{R}^2)$*, with* $\|\mathcal{C}_{\mathrm{ann}}\|_2 \lesssim 1$*.*

We remark that for $2 < p < \infty$ the only condition needed for $\|\mathcal{C}_{\mathrm{ann}}\|_p \lesssim 1$ is measurability of v, a result from [2].

Lemma 6. *Assume that* $v \in C^{1+\eta}$ *and* $\|v\|_{C^{1+\eta}} \leq 1$ *for some* $\eta > 0$*, and again that* v *satisfies conjecture 1. Then* $\|\mathcal{C}\|_2 \lesssim 1$ *and additionally for all* $l \in \mathbb{Z}$ *we have:*

$$\Big\| \sum_{j=-\infty}^{+\infty} \sum_{\substack{s\in\mathcal{AT} \\ \mathrm{ann}(s)=2^j \\ \mathrm{scl}(s)=2^l}} \langle f, \varphi_s \rangle \phi_s \Big\|_2 \lesssim \left(1 + \log(1 + 2^{-l}\|v\|_{C^{1+\eta}})\right).$$

The operators $\mathcal{C}_{\mathrm{ann}}$ and $\mathcal{C}$ are constructed from a kernel which is a smooth analogue of the truncated kernel p.v.$\frac{1}{t}\mathbf{1}_{\{|t|\leq 1\}}$. In the proof of theorem 3 we pass to a smooth kernel in the following way. One can choose Schwartz kernels $(\psi_{(1+\kappa)^n})_{n\in\mathbb{Z}}$ such that for

$$K(t) := \sum_{n\in\mathbb{Z}} a_n (1+\kappa)^n \psi_{(1+\kappa)^n}((1+\kappa)^n t)$$

we have p.v.$\frac{1}{t}\mathbf{1}_{\{|t|\leq 1\}} = K(t) - \overline{K(t)}$. Here $|a_n| \lesssim 1$ for $n \geq 0$, and $|a_n| \lesssim (1+\kappa)^{-n}$ for $n < 0$. The main part of the sum is for $n \geq \max(0, \|v\|_{C^{1+\eta}})$ and corresponds to the operator $\mathcal{C}$. For the part of the sum where $n < \max(0, \|v\|_{C^{1+\eta}})$ we use rapid decay of coefficients a_n and the estimate from lemma 6.

References

[1] Lacey M. T. and Li X., *On a Conjecture of E. M. Stein on the Hilbert Transform on Vector Fields*, To appear in Memoirs AMS, `arXiv:0704.0808v3 [math.CA]`

[2] Lacey M. T. and Li X., *Maximal theorems for the directional Hilbert transform on the plane*, Trans. Amer. Math. Soc., **358**, 2006, no. 9, 4099–4117 (electronic)

Vjekoslav Kovač, UCLA
email: **vjekovac@math.ucla.edu**

8 Return Times Theorem via Time-Frequency Fourier Analysis

after Demeter, Lacey, Tao and Thiele [2]
A summary written by Patrick LaVictoire

Abstract

In this first of two summaries on [2], we reduce the extension of Bourgain's Return Times Theorem to a theorem in time-frequency Fourier analysis regarding a model operator; this is complementary to the presentation of M. Bateman.

8.1 Introduction

This summary and that of M. Bateman will together cover the result of Demeter, Lacey, Tao and Thiele [2] extending the Return Times Theorem beyond the range of exponents implied by duality.

We consider dynamical systems $\mathbf{X} = (X, \Sigma, \mu, \tau)$, where (X, Σ, μ) is a probability space and $\tau : X \to X$ is a measure-preserving transformation. We wish to examine the range of exponents $1 \leq p, q \leq \infty$ for which the following theorem is valid:

Theorem 1 (Return Times Theorem for Exponents p, q)**.** *Let $\mathbf{X} = (X, \Sigma, \mu, \tau)$ be a dynamical system. For each function $f \in L^p(X)$, there exists 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 $g \in L^q(Y)$ and each $x \in X_0$, the averages*

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

converge ν-almost everywhere in Y.

Bourgain [1] proved this theorem originally for $p = q = \infty$, which implies the entire duality range $\frac{1}{p} + \frac{1}{q} \leq 1$ by an application of Hölder's inequality to the relevant maximal inequality. Assani, Buczolich and Mauldin [4] have produced a negative result for $p = q = 1$.

The main result of [2] is an extension of the Return Times Theorem to all $p > 1, q \geq 2$ by a method of time-frequency Fourier analysis. In this summary, we will show how the main theorem reduces to an inequality for a model operator (Theorem 5), working backward to show that a series of intermediate results each imply Theorem 1; the complementary chapter by M. Bateman will then prove Theorem 5. In some places, we will follow the structure of a recent preprint by Demeter [3] which uses similar techniques to extend the range of exponents to all $1 \leq p, q \leq \infty$ with $\frac{1}{p} + \frac{1}{q} < \frac{3}{2}$.

8.2 From Ergodic Theory to the Real Line

8.2.1 Reduction to a Maximal Ergodic Inequality

In many problems of ergodic theory, an a.e. convergence result typically reduces to a result for convergence for a dense class of functions, coupled with a maximal inequality. Since we already have Theorem 1 for $p = q = \infty$, we claim

Lemma 2. *Suppose that for p and q we have the maximal inequality*

$$\left\| \sup_{\mathbf{Y}} \sup_{\|g\|_{L^q(Y)}=1} \left\| \sup_N \left| \frac{1}{N} \sum_{n=1}^{N} f(\tau^n x) g(\sigma^n y) \right| \right\|_{L^q_y(Y)} \right\|_{L^p_x(X)} \lesssim_{p,q} \|f\|_{L^p(X)}. \quad (1)$$

Then Theorem 1 holds for that p and q.

This would be proved by a standard convergence argument akin to the derivation of the Lebesgue Differentiation Theorem from the Hardy-Littlewood Maximal Inequality, except that the supremum over $\mathbf{Y}$ in the left-hand side of (1) raises the question of measurability. This is established by an application of the Conze principle (Theorem 4.2 in [2]).

It is quickly verified that Hölder's inequality and the maximal ergodic theorem imply (1) within the duality range $1 \leq p, q \leq \infty$, $\frac{1}{p} + \frac{1}{q} \leq 1$.

8.2.2 Transfer to a Maximal Inequality on $\mathbb{R}$

By a standard transfer argument, we will reduce (1) to the following maximal inequality on $\mathbb{R}$:

Lemma 3. *For each* $1 < p < \infty$ *and each* $f \in L^p(\mathbb{R})$,

$$\left\| \sup_{\|g\|_{L^2(\mathbb{R})}=1} \left\| \sup_{t>0} \frac{1}{2t} \int_{-t}^{t} |f(x+y)g(z+y)|dy \right\|_{L^2_z(\mathbb{R})} \right\|_{L^p_x(\mathbb{R})} \lesssim \|f\|_{L^p(\mathbb{R})}. \quad (2)$$

By considering step functions, (2) directly implies an analogous inequality on $\mathbb{Z}$: for each finitely supported $\phi : \mathbb{Z} \to \mathbb{R}^+$,

$$\left\| \sup_{\substack{\psi : \mathbb{Z} \to \mathbb{R}^+ \\ \|\psi\|_{\ell^2(\mathbb{Z})} = 1}} \left\| \sup_{N} \frac{1}{N} \sum_{b=1}^{N} \phi(a+b)\psi(c+b) \right\|_{\ell^2_c(\mathbb{Z})} \right\|_{\ell^p_a(\mathbb{Z})} \lesssim \|\phi\|_{\ell^p(\mathbb{Z})}. \quad (3)$$

Now let $\mathbf{X}$ and $\mathbf{Y}$ be two dynamical systems, where we assume $\mathbf{Y}$ is ergodic. (This is no restriction: the Conze principle implies that we can replace the supremum over $\mathbf{Y}$ in (1) with any single nonatomic ergodic $\mathbf{Y}$.) Take positive functions $f \in L^p(X)$, $g \in L^q(Y)$, and fix $x \in X$, $y \in Y$. If we fix a sufficiently large $K > 0$ and define $\phi, \psi : \mathbb{Z} \to \mathbb{Z}$ by

$$\phi(n) := \begin{cases} f(\tau^n x) & 0 \le n \le K \\ 0 & \text{otherwise} \end{cases} \qquad \psi(n) := \begin{cases} g(\sigma^n y) & 0 \le n \le K \\ 0 & \text{otherwise,} \end{cases}$$

then applying (3) to these functions and integrating over y and x proves (1).

8.2.3 Toward the Model Operator

Finally, we will find it useful to deal with smooth kernels rather than characteristic functions of intervals. Thus for $f \in L^\infty(\mathbb{R})$ and a kernel $K \in L^2(\mathbb{R})$ with $K \ge 0$, $K(0) > 0$ and supp $\hat{K} \subseteq [-1, 1]$, we define

$$Rf(x) \quad := \quad \sup_{\|g\|_{L^2(\mathbb{R})}=1} \left\| \sup_{k \in \mathbb{Z}} 2^{-k} \left| \int f(x+y)g(z+y)K(y2^{-k})dy \right| \right\|_{L^2_z(\mathbb{R})}.$$

Since K is C^∞, it majorizes some multiple of a characteristic function $1_{[-t,t]}$. Therefore the inequality $\|Rf\|_{L^p(\mathbb{R})} \lesssim_p \|f\|_{L^p(\mathbb{R})}$ for all $p > 1$ will immediately imply (2) and thus the Return Times Theorem for all $p > 1$, $q \ge 2$.

Hölder's inequality and the L^2 boundedness of the Hardy-Littlewood maximal operator prove that $\|Rf\|_{L^2(\mathbb{R})} \lesssim_p \|f\|_{L^2(\mathbb{R})}$, so by the Marcinkiewicz Interpolation Theorem we will only need a restricted weak type inequality in L^1. Therefore the main result of [2] reduces to the following theorem:

Theorem 4. *Let $K \in L^2(\mathbb{R})$ with $K \geq 0$, $K(0) > 0$ and supp $\hat{K} \subseteq [-1, 1]$. Then for all $1 < p < \infty$, all $F \subset \mathbb{R}$ with $|F| < \infty$, and all $\lambda \leq 1$,*

$$m\{x : R1_F(x) > \lambda\} \lesssim_p \frac{|F|}{\lambda}. \tag{4}$$

8.3 Discretization of $R1_F$

8.3.1 Decomposition and Restatement as Multiplier Norm

We will decompose $R1_F$ by a Gabor basis expansion. We take a Schwartz function φ such that supp $\hat{\varphi} \subseteq [0, 1]$ and

$$\sum_{l \in \mathbb{Z}} |\hat{\varphi}(\xi - \frac{l}{2})|^2 \equiv C,$$

and then define

$$\varphi_{k,m,l}(x) := 2^{-\frac{k}{2}} \varphi(2^{-k}x - m) e^{2\pi i 2^{-k} x l}.$$

By considering the inversion formula for Fourier series (which applies since supp $\hat{\varphi}_{0,m,l} \subseteq [-l, 1-l]$), we see that $\sum_{m \in \mathbb{Z}} \langle \hat{f}, \hat{\varphi}_{0,m,\frac{l}{2}} \rangle \hat{\varphi}_{0,m,\frac{l}{2}} \equiv C' \hat{f} |\hat{\varphi}_{0,0,\frac{l}{2}}|^2$ for each $f \in L^2$ and $l \in \mathbb{Z}$, which further implies (by dilation invariance in k) that

$$\sum_{m,l \in \mathbb{Z}} \langle f, \varphi_{k,m,\frac{l}{2}} \rangle \varphi_{k,m,\frac{l}{2}} \equiv C'' f \qquad \forall f \in L^2.$$

We will choose C so that $C'' = 1$. Using this expansion and the Fourier transform, $R1_F$ can be expressed as a multiplier norm; for a sequence of functions $m_k(\theta) : \mathbb{R} \to \mathbb{R}$ we define

$$\|(m_k(\theta))_{k \in \mathbb{Z}}\|_{M^*_{q,\theta}(\mathbb{R})} := \sup_{\|h\|_q = 1} \left\| \sup_k \left| \int m_k(\theta) \hat{h}(\theta) e^{2\pi i \theta x} d\theta \right| \right\|_{L^q_x(\mathbb{R})}$$

and rewrite

$$R1_F(x) = \left\| \left(\sum_{m,l \in \mathbb{Z}} \langle 1_F, \varphi_{k,m,l/2} \rangle \phi_{k,m,l/2}(x, \theta) \right)_{k \in \mathbb{Z}} \right\|_{M^*_{2,\theta}(\mathbb{R})}$$

where

$$\phi_{k,m,l/2}(x,\theta) := \mathcal{F}[\varphi_{k,m,l/2}(x+\cdot)2^{-k}K(\cdot\, 2^{-k})](\theta).$$

Note that $|\phi_{k,m,l/2}(x,\theta)| \leq \|\varphi_{k,m,l/2}(x+\cdot)2^{-k}K(\cdot\, 2^{-k})\|_{L^1(\mathbb{R})}$, and therefore $\phi_{k,m,l/2}(x,\theta) \leq C_k\varphi_{k,m,l/2}(x)$ for x large; also,

$$\begin{aligned} \text{supp}_\theta\ \phi_{k,m,l/2}(x,\theta) &\subseteq \text{supp}\ (\hat{\varphi}_{k,m,l/2} * \hat{K}(\cdot\, 2^k)) \\ &\subseteq [l2^{-k},(l+1)2^{-k}] + [-2^{-k},2^{-k}]. \end{aligned}$$

Similarly, the Fourier transform $\mathcal{F}_x[\phi_{k,m,l/2}(x,\theta)](\xi) = \hat{\varphi}(\xi)\hat{K}(2^k(\theta-\xi))$ has support contained in $[l2^{-k},(l+1)2^{-k}]$. These localization properties will be essential to the argument, as our model operator will act on functions supported in time and frequency on these shifted dyadic 'tiles'. We also make use of elementary regularity properties of $\phi_{k,m,l/2}$ and $\varphi_{k,m,l/2}(x)$, as shown in the next section.

8.3.2 Tiles

The final stage in the reduction of this problem is the introduction of tiles in 'time' and 'frequency' variables. We define the collection $\mathbf{S}_{univ}$ of all tiles $s = I_s \times \omega_s \subset \mathbb{R}^2$ with area 1, where I_s and ω_s are dyadic intervals; we refer to these as the time and frequency components of s, respectively.

We define a collection of tiles $\mathbf{S} \subseteq \mathbf{S}_{univ}$ to be *convex* if $s, s'' \in \mathbf{S}$ implies that $s' \in \mathbf{S}$ for every $s' \in \mathbf{S}_{univ}$ with $I_s \subseteq I_{s'} \subseteq I_{s''}$ and $\omega_{s''} \subseteq \omega_{s'} \subseteq \omega_s$. We may now state the theorem in time-frequency analysis to which the improved Return Times Theorem of [2] reduces:

Theorem 5. *Let* $\mathbf{S}$ *be a convex finite collection of tiles. Say that we have sets of Schwartz functions* $\{\phi_s : s \in \mathbf{S}\}$ *and* $\{\varphi_s : s \in \mathbf{S}\}$ *such that*

$$\textit{supp}_\theta\ \phi_s(x,\theta) \quad \subseteq \quad \omega_s \qquad \forall x \tag{5}$$

$$\textit{supp}_\xi\ \mathcal{F}_x(\phi_s(x,\theta))(\xi) \quad \subseteq \quad \omega_s \qquad \forall \theta \tag{6}$$

$$\sup_{c\in\omega_s}\left\|\frac{\partial^n}{\partial\theta^n}\frac{\partial^m}{\partial x^m}[\phi_s(x,\theta)e^{-2\pi icx}]\right\|_{L^\infty_\theta(\mathbb{R})} \quad \lesssim_{n,m,M} \quad |I_s|^{n-m-1/2}\chi^M_{I_s}(x) \tag{7}$$

for all $n, m, M \geq 0$*, uniformly in* s*; and such that*

$$\textit{supp}\ \hat{\varphi}_s \quad \subseteq \quad \omega_s \tag{8}$$

$$\sup_{c\in\omega_s}\left|\frac{\partial^n}{\partial x^n}[\varphi_s(x)e^{-2\pi icx}]\right\| \quad \lesssim_n, M \quad |I_s|^{-n-1/2}\chi^M_{I_s}(x) \tag{9}$$

uniformly in s. Then for each measurable $F \subset \mathbb{R}$ *with* $|F| < \infty$ *and each* $0 < \lambda \leq 1$,

$$m \left\{ x : \left\| \left(\sum_{s \in \mathbf{S}: |I_s| = 2^k} \langle 1_F, \varphi_s \rangle \phi_s(x, \theta) \right)_{k \in \mathbb{Z}} \right\|_{M^*_{2,\theta}(\mathbb{R})} > \lambda \right\} \lesssim \frac{|F|}{\lambda} \tag{10}$$

with the implicit constant depending only on the constants in (7) and (9).

At this point, the reader is directed to M. Bateman's summary of the second part of [2], for an outline of the proof of this theorem.

The author is indebted to C. Demeter for valuable recommendations on the organization of the summary and presentation, and in particular for passing along a copy of [3].

References

[1] Bourgain, Jean. *Return Times of Dynamical Systems.* C. R. Acad. Sci. Paris t.**306**, Série I (1988), no. 12, 483–485.

[2] Demeter, Ciprian; Lacey, Michael; Tao, Terence; Thiele, Christoph. *Breaking the Duality in the Return Times Theorem.* Duke Math. J. **143** (2008), no. 2, 281–355.

[3] Demeter, Ciprian. *Improved Range in the Return Times Theorem.* Available at http://arxiv.org/abs/0904.1254

[4] Assani, Idris; Buczolich, Zoltán; Mauldin, R. Daniel. *An* L^1 *counting problem in ergodic theory.* J. Anal. Math. **95** (2005), 221–241.

Patrick LaVictoire, UC Berkeley
email: **patlavic@math.berkeley.edu**

9 Sum rules and spectral measures of Schrödinger operators with L^2 potentials

after Rowan Killip and Barry Simon [1]
A summary written by Helge Krüger

Abstract

We prove the L^2 estimate for the Hilbert transform on Lipschitz vector fields in the plane, assuming the conjectured estimate for the Lipschitz Kakeya maximal function, presented in the first part of this exposition.

9.1 Some Notation

In order to set notation, we let $\mathbb{R}^+ = [0, \infty)$ be the half line. We denote by $H^2(\mathbb{R}^+)$ the Sobolev space of twice (weakly) differentiable functions. We say that $V \in L^2_{\text{loc}}(\mathbb{R}^+)$ if for every $a > 0$,

$$\int_0^a |V(x)|^2 dx < \infty.$$

We introduce a domain by

$$\mathfrak{D}(H) = \{f \in H^2(\mathbb{R}^+) : \quad f(0) = 0\}. \tag{1}$$

Since, the embedding $H^2(\mathbb{R}^+) \to L^\infty(\mathbb{R}^+)$ is continuous, we have that $Vf \in L^2(\mathbb{R}^+)$ for $f \in \mathfrak{D}(H)$[17] and $V \in L^2_{\text{loc}}(\mathbb{R}^+)$. Assume furthermore that V is real-valued. Hence, we can define an operator H with domain $\mathfrak{D}(H) \subseteq L^2(\mathbb{R}^+)$ by

$$H = -\frac{d^2}{dx^2} + V. \tag{2}$$

We say that V is *limit-point at* ∞ if H defines a self-adjoint operator, and we will assume from now on that V is limit-point. Furthermore note that V is commonly known as *the potential*.

Next, since H is self-adjoint, we can define its *spectrum* by

$$\sigma(H) = \{z \in \mathbb{C} : \quad (H - z)^{-1} \text{ is not a bounded operator}\}^{18}, \tag{3}$$

[17] Here, we used $V \in L^2_{\text{loc}}(\mathbb{R}^+)$. The assumption $V \in L^1_{\text{loc}}(\mathbb{R}^+)$ would be sufficient to define H, however then one has to work with *quadratic forms* (see Reed–Simon).

[18] As an operator $L^2(\mathbb{R}^+) \to L^2(\mathbb{R}^+)$

from self-adjointness we know that $\sigma(H) \subseteq \mathbb{R}$. Furthermore, if $V \in L^2(\mathbb{R}^+)$ one can show that the operator

$$V(-\frac{d^2}{dx^2} + \mathrm{i})^{-1}$$

is compact. This is commonly known as V is *relatively compact*.

Lemma 1. *If $V \in L^2(\mathbb{R}^+)$, then there are E_j (possibly 0) such that $\lim_{j\to\infty} E_j = 0$ (if there are infinitely many), such that*

$$\sigma(H) = [0, \infty) \cup \{E_j\}. \tag{4}$$

Proof. Follows from the above compactness and $\sigma(-\frac{d^2}{dx^2}) = [0, \infty)$. □

We call the E_j the eigenvalues and $[0, \infty)$ the essential spectrum.

9.2 The spectral measure

Since $\sigma(H) \subseteq \mathbb{R}$, we may find for $z \in \mathbb{C}\backslash\mathbb{R}$ a solution ψ of $-\psi'' + V\psi = z\psi$, $\psi \in L^2(\mathbb{R}^+)$, and $\psi(0) = 1$.[19] Given this solution, we introduce the m-function by

$$m(z) = \psi'(0). \tag{5}$$

Next, one can show that for $\mathrm{Im}(z) > 0$, one also has that $\mathrm{Im}(m(z)) > 0$. Hence, m is a *Herglotz function*. This implies that there exists a measure ρ called the *spectral measure*, such that

$$\int \frac{d\rho(E)}{1 + E^2} < \infty \tag{6}$$

and

$$m(z) = \int \left(\frac{1}{E - z} - \frac{E}{1 + E^2} \right) d\rho(E) + \mathrm{Re}(m(i)). \tag{7}$$

Furthermore, the boundary values

$$m(E + \mathrm{i}0) = \lim_{\varepsilon \downarrow 0} m(E + \mathrm{i}\varepsilon) \tag{8}$$

exist for almost every $E \in \mathbb{R}$. The importance of the spectral measure comes from the following two facts.

[19]If $\psi(0) = 0$, then $\psi \in \mathfrak{D}(H)$, so z is an eigenvalue, which is a contradiction. So we must have $\psi(0) \neq 0$, and we can achieve $\psi(0) = 1$ by multiplication by a constant.

Theorem 2 (Spectral Theorem[20]). *H on $L^2(\mathbb{R}^+)$ is unitarily equivalent to multiplication by λ in $L^2(\mathbb{R}, \rho)$.*

Furthermore, we have that

Theorem 3 (Borg–Marchenko). *The map from the limit-point $V \in L^2_{\text{loc}}(\mathbb{R}^+)$ to their spectral measure ρ is injective.*

One of the main concerns of spectral theory is to understand the spectral measures ρ. Consider the Lebesgue decomposition

$$\rho = \rho^{\text{ac}} + \rho^{\text{sc}} + \rho^{\text{pp}} \tag{9}$$

a natural question is which of these parts arise?[21] A partial answer is, that if $V \in L^2_{\text{loc}}(\mathbb{R}^+) \cap L^1(\mathbb{R}^+)$ then $\rho = \rho^{\text{ac}} + \rho^{\text{pp}}$ with $\text{supp}(\rho^{\text{pp}}) \subseteq (-\infty, 0]$ (this follows from the theory of trace class scattering[22]).

So what happens if $V \in L^2(\mathbb{R}^+)$? The Wigner–von Neumann example[23] with

$$V(x) = \frac{\sin(6x)}{x}(1 + o(1)),$$

which has an embedded eigenvalue shows that this situation has to be more complicated.

9.3 The Killip–Simon Theorem

Since in the case $V = 0$, the solution $\psi(z, x) = \mathrm{e}^{-\mathrm{i}\sqrt{z}x}$, it turns out useful instead of working in z coordinate to work in k coordinates, where

$$z = k^2.$$

We define

$$w(k) = m(k^2). \tag{10}$$

We furthermore define $\rho_0(E)$ by

$$d\rho_0(E) = \frac{1}{\pi}\chi_{[0,\infty)}(E)\sqrt{E}dE, \tag{11}$$

[20] Think of this as the analog to matrix diagonalization.

[21] See the paper by Denisov and Kiselev in Simon's birthday Festschrift.

[22] See Reed–Simon III. In fact then the pure point part corresponds to the eigenvalues $\{E_j\}$

[23] This can be found in Reed–Simon IV in Section XIII.13.

which corresponds to the ρ in the case $V = 0$. We introduce a measure ν on $(1, \infty)$ by

$$\frac{2}{\pi} \int f(k^2) k d\nu(k) = \int f(E)(d\rho(E) - d\rho_0(E)), \tag{12}$$

where $E = k^2$. We note that

$$\frac{d\nu}{dk} = \mathrm{Im}(w(k + \mathrm{i}0)) - k \tag{13}$$

for almost every $k \in (1, \infty)$. We furthermore, define a function

$$F(q) = \frac{1}{\sqrt{\pi}} \int_1^\infty \frac{\mathrm{e}^{-(q-p)^2}}{p} d\nu(p). \tag{14}$$

Theorem 4. *A positive measure ρ on $\mathbb{R}$ is the spectral measure associated to $V \in L^2(\mathbb{R}^+)$ if and only if*

1. *(Weyl)* $\mathrm{supp}(\rho) = [0, \infty) \cup \{E_j\}$.

2. *(Local Solubility)*

$$\int_0^\infty |F(q)|^2 dq < \infty. \tag{15}$$

3. *(Lieb–Thirring)*[24]

$$\sum_j |E_j|^{3/2} < \infty. \tag{16}$$

4. *(Strong Quasi–Szegő)*[25]

$$\int \log \left(\frac{|w(k + \mathrm{i}0) + k|^2}{4k \mathrm{Im}(w(k + \mathrm{i}0))} \right) k^2 dk < \infty. \tag{17}$$

The meaning of the conditions (i) to (iv) is as follows.

1. guaranties the right support.

2. implies that $V \in L^2_{\mathrm{loc}}$.

[24]Lieb and Thirring derived inequalities of this form in their proof of the stability of matter.

[25]Szegő showed that a similar condition is equivalent to the Verblunsky coefficients being in ℓ^2 in the case of orthogonal polynomials on the unit circle.

(iii, iv) imply that $V \in L^2$ assuming $V \in L^2_{\text{loc}}$, since under this assumption the sum rule

$$\frac{1}{8}\int_0^\infty V(x)^2 dx = \frac{2}{3}\sum_j (E_j^0)^{3/2} + \frac{1}{\pi}\int \log\left(\frac{|w(k+\mathrm{i}0)+k|^2}{4k\mathrm{Im}(w(k+\mathrm{i}0))}\right) k^2 dk \tag{18}$$

holds, where E_j^0 are the eigenvalues of the operator L_0, which is the extension of H to an operator on $L^2(\mathbb{R})$. From general results, one has that

$$\sum_j (E_j^0)^{3/2} \approx \sum_j |E_j|^{3/2}. \tag{19}$$

In the following, we will try to prove this part.

These conditions are unsatisfactory in order to understand what spectral measures ρ can arise, since the function F is a complicated object. In order to correct this, introduce the short-range part of the Hardy–Littlewood maximal function of ν by

$$M_s\nu(x) = \sup_{0<L\leq 1} \frac{1}{2L}|\nu|([x-L, x+L]). \tag{20}$$

Then one can show that the previous theorem is equivalent to

Theorem 5. *A positive measure ρ on $\mathbb{R}$ is the spectral measure associated to $V \in L^2(\mathbb{R}^+)$ if and only if*

1. *(Weyl)* $\text{supp}(\rho) = [0,\infty) \cup \{E_j\}$.
2. *(Normalization)*[26]
$$\int \log\left(1 + \left(\frac{M_s\nu(k)}{k}\right)^2\right) k^2 dk < \infty. \tag{21}$$
3. *(Lieb–Thirring)*
$$\sum_j |E_j|^{3/2} < \infty. \tag{22}$$

[26] In analogy to the paper of Killip and Simon on the discrete case [3]. There the condition is just $\rho(\mathbb{R}) = 1$!

4. (Quasi–Szegő)

$$\int \log\left(\frac{1}{4}\frac{d\rho}{d\rho_0} + \frac{1}{2} + \frac{1}{4}\frac{d\rho_0}{d\rho}\right)\sqrt{E}dE < \infty. \tag{23}$$

This theorem shows in particular, that if $V \in L^2(\mathbb{R}^+)$, then ρ^{ac} is supported on $\mathbb{R}^+$ (so as the free operator), which is a result by Deift and Killip [1]. Furthermore, one sees that the only obstruction to constructing measures with a singular part is the normalization condition (ii).

9.4 A historical note

Killip and Simon have first proven their theorem in the case of Jacobi operators, which are discrete operators acting $\ell^2(\mathbb{Z}_+) \to \ell^2(\mathbb{Z}_+)$. This is somewhat easier, since one only has to worry about the condition '$V \in L^2$' and not about being locally in L^2. See [3] for details.

References

[1] P. Deift, R. Killip, *On the absolutely continuous spectrum of one-dimensional Schrodinger operators with square summable potentials*, Commun. Math. Phys. **203** (1999) 341-347.

[2] R. Killip, B. Simon, *Sum rules and spectral measures of Schrödinger operators with L^2 potentials*, Ann. of Math. (to appear).

[3] R. Killip, B. Simon, *Sum rules for Jacobi matrices and their applications to spectral theory*, Ann. of Math. **158** (2003), 253-321.

10 The bilinear Hilbert transform along a parabola

after X. Li [3]
A summary written by Victor Lie

Abstract

We summarize the proof in [3] of the $L^2 \times L^2 \to L^1$ - boundedness of the bilinear Hilbert transform along a parabola.

10.1 Introduction

The paper that we intend to present here treats the problem of providing bounds for the bilinear Hilbert transform along a parabola[27] - denoted in what follows with $H_{\mathcal{P}}$.

Let us start by presenting the definition of $H_{\mathcal{P}}$.

$$H_{\mathcal{P}} : S(\mathbb{R}) \times S(\mathbb{R}) \longmapsto S'(\mathbb{R})$$

$$H_{\mathcal{P}}(f,g)(x) := \text{p.v.} \int_{\mathbb{R}} f(x-t)g(x-t^2)\frac{dt}{t} \,.$$

The main result of the paper is given by

Main Theorem.[28] *The bilinear Hilbert transform along the parabola, $H_{\mathcal{P}}$, is a bounded operator from $L^2(\mathbb{R}) \times L^2(\mathbb{R})$ to $L^1(\mathbb{R})$.*

10.2 The analysis of the multiplier

If viewed in a multiplier setting, we have:

$$H_{\mathcal{P}}(f,g)(x) := \text{p.v.} \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{f}(\xi)\hat{g}(\eta) m(\xi,\eta) e^{i\xi x} e^{i\eta x} d\xi d\eta \,.$$

where

$$m(\xi,\eta) = \int_{\mathbb{R}} e^{-i\xi t}\, e^{-i\eta t^2}\, \frac{dt}{t} \,.$$

[27]The parabola here can actually be replaced by any curve γ of the form $\gamma(t) = (t, t^d)$ for $2 \leq d \in \mathbb{N}$.

[28]One can extend this theorem to obtain boundedness in the general local L^2 case.

Our goal in this section is to make a careful analysis of the symbol m; we will use a different decomposition from that appearing in [3] which seems more natural, but the most relevant pieces will be treated following Li.

As usual we start by decomposing the kernel $\frac{1}{t}$ as follows:

$$\frac{1}{t} = \sum_{j\in\mathbb{Z}} \rho_j(t) \qquad \forall\ t \in \mathbb{R}^* ,$$

where ρ is an odd C^∞ function with $\operatorname{supp}\rho \subseteq \left\{t\in\mathbb{R} \mid \frac{1}{4} < |t| < 1\right\}$ and $\rho_j(t) := 2^j \rho(2^j t)$ (with $j \in \mathbb{Z}$).

Consequently,

$$m(\xi,\eta) = \sum_{j\in\mathbb{Z}} m_j(\xi,\eta)$$

with

$$m_j(\xi,\eta) = \int_{\mathbb{R}} e^{-i\xi t}\, e^{-i\eta t^2}\, \rho_j(t) dt .$$

Using Parseval, we notice that our symbol m_j obeys the following key identity:[29]

$$\int_{\mathbb{R}} e^{-i\,\frac{\xi}{2^j}\,t}\, e^{-i\,\frac{\eta}{2^{2j}}\,t^2}\, \rho(t)\, dt = \frac{2^j}{\sqrt{|\eta|}}\, e^{i\frac{\xi^2}{4\eta}} \int_{\mathbb{R}} e^{i\,\frac{2^{2j}}{\eta}\,u^2}\, e^{i\,u\,\frac{2^j\xi}{2\eta}}\, \hat{\rho}(u)\, du . \qquad (1)$$

Relation (1) suggests the further analysis of the symbol m_j relative to the size of the terms $\frac{\xi}{2^j}$ and $\frac{\eta}{2^{2j}}$. For this we are invited to split m_j as follows.

Let ν_0, ν_1, ν_2 be (even) positive smooth functions satisfying: $\nu_0 \in C_0^\infty(\mathbb{R})$ with $\operatorname{supp}\nu_0 \subset (-9/10, 9/10)$, $\nu_1 \in C_0^\infty(\mathbb{R})$ with $\operatorname{supp}\nu_1 \subset \{x \mid \frac{1}{2} < |x| < 2\}$, $\nu_2 \in C^\infty(\mathbb{R})$ with $\operatorname{supp}\nu_2 \subset \{x \mid |x| > 3/2\}$ and such that

$$\nu_0 + \nu_1 + \nu_2 = 1 .$$

Now set $\nu_{j,k}(x) := \nu_k(2^{-j}x)$ with $k \in \{0,1,2\}$.

Then, each component m_j of the multiplier m is expressed as:

$$m_j = \sum_{k,l=0}^{2} m_j^{kl}$$

where

$$m_j^{kl}(\xi,\eta) := m_j(\xi,\eta)\, \nu_{j,k}(\xi)\, \nu_{2j,l}(\eta) .$$

[29] Here we ignore the constants.

This last relation can be written in a more explicit form:

$$m_j^{kl}(\xi,\eta) = \left(\int_{\mathbb{R}} e^{-i\,\frac{\xi}{2^j}\,t}\, e^{-i\,\frac{\eta}{2^{2j}}\,t^2}\,\rho(t)\,dt\right)\,\nu_k(\frac{\xi}{2^j})\,\nu_l(\frac{\eta}{2^{2j}})\,. \tag{2}$$

Now using (1), the mean zero property of the function ρ, and Taylor expansions, one can easily show that all the symbols m_j^{kl} excepting m_j^{22} can be essentially reduced[30] to the study of the symbols having the form

$$u_j(\xi,\eta) := \psi(\frac{\xi}{2^j})\varphi(\frac{\eta}{2^{2j}})\,,$$

where ψ, η are smooth compactly supported functions and at least one of them has mean zero.

Now taking $u = \sum_j u_j$ and defining

$$\mathcal{V}(f,g)(x) := \text{p.v.}\int_{\mathbb{R}}\int_{\mathbb{R}} \hat{f}(\xi)\hat{g}(\eta)u(\xi,\eta)e^{i\xi x}e^{i\eta x}d\xi d\eta\,,$$

one may apply the classical paraproduct theory[31] to conclude:

Theorem 1. *For any* $\frac{1}{p}+\frac{1}{q}=\frac{1}{r}$ *with* $p,q>1$ *and* $r>\frac{1}{2}$ *we have*

$$\|\mathcal{V}(f,g)\|_r \lesssim_{p,q,r} \|f\|_p\,\|g\|_q\ .$$

Thus, Theorem 1 reasonably easily dispenses with our concerns relative to the boundedness properties of the (multilinear) operators given by the symbols

$$m_{kl}(\xi,\eta) := \sum_j m_j^{kl}(\xi,\eta)\,,$$

where as mentioned before $k,l\in\{0,1,2\}$ but k,l are not simultaneously equal to 2.

Now we turn our attention towards the last component in our decomposition of m_j, namely m_j^{22}. Using again the "duality formula" (1) we remark that

$$m_j^{22}(\xi,\eta) = \frac{2^j}{\sqrt{|\eta|}}\,e^{i\frac{\xi^2}{4\eta}}\,\rho(\frac{2^j\xi}{2\eta})\,\left(1-\nu_2(\frac{\xi}{2^j})\right)\,\left(1-\nu_2(\frac{\eta}{2^{2j}})\right)+\text{Error term} \tag{3}$$

[30] As linear combinations of u_j 's with l^1-summable coefficients.

[31] As the reader may note, to establish the Main Theorem one only needs the conclusion of Theorem 1 for $p=q=2$. In the general local L^2 setting, one may need uniform paraproduct estimates; see [4].

where the error term is given by[32]

$$\left(\frac{2^j}{\sqrt{|\eta|}}\right)^3 e^{i\frac{\xi^2}{4\eta}}\,\tilde{\rho}(\frac{2^j\xi}{2\eta})\,\left(1-\nu_2(\frac{\xi}{2^j})\right)\,\left(1-\nu_2(\frac{\eta}{2^{2j}})\right)\,. \tag{4}$$

This error term can be easily treated due to the extra decay offered by the cubic expression in (4).

We let $v_j(\xi,\eta)$ be the main term in (3). Set ϕ a smooth compactly supported function with $\operatorname{supp}\phi \subset \{x \,|\, \frac{1}{10} < |x| < 10\}$. Then, we rewrite[33] the main term v_j as

$$v_j(\xi,\eta) = \sum_{m\geq 0} \frac{1}{2^{m/2}}\, e^{i\frac{\xi^2}{4\eta}}\, \phi(\frac{\xi}{2^{m+j}})\, \phi(\frac{\eta}{2^{m+2j}}) = \sum_{m\geq 0} v_{j,m}(\xi,\eta)\,. \tag{5}$$

Now we define our essential pieces as

$$T_{j,m}(f,g)(x) := \int_{\mathbb{R}}\int_{\mathbb{R}} \hat{f}(\xi)\hat{g}(\eta)v_{j,m}(\xi,\eta)e^{i\xi x}e^{i\eta x}d\xi d\eta\,. \tag{6}$$

It remains now to show (and this constitutes the most difficult part of our result) that the operators $T_{j,m}$ obey the condition

$$\left\|\sum_{j\in\mathbb{Z},m\geq 0} T_{j,m}(f,g)\right\|_1 \lesssim \|f\|_2\,\|g\|_2\ . \tag{7}$$

The methods for proving (7) will be described in the next section.

10.3 The proof - key argument

¿From the above description our main theorem is reduced to the task of obtaining good bounds for each operator $T_{j,m}$. This aim is attained through the following:

Theorem 2. *There exists $\epsilon \in (0,1)$ such that*

$$\|T_{j,m}(f,g)\|_1 \lesssim \max\{2^{-\epsilon m},\, 2^{-\epsilon|m-|j||}\}\,\|f\|_2\,\|g\|_2\ . \tag{8}$$

[32]Here $\tilde{\rho}$ is a smooth function (but not compactly supported) "mostly" concentrated in the interval $[-1,1]$.

[33]Here we use the good localization of ρ, i.e. supp $\rho \subseteq \left\{t\in\mathbb{R}\,|\,\frac{1}{4}<|t|<1\right\}$.

The proof of Theorem 2 is realized in two steps. The first step consists of proving the estimate (8) for the case $|j| < m$. In this situation, one uses the standard TT^*- method; the summarized result is contained in the following

Proposition 3. *For $|j| < m$ we have that*

$$\|T_{j,m}(f,g)\|_1 \lesssim 2^{\frac{|j|-m}{8}} \|f\|_2 \|g\|_2 \, . \tag{9}$$

The second step brings the true "caviar" of the proof. Indeed in closing the estimate for the case $|j| \geq m$, Li had the nice idea of making use of the $\sigma-$uniformity concept described below. While this uniformity concept is not original, appearing in various forms in fields such as ergodic theory and additive combinatorics (see *e.g.* [2]), its appearance in our context is rather surprising. (For the sake of truth though, one must say that this concept was previously used in a related setting in the paper [1].)

Let us now describe this concept as it is used in our context.

Set $\sigma \in (0, 1]$ and let $\mathcal{Q}$ be a family of real-valued measurable functions. Also set I a bounded interval.

Definition. *A function $f \in L^2(I)$ is $\sigma-$uniform in $\mathcal{Q}$ if*

$$\left| \int_I f(\xi)\, e^{-i\, q(\xi)}\, d\xi \right| \leq \sigma \|f\|_{L^2(I)}$$

for all $q \in \mathcal{Q}$.

Lemma 4. *Let L be a bounded sub-linear functional from $L^2(I)$ to $\mathbb{C}$, and let S_σ be the collection of all functions that are $\sigma-$uniform in $\mathcal{Q}$. Set*

$$U_\sigma = \sup_{f \in S_\sigma} \frac{|L(f)|}{\|f\|_{L^2(I)}} \quad \& \quad Q = \sup_{q \in \mathcal{Q}} |L(e^{i\,q})| \, .$$

Then for all functions in $L^2(I)$ we have

$$|L(f)| \leq \max\{U_\sigma,\, 2\sigma^{-1} Q\} \|f\|_{L^2(I)} \, . \tag{10}$$

The moral of the $\sigma-$uniformity concept introduced above and the way in which it is used in Lemma 4 can be described as follows: given a function f, then either the "q-Fourier" coefficients are uniformly small (so morally f is orthogonal to the family $e^{i\,\mathcal{Q}} := \{e^{iq}\}_{q \in \mathcal{Q}}$) or f must resemble a singleton of the form $e^{i\,q}$ for some $q \in \mathcal{Q}$. Thus to control the size of L it is enough to know the behavior of L on the classes: S_σ and $e^{i\,\mathcal{Q}}$.

Once we have defined this concept, using the lemma above one proves the second step needed for our theorem:

Proposition 5. *Let $|j| \geq m$. Then we have that*

$$\|T_{j,m}(f,g)\|_1 \lesssim \max\{2^{-\frac{m}{10}},\, 2^{\frac{m-|j|}{2}}\}\, \|f\|_2\, \|g\|_2\ . \tag{11}$$

As expected, this last proposition borrows from the structure described in Lemma 4: firstly, one needs to identify the form of the family $\mathcal{Q}$ and to obtain good estimates for the L^1-norm of $T_{j,m}(f,g)$ when one of the functions $\hat{f}$ and $\hat{g}$ is σ-uniform with respect to $\mathcal{Q}$; secondly, one must control the expressions like $T_{j,m}(\check{e^{iq}}, g)$ (or $T_{j,m}(f, \check{e^{iq}})$) when $q \in \mathcal{Q}$.

Finally, we give a short description of the steps followed for proving Proposition 5.

First we notice that it is enough to prove our result in the case $j \geq m$, (the other case $-j \geq m$ has a similar treatment). Next, one makes use of the scaling symmetry and "moves" the problem inside of the unit interval; more exactly, one defines the operator

$$B_{j,m}(f(\cdot), g(\cdot))(x) := 2^{-\frac{j}{2}}\, T_{j,m}\left(f(2^{m+j}\cdot), g(2^{m+2j}\cdot)\right)\left(\frac{x}{2^{2j+m}}\right),$$

and notices that relation (11) is equivalent to

$$\|B_{j,m}(f,g)\|_1 \lesssim \max\{2^{-\frac{m}{10}},\, 2^{\frac{m-|j|}{2}}\}\, \|f\|_2\, \|g\|_2\ . \tag{12}$$

The advantage in working with $B_{j,m}$ instead of $T_{j,m}$ is that now all the information of the multiplier is translated inside of the unit[34] cube, as can easily be observed from the following formula:

$$B_{j,m}(f,g) = 2^{-\frac{m+j}{2}} \int_{\mathbb{R}}\int_{\mathbb{R}} \hat{f}(\xi)\, \hat{g}(\eta)\, e^{i(2^{-j}\xi+\eta)\,x}\, e^{i\,\frac{\xi^2}{4\eta}2^m}\, \phi(\xi)\, \phi(\eta)\, d\xi\, d\eta\,. \tag{13}$$

Next one defines the trilinear form associated to $B_{j,m}$ and given by

$$\Lambda_{j,m}(f,g,h) := \int_{\mathbb{R}} B_{j,m}(f,g)(x)\, h(x)\, dx\,.$$

Now, given the form of (13) and our plans of attacking the proof along the lines of Lemma 4, we define the family $\mathcal{Q}$ as being

$$\mathcal{Q} := \{a\xi^2 + b\xi \,|\, 2^{m-100} \leq |\,a| \leq 2^{m+100}\ \&\ b \in \mathbb{R}\}\,. \tag{14}$$

With this done, we are now ready to state the two lemmas on which the proof of (12) is based.

[34] This is verbatim true if we think at the support of ϕ as being placed inside $\{x \,|\, \frac{1}{2} < |x| < 1\}$ instead of $\{x \,|\, \frac{1}{10} < |x| < 10\}$.

Lemma 6. *Let $\hat{f}\phi \in L^2([0,1])$ be $\sigma-$uniform in $\mathcal{Q}$. Then*

$$|\,\Lambda_{j,m}(f,g,h)| \lesssim \max\{2^{-m}, 2^{m-j}, \sigma\}\|f\|_2\|g\|_2\|h\|_\infty\,. \tag{15}$$

Lemma 7. *Let $q \in \mathcal{Q}$. Then*

$$|\,\Lambda_{j,m}((\check{e^{iq}\phi}), g, h)| \lesssim 2^{-\frac{m}{5}}\|g\|_2\|h\|_\infty\,. \tag{16}$$

The proof of Lemma 6 involves the mean value theorem, interpolation techniques and the use of the $\sigma-$uniformity. Lemma 7 uses a variant of van der Corput exploiting the highly oscillatory behavior of the multiplier in $\Lambda_{j,m}((\check{e^{iq}\phi}), g, h)$.

This being said, we apply Lemma 4 to the function $\hat{f}\phi$ and the functional $L(f) = (L_{g,h}(f)) = \Lambda_{j,m}(f,g,h)$, and exploit the statements of Lemma 6 and 7. Then we have

$$|\,\Lambda_{j,m}(f,g,h)| \lesssim \max\{\max\{2^{-m}, 2^{m-j}, \sigma\},\, 2\sigma^{-1}2^{-\frac{m}{5}}\}\|f\|_2\|g\|_2\|h\|_\infty\,.$$

Now, by properly choosing σ we obtain (12), thus ending the proof of Proposition 5.

References

[1] M. Christ, X. Li, T. Tao, and C. M. Thiele. On multilinear oscillatory integrals, nonsingular and singular. *Duke Math. J.* **130** (2005), 321–351.

[2] W. T. Gowers. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.* **8** (1998), 529–551.

[3] X. Li. Bilinear Hilbert transforms along curves I. The monomial case. Preprint, arXiv:0805.0107v2 [math.CA] (2008).

[4] X. Li. Uniform estimates for some paraproducts. To appear, *N. Y. Math. J.*

Victor Lie, UCLA
email: **vilie@ucla.edu**

11 The Brascamp-Lieb inequalities: finiteness, structure and extremals

after J. Bennett, A. Carbery, M. Christ, and T. Tao [1]
A summary written by Kabe Moen

Abstract

Fundamental inequalities in analysis such as Hölder's, Young's, and the Loomis-Whitney inequalities naturally fall into the more general framework of the Brascamp-Lieb inequalities. We summarize some of the results in [1]. Heat flow methods adapted to the multilinear setting are used to give a new proof of the geometric Brascamp-Lieb inequality. We also address issues such as the finiteness of the constant, extremals, and a gaussian extremals.

11.1 Introduction

Let $m \geq 1$, $(B_1, \ldots, B_m)$ be an m-tuple of surjective linear transformations with each $B_j : \mathbb{R}^n \to \mathbb{R}^{n_j}$, and $(p_1, \ldots, p_m)$ be an m-tuple of exponents, with each $p_j \in [0, \infty)$. The Brascamp-Lieb inequalities arise when considering conditions for which

$$\sup_{f_1,\ldots,f_m} \frac{\int_{\mathbb{R}^n} \prod_{j=1}^m (f_j \circ B_j)^{p_j} \, dx}{\prod_{j=1}^m \left(\int_{\mathbb{R}^{n_j}} f_j \, dx\right)^{p_j}} < \infty, \tag{1}$$

where the supremum is over all nonnegative measurable functions $f_j : \mathbb{R}^{n_j} \to [0, +\infty)$ with $0 < \int_{\mathbb{R}^{n_j}} f_j \, dx < \infty$.

Example 1 (Hölder's Inequality). *If $n = n_j$, $B_j = \mathrm{Id}$, $p_1 + \cdots p_m = 1$, then* (1) *is essentially a restatement of Hölder's inequality.*

Example 2 (Young's convolution inequality). *If $n = 2d$, $B_j : \mathbb{R}^{2d} \to \mathbb{R}^d$, for $j = 1, 2, 3$ are given by $B_1(x, y) = y, B_2(x, y) = x - y, B_3(x, y) = x$, and $p_1 + p_2 + p_3 = 2$, then* (1) *is a reformulation Young's convolution inequality.*

Example 3 (Loomis-Whitney inequality). *Suppose $n = m$, $n_j = n - 1$ for each $1 \leq j \leq n$, B_j are the orthogonal projections $B_j(x_1, \ldots, x_n) = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$ and $p_j = 1/(n-1)$. In this case* (1) *can be interpreted as the Loomis-Whitney inequality.*

11.2 Notation and preliminaries

11.2.1 Definitions

In order to take advantage of invariance of isometries, restrictions, and orthogonal subspaces of Hilbert space, we work on finite dimensional real Hilbert spaces, usually denoted by H, H_j, etc. For an integer $m \geq 1$, we define an m-transformation to be

$$\mathbf{B} = (H, (H_j)_{1 \leq j \leq m}, (B_j)_{1 \leq j \leq m})$$

where for each $1 \leq j \leq m$, $B_j : H \to H_j$. We say that an m-transformation is non-degenerate if $B_j H = H_j$ and $\bigcap_{j=1}^m \ker B_j = \{0\}$. An m-exponent, denoted $\mathbf{p}$, is an m-tuple of exponents $p_1, \ldots, p_m \in [0, \infty]$. We define a Brascamp-Lieb datum to be a pair $(\mathbf{B}, \mathbf{p})$ where $\mathbf{B}$ is an m-transformation and $\mathbf{p}$ is an m-exponent. If $(\mathbf{B}, \mathbf{p})$ is a Brascamp-Lieb datum, an input, denoted $\mathbf{f}$, for $(\mathbf{B}, \mathbf{p})$ is m-tuple of functions $f_j : H_j \to [0, \infty)$ such that $0 < \int_{H_j} f_j \, dx < \infty$.

Definition 4 (Brascamp-Lieb constant)**.** *Define the quantity* $0 \leq BL(\mathbf{B}, \mathbf{p}, \mathbf{f}) \leq \infty$*, by the formula*

$$\mathrm{BL}(\mathbf{B}, \mathbf{p}, \mathbf{f}) := \frac{\int_H \prod_{j=1}^m (f_j \circ B_j)^{p_j} \, dx}{\prod_{j=1}^m \left(\int_{H_j} f_j \, dx \right)^{p_j}} \tag{2}$$

and define the Brascamp-Lieb constant

$$\mathrm{BL}(\mathbf{B}, \mathbf{p}) := \sup \mathrm{BL}(\mathbf{B}, \mathbf{p}, \mathbf{f}) \tag{3}$$

where the supremum is over all inputs $\mathbf{f}$.

Of course one can define the Brascamp-Lieb constant by (3) when $(\mathbf{B}, \mathbf{p})$ is degenerate, however in this case $\mathrm{BL}(\mathbf{B}, \mathbf{p}) = \infty$. Thus, we shall restrict to non-degenerate data. One of the fundamental inputs are gaussian inputs. More specifically, given any positive definite linear transformation $A : H \to H$ one has the well known formula

$$\int_H \exp(-\pi \langle Ax, x \rangle) \, dx = (\det A)^{-1/2}. \tag{4}$$

Given a Brascamp-Lieb datum $(\mathbf{B}, \mathbf{p})$ we define a gaussian input to be an m-tuple $\mathbf{A} = (A_j)_{1 \leq j \leq m}$ of positive definite linear transformations $A_j : H_j \to$

H_j. Plugging the datum $(\exp(-\pi\langle A_j x, x\rangle))_{1\le j\le m}$ into (2) and using formula (4) we define $0 < \mathrm{BL}_{\mathbf{g}}(\mathbf{B},\mathbf{p},\mathbf{A}) < \infty$ by

$$\begin{aligned}\mathrm{BL}_{\mathbf{g}}(\mathbf{B},\mathbf{p},\mathbf{A}) \quad &:= \quad \mathrm{BL}(\mathbf{B},\mathbf{p},(\exp(-\pi\langle A_j x, x\rangle))_{1\le j\le m}) \\ &= \quad \left(\frac{\prod_{j=1}^{m}(\det_{H_j} A_j)^{p_j}}{\det_H\left(\sum_{j=1}^{m} p_j B_j^* A_j B_j\right)}\right)^{1/2} \end{aligned} \tag{5}$$

Definition 5 (Gaussian Brascamp-Lieb constant)**.** *Let* $(\mathbf{B},\mathbf{p})$ *be a Brascamp-Lieb datum, then we define the gaussian Brascamp-Lieb constant to be*

$$\mathrm{BL}_{\mathbf{g}}(\mathbf{B},\mathbf{p}) := \sup \mathrm{BL}_{\mathbf{g}}(\mathbf{B},\mathbf{p},\mathbf{A})$$

where the supremum is over all gaussian inputs $\mathbf{A}$*.*

Clearly, we have

$$\mathrm{BL}_{\mathbf{g}}(\mathbf{B},\mathbf{p}) \le \mathrm{BL}(\mathbf{B},\mathbf{p}).$$

However, as Lieb [3] showed we actually have equality (see Theorem 11 below). For the Brascamp-Lieb constants one issue is determining when they are finite, another issue is determining when there are extremisers.

Definition 6 (Extremisability)**.** *A Brascamp-Lieb datum is said to extremisable if* $\mathrm{BL}(\mathbf{B},\mathbf{p})$ *is finite and there exists an input* $\mathbf{f}$ *for which* $\mathrm{BL}(\mathbf{B},\mathbf{p}) = \mathrm{BL}(\mathbf{B},\mathbf{p},\mathbf{f})$*. The datum* $(\mathbf{B},\mathbf{p})$ *is said to be gaussian-extremisable if there exists a gaussian input* $\mathbf{A}$ *for which* $\mathrm{BL}_{\mathbf{g}}(\mathbf{B},\mathbf{p}) = \mathrm{BL}_{\mathbf{g}}(\mathbf{B},\mathbf{p},\mathbf{A})$*.*

Finally, we introduce a special Brascamp-Lieb datum which plays an important role in determining general Brascamp-Lieb constants.

Definition 7 (Geometric Brascamp-Lieb data)**.** *A Brascamp-Lieb datum* $(\mathbf{B},\mathbf{p})$ *is said to be geometric if* $H_1,\ldots,H_m$ *are subspaces of* H*,* $B_j : H \to H_j$ *are orthogonal projections, and*

$$\sum_{j=1}^{m} p_j B_j^* B_j = \mathrm{Id}_H. \tag{6}$$

First, notice that geometric Brascamp-Lieb datum are always non-degenrate. Also notice that the geometric Brascamp-Lieb inequality generalizes Examples 1 and 3 significantly.

11.2.2 Main results

In this section we provide the main results that address some of the issues concerning Brascamp-Lieb constants. Specifically, given a Brascamp-Lieb datum, $(\mathbf{B}, \mathbf{p})$, when is its Brascamp-Lieb constant finite? When does it have extremisers? When does it have gaussian extremisers? We first start with the geometric case of the Brascamp-Lieb constants. This case lays the foundation for heat flow techniques and leads to more general Brascamp-Lieb inequalities.

Theorem 8. *Let $(\mathbf{B}, \mathbf{p})$ be a geometric Brascamp-Lieb datum. Then*

$$\mathrm{BL}(\mathbf{B}, \mathbf{p}) = \mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) = 1$$

and $(\mathbf{B}, \mathbf{p})$ is gaussian-extremisable (hence also extremisable).

Once we have established Theorem 8 we may use it to characterize gaussian extremisers.

Definition 9. *Two m-transformations $\mathbf{B} = (H, (H_j)_{1\leq j\leq m}, (B_j)_{1\leq j\leq m})$ and $\mathbf{B}' = (H', (H_j')_{1\leq j\leq m}, (B_j')_{1\leq j\leq m})$ are said to be equivalent if there exist invertible linear transformations $C : H' \to H$, and $C_j : H_j' \to H_j$ such that $B_j' = C_j^{-1} B_j C$ for all j; we refer to C and C_j as intertwining transformations.*

We call two Brascamp-Lieb data $(\mathbf{B}, \mathbf{p})$ and $(\mathbf{B}', \mathbf{p}')$ equivalent if $\mathbf{B}$ and $\mathbf{B}'$ are equivalent and $\mathbf{p} = \mathbf{p}'$. Using a change of variables we have the following relationship between Brascamp-Lieb constant for equivalent data

$$\mathrm{BL}(\mathbf{B}', \mathbf{p}') = \frac{\prod_{j=1}^{m} |\det C_j|^{p_j}}{|\det C|} \mathrm{BL}(\mathbf{B}, \mathbf{p})$$

and similar with gaussian Brascamp-Lieb constants.

Theorem 10. *Let $(\mathbf{B}, \mathbf{p})$ be a Brascamp-Lieb datum and $\mathbf{A}$ a gaussian input for $(\mathbf{B}, \mathbf{p})$. Let $M : H \to H$ be the positive semi-definite transformation $M := \sum_{j=1}^{m} p_j B_j^* A_j B_j$. Then $\mathbf{A}$ is a gaussian extremiser for $(\mathbf{B}, \mathbf{p})$ if and only if $(\mathbf{B}, \mathbf{p})$ is equivalent to a geometric Brascamp-Lieb data $(\mathbf{B}', \mathbf{p}')$ with intertwining maps $C = M^{-1/2}$ and $C_j = A_j^{-1/2}$, and*

$$\mathrm{BL}(\mathbf{B}, \mathbf{p}) = \mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p})$$

In determining if a Brascamp-Lieb constant is finite one of the tools is the following theorem by Lieb [3].

Theorem 11. *Let* $(\mathbf{B}, \mathbf{p})$ *be a Brascamp-Lieb datum, then* $\mathrm{BL}(\mathbf{B}, \mathbf{p}) = \mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p})$*.*

Thus, in order to understand a Brascamp-Lieb constant, we may reduce the task to understanding the gaussian Brascamp-Lieb constant. However, Theorem 11 does not resolve everything. For instance, it does clarify when conditions exist for either $\mathrm{BL}(\mathbf{B}, \mathbf{p})$ or $\mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p})$ to be finite. Also, this theorem does not answer the question of when extremal exist. In order to address these question we first need a definition which was first introduced in [2].

Definition 12 (Critical subspace and simplicity)**.** *Let* $(\mathbf{B}, \mathbf{p})$ *be a Brascamp-Lieb datum. A critical subspace* V *for* $(\mathbf{B}, \mathbf{p})$ *is a non-zero proper subspace of* H *such that*

$$\dim(V) = \sum_{j=1}^{m} p_j \dim(B_j V).$$

The datum $(\mathbf{B}, \mathbf{p})$ *is simple if it has no critical subspaces.*

Theorem 13. *Let* $(\mathbf{B}, \mathbf{p})$ *be a Brascamp-Lieb datum. Then* $\mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p})$ *is finite if and only if we have the scaling condition*

$$\dim(H) = \sum_{j=1}^{m} p_j \dim(H_j) \tag{7}$$

and the dimension condition

$$\dim(V) \leq \sum_{j=1}^{m} p_j \dim(B_j V) \qquad \text{for all subspaces } V \subseteq H. \tag{8}$$

Furthermore, if $(\mathbf{B}, \mathbf{p})$ *is simple, then it is gaussian-extremisable.*

By combining Theorems 11 and 13 we have the following corollary.

Corollary 14. *Let* $(\mathbf{B}, \mathbf{p})$ *be a Brascamp-Lieb datum. Then the following three statements are equivalent:*

(1) $\mathrm{BL}(\mathbf{B}, \mathbf{p}) < \infty$.

(2) $\mathrm{BL}_{\mathbf{g}}(\mathbf{B},\mathbf{p}) < \infty$.

(3) (7) *and* (8) *hold.*

Furthermore, if any of (1)-(3) hold, and $(\mathbf{B},\mathbf{p})$ *is simple, then* $(\mathbf{B},\mathbf{p})$ *is gaussian-extremisable.*

11.3 Outline of proofs

We briefly outline the proofs for some of the results in the previous section. We start with a multilinear monotonicity lemma that is crucial to the proof of Theorem 8. In turn, Theorem 8 leads to the characterization of gaussian-extremisers given in Theorem 10. Finally, we combine Theorem 10, Theorem 13 and a factorisation method (Lemma 19) to provide a short proof of Lieb's Theorem (Theorem 11). We mainly focus on the multilinear heat flow techniques. It is for this reason that we do not provide proofs of many of the results including Lemmas 15, 19 and Theorems 10, 13. First we give the linear version from which we will obtain the multilinear version.

Lemma 15. *Let* $I \subset \mathbb{R}^+$ *be a time interval,* H *be a Euclidean space,* $u : I \times H \to \mathbb{R}^+$ *be smooth function, and* $\vec{v} : I \times H \to H$ *be a smooth vector field, such that* $u\vec{v}$ *is rapidly decreasing at spatial infinity, locally uniform on* I*. Suppose that we have the transport inequality*

$$\partial_t u(t,x) + \mathrm{div}\,(\vec{v}(t,x)u(t,x)) \geq 0 \tag{9}$$

for all $(t,x) \in I \times H$*. Then the quantity* $Q(t) := \int_H u(t,x)\ dx$ *is non-decreasing in time.*

Lemma 16. *Let* $I \subseteq \mathbb{R}^+$ *be a time interval,* $p_1, \ldots, p_m$ *be positive exponents,* $u_j : \mathbb{R}^+ \times H \to \mathbb{R}^+$ *be a smooth functions, and* $\vec{v}_j : \mathbb{R}^+ \times H \to$ *be smooth vector fields. Suppose* $\vec{v} : \mathbb{R}^+ \times H \to H$ *is a smooth vector field such that* $\vec{v}\prod_{j=1}^m u_j^{p_j}$ *is rapidly decreasing space locally uniform on* I*, and the following inequalities hold:*

$$\begin{aligned}
\partial_t u_j(t,x) + \mathrm{div}(\vec{v}_j(t,x)u_j(t,x)) &\geq 0 \qquad 1 \leq j \leq m && (10)\\
\mathrm{div}\left(\vec{v} - \sum_{j=1}^m p_j v_j\right) &\geq 0 && (11)\\
\sum_{j=1}^m p_j \langle \vec{v} - \vec{v}_j, \nabla \log u_j \rangle &\geq 0. && (12)
\end{aligned}$$

Then the quantity $Q(t) := \int_H \prod_{j=1}^m u_j(t,x)^{p_j}\ dx$ *is non-decreasing in time.*

Remark 17. *In practice* u_j *will be solutions of the heat equation (where the name heat flow comes from)*

$$\partial_t u_j(t,x) = \Delta u_j(t,x). \tag{13}$$

Notice, in this case we may rewrite the heat equation (13) *as the transport equation*

$$\partial_t u_j + \operatorname{div}(\vec{v}_j u_j) = 0 \tag{14}$$

where $\vec{v}_j = -\nabla \log u_j$*. Equation* (14) *shows that condition* (10) *is satisfied, and setting*

$$\vec{v} = \sum_{j=1}^m p_j \vec{v}_j$$

condition (11) *is also satisfied.*

We are now ready to prove Theorem 8 using the monotonicity Lemma (16).

Proof of Theorem 8. Suppose $(\mathbf{B}, \mathbf{p})$ is a geometric Brascamp-Lieb datum. Observe that $\mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) \geq 1$ by testing the gaussian input $(\mathrm{Id}_{H_j})_{1\leq j\leq m}$. We aim to show $\mathrm{BL}(\mathbf{B}, \mathbf{p}) \leq 1$, this will imply $\mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) = \mathrm{BL}(\mathbf{B}, \mathbf{p}) = 1$ and $(\exp(-\pi\|x\|^2_{H_j}))_{1\leq j\leq m}$ is an extremiser. Assume $\mathbf{f}$ is an input with each f_j positive, smooth and rapidly decreasing. Let $u_j : \mathbb{R}^+ \times H \to \mathbb{R}^+$ be the solution to the heat equation (13) with initial condition $u_j(0,x) = f_j \circ B_j(x)$. Because of the orthogonality of B_j, we may write the explicit solution of (13),

$$u_j(t,x) = \frac{1}{(4\pi t)^{\dim(H_j)/2}} \int_{H_j} e^{-\|B_j x - y\|^2_{H_j}/4t} f_j(y)\ dy.$$

As noted in Remark 17, setting $\vec{v}_j = -\nabla \log u_j$ and $\vec{v} = \sum_{j=1}^m p_j \vec{v}_j$ implies conditions (10) and (11) of Lemma 16 are satisfied. It is a technical condition to verify $\vec{v} \prod_j u_j^{p_j}$ is rapidly decreasing and that condition (12) is satisfied. Thus, $Q(t) = \int_H \prod_{j=1}^m u_j(t,x)^{p_j}\ dx$ is non-decreasing, so in particular $\limsup_{t\to 0^+} Q(t) \leq \liminf_{t\to\infty} Q(t)$. By Fatou's lemma we have,

$$\int_H \prod_{j=1}^m (f_j \circ B_j)^{p_j}\ dx \leq \limsup_{t\to 0^+} Q(t).$$

On the other hand by a change of variables $x = t^{1/2}w$ we have

$$\liminf_{t\to\infty} Q(t) \quad = \quad \prod_{j=1}^{m} \left(\int_{H_j} f_j \ dy \right)^{p_j} .$$

□

We now give a short proof of Theorem 11. First we give a definition and a lemma.

Definition 18. *Let* $\mathbf{B} = (H, (H_j)_{1\leq j\leq m}, (B_j)_{1\leq j\leq m})$ *be an* m*-transformation and* V *a subspace of* H*. We define the restriction* $\mathbf{B}_V$ *of* $\mathbf{B}$ *to* V *to be*

$$\mathbf{B}_V = (H, (B_j V)_{1\leq j\leq m}, (B_j|_V)_{1\leq j\leq m}).$$

Lemma 19. *Suppose* $(\mathbf{B}, \mathbf{p})$ *has a critical subspace* V*, then*

$$\mathrm{BL}(\mathbf{B}, \mathbf{p}) = \mathrm{BL}(\mathbf{B}_V, \mathbf{p})\mathrm{BL}(\mathbf{B}_{V^\perp}, \mathbf{p}) \tag{15}$$

and

$$\mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) = \mathrm{BL}_{\mathbf{g}}(\mathbf{B}_V, \mathbf{p})\mathrm{BL}_{\mathbf{g}}(\mathbf{B}_{V^\perp}, \mathbf{p}). \tag{16}$$

Proof of Theorem 11. We induct on the dimension of H. When $\dim(H) = 0$ there is nothing to prove. Suppose $n > 0$ and $\mathrm{BL}_{\mathbf{g}}(\tilde{\mathbf{B}}, \tilde{\mathbf{p}}) = \mathrm{BL}(\tilde{\mathbf{B}}, \tilde{\mathbf{p}})$ for any Brascamp-Lieb datum, $(\tilde{\mathbf{B}}, \tilde{\mathbf{p}})$ with $\dim(\tilde{H}) < n$. Let $(\mathbf{B}, \mathbf{p})$ be a Brascamp-Lieb datum with $\dim(H) = n$. We claim that $\mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) = \mathrm{BL}(\mathbf{B}, \mathbf{p})$. Since $\mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) = \infty$ implies $\mathrm{BL}(\mathbf{B}, \mathbf{p}) = \infty$ we may suppose $\mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) < \infty$. First suppose $(\mathbf{B}, \mathbf{p})$ is simple, then since $\mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) < \infty$ by Theorem 13 the datum is gaussian-extremisable. The claim now follows from Theorem 10. Now suppose $(\mathbf{B}, \mathbf{p})$ is not simple, i.e. there is a critical subspace V. We may use Lemma 19 to split the datum $(\mathbf{B}, \mathbf{p})$ into $(\mathbf{B}_V, \mathbf{p})$ and $(\mathbf{B}_{V^\perp}, \mathbf{p})$ with $\mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p})$ and $\mathrm{BL}(\mathbf{B}, \mathbf{p})$ factoring accordingly. By the inductive hypothesis we have

$$\mathrm{BL}(\mathbf{B}_V, \mathbf{p}) = \mathrm{BL}_{\mathbf{g}}(\mathbf{B}_V, \mathbf{p}) \quad \text{and} \quad \mathrm{BL}(\mathbf{B}_{V^\perp}, \mathbf{p}) = \mathrm{BL}_{\mathbf{g}}(\mathbf{B}_{V^\perp}, \mathbf{p})$$

and the claim follows from (15) and (16). Thus closing the induction and completing the proof. □

References

[1] J. Bennett, A. Carbery, M. Christ, T. Tao, *The Brascamp-Lieb inequalities: finiteness, structure and extremals*, GAFA **17** (2008), 1343–1415.

[2] E. Carlen, C. Lieb, M. Loss, *A sharp analog of Young's inequality on* S^N *and related entropy inequalities*, Jour. Geom. Anal. **14** (2004), 487–520.

[3] E. Lieb, *Gaussian kernels have only Gaussian maximizers*, Invent. Math. **102** (1990), 179–208.

Kabe Moen, Washinton University in St. Louis
email: **kabemoen@gmail.edu**

12 A variational norm Carleson theorem

After R. Oberlin, A. Seeger, T. Tao, C. Thiele, and J. Wright [9]
A summary written by Richard Oberlin

Abstract

We give a variational-norm version of Carleson's theorem, and a related estimate for curves in Lie groups which is pertinent to nonlinear Schrödinger and Dirac operators.

12.1 Introduction

Given a Schwartz function f on $\mathbb{R}$ and real numbers ξ and x, we consider the partial Fourier-summation operator $\mathcal{S}[f](\xi, x) = \int_{-\infty}^{\xi} \hat{f}(\xi')e^{2\pi i \xi' x}\, d\xi'$ where $\hat{f}$ denotes the Fourier transform of f. The Carleson-Hunt theorem [2], [6] (after restatement) tells us that for $1 < p < \infty$

$$\|\mathcal{S}[f]\|_{L^p_x(L^\infty_\xi)} \leq C_p \|f\|_{L^p}. \tag{1}$$

Once this bound is known, a standard density argument allows one to define $\mathcal{S}[f]$ for $f \in L^p$ as a continuous function in ξ for almost every x; thus defined the bound extends to L^p.

We are interested in replacing the L^∞ norm above by a stronger (at least in the current context) variational norm. Given a function g on $\mathbb{R}$ and $1 \leq r < \infty$ let

$$\|g\|_{V^r} = \sup_{N, \xi_0 < \ldots < \xi_N} \left(\sum_{j=1}^{N} |g(\xi_j) - g(\xi_{j-1})|^r \right)^{1/r}.$$

The main result is:

Theorem 1. *Suppose* $2 < r < \infty$ *and* $r' < p < \infty$. *Then*

$$\|\mathcal{S}[f]\|_{L^p_x(V^r_\xi)} \leq C_{p,r} \|f\|_{L^p}. \tag{2}$$

A restricted-weak-type estimate at $p = r'$ is also obtained.

The exponents above are sharp in the following sense. By using the fact that a related variational estimate for dyadic martingales does not hold for $r \leq 2$ [10], together with a square function argument (see Section 12.2), one sees that (2) does not hold for $r \leq 2$. The condition $p > r'$ is also necessary; this can be seen by considering a Schwartz function ψ with $\hat{\psi} = 1$ on $[-1, 1]$ and $\hat{\psi}$ supported on $[-2, 2]$. For this example, one may explicitly compute $\mathcal{S}[\psi](\xi, x) - \mathcal{S}[\psi](-\xi, x)$ for $|\xi| \leq 1$ and analyze the variation.

Before discussing the proof, we give two motivations for considering the V^r norm. First, recall that L^∞ bounds such as (1) imply pointwise almost everywhere convergence once such convergence is known for a dense subclass of functions (for example, the continuity in ξ of $\mathcal{S}[f]$ for $f \in L^p$ which follows from the same fact for Schwartz f). The stronger V^r bounds allow one to deduce pointwise almost everywhere convergence without any previously known convergence. This could be useful, say, when transferring estimates into the setting of ergodic theory, where one may not find a dense subclass with trivial convergence; see [1]. Second, as we will see, one may deduce bounds for a certain r-variational length of a curve in a Lie group from the corresponding length of its "trace". It had been hoped that this would allow the deduction of a (variational) bound for the "nonlinear Carleson operator" from a variational bound for the standard (i.e. linear) Carleson operator. Unfortunately, this approach does not seem to work, since the correspondence holds only for $r < 2$ and the bound for the standard Carleson operator only holds for $r > 2$.

12.2 Lacunary version

We first recall a proof of a lacunary version of the Carleson-Hunt theorem

$$\| \sup_{k \in \mathbb{Z}} |\mathcal{S}[f](2^k, \cdot)| \|_{L^p} \leq C_p \|f\|_{L^p}. \tag{3}$$

Using the boundedness of the Hilbert transform, one sees that the above bound follows from the even version

$$\| \sup_{k \in \mathbb{Z}} |\mathcal{S}[f](2^k, \cdot) - \mathcal{S}[f](-2^k, \cdot)| \|_{L^p} \leq C_p \|f\|_{L^p}. \tag{4}$$

Letting ψ be as in the previous subsection and $\psi_k(x) = 2^{-k}\psi(2^{-k}x)$, recall the standard square function estimate

$$\| \left(\sum_{k \in \mathbb{Z}} |\mathcal{S}[f](2^k, \cdot) - \mathcal{S}[f](-2^k, \cdot) - \psi_{-k} * f|^2 \right)^{1/2} \|_{L^p} \leq C_p \|f\|_{L^p} \tag{5}$$

where $*$ denotes convolution. Then (4) follows from (5) and the bound

$$\| \sup_{k\in\mathbb{Z}} |\psi_k * f| \|_{L^p} \le C_p \|f\|_{L^p},$$

which in turn follows from the standard L^p bound for the Hardy-Littlewood maximal operator $\mathcal{M}$.

Moving to the variational version of (3), we want to bound

$$\| \sup_{N\in\mathbb{N}, k_0<\ldots<k_N\in\mathbb{Z}} \left(\sum_{j=1}^{N} |\mathcal{S}[f](2^{k_j},\cdot) - \mathcal{S}[f](2^{k_{j-1}},\cdot)|^r \right)^{1/r} \|_{L^p} \le C_p \|f\|_{L^p}. \quad (6)$$

Using the argument in the previous paragraph (here it is important that $r \ge 2$) one sees that the crucial step is estimating

$$\| \sup_{N\in\mathbb{N}, k_0<\ldots<k_N\in\mathbb{Z}} \left(\sum_{j=1}^{N} |\psi_{k_j} * f - \psi_{k_{j-1}} * f|^r \right)^{1/r} \|_{L^p} \le C_p \|f\|_{L^p}. \quad (7)$$

The bound above then follows from a known L^p bound for a variational version of $\mathcal{M}$ [7] (or alternatively from a variational estimate for dyadic martingales). In fact, for the lacunary bound (6), we end up with the range of exponents $1 < p < \infty$ instead of $p > r'$ as in the general bound (2).

12.3 General version

12.3.1 Discretization

After linearizing and dualizing the V^r norm, and using a partition of unity argument to discretize (as in, say, [11]) we reduce the problem (2) to that of proving bounds for a finite number of model operators each of the form

$$\| \sum_{P\in\mathbf{P}} \langle f, \phi_P \rangle \phi_P a_P \|_{L^p} \le C_{p,r} \|f\|_{L^p}. \quad (8)$$

The sum above is over *multitiles* P in a finite collection $\mathbf{P}$. A multitile is a subset $I \times \omega$ of $\mathbb{R}^2$. The time interval I is dyadic, and the frequency set ω is the union of three intervals $\omega_l, \omega_u, \omega_h$. We describe the nature of the frequency intervals for a typical model operator: the intervals ω_l, ω_u are dyadic half open intervals of the form $[a, b)$ with $\omega_l = \omega_u - 2|\omega_u|$ and $|\omega_u||I| = 1/2$, the interval ω_h is the infinite interval $[\sup \omega_u + |\omega_u|, \infty)$.

The functions ϕ_P are L^2 normalized Schwartz functions adapted to the multitiles P in the following sense: $\hat{\phi}_P$ is supported in $\frac{11}{10}\omega_u$ (here we dilate around the center of ω_u), and $e^{-2\pi i c(\omega_u)x}\phi_P(x)$ is adapted to I, that is

$$|\frac{d^n}{dx^n}[e^{-2\pi i c(\omega_u)\cdot}\phi_P](x)| \leq C|I|^{-(1/2+n)}(1+|x-c(I)|/|I|)^{-N}$$

for some large large N and for $n = 0, \ldots, N'$ for some large N'. (above $c(\omega_u)$ denotes the center of ω_u).

Finally, the functions a_P are obtained from the linearization/dualization process as follows. Fix some large M, real valued functions $\xi_0, \ldots, \xi_M$ on $\mathbb{R}$ satisfying $\xi_0(x) < \xi_1(x) < \ldots < \xi_M(x)$ for every x, and complex-valued functions $a_1, \ldots, a_M$ on $\mathbb{R}$ satisfying $\left(\sum_{j=1}^M |a_j(x)|^{r'}\right)^{1/r'} \leq 1$ for every x. Then, for each multitile P and $x \in \mathbb{R}$ there is at most one j with $\xi_{j-1} \in \omega_l$ and $\xi_j \in \omega_h$; if such a j exists, we set $a_P(x) = a_j(x)$ and otherwise we set $a_P(x) = 0$. We then want to prove (8) with $C_{p,r}$ independent of M and the functions $\xi_0, \ldots, \xi_M, a_1, \ldots, a_M$.

12.3.2 Bound for trees

We first consider a version of (8) with the (basically) arbitrary collection of multitiles $\mathbf{P}$ replaced by a type of collection T called a *tree* which we now define.

Let $C_2 = 2$. Given a dyadic interval I_T and a frequency $\xi_T \in \mathbb{R}$ consider the interval $\omega_T = (\xi_T - (C_2-1)/(4|I_T|), \xi_T + (C_2-1)/(4|I_T|))$. If T is a collection of multitiles satisfying $I \subset I_T$ and $\omega_T \subset \omega_m$ for every $P \in T$, where ω_m is the convex hull of $C_2\omega_u \cup C_2\omega_l$, then we say that that (T, I_T, ξ_T) is a tree. The tree is said to be *l-overlapping* if for every $P \in T$ we have $\xi_T \in C_2\omega_l$, we will say that it is *l-lacunary* if for every $P \in T$ we have $\xi_T \notin C_2\omega_l$. Of course every tree can be written as the union of an l-overlapping tree and an l-lacunary tree, each of which have the same top data (ξ_T, I_T) as the original tree. This allows us to reduce the proof of

$$\|\sum_{P\in T}\langle f, \phi_P\rangle\phi_P a_P\|_{L^p} \leq C_{p,r}\|f\|_{L^p} \tag{9}$$

where T is a tree, to the cases where T is l-lacunary or l-loverlapping.

If T is l-overlapping, then for each $P \in T$, we have the frequency support of ϕ_P contained in an interval of size $\approx 1/|I|$ distance $\approx 1/|I|$ away from ξ_T.

This lacunarity of the frequency supports allows us to obtain (9) from (6). Indeed, for each x, unraveling the definition of a_P gives

$$|\sum_{P\in T}\langle f,\phi_P\rangle\phi_P(x)a_P(x)| \leq (\sum_{k=1}^{M}|\sum_{\substack{P\in T\\ \xi_{k-1}(x)\in\omega_l\\ \xi_k(x)\in\omega_h}}\langle f,\phi_P\rangle\phi_P(x)|^r)^{1/r}.$$

A geometric argument shows that one may apply (9) to bound the L^p norm of the right side above by $\|\sum_{P\in T}\langle f,\phi_P\rangle\phi_P\|_{L^p}$, which can be controlled by $\|f\|_{L^p}$.

If T is l-lacunary, the situation is a little simpler. A geometric argument shows that for each x there is at most one $j(x)$ such that there is a $P\in T$ with $|I| = 2^{j(x)}$ and $a_P(x)\neq 0$. This implies that

$$|\sum_{P\in T}\langle f,\phi_P\rangle\phi_P(x)a_P(x)| \leq \sum_{\substack{P\in T\\ |I|=2^{j(x)}}}|\langle f,\phi_P\rangle\phi_P(x)|.$$

The right side above can be bounded by $\mathcal{M}[f](x)$ and applying the L^p bound for $\mathcal{M}$ gives (9).

To obtain the bound for arbitrary collections of multitiles (8), we would like to decompose $\mathbf{P}$ into the union of trees T and estimate each tree individually. The bound (9) is too crude for this purpose– one would like to improve the right side $\|f\|_{L^p}$ to a quantity which captures only the parts of f and a_P corresponding to the phase-space support of T. This is (at least sort of morally) accomplished by considering the quantities *energy* and *density*.

The energy of a collection of multitiles Q is

$$\sup_T(\frac{1}{|I_T|}\sum_{P\in T}|\langle f,\phi_P\rangle|^2)^{1/2}$$

where the sup is over all l-overlapping trees $T\subset Q$. The density of Q is

$$\sup_T(\int_E(1+|x-c(I_T)|/|I_T|)^{-4}\sum_{k=1}^{M}|a_k(x)|^{r'}1_{\omega_T}(\xi_{k-1}(x))\;dx)^{1/r'}$$

where the sup is over all nonempty trees T contained in Q, 1 denotes the characteristic function of a set, and 1_E is a (fixed depending on f, independent of Q) dualizing function.

Using the ideas from the proof of (9) together with a careful localization argument (essentially as in [8]), one then proves

$$\|1_E \sum_{P\in T} \langle f, \phi_P\rangle \phi_P a_P\|_{L^p} \leq C_{p,r}\, \text{energy}(T)(\text{density}(T))^{\min(1,r'/p)}|I_T|^{1/p}. \quad (10)$$

12.3.3 Bound for arbitrary collections of tiles

We now want to obtain (8) from (10); for simplicity assume $f = 1_F$ where F has finite measure, also consider a dualizing function 1_E where $|E| \approx 1$ (we can at least assume this after rescaling).

The more complicated case is $|F| \geq 1$ We decompose

$$\mathbf{P} = \bigcup_{k\geq 0} \bigcup_{T\in \mathbf{T}_k} T$$

where the union above is disjoint and, for each k, $\mathbf{T}_k$ is a collection of trees which satisfy

$$\text{energy}(T) \leq 2^{-k/2}|F|^{1/2}, \text{density}(T) \leq 2^{-k/r'} \text{ for every } T \in \mathbf{T}_k \quad (11)$$

$$\|\sum_{T\in\mathbf{T}_k} 1_{I_T}\|_{L^1} \leq C2^k \quad (12)$$

$$\|\sum_{T\in\mathbf{T}_k} 1_{I_T}\|_{BMO} \leq C2^k|F|^{-1}. \quad (13)$$

This decomposition is obtained through energy/density decrement lemmas such as in [8], with the additional consideration of the BMO norm. In reality we are not actually able to obtain (13) from the density lemma, but through some additional decompositional trickery we may pretend it holds.

Then

$$\|1_E \sum_{P\in\mathbf{P}} \langle f, \phi_P\rangle \phi_P a_P\|_{L^1} \leq \sum_{k\geq 0} \|1_E \sum_{T\in\mathbf{T}_k} \sum_{P\in T} \langle f, \phi_P\rangle \phi_P a_P\|_{L^1}.$$

Through the use of exceptional sets, we can basically pretend the BMO norm in (13) is an L^∞ norm. Ignoring some additional technical complications with Schwartz tails, we assume that each ϕ_P is supported on I. Then we obtain, for each k,

$$\|1_E \sum_{T\in\mathbf{T}_k} \sum_{P\in T} \langle f, \phi_P\rangle \phi_P a_P\|_{L^1} \leq C2^{k/r}|F|^{-1/r}\|1_E(\sum_{T\in\mathbf{T}_k} |\sum_{P\in T}\langle f, \phi_P\rangle \phi_P a_P|^{r'})^{1/r'}\|_{L^1}$$

Estimating L^1 by $L^{r'}$ and applying (10) with $p = r'$, we bound the right side above by

$$C2^{k/r}|F|^{-1/r}2^{-k/2}|F|^{1/2}2^{-k/r'}(\sum_{T\in\mathbf{T}_k}|I_T|)^{1/r'}$$

which, by (12) is

$$\leq C2^{-k(\frac{1}{2}-\frac{1}{r})}|F|^{1/2-1/r}$$

Summing over k gives

$$\|1_E\sum_{P\in\mathbf{P}}\langle f,\phi_P\rangle\phi_P a_P\|_{L^1} \leq C|F|^{1/2-1/r}.$$

This proves the desired bound when $p < (1/2 - 1/r)^{-1}$. This restriction in p may be lifted by using the monotonicity of the V^r norms.

The bound when $|F| < 1$ is essentially as in [5]. An exceptional set where $\mathcal{M}[f]$ is large is defined. Throwing out multitiles whose time support is contained inside this exceptional set, we obtain a good bound on the energy of the remaining tiles using a square function argument. Then we apply (10) with $p = 1$ (in this case no BMO estimates are needed), and everything adds up.

12.4 Variational norms on Lie groups

Let G be a connected finite-dimensional Lie group with Lie algebra $\mathfrak{g}$. We begin by recalling the construction of a left-invariant metric on G. Letting $\|\cdot\|_{\mathfrak{g}}$ be any norm on $\mathfrak{g}$, push forward with left-multiplication to define a norm on each tangent space T_gG of the group $\|x\|_{T_gG} = \|g^{-1}x\|_{\mathfrak{g}}$ (here g^{-1} denotes the inverse of the differential of the map from $G \to G$ defined by left-multiplication by g – this notation is natural when G is a matrix group). The induced norm is left-invariant i.e. $\|hx\|_{T_{hg}G} = \|x\|_{T_gG}$. We define the *length* of a continuously differentiable path $\gamma : [a, b] \to G$ by

$$|\gamma| = \int_a^b \|\gamma'(t)\|_{T_{\gamma(t)}G}\, dt \tag{14}$$

and the metric $d(g, g')$ on G

$$d(g, g') = \inf_{\gamma:\gamma(a)=g,\gamma(b)=g'} |\gamma|. \tag{15}$$

Now we consider variational lengths of curves. For $1 \leq r < \infty$ and a continuous path $\gamma : [a, b] \to G$ define the r-variation of γ

$$\|\gamma\|_{V^r} = \sup_{N, t_0 < \ldots < t_N \in [a,b]} \left(\sum_{j=1}^{N} d(\gamma(t_{j-1}), \gamma(t_j))^r\right)^{1/r}. \tag{16}$$

This is extended to $r = \infty$ in the usual way, replacing the sum by a sup. When γ is differentiable, the V^1 norm coincides with the length $|\gamma|$ defined earlier; the V^∞ norm is just the diameter of the curve.

Associated to each continuously differentable curve $\gamma : [a, b] \to G$ is the *left trace* $\gamma_l : [a, b] \to \mathfrak{g}$ defined

$$\gamma_l(t) = \int_a^t \gamma(s)^{-1} \gamma'(s)\, ds.$$

Note that one can recover the original curve from the trace, by solving the ordinary differential equation

$$\gamma'(t) = \gamma(t)\gamma_l'(t)$$

with the initial condition $\gamma(a)$. Using $\|\cdot\|_{\mathfrak{g}}$ in place of $\|\cdot\|_{T_{\gamma(t)}G}$, the definitions (14),(15),(16) yield notions of length and r-variation for curves in $\mathfrak{g}$. We then have

$$\|\gamma_l\|_{V^1} = \|\gamma\|_{V^1}. \tag{17}$$

The main result here is that (17) can be (sort of) extended to $1 \leq r < 2$:

Theorem 2. *Suppose* $1 \leq r < 2$ *and* $\gamma : [a, b] \to G$ *is smooth. Then*

$$\|\gamma\|_{V^r} \leq \|\gamma_l\|_{V^r} + C \min(\|\gamma_l\|_{V^r}^2, \|\gamma_l\|_{V^r}^r)$$

and

$$\|\gamma_l\|_{V^r} \leq \|\gamma\|_{V^r} + C \min(\|\gamma\|_{V^r}^2, \|\gamma\|_{V^r}^r)$$

Now a few words describing the proof. An induction on scales argument using the Baker-Campbell-Hausdorff formula allows one to basically reduce matters to comparing $d(\gamma(t_{j-1}), \gamma(t_j))$ with $d(\gamma_l(t_{j-1}), \gamma_l(t_j))$ where t_{j-1} and t_j are close together and "close" is allowed to depend on γ. Using left-invariance, we have $d(\gamma(t_{j-1}), \gamma(t_j)) = d(I, \gamma(t_{j-1})^{-1}\gamma(t_j))$ (here $I \in G$ is the identity). In the special case that γ_l is a straight line on $[t_{j-1}, t_j]$ we have $\gamma(t_{j-1})^{-1}\gamma(t_j) = \exp(\gamma_l(t_j) - \gamma_l(t_{j-1}))$ and so $d(1, \gamma(t_{j-1})^{-1}\gamma(t_j)) = d(\gamma_l(t_{j-1}), \gamma_l(t_j))$. But, any differentiable function is locally linear, so for t_{j-1} and t_j close-together γ_l is almost a straight line on $[t_{j-1}, t_j]$. Using stability of solutions to ODE's, almost is good enough.

Here is the motivation for Theorem 2. Fix a Lie group G and any two vectors $w, v \in \mathfrak{g}$. Given a function f on $\mathbb{R}$ and a frequency $\xi \in \mathbb{R}$ we can define a curve γ in G via the formula

$$\gamma_l'(t) = \mathrm{Re}(e^{-2\pi i\xi t} f(t))w + \mathrm{Im}(e^{-2\pi i\xi t} f(t))v$$

and initial condition $\gamma(0) = I$. Then, we can define a "nonlinear Fourier summation operator" $\mathcal{NC}[f](\xi, x) = \gamma(x)$. The trace γ_l is identified with the usual partial Fourier transform $\mathcal{C}[f](\xi, x) = \int_0^x e^{-2\pi i\xi t} f(t)\ dt$. Theorem 2 then allows one to deduce the bound

$$\|1_{|\mathcal{NC}[f]|\leq 1}\mathcal{NC}[f]\|_{L_\xi^{p'}(V_x^r)} \leq C\|f\|_{L^p(\mathbb{R})}$$

for $r < 2$ from the variational version of the Menshov-Paley-Zygmund theorem

$$\|\mathcal{C}[f]\|_{L_\xi^{p'}(V_x^r)} \leq C'\|f\|_{L^p} \tag{18}$$

which holds for $1 \leq p \leq 2$ and $r > p'$.

For $p = 2$, the bound (18) is a restatement of the $p = 2$ case of Theorem 1. For $p < 2$, it follows by interpolating $p = 2$ with a trivial estimate at $p = 1$. Alternately, one can apply a variational version of the Christ-Kiselev lemma (which can be proven using the usual Christ-Kiselev lemma from [3]) to deduce (18) for $p < 2$ from the Hausdorff-Young inequality.

A case of particular interest above is when $G = SU(1, 1)$,

$$w = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \text{ and } v = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

Here, one obtains estimates which give a variational/Dirac operator version of the Christ-Kiselev theorem from [4].

References

[1] Jean Bourgain, *Pointwise ergodic theorems for arithmetic sets*, Inst. Hautes Études Sci. Publ. Math. (1989), no. 69, 5–45, With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein.

[2] Lennart Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135–157.

[3] Michael Christ and Alexander Kiselev, *Maximal functions associated to filtrations*, J. Funct. Anal. **179** (2001), no. 2, 409–425.

[4] Michael Christ and Alexander Kiselev, *WKB asymptotic behavior of almost all generalized eigenfunctions for one-dimensional Schrödinger operators with slowly decaying potentials*, J. Funct. Anal. **179** (2001), no. 2, 426–447.

[5] Loukas Grafakos, Terence Tao, and Erin Terwilleger, *L^p bounds for a maximal dyadic sum operator*, Math. Z. **246** (2004), no. 1-2, 321–337.

[6] Richard A. Hunt, *On the convergence of Fourier series*, Orthogonal Expansions and their Continuous Analogues (Proc. Conf., Edwardsville, Ill., 1967), Southern Illinois Univ. Press, Carbondale, Ill., 1968, pp. 235–255.

[7] Roger L. Jones, Robert Kaufmann, Joseph Rosenblatt, and Máté Wierdl, *Oscillation in ergodic theory*, Ergodic Theory Dynam. Systems **18** (1998), no. 4, 889–935.

[8] Michael Lacey and Christoph Thiele, *A proof of boundedness of the Carleson operator*, Math. Res. Lett. **7** (2000), no. 4, 361–370.

[9] Richard Oberlin, Andreas Seeger, Terence Tao, Christoph Thiele, and Jim Wright, *A variational norm Carleson theorem*, preprint.

[10] Jinghua Qian, *The p-variation of partial sum processes and the empirical process*, Ann. Probab. **26** (1998), no. 3, 1370–1383.

[11] Christoph Thiele, *Wave packet analysis*, CBMS Regional Conference Series in Mathematics, vol. 105, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2006.

Richard Oberlin, UCLA
email: **richard.oberlin@gmail.com**

13 On multilinear oscillatory integrals, non-singular and singular

after M. Christ, X. Li, T. Tao and C. Thiele [2]
A summary written by Diogo Oliveira e Silva

Abstract

We explore the relationship between decay estimates for certain multilinear oscillatory integrals and nondegeneracy of the corresponding polynomial phase.

13.1 Introduction

The classical theory of oscillatory integrals of the second kind [3] establishes the boundedness of operators of the form

$$(T_\lambda f)(\xi) = \int_{\mathbb{R}^m} e^{i\lambda\Phi(x,\xi)} f(x)\psi(x,\xi)dx.$$

Here ψ is a fixed smooth function of compact support in x and ξ, the phase Φ is real-valued and smooth and the Hessian of Φ is nonvanishing on the support of ψ. In this case,

$$\|T_\lambda(f)\|_{L^2(\mathbb{R}^m)} \leq C\lambda^{-m/2}\|f\|_{L^2(\mathbb{R}^m)}.$$

Aiming at similar results in a somewhat different context, we start by considering multilinear oscillatory operators of the form

$$\Lambda_\lambda(f_1, \ldots, f_n) = \int_{\mathbb{R}^m} e^{i\lambda P(x)} \prod_{j=1}^{n} f_j \circ \pi_j(x)\eta(x)dx.$$

Here $P : \mathbb{R}^m \to \mathbb{R}$ is a real-valued polynomial, $\pi_j : \mathbb{R}^m \to V_j$ are orthogonal projections onto some subspaces V_j of $\mathbb{R}^m$, $f_j : V_j \to \mathbb{C}$ are locally integrable functions with respect to Lebesgue measure on V_j, and $\eta \in C_0^1(\mathbb{R}^m)$ is compactly supported. We assume that all the subspaces V_j have the same dimension, which we denote by κ.

Definition 1. *A polynomial P has the* power decay property with respect to $\{V_j\}_{j=1}^n$ in an open set $U \subset \mathbb{R}^m$ *if for any $\eta \in C_0^1(U)$, there exist $\epsilon > 0$ and $C < \infty$ such that*

$$|\Lambda_\lambda(f_1, \dots, f_n)| \leq C(1 + |\lambda|)^{-\epsilon} \prod_{j=1}^{n} \|f_j\|_{L^\infty(V_j)}, \quad \forall f_j \in L^\infty(V_j), \forall \lambda \in \mathbb{R}. \quad (1)$$

The goal is to characterize data $(P, \{V_j\}_j)$ for which the power decay property holds, and this is accomplished to a significant but incomplete extent in [2].

Observe that, if $P = \sum_j f_j \circ \pi_j$ for some measurable functions f_j, then (1) cannot hold. This motivates the following definition:

Definition 2. *A polynomial P is* degenerate with respect to $\{V_j\}_{j=1}^n$ *if $P = \sum_{j=1}^n p_j \circ \pi_j$ for some polynomials $p_j : V_j \to \mathbb{R}$. Otherwise P is said to be* nondegenerate. *If $n = 0$, P is degenerate if and only if P is constant.*

An important definition is not complete without a good example:

Example 3. *In $\mathbb{R}^3$, let $P(x) = x_3^2$ and $L = \frac{\partial^2}{\partial x_3^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2}$. Let $n \geq 1$ be arbitrary. For $1 \leq j \leq n$, take light-cone unit vectors $v_j = (v_j^1, v_j^2, v_j^3) \in \mathbb{R}^3$ such that $(v_j^3)^2 = (v_j^1)^2 + (v_j^2)^2$. Let π_j be the orthogonal projection onto $V_j :=$ span (v_j), that is, $\pi_j(x) = x \cdot v_j = x_1 v_j^1 + x_2 v_j^2 + x_3 v_j^3$. One readily checks that P is nondegenerate with respect to $\{V_j\}_{j=1}^n$, for every $n \in \mathbb{N}$. This is surprising in view of the following fact: in $\mathbb{R}^2$, any polynomial $Q : \mathbb{R}^2 \to \mathbb{R}$ of degree two is degenerate with respect to any family of three or more mappings of the form $\pi_j(x) = x \cdot w_j$ (where none of the w_j is a multiple of any of the others).*

The nondegeneracy condition is to replace the hypothesis of a nonvanishing derivative in the result about oscillatory integrals of the second kind. The question is then the following:

Question 4. *Is the power decay property equivalent to nondegeneracy?*

The cases $n = 0, 1$ fall into the theory of oscillatory integrals of the first kind. In the case $n = 2$ one is dealing with bilinear forms $\langle T_\lambda(f_1), f_2 \rangle$, where the associated operators T_λ are of the form discussed above. If $n \geq 3$ and $m < n\kappa$, however, there arises a class of singular oscillatory integrals which have no direct analogues in the bilinear case.

13.2 Algebraic aspects of nondegeneracy

We begin our discussion by introducing a new definition:

Definition 5. *A polynomial P is* simply nondegenerate with respect to $\{V_j\}_{j=1}^n$ *if there exists a differential operator of the form $L = \prod_{j=1}^n w_j \cdot \nabla$ with $w_j \in V_j^\perp$ and such that $L(P)$ does not vanish identically.*

Simple nondegeneracy implies nondegeneracy, but the converse does not hold in general (example 3). The converse does hold, however, in the following two special cases:

(i) If $\kappa = m - 1$;

(ii) If $n(m - \kappa) \leq m$ and the V_j's are in general position[35].

We have the following characterization of nondegeneracy:

Lemma 6. *Let P be a real-valued polynomial of degree d. Then:*

(i) P is nondegenerate if and only if there exists a constant-coefficient partial differential operator L such that $L(P) \neq 0$ but $L(\sum_j p_j \circ \pi_j) = 0$ for every polynomial p_j of degree d;

(ii) P is degenerate if and only if $P = \sum_j h_j \circ \pi_j$ for some distributions $h \in \mathcal{D}'(V_j)$;

(iii) P is nondegenerate if and only if one of its homogeneous summands is nondegerate.

In the case of homogeneous polynomials, we can refine our characterization as follows:

Lemma 7. *Let P be a homogeneous polynomial of degree d. Then:*

(i') P is nondegenerate if and only if there exists a constant-coefficient partial differential operator L, homogeneous of degree d, such that $L(P) \neq 0$ but $L(\sum_j p_j \circ \pi_j) = 0$ for every polynomial p_j of degree d;

(ii') P is degenerate if and only if $P = \sum_j p_j \circ \pi_j$ for some homogeneous polynomials p_j of degree d.

[35]In this context, a family of subspaces $\{V_j\}_{j=1}^n$ of $\mathbb{R}^m$ of dimension κ is said to be in *general position* if any subfamily of cardinality $k \geq 1$ spans a subspace of dimension $\min\{k\kappa, m\}$.

13.3 Main results

13.3.1 Further definitions

Let $\mathcal{P}_{\le d}$ denote the finite-dimensional vector space of polynomials in $\mathbb{R}^m$ of degree $\le d$, endowed with a metric $\|\cdot\|$. Given d, the norm of P with respect to $\{V_j\}_{j=1}^n$ is defined to be

$$[P]_{d,\{V_j\}_j} := \inf_{\deg p_j \le d} \|P - \sum_{j=1}^n p_j \circ \pi_j\|.$$

This indeed defines a norm on the quotient space $\mathcal{P}_{\le d}$ modulo degenerate polynomials.

Definition 8. *A family of polynomials* $\{P_\alpha\}_\alpha$ *is* uniformly nondegenerate with respect to $\{V_j\}_{j=1}^n$ *if there exist* $d < \infty$ *and* $c > 0$ *such that*

$$\sup_\alpha \deg P_\alpha \le d \quad \textit{and} \quad \inf_\alpha [P_\alpha]_{d,\{V_j\}_j} \ge c.$$

Definition 9. *The collection* $\{V_j\}_{j=1}^n$ *has the* power decay property *if every polynomial* P *which is nondegenerate with respect to* $\{V_j\}_j$ *has the power decay property* (1) *in every open set* U. *The power decay is* uniform *if* (1) *holds with uniform constants* C, ϵ *for any family of polynomials which are uniformly nondegenerate with respect to* $\{V_j\}_j$.

13.3.2 Decay for nonsingular multilinear oscillatory integrals

The first result states that a simply nondegenerate polynomial has the power decay property in every open set. More precisely:

Theorem 10. *Fix* $d \in \mathbb{N}$ *and* $c > 0$. *Consider the operator* $L = \prod_j w_j \cdot \nabla$ *with* $w_j \in V_j^\perp$ *and* $\|w_j\| = 1$. *Then there exist* $C < \infty$ *and* $\epsilon > 0$ *with the following property: if* P *is a polynomial such that* $\deg P \le d$ *and* $\max_{|x|\le 1} |L(P)(x)| \ge c$, *then* (1) *holds.*

As corollaries, we get that families of codimension one subspaces have the uniform power decay property, as do families of small[36] codimension subspaces in general position.

The second result tells us that the same conclusion still holds in the one-dimensional case, provided we do not consider "too many" subspaces:

[36]Here, "small" means that $n(m-\kappa) \le m$.

Theorem 11. *If $n < 2m$, then any family $\{V_j\}_j$ of one-dimensional subspaces which lie in general position has the uniform power decay property. Moreover under these assumptions*

$$|\Lambda_\lambda(f_1, \ldots, f_n)| \leq C(1 + |\lambda|)^{-\epsilon} \prod_{j=1}^{n} \|f_j\|_{L^2(\mathbb{R})}, \quad \forall f_j \in L^2(\mathbb{R}), \forall \lambda \in \mathbb{R} \qquad (2)$$

uniformly for all polynomials P which are uniformly nondegenerate with respect to $\{V_j\}_j$.

The rest of this paper is devoted to outlining the proofs of theorem 10 (the simply nondegenerate case) and theorem 11 (the case $\kappa = 1$). Both proofs are by induction on the number of subspaces, the base case $n = 0$ being a straightforward consequence of the well-known theory of oscillatory integrals of the first kind.

13.4 The simply nondegenerate case

In this section we sketch the proof of the fact that families of codimension one subspaces have the uniform power decay property. This turns out to be equivalent to theorem 10.

We start by expressing $\Lambda_\lambda(f_1, \ldots, f_n) = \langle T_\lambda(f_1, \ldots, f_{n-1}), \bar{f}_n \rangle$. By Cauchy-Schwarz, it is enough to show the existence of $C < \infty$ and $\epsilon > 0$ such that

$$\|T_\lambda(f_1, \ldots, f_{n-1})\|_2 \leq C\lambda^{-\epsilon} \prod_{j=1}^{n-1} \|f_j\|_\infty \text{ for } |\lambda| \geq 1.$$

Choose coordinates $x = (y, z) \in \mathbb{R}^{m-1} \times \mathbb{R}$ in such a way that $V_n = \{z = 0\}$, and define $Q_\zeta(y, z) := P(y, z) - P(y, z + \zeta)$, $F_j^\zeta(\pi_j(y, z)) := f_j(\pi_j(y, z))\bar{f}_j(\pi_j(y, z + \zeta))$ and $\tilde{\eta}_\zeta(y, z) := \eta(y, z)\bar{\eta}(y, z + \zeta)$. Then

$$\begin{aligned} \|T_\lambda(f_1, \ldots, f_{n-1})\|_2^2 &= \int_{\mathbb{R}} \Big(\int_{\mathbb{R}^m} e^{i\lambda Q_\zeta(x)} \prod_{j=1}^{n-1} F_j^\zeta(\pi_j(x)) \tilde{\eta_\zeta}(x) dx \Big) d\zeta \\ &=: \int_{\mathbb{R}} \Lambda_\lambda^\zeta(F_1^\zeta, \ldots, F_{n-1}^\zeta) d\zeta. \end{aligned}$$

Since P is nondegenerate and $\kappa = m - 1$, P is simply nondegenerate. Let $L = \prod_{j=1}^{n} w_j \cdot \nabla$ be such that $\max_{|x| \leq 1} |L(P)(x)| \geq c > 0$ and $L' := \prod_{j<n} w_j \cdot \nabla$. The key idea is to define the sublevel set

$$E_\rho := \{\zeta \in \mathbb{R} : \max_{|x|\le 1} |L' Q_\zeta(x)| \le \rho\}$$

and to prove the bound $|E_\rho| \le C\rho^\delta$ for some $C < \infty$ and $\delta > 0$. For this purpose, note that the hypothesis on L and the fact that $w_n \cdot \nabla = \partial_{x_m}$ together imply that $\sup_{(x,\zeta)} |\partial_\zeta(L'Q_\zeta(x))| \ge c$. The desired bound follows from a well-known sublevel set estimate [1], provided x, ζ are restricted to lie in a fixed bounded set.

If $\zeta \notin E_\rho$, we use the induction hypothesis to conclude that

$$|\Lambda_\lambda^\zeta(F_1^\zeta, \dots, F_{n-1}^\zeta)| \le C(1 + |\lambda|\rho)^{-\epsilon'} \prod_{j<n} \|f_j\|_\infty^2.$$

Putting everything together, we have that

$$\begin{aligned} \|T_\lambda(f_1, \dots, f_{n-1})\|_2^2 = & \int_{\mathbb{R}\setminus E_\rho} \Lambda_\lambda^\zeta(F_1^\zeta, \dots, F_{n-1}^\zeta) d\zeta + \int_{E_\rho} \Lambda_\lambda^\zeta(F_1^\zeta, \dots, F_{n-1}^\zeta) d\zeta \\ \le & C(1 + |\lambda|\rho)^{-\epsilon'} \prod_{j<n} \|f_j\|_\infty^2 + C|E_\rho| \prod_{j<n} \|F_j^\zeta\|_\infty \\ \le & C((|\lambda|\rho)^{-\epsilon'} + \rho^\delta) \prod_{j=1}^{n-1} \|f_j\|_\infty^2. \end{aligned}$$

Choosing $\rho = |\lambda|^{-\frac{\epsilon'}{\epsilon'+\delta}}$ yields the desired bound.

13.5 The case $\kappa = 1$

This section is devoted to presenting the main ideas behind the proof of the statement that, for a family $\{V_j\}_j$ of one-dimensional subspaces which lie in general position, the estimate (2) holds provided $n < 2m$. This last condition turns out to be necessary as well.

Since the case $n < m$ follows by a simple argument from the case $n = m$ and the theorem is already known in a more precise form when $n = m$ [3], we may assume without loss of generality that $m < n < 2m$.

Letting $A(\lambda)$ denote the best constant for which

$$|\Lambda_\lambda(f_1, \dots, f_n)| \le A(\lambda) \prod_j \|f_j\|_2,$$

it is enough to show that $|A(\lambda)| \le C|\lambda|^{-\epsilon}$ for some $\epsilon > 0$.

In what follows we may assume that f_1 is λ-uniform in the sense that its generalized Fourier coefficients satisfy the bounds

$$\Big| \int f_1(t) e^{-iq(t)} dt \Big| \le C|\lambda|^{-\tau} \|f_1\|_2 \text{ uniformly for all real-valued polynomials } q.$$

Indeed, f_1 could otherwise be decomposed in L^2 into its projection onto e^{iq} plus an orthogonal vector in such a way that

$$\|f_1 - ce^{iq}\|_2 \le (1 - |\lambda|^{-2\tau})\|f_1\|_2,$$

and the desired conclusion would then follow from this and the inductive hypothesis.

We lose no generality in assuming that $\|f_j\|_2 \le 1$ for every j.

Endowing $\mathbb{R}^m$ with suitable coordinates is once again an important technical point of the proof. For this purpose, let e_1 be a unit vector orthogonal to the span of $\{V_j\}_{j=2}^m$, and e_2 be a unit vector orthogonal to the span of $\{V_j\}_{j=m+1}^n$ and not orthogonal to V_1. Then e_1 and e_2 are linearly independent, and so we write $\mathbb{R}^m \ni x = t_1 e_1 + t_2 e_2 + y$ (that is, $x = (t, y) \in \mathbb{R}^2 \times \mathbb{R}^{m-2}$).

Defining $P^y(t) := P(t, y)$,

$$F_1^y(t_2) := \prod_{j=2}^m f_j(\pi_j(t, y)),$$

$$F_2^y(t_1) := \prod_{j=m+1}^n f_j(\pi_j(t, y)) \quad \text{and} \quad G^y(\pi(t)) := f_1(\pi_1(t, y)),$$

we have that

$$\Lambda_\lambda(f_1, \ldots, f_n) = \int_{\mathbb{R}^{m-2}} \Big(\int_{\mathbb{R}^2} e^{i\lambda P^y(t)} F_1^y(t_2) F_2^y(t_1) G^y(\pi(t)) \eta(t, y) dt \Big) dy$$
$$=: \int_{\mathbb{R}^{m-2}} \Lambda_\lambda^y dy.$$

Moreover the assumptions, Fubini's theorem and Cauchy-Schwarz together imply that

$$\int_{\mathbb{R}^{m-2}} \|F_1^y\|_2 \|F_2^y\|_2 \|G^y\|_2 dy < \infty. \tag{3}$$

The last important step consists of introducing a set of "bad" parameters, denoted $\mathcal{B}$, consisting of all y for which P^y has small norm in the quotient space $\mathcal{P}_{\leq d}$ modulo degenerate polynomials with respect to the three projections $t \mapsto t_1, t_2, \pi(t)$. More precisely, $y \in \mathcal{B}$ if P^y can be decomposed as

$$P^y(t) = Q_1(t_1) + Q_2(t_2) + Q_3(\pi(t)) + R(t),$$

for some polynomials Q_j and R of degree $\leq d$, with the additional requirement that $\|R\| \leq |\lambda|^{-\rho}$ (here $\|\cdot\|$ denotes a given norm on $\mathcal{P}_{\leq d}$ and ρ is a small parameter to be chosen later on).

On $\mathcal{B}^{\complement}$, we can use theorem 10 with $m = 2$ and $n = 3$, interpolate, and appeal to (3) to conclude that

$$\int_{y \notin \mathcal{B}} |\Lambda_\lambda^y| dy \leq C|\lambda|^{-(1-\rho)\tilde{\rho}} \text{ for some } \tilde{\rho} > 0.$$

Despite that fact that the set of bad parameters might have full measure (example 3), we will be in good shape if we show that if ρ is small enough, then there exists $\tilde{\epsilon} > 0$ such that

$$|\Lambda_\lambda^y| \leq C|\lambda|^{-\tilde{\epsilon}} \|F_1^y\|_2 \|F_2^y\|_2 \|G^y\|_2 \text{ uniformly for all } y \in \mathcal{B}.$$

This is a nice exercise in Fourier analysis and involves a clever use of the uniformity condition on f_1. I will omit the details for now and present them at the summer school.

References

[1] A. Carbery, M. Christ and J. Wright, *Multidimensional van der Corput and sublevel set estimates.* J. Amer. Math. Soc. **12** (1999), no.4, 981-1015.

[2] M. Christ, X. Li, T. Tao and C. Thiele, *On multilinear oscillatory integrals, nonsingular and singular.* Duke Math. J. **130** (2005), no. 2, 321-351.

[3] E. M. Stein, *Harmonic Analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, 43. Princeton University Press, Princeton, NJ, 1993.

[4] T. Tao and V. H. Vu, *Additive Combinatorics*, Cambridge studies in advanced mathematics, 105. Cambridge University Press, Cambridge, 2006.

Diogo Oliveira e Silva, UC Berkeley
email: **dosilva@math.berkeley.edu**

14 Bi-parameter paraproducts

after C. Muscalu, J. Pipher, T. Tao and C. Thiele [1]
A summary written by Katharine Ott

Abstract

The authors prove the bi-parameter analogue of the Coifman-Meyer theorem.

14.1 Introduction

Let $m(= m(\gamma))$ in $L^\infty(\mathbb{R}^2)$ be a bounded function, smooth away from the origin, satisfying $|\partial^\beta m(\gamma)| \lesssim \frac{1}{|\gamma|^{|\beta|}}$ for sufficiently many β. Denote by $T_m^{(1)}$ the bi-linear operator

$$T_m^{(1)}(f,g)(x) = \int_{\mathbb{R}^2} m(\gamma)\widehat{f}(\gamma_1)\widehat{g}(\gamma_2)e^{2\pi i x(\gamma_1+\gamma_2)}d\gamma, \tag{1}$$

where f, g are Schwartz functions on the real line $\mathbb{R}$. Then the classical Coifman-Meyer theorem states the following.

Theorem 1. *The operator $T_m^{(1)}$ defined in (1) maps $L^p \times L^q \to L^r$ boundedly provided that $1 < p, q \leq \infty$, $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$ and $0 < r < \infty$.*

Now consider the bi-parameter analogue of $T_m^{(1)}$. Let $m(= m(\xi,\eta))$ in $L^\infty(\mathbb{R}^4)$ be a bounded function that is smooth away from the subspaces $\{\xi = 0\} \cup \{\eta = 0\}$ and satisfying

$$|\partial_{\xi_1}^{\alpha_1}\partial_{\xi_2}^{\alpha_2}\partial_{\eta_1}^{\beta_1}\partial_{\eta_2}^{\beta_2} m(\xi,\eta)| \leq \frac{1}{|(\xi_1,\eta_1)|^{\alpha_1+\beta_1}|(\xi_2,\eta_2)|^{\alpha_2+\beta_2}} \tag{2}$$

for sufficiently many multi-indices α, β. Denote by T_m the bi-linear operator given by

$$T_m(f,g)(x) = \int_{\mathbb{R}^4} m(\xi,\eta)\widehat{f}(\xi)\widehat{g}(\eta)e^{2\pi i x(\xi+\eta)}d\xi d\eta, \tag{3}$$

where f, g are Schwartz functions on the plane $\mathbb{R}^2$. The main result of the paper is the following.

Theorem 2. *Let m be a symbol in $\mathbb{R}^4$ satisfying (2). Then the bilinear operator T_m defined by (3) maps $L^p \times L^q \to L^r$ boundedly as long as $1 < p, q \leq \infty$, $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$ and $0 < r < \infty$.*

This summary proceeds as follows. In the next section the operator T_m will be discretized and reduced to finitely many bi-parameter paraproducts. At this stage the proof of the main result reduces to a proposition regarding the size of each paraproduct. The third section summarize the main steps in proving Proposition 3. As noted in [1], the n-linear analogue of the main theorem is true for $n \geq 1$ (see also [2]).

14.2 Bi-parameter paraproducts

The first step toward proving Theorem 2 is to separate the operator T_m into smaller pieces well suited to its bi-parameter structure. The operator T_m can be discretized into finitely many operators of the form

$$\Pi_{\mathbf{\vec{P}}}^{\vec{j}}(f,g) = \sum_{\vec{P} \in \mathbf{\vec{P}}} \frac{1}{|I_{\vec{P}}|^{1/2}} \langle f, \Phi^1_{\vec{P}_1} \rangle \langle g, \Phi^2_{\vec{P}_2} \rangle \Phi^3_{\vec{P}_3}, \tag{4}$$

where $\vec{j} = (j^{'}, j^{''}) \in \{1,2,3\}^2$. Due to the symmetry in (4) is suffices to analyze only one case; consider $\vec{j} = (1,2)$. Also, in the equation above f, g are complex-valued measurable functions on $\mathbb{R}^2$, $\mathbf{\vec{P}}$ is a collection of bi-parameter tiles corresponding to lattice points $\vec{k} = (k^{'}, k^{''}), \vec{l} = (l^{'}, l^{''})$, $\vec{k}, \vec{l} \in \mathbb{Z}^2$, and $I_{\vec{P}} = I_{P'} \times I_{P''}$ with $P^{'}, P^{''}$ dyadic intervals. Specifically, $\vec{P}_i = (P_i^{'}, P_i^{''})$ and

$$\vec{P}_1 = (2^{-k^{'}}[l^{'}, l' + 1] \times 2^{k'}[-\frac{1}{4}, \frac{1}{4}], 2^{-k^{''}}[l^{''}, l^{''} + 1] \times 2^{k^{''}}[\frac{3}{4}, \frac{5}{4}]),$$

$$\vec{P}_2 = (2^{-k^{'}}[l^{'}, l' + 1] \times 2^{k'}[\frac{3}{4}, \frac{5}{4}], 2^{-k^{''}}[l^{''}, l^{''} + 1] \times 2^{k^{''}}[-\frac{1}{4}, \frac{1}{4}]),$$

$$\vec{P}_3 = (2^{-k^{'}}[l^{'}, l' + 1] \times 2^{k'}[-\frac{7}{4}, -\frac{1}{4}], 2^{-k^{''}}[l^{''}, l^{''} + 1] \times 2^{k^{''}}[-\frac{7}{4}, -\frac{1}{4}]),$$

and $|I_{\vec{P}}| := |I_{\vec{P}_1}| = |I_{\vec{P}_2}| = |I_{\vec{P}_3}| = 2^{-k^{'}} 2^{-k^{''}}$. Then the functions $\Phi^i_{\vec{P}_i} = \Phi^i_{P_i^{'}} \otimes \Phi^i_{P_i^{''}}$, $i = 1,2,3$, are $L^2(\mathbb{R}^2)$-normalized bump functions adapted to $\vec{P}_i$. In particular, if $i \neq j^{'}$, then $\int_{\mathbb{R}} \Phi^i_{P_i^{'}}(x)dx = 0$ and if $i \neq j^{''}$, $\int_{\mathbb{R}} \Phi^i_{P_i^{''}}(x)dx = 0$.

Denote by $\Lambda_{\vec{\mathbf{P}}}(f,g,h)$ the trilinear form given by

$$\begin{aligned}\Lambda_{\vec{\mathbf{P}}}(f,g,h) &= \int_{\mathbb{R}^2} \Pi_{\vec{\mathbf{P}}}(f,g)(x)h(x)dx \\ &= \sum_{\vec{P}\in\vec{\mathbf{P}}} \frac{1}{|I_{\vec{P}}|^{1/2}} \langle f, \Phi^1_{\vec{P}_1}\rangle\langle g, \Phi^2_{\vec{P}_2}\rangle\langle h, \Phi^3_{\vec{P}_3}\rangle. \qquad (5)\end{aligned}$$

The following proposition is the key ingredient in the proof of the main result. Once it is proved, the statement of Theorem 2 follows by appealing to the symmetry of the paraproduct and utilizing multilinear interpolation.

Proposition 3. *Let $1 < p,q < \infty$ be two numbers arbitrarily close to 1. Also, let $f \in L^p$, $g \in L^q$ such that $||f||_p = ||g||_q = 1$, and $E_3 \subseteq \mathbb{R}^2$ with $|E_3| = 1$. Then, there exists a subset $E_3^{'} \subseteq E_3$, $|E_3^{'}| \sim 1$ such that, for $h := \chi_{E_3^{'}}$,*

$$\left|\Lambda_{\vec{\mathbf{P}}}(f,g,h)\right| \lesssim 1. \qquad (6)$$

To construct the set $E_3^{'}$ with the desired properties, define the *maximal-square function*, the *square-maximal function* and the *square-square function* as follows. Let $I, J \subseteq \mathbb{R}$ be dyadic intervals and denote $I_{\lambda,t} = 2^{-\lambda}[t,t+1]$. Next, let $\Phi^i_{R_{\vec{\lambda},\vec{t_i}}} = \Phi^i_{I_{\lambda^{'},t_i^{'}}} \otimes \Phi^i_{J_{\lambda^{''},t_i^{''}}}$. For $(x^{'},x^{''}) \in \mathbb{R}^2$ define

$$\mathrm{MS}(f)(x^{'},x^{''}) := \sup_I \frac{1}{|I|^{1/2}} \left(\sum_{J:R=I\times J\in\vec{\mathbf{P}}} \sup_{\vec{\lambda},\vec{t_1}} \frac{|\langle f, \Phi^1_{R_{\vec{\lambda},\vec{t_1}}}\rangle|^2}{|J|} \chi_J(x^{''}) \right)^{1/2} \chi_I(x^{'}),$$

$$\mathrm{SM}(g)(x^{'},x^{''}) := \left(\sum_I \frac{\sup_{J:R=I\times J\in\vec{\mathbf{P}}} \sup_{\vec{\lambda},\vec{t_2}} \frac{|\langle g, \Phi^2_{R_{\vec{\lambda},\vec{t_2}}}\rangle|^2}{|J|} \chi_J(x^{''})}{|I|} \chi_I(x^{'}) \right)^{1/2},$$

and

$$\mathrm{SS}(h)(x^{'},x^{''}) := \left(\sum_{R\in\vec{\mathbf{P}}} \sup_{\vec{\lambda},\vec{t_3}} \frac{|\langle h, \Phi^3_{R_{\vec{\lambda},\vec{t_3}}}\rangle|^2}{|R|} \chi_R(x^{'},x^{''}) \right)^{1/2}.$$

Finally, recall the bi-parameter Hardy-Littlewood maximal function

$$MM(F)(x^{'},x^{''}) := \sup_{(x^{'},x^{''})\in I\times J} \frac{1}{|I||J|} \int_{I\times J} |F(y^{'},y^{''})| dy^{'} dy^{''}.$$

It can be shown that all four of these operators MS, SM, SS and MM are bounded on $L^p(\mathbb{R}^2)$, $1 < p < \infty$. Then set

$$\begin{aligned}\Omega_0 &= \{x \in \mathbb{R}^2 : MS(f)(x) > C\} \cup \{x \in \mathbb{R}^2 : SM(g)(x) > C\} \\ & \quad \cup \{x \in \mathbb{R}^2 : MM(f)(x) > C\} \cup \{x \in \mathbb{R}^2 : MM(g)(x) > C\}. \quad (7)\end{aligned}$$

Also, define

$$\Omega = \{x \in \mathbb{R}^2 : MM(\chi_{\Omega_0})(x) > \tfrac{1}{100}\}, \quad (8)$$

and

$$\tilde{\Omega} = \{x \in \mathbb{R}^2 : MM(\chi_\Omega)(x) > \tfrac{1}{2}\}.$$

Fix $C > 0$ large enough such that $|\tilde{\Omega}| < \frac{1}{2}$. Then define $E_3' := E_3 \cap \tilde{\Omega}^c$ and observe that $|E_3'| \sim 1$ as desired.

14.3 Proof of Proposition 3

Recall that the goal is to prove $|\Lambda_{\vec{\mathbf{P}}}(f, g, h)| \lesssim 1$. It is equivalent to show that

$$\sum_{\vec{P} \in \vec{\mathbf{P}}} \frac{1}{|I_{\vec{P}}|^{1/2}} |\langle f, \Phi_{\vec{P}_1}\rangle||\langle g, \Phi_{\vec{P}_2}\rangle||\langle h, \Phi_{\vec{P}_3}\rangle| \lesssim 1. \quad (9)$$

Using the definition of Ω given in (8), split the preceding sum as

$$\sum_{I_{\vec{P}} \cap \Omega^c \neq \varnothing} + \sum_{I_{\vec{P}} \cap \Omega^c = \varnothing} =: I + II. \quad (10)$$

14.3.1 Estimates for term I

Since $I_{\vec{P}} \cap \Omega^c \neq \varnothing$, it follows that $|I_{\vec{P}} \cap \Omega_0^c| > \frac{99}{100}|I_{\vec{P}}|$.

The next step in estimating term I is to define three decomposition procedures, one for each function f, g and h. First, define a sequence of sets

$$\Omega_1 = \left\{x \in \mathbb{R}^2 : \mathrm{MS}(f)(x) > \tfrac{C}{2^1}\right\}, \mathbf{T_1} = \{\vec{P} \in \vec{\mathbf{P}} : |I_{\vec{P}} \cap \Omega_1| > \tfrac{1}{100}|I_{\vec{P}}|\},$$

$$\Omega_2 = \left\{\mathrm{MS}(f)(x) > \tfrac{C}{2^2}\right\}, \mathbf{T_2} = \{\vec{P} \in \vec{\mathbf{P}} \setminus \mathbf{T_1} : |I_{\vec{P}} \cap \Omega_2| > \tfrac{1}{100}|I_{\vec{P}}|\},$$

and so on. The constant $C > 0$ is the same constant appearing in the definition of E_3'. Since there are finitely many tiles $\vec{P} \in \vec{\mathbf{P}}$ this algorithm stops eventually, producing sets $\{\Omega_n\}$ and $\{\mathbf{T_n}\}$ such that $\vec{\mathbf{P}} = \bigcup_n \mathbf{T_n}$.

Simultaneously, define another sequence of sets $\{\Omega_n^{'}\}$ and $\{\mathbf{T_n^{'}}\}$ by replacing $MS(f)(x)$ with $SM(g)(x)$ above. Then $\vec{\mathbf{P}} = \bigcup_n \mathbf{T_n^{'}}$ as well. Finally it is left to find the appropriate decomposition for the function h. First construct the analogue of the set Ω_0 as in (7) for h as follows. Pick $N > 0$ a big enough integer so that for every $\vec{P} \in \vec{\mathbf{P}}$ we have $|I_{\vec{P}} \cap \Omega_{-N}^{''}| > \frac{99}{100}|I_{\vec{P}}|$, where $\Omega_{-N}^{''} = \{x \in \mathbb{R}^2 : SS(h)(x) > C2^N\}$. Then mimicking the previous decompositions, define

$$\Omega_{-N+1}^{''} = \left\{x \in \mathbb{R}^2 : SS(h)(x) > \tfrac{C2^N}{2^1}\right\}$$

and

$$\mathbf{T_{-N+1}^{''}} = \{\vec{P} \in \vec{\mathbf{P}} : |I_{\vec{P}} \cap \Omega_{-N+1}^{''}| > \tfrac{1}{100}|I_{\vec{P}}|\},$$

and so on, constructing sets $\{\Omega_n^{''}\}$ and $\{\mathbf{T_n^{''}}\}$ such that $\vec{\mathbf{P}} = \bigcup_n \mathbf{T_n^{''}}$.

Write term I as

$$\sum_{n_1,n_2>0,n_3>-N} \sum_{\vec{P}\in\mathbf{T_{n_1,n_2,n_3}}} \frac{1}{|I_{\vec{P}}|^{3/2}} |\langle f, \Phi_{\vec{P}_1}\rangle||\langle g, \Phi_{\vec{P}_2}\rangle||\langle h, \Phi_{\vec{P}_3}\rangle||I_{\vec{P}}|, \tag{11}$$

where $\mathbf{T_{n_1,n_2,n_3}} := \mathbf{T_{n_1}} \cap \mathbf{T_{n_2}^{'}} \cap \mathbf{T_{n_3}^{''}}$. Using the fact that $\vec{P} \in \mathbf{T_{n_1,n_2,n_3}}$ means that $\vec{P}$ has not been selected in the previous $n_1 - 1, n_2 - 1, n_3 - 1$ steps, respectively, it follows that $|I_{\vec{P}} \cap \Omega_{n_1-1}^c \cap \Omega_{n_2-1}^{'c} \cap \Omega_{n_3-1}^{''c}| > \frac{97}{100}|I_{\vec{P}}|$.

Then (11) is smaller than

$$\begin{aligned} &\lesssim \sum_{\substack{n_1,n_2>0\\ n_3>-N}} \int_{\Omega_{n_1-1}^c\cap\Omega_{n_2-1}^{'c}\cap\Omega_{n_3-1}^{''c}\cap\mathbf{T_{n_1,n_2,n_3}}} MS(f)(x)SM(g)(x)SS(h)(x)dx \\ &\lesssim \sum_{\substack{n_1,n_2>0\\ n_3>-N}} 2^{-n_1}2^{-n_2}2^{-n_3}|\Omega_{\mathbf{T_{n_1,n_2,n_3}}}|, \end{aligned} \tag{12}$$

where $\Omega_{\mathbf{T_{n_1,n_2,n_3}}} := \bigcup_{\vec{P}\in\mathbf{T_{n_1,n_2,n_3}}} I_{\vec{P}}$.

On the other hand, using $|\Omega_{\mathbf{T_{n_1,n_2,n_3}}}| \le |\Omega_{\mathbf{T_{n_i}}}|$ for each $i = 1, 2, 3$, one can deduce that $|\Omega_{\mathbf{T_{n_1,n_2,n_3}}}| \lesssim 2^{n_1 p}$, $|\Omega_{\mathbf{T_{n_1,n_2,n_3}}}| \lesssim 2^{n_2 q}$, and $|\Omega_{\mathbf{T_{n_1,n_2,n_3}}}| \lesssim 2^{n_3\alpha}$ for every $\alpha > 1$. It follows that

$$|\Omega_{\mathbf{T_{n_1,n_2,n_3}}}| \lesssim 2^{n_1 p\theta_1} 2^{n_2 q\theta_2} 2^{n_3\alpha\theta} \tag{13}$$

for any $0 \le \theta_1, \theta_2, \theta_3 < 1$ such that $\theta_1 + \theta_2 + \theta_3 = 1$.

Now split (12) into two terms, a sum over $n_3 > 0$ and a sum over $0 > n_3 > -N$. In each of these two terms apply (13) with an appropriate choice for $\theta_1, \theta_2, \theta_3$ and it follows that the sum in (12) is $O(1)$.

14.3.2 Estimates for term II

The sum in term II runs over all tiles $I_{\vec{P}}$ such that $I_{\vec{P}} \subseteq \Omega$. For every such $\vec{P}$ there exists a maximal dyadic rectangle R such that $I_{\vec{P}} \subseteq R \subseteq \Omega$. Collect all the distinct maximal rectangles into a set called $R_{\max}$. Then for $d \in \mathbb{Z}$, $d \geq 1$, let $R^d_{\max}$ denote the set of all $R \in R_{\max}$ such that $2^d R \subseteq \tilde{\Omega}$ and d is maximal with this property.

Then term II is bounded by

$$\sum_{d \geq 1} \sum_{R \in R^d_{\max}} \sum_{I_{\vec{P}} \subseteq R \cap \Omega} \frac{1}{|I_{\vec{P}}|^{1/2}} |\langle f, \Phi_{\vec{P}_1} \rangle| |\langle g, \Phi_{\vec{P}_2} \rangle| |\langle h, \Phi_{\vec{P}_3} \rangle|.$$

The estimate of term II reduces to the following claim: For every $R \in R^d_{\max}$,

$$\sum_{I_{\vec{P}} \subseteq R \cap \Omega} \frac{1}{|I_{\vec{P}}|^{1/2}} |\langle f, \Phi_{\vec{P}_1} \rangle| |\langle g, \Phi_{\vec{P}_2} \rangle| |\langle h, \Phi_{\vec{P}_3} \rangle| \lesssim 2^{-Nd} |R|. \tag{14}$$

An application of Journé's lemma gives that for every $\epsilon > 0$, $\sum_{R \in R^d_{\max}} |R| \lesssim 2^{\epsilon d} |\Omega|$. Assuming the claim holds, combining this result with (14) gives

$$\sum_{d \geq 1} \sum_{R \in R^d_{\max}} 2^{-Nd} |R| = \sum_{d \geq 1} 2^{-Nd} \sum_{R \in R^d_{\max}} |R| \lesssim \sum_{d \geq 1} 2^{-Nd} 2^{\epsilon d} \lesssim 1, \tag{15}$$

which completes the estimate of term II. The proof of the claim relies on a splitting argument, the details of which are omitted in this short summary.

As a final note, the authors remark in [1] that the estimate for term II can be replaced by a simpler argument which avoids the use of Journé's lemma. The details of this simplified argument, which also extends Theorem 2 to the multi-parameter setting, can be found in [2].

References

[1] Muscalu, C., Pipher, J., Tao, T. and Thiele, T., *Bi-parameter paraproducts.* Acta. Math., 193 (2004), pp. 269-296.

[2] Muscalu, C., Pipher, J., Tao, T. and Thiele, T., *Multi-parameter paraproducts.* Revista M. Ibero., Vol, 22, Number 3 (2006), pp. 963-976.

Katharine Ott, University of Kentucky
email: **katharine.ott@uky.edu**

15 WKB asymptotic behavior of almost all generalized eigenfunctions for one-dimensional Schrödinger operators with slowly decaying potentials

after M. Christ and A. Kiselev [1]
A summary written by Eyvindur Ari Palsson

Abstract

We consider slowly decaying potentials V where $V = V_1 + V_2$, $V_1 \in L^p(\mathbb{R})$, V_2 is bounded from above with $A = \limsup_{x\to\infty} V(x)$ and $V_2' \in L^p(\mathbb{R}$ for some $1 \le p < 2$. We then prove that solutions of the differential equation $-\frac{d^2u}{dx^2} + V(x)u = Eu$ have WKB asymptotic behavior for a.e. $E > A$. This implies that Schrödinger operators with such slowly decaying potentials have absolutely continuous spectrum on $]A, \infty[$.

15.1 Introduction

Let $D = \frac{d}{dx}$, where $x \in \mathbb{R}$. We will be studying a time-independent Schrödinger operator on the real line

$$H_V = -D^2 + V$$

where V is a real valued potential. A quantum mechanical interpretation of $H_0 = -D^2$, the free Hamiltonian, is that it describes the behavior of a free electron, while $H_V = H_0 + V$ describes one electron that interacts with an external electrical field which is described by the potential V. If V is sufficiently small then we expect the spectrum of H_V to resemble that of H_0.

Any finite measure μ can be decomposed as $\mu = \mu_{\text{ac}} + \mu_{\text{sc}} + \mu_{\text{pp}}$ where μ_{ac} is absolutely continuous with respect to the Lebesgue measure, μ_{sc} is singular with respect to the Lebesgue measure and contains no atoms and μ_{pp} is singular with respect to the Lebesgue measure and is a countable linear combination of Dirac masses. To any self-adjoint operator H on a Hilbert space $\mathcal{H}$ and any vector $\phi \in \mathcal{H}$ there is associated a spectral measure μ_ϕ that satisfies

$$\langle f(H)\phi, \phi\rangle = \int_{\mathbb{R}} f(\lambda)d\mu_\phi(\lambda)$$

for any Borel measurable, bounded function f. We say that H has absolutely continuous (ac) spectrum if there exists some $\phi \neq 0$ such that $(\mu_\phi)_{\rm ac} \neq 0$ and purely absolutely continuous spectrum if $u_\phi = (\mu_\phi)_{\rm ac}$ for all ϕ. The spectrum for the free Hamiltonian H_0 is $\mathbb{R}^+$ and it has a purely absolutely continuous spectrum.

A generalized eigenfunction is defined, in this paper, as any solution of the ordinary differential equation

$$(-D^2 + V - E)u = 0.$$

The L^2 solutions of this equation give the eigenfunctions that dictate the point spectrum. It is also well known that boundedness of generalized eigenfunctions implies positivity of the derivative of the spectral measure which then implies that the spectral measure has a nonzero absolutely continuous component. Thus it is of interest to explore the asymptotic behavior of solutions to $H_V u = Eu$ to see if H_V has an absolutely continuous spectrum like H_0.

15.2 Results

We define the space $l^p(L^q)(\mathbb{R})$ by saying that a function f belongs to it if

$$\|f\|_{l^p(L^q)(\mathbb{R})} = \left(\sum_{n=-\infty}^{\infty} \left(\int_n^{n+1} |f(x)|^q dx \right)^{p/q} \right)^{1/p} < \infty.$$

Note that $L^r(\mathbb{R})$ is contained in $l^p(L^1)(\mathbb{R})$ for $1 \leq r \leq p$.

Denote by $\tilde{u}_\pm(x, E)$ the unique solutions of $(-D^2 + V - E)u = 0$ that satisfy $\tilde{u}_\pm(0, E) = 1$ and $\frac{d\tilde{u}_\pm}{dx}(0, E) = \pm i\sqrt{E}$.

Theorem 1. *Let $1 \leq p < 2$, and let $V = V_1 + V_2$, where $V_1 \in l^p(L^1)(\mathbb{R})$, V_2 is bounded with $A = \limsup_{x\to\infty} V_2(x)$, and $V_2' \in l^p(L^1)(\mathbb{R})$. Then for almost every $E \in]A, \infty[$, there exist two solutions of $(-D^2 + V - E)u = 0$ with the WKB-type asymptotic behavior*

$$u_\pm(x, E) - \frac{E}{\sqrt{E - V_2(x)}} e^{\pm i \int_0^x (\sqrt{E - V_2(y)} - \frac{V_1(y)}{2\sqrt{E - V_2(y)}}) dy} \to 0 \text{ as } x \to \infty$$

Moreover

$$\int_a^b \log(1 + \sup_x |\tilde{u}_\pm(x, E)|) dE < \infty$$

for any $A < a < b < \infty$.

Look at a generalized eigenfunction u_E of $H_V = H_0 + V$ for some fixed $E \neq 0$ with sufficient conditions on the potential V. The idea behind the WKB approximation method is to set $u_E(x) \sim e^{i\psi(x)}$ and then either expand ψ or ψ' as power series. We thus get equations for the power series and if we make the assumption that the power series decay fast enough then the last step is to make the approximation that we can drop all but finitely many terms from the power series. When $V_1 = 0$ or $V_2 = 0$ in the above theorem we recover well known WKB approximation formulas. The method is named after Wentzel, Kramers and Brillouin.

The first half of the next theorem is a direct corollary of the previous theorem using the results mentioned in the introduction. The second half will be deduced by using a well known argument that is called the wide barriers-type argument.

Theorem 2. *Consider a half-line operator on $L^2[0,\infty[$, $-D^2+V-E$, with some self-adjoint boundary condition at the origin. Assume that $V = V_1 + V_2$ where $V_1 \in l^p(L^1)[0,\infty[$, V_2 is bounded with $A = \limsup_{x\to\infty} V_2(x)$ and $V_2' \in l^p(L^1)[0,\infty[$ for some $1 \leq p < 2$. Then $]A,\infty[$ is an essential support of the absolutely continuous spectrum of the operator $-D^2+V$. Moreover, the essential spectrum coincides with $[A_1,\infty]$, where $A_1 = \liminf_{x\to\infty} V_2(x)$, and is purely singular on $[A_1, A]$.*

These results are rather sharp in the sense that there exists $V \in L^p$, $V' \in L^r$ for any $p > 2$, $r > 2$ such that $H_V = -D^2 + V$ has purely singular spectrum on $\mathbb{R}^+$ and thus has a spectrum that does not resemble the one that H_0, the free Hamiltonian, has.

The last theorem shows that the potential can depend on the energy. This is not directly related to spectral theory but is rather a natural ODE application.

Theorem 3. *Suppose that $1 \leq p < 2$ and that $W(x,E)$ is real-valued and that*

$$\frac{\partial^j W}{\partial E^j} \in L^p(\mathbb{R})$$

uniformly in $E \in J$ for $j = 0, 1$. Suppose further that the derivatives $\frac{\partial^j W}{\partial E^j}$ for $j = 2, 3$ satisfy

$$|\partial_E^j \int_y^x W(t,E)dt| = o(|x-y|)$$

as $x, y \to \infty$ uniformly in $E \in J$. Then for almost every $E \in J$ there exist linearly independent, bounded solutions $u_\pm(x, E)$ of

$$u''(x) = W(x, E) - Eu$$

with WKB asymptotic behavior as $x \to \infty$.

15.3 Idea of proof of the main theorem

We rewrite the equation $(-D^2 + V - E)u = 0$ as a system

$$y' = \begin{pmatrix} 0 & 1 \\ V - E & 0 \end{pmatrix} y$$

with $y = \begin{pmatrix} u \\ u' \end{pmatrix}$, $E > A$. Applying some transformations a couple of times and iterating the system starting from the vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ we obtain a series representation for one of the solutions of the original equation. This series representation immediately yields our result provided that we can show summability of the following infinite series

$$\sum_{n=0}^{\infty} B_{2n}((V_1, V_2'); \ldots ; (V_1, V_2'))(x, E) \text{ and } \sum_{n=1}^{\infty} B_{2n-1}((V_1, V_2'); \ldots ; (V_1, V_2'))(x, E)$$

where B_n is a multilinear operator. To define B_n we first need to define $\xi(x, E) = \sqrt{E - V_2}$, then we define the kernels

$$S_1(x, E) = -\frac{i}{2\xi(x, E)} e^{-i \int_0^x (2\xi(t,E) - \frac{V_1(t)}{\xi(t,E)})dt}$$

$$S_2(x, E) = -\frac{1}{4\xi(x, E)^2} e^{-i \int_0^x (2\xi(t,E) - \frac{V_1(t)}{\xi(t,E)})dt}$$

and the corresponding integral operators

$$(S_1 f)(E) = \int_0^x S_1(x, E) f(x) dx \quad \text{and} \quad (S_2 f)(E) = \int_0^x S_2(x, E) f(x) dx.$$

The multilinear operators are then defined as

$$
\begin{aligned}
&B_n((f_1^1, f_1^2); (f_2^1, f_2^2); \ldots; (f_n^1, f_n^2))(x, E) \\
&\quad = \int_x^\infty \int_{t_1}^\infty \cdots \int_{t_{n-1}}^\infty \prod_{j=1}^n [\tilde{S}_1(t_j, E) f_j^1(t_j) + \tilde{S}_2(t_j, E) f_j^2(t_j)] dt_j
\end{aligned}
$$

where $\tilde{S}_i(t_j, E)$ is equal to $\bar{S}_i(t_j, E)$ if $n-j$ is even and to $S_i(t_j, E)$ otherwise. The main goal in the proof is then to get good estimates on these multilinear operators.

References

[1] M. Christ and A. Kiselev, "WKB asymptotic behavior of almost all generalized eigenfunctions for one-dimensional Schrödinger operators with slowly decaying potentials", *J. Funct. Anal.* **179** (2001), 426–447.

[2] M. Christ and A. Kiselev, "One-dimensional Schrödinger operators with slowly decaying potentials: spectra and asymptotics", lecture notes for the 2002 University of Arkansas Spring lecture series, acquired June 10, 2009 (http://www.math.berkeley.edu/˜mchrist/preprints.html).

[3] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, 4 volumes,* Academic Press, New York, 1978.

Eyvindur Ari Palsson, Cornell
email: **eap48@cornell.edu**

16 L^p estimates on the bilinear Hilbert transform

after M. Lacey and C. Thiele [3]
A summary written by Maria Carmen Reguera

Abstract

We summarize the proof of the L^p boundedness of the bilinear Hilbert transform restricted to the local L^2 case given by M. Lacey and C. Thiele in [3]

16.1 Introduction

The bilinear Hilbert transform, defined for $f, g \in \mathcal{S}$ as

$$H(f,g)(x) = p.v. \int_{\mathbb{R}} f(x+y)g(x-y)\frac{dy}{y},$$

was first considered by A.P. Calderón in the 60's when studying boundedness of Cauchy integrals along curves. He conjectured that $H : L^2 \times L^2 \longmapsto L^1$, it wasn't until the late 90's that this conjecture was proven by M. Lacey and C. Thiele, see [3], [4]. Their main theorem states,

Theorem 1. *Let H be the bilinear Hilbert transform, then H extends to a bounded operator on $L^{p_1} \times L^{p_2} \longmapsto L^p$ for*

$$\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p} \qquad 1 < p_1, p_2 \leq \infty \qquad 2/3 < p < \infty.$$

In this paper we are only concerned with the case $2 < p_1, p_2 < \infty$ and $1 < p < 2$, this is the so called local L^2 case. It is significant that boundedness of this operator is unknown for $1/2 < p < 2/3$.

16.2 Time-frequency analysis: The Model Sum

The bilinear Hilbert transform satisfies the following symmetries: it is translation invariant, dilation invariant and modulation invariant. A decomposition of this operator has to be consistent with those symmetries, it is here that the time-frequency analysis comes into play. Let us consider the trilinear form $\Gamma(f_1, f_2, f_3) := \langle H(f_1, f_2), f_3 \rangle$. Since the distribution $p.v.\frac{1}{t}$ is a combination of δ_0 and $k(t)$ where $\widehat{k}(t) = 1_{(0,\infty)}$, and the result is trivial for δ_0, we can reduce the problem to study boundedness for

$$T(f_1, f_2)(x) = \int_{\mathbb{R}} f_1(x+y) f_2(x-y) k(y) dy.$$

We decompose $k(t) = \sum_k \psi_k$, $\psi_k = 2^{-k/2}\psi(2^{-k}t)$where ψ is an L^2 normalized Schwartz function, $\widehat{\psi}$ supported on $[1/2, 2]$. Moreover, heuristically at each scale k, we have

$$f_j = \sum_{|I_s|=2^k} \langle f_j, \phi_s \rangle \phi_s,$$

where $\phi_s(x) := |I_s|^{-1/2}\phi(\frac{x-c(I_s)}{|I_s|})e^{2\pi i c(w_s)x}$, $c(I_s)$ is the center of I_s, and $c(w_s)$ is the center of w_s, w_s an interval of lenght 2^{-k}. ϕ is again a Schwartz function normalized in L^2, $\widehat{\phi}$ supported in $[-1/2, 1/2]$ and such that $\sum_{l\in\mathbb{Z}} \left|\widehat{\phi}(\xi - l)\right|^2 = C$.

Our trilinear form looks like

$$\sum_k \sum_{s_1,s_2,s_3,|I_{sj}|=2^k} \prod_{j=1}^{3} \langle f_j, \phi_{sj} \rangle \langle T(\phi_{s1}, \phi_{s2}), \phi_{s3} \rangle,$$

actually it can be reduced further, if $\widehat{\phi_{s1}}$ is supported in w_{s1}, the supports of $\widehat{\phi_{s2}}$ and $\widehat{\phi_{s3}}$ are essentially localized on $w_s + 2^{-k}$ and $w_s + c(w_s) + 2^{-k}$ respectively. Moreover since ϕ_s are Schwartz functions highly localized on I_s, we can assume that $I_{si} = I_s$ for all $i = 1, 2, 3$. These considerations allow us to diagonalize the triple sum in s_i, and the relationship between the supports in frequency will lead to the use of the so called tri-tiles.

Definition 2. *A collection S of rectangles $s = I_s \times \omega_s$ are called tiles if $|s| \leq |s'| \leq 4|s|$ for all rectangles $s, s' \in S$, moreover $I(S)$ and $\omega(S)$ are grids. A collection of tiles S are called tri-tiles if each $s = I_s \times \omega_s \in S$ is a union of three tiles $s_j = I_s \times \omega_{sj}$, $j = 1, 2, 3$ satisfying that $w_{sj} \cap w_{si} = \emptyset$ if $i \neq j$, for all $s \in S$, $\xi_1 < \xi_2 < \xi_3$ for all $\xi_j \in w_{sj}$ and $w_1(S) \cup w_2(S) \cup w_3(S)$ is a grid.*

Definition 3. *A collection of functions $\{\phi_{sj} : s \in S, 1 \leq j \leq 3\}$ are adapted to the collection of tri-tiles S if for all $s \in S$ and all j, $||\phi_{sj}||_{L^2} \leq 1$ and*

1. *if* $w_{sj} \cap w_{s'j} = \emptyset$ *then* $\langle \phi_{sj}, \phi_{s'j} \rangle = 0$,

2. $|\phi_{sj}(x)| \leq \frac{C_0}{\sqrt{|I_s|}} \left(1 + \frac{|x - c(I_s)|}{|I_s|}\right)^{-10}$.

By previous considerations, we can argue that the study of the trilinear form Γ can be reduced to the study of the following model sum,

$$\Lambda(f_1, f_2, f_3) := \sum_{s \in S} \frac{1}{\sqrt{|I_s|}} \prod_{j=1}^{3} \langle f_j, \phi_{sj} \rangle, \tag{1}$$

where S is a collection of tri-tiles.

Let S be a finite collection of tri-tiles, $f_1, f_2, f_3 \in \mathcal{S}$, with $||f_i||_{p_i} = 1$ and $1/p_1 + 1/p_2 + 1/p_3 = 1$, we want to prove

$$\Lambda(f_1, f_2, f_3) \leq ||f_1||_{L^{p_1}} ||f_2||_{L^{p_2}} ||f_3||_{L^{p_3}} \tag{2}$$

We are actually going to consider a slightly worse operator,

$$\mathcal{L}_S(f_1, f_2, f_3) := \sum_{s \in S} |I_s|^{-3/2} \prod_{i=1}^{3} \langle f_i, \phi_{si} \rangle 1_{I_s}(x). \tag{3}$$

The L^1 norm of this operator is equal to the trilinear form in (1). Using the fact that the operator $\mathcal{L}$ is dilation invariant and interpolation, it is enough to prove that for all r_1, r_2, r_3 with $|r_i - p_i| < \epsilon$ and all functions $f_1, f_2, f_3 \in \mathcal{S}$, with $||f_i||_{r_i} = 1$, there exist $t > 1$ and a set E such that $|E| \leq C$ and

$$||\mathcal{L}_S(f_1, f_2, f_3)||_{L^t(E^c)} \leq C.$$

16.3 Main Lemma

We consider the following ordering of the tiles, $s \ll s'$ if $I_s \subset I_{s'}$ and $w_{s'} \subset w_s$. We say that a collection of tri-tiles T is a tree with top t if $s \ll t$ for every $s \in T$. We also say that T is an i-tree if $w_t \subset w_{si}$ for every $s \in T$. Notice that for any i-tree T, $I_{sj} \times w_{sj} \cap I_{s'j} \times w_{s'j} = \emptyset$ for all $s \in T$, $j \neq i$. The importance of considering i-trees comes from the similarity of the trilinear form with a paraproduct, we have two terms with cancellation and one without it,

$$\sum_{s \in T} |I_s|^{-1/2} \prod_{j=1}^{3} \langle f_j, \phi_{sj} \rangle \leq \sup_{s \in T} \frac{\langle f_i, \phi_{si} \rangle}{|I_s|^{-1/2}} \prod_{j \neq i} \left(\sum_{s \in T} |\langle f_j, \phi_{sj} \rangle|^2 \right)^{1/2}. \tag{4}$$

The following lemma is the main result of the paper,

Lemma 4. *Let S be a collection of tri-tiles and let $\mathcal{L}_S$ as in (3). Then there exist $K, C > 0$, $t > 1$, $S_{k,i,j}$ collection of i-trees T for $k \geq 0$, $1 \leq i, j \leq 3$ and S_0 collection of tiles such that the following properties hold:*

$$S = S_0 \cup \bigcup_{k \geq 0} \bigcup_{i,j=1}^{3} S_{k,i,j}, \qquad \left| \bigcup_{s \in S_0} I_s \right| \leq K,$$

we have the following pointwise estimate for the $\mathcal{L}_S$ function,

$$\sum_{s \in T} \mathcal{L}_T(x) \leq 2^{-k/r + k\epsilon} \quad \forall T \, : \, t \in S_k^*, \tag{5}$$

where $\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} = \frac{1}{r}$ and S_k^* stands for the set of tops of maximal trees in S_k. Moreover, if we define $N_{k,i,j} = \sum_{t \in S_{k,i,j}^*} 1_{I_t}$ to be the counting function, then

$$\|N_{k,i,j}\|_t \leq K 2^{k/t + kC\epsilon} \tag{6}$$

Remark 5. *For the proof of the lemma we will use $t = \frac{1}{2} \min p_i - \eta$, where η is sufficiently small. The proof of this lemma immediately gives the proof of the main theorem. For that, we notice that r is essentially 1 and t is fixed and essentially $p/2$, what makes the sum in k be finite.*

Next we indicate some of the ideas in the proof, for a complete proof we refer the reader to [3]. An efficient organization of the tiles that will allow us to exploit orthogonality is the key ingredient of this proof. Let us consider the case $i \neq j$, the case $i = j$ is slighty different but follows the same philosophy. The sets $S_{k,i,j}$ are collections of i-trees that satisfy upper and lower bounds, depending on k, for

$$\Delta(T, f_j) := \left(\frac{1}{|I_t|} \sum_{s \in T} |\langle f_j, \phi_{sj} \rangle|^2 \right)^{1/2}. \tag{7}$$

This collection of trees is defined inductively in a way that would provide an extra disjointness condition among trees, we are talking about the following, for $s, s' \in S_{k,i,j}$, with some exceptions,

$$\text{If } w_{sj} \subsetneq w_{s'j}, \text{ then } I_t \cap I_{s'} = \emptyset, \tag{8}$$

where t is the top of the tree T that s belongs to. Collections with large $\Delta(T, f_j)$ will define S_0. The bound in (5) comes from the estimate for each tree (4), the above mentioned upper bounds and a John-Nirenberg type argument. The estimate for the counting function uses the lower bounds and extra orthogonality as defined in (8).

References

[1] Lacey, M., *On the bilinear Hilbert transform.* Proc. ICM, vol II (Berlin, 1998), 1998, pp. 647-656.

[2] Lacey, M. and Thiele, C., *L^p estimates for the bilinear Hilbert transform.* Proc. Nat. Acad. Sci. U.S.A. 94 (1997), pp. 33-35.

[3] Lacey, M. and Thiele, C., *L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$.* Ann. Math. 146 (1997), pp. 693-724.

[4] Lacey, M. and Thiele, C., *On Calderón's conjecture* Ann. Math. (2) 149 (1999), pp. 475-496.

[5] Thiele, C., *Wave packet analysis* CBMS Regional Conference Series in Mathematics, 105. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006.

Maria Carmen Reguera, Georgia Institute of Technology
email: **mreguera@math.gatech.edu**

17 On certain elementary trilinear operators

after Michael Christ [1]
A summary written by Prabath Silva

Abstract

Here we summarize [1], which gives both positive and negative results on the boundedness of a trilinear operator.

17.1 Introduction

Here we consider the family of trilinear operators

$$T(f_1, f_2, f_3)(x) = \int_{|t|\leq 1} \prod_{j=1}^{3} f_j(S_j(x,t))dt \tag{1}$$

where f_j are locally integrable functions on $\mathbb{R}^d$ and $S_j : \mathbb{R}^{d+d} \to \mathbb{R}^d$ are surjective linear mappings.

We exclude degenerated cases by assuming the mappings $(x,t) \mapsto (S_i(x,t), S_j(x,t))$ are invertible for $i, j = 0, 1, 2, 3; i \neq j$ where $S_0(x,t) = x$.

There is a trivial range for T coming from Fubini's Theorem. We have T maps $L^1 \otimes L^1 \otimes L^\infty$ to L^1 and $L^1 \otimes L^\infty \otimes L^\infty$ to L^∞. By using multilinear interpolation we get that T maps $L^p \otimes L^p \otimes L^p$ to L^q where $3/2 \leq p \leq 3$ and $1/q + 1 = 1/p$. When we consider the situation $1 < p < 3/2$ the interconnections between $\{S_j\}_{j=1,2,3}$ come in to play.

Definition 1. *Suppose $\{S_j\}$ is non degenerate. The collection of linear maps $S_j : j = 0, 1, 2, 3$ is called rationally commensurate if there exist linear automorphisms h_j of $\mathbb{R}^d$ such that the vector subspace of endomorphisms of $\mathbb{R}^{2d}$ generated by the $\mathbb{Q}$ linear combinations of $\{h_j \circ S_j : 0 \leq j \leq 3\}$ has dimension 2 over $\mathbb{Q}$.*

In the one dimensional case, T can be reduced to the canonical forms $T(f,g,h)(x) = \int f(x+t)g(x-t)h(t)dt$ and $T(f,g,h)(x) = \int f(x+t)g(x-t)h(x-\theta t)dt$. The first form is rational commensurate, and the second form is rational commensurate if and only if θ is rational.

17.2 Main Theorem

Theorem 2. *Suppose that $\{S_j\}$ is nondegenerate.*
If $\{S_j\}$ is rationally commensurate, then there exist $p < 3/2$ such that the trilinear operator

$$\int_{\mathbb{R}^d} \prod_{j=1}^{3} f_j(S_j(x,t))dt \tag{2}$$

maps $L^p \otimes L^p \otimes L^p$ boundedly to L^q, where $1 + q^{-1} = 3p^{-1}$

If $\{S_j\}$ is not rationally commensurate, then for any $p < 3/2$, there exist nonnegative functions $f_j \in L^p$ and a set $E \subset \mathbb{R}^d$ of positive measure such that $T(f_1, f_2, f_3)(x) = +\infty$ for all $x \in E$.

Given $p < 3/2$, there exist rationally commensurate $\{S_j\}$, non negative functions $f_j \in L^p$,and a set E as above, such that $T(f_1, f_2, f_3)(x) = +\infty$ for all $x \in E$.

For any $\{S_j\}$ there exist $p > 1$, functions $f_j \in L^p$ and a set E as above, such that $T(f_1, f_2, f_3)(x) = +\infty$ for all $x \in E$.

We prove the boundedness of a larger operator (2) than T. From this the boundedness of T as a map from $L^p \otimes L^p \otimes L^p$ to L^q is immediate.When we get the boundedness for some $1 < p_0 < 3/2$ then by interpolation we get boundedness for $p_0 \leq p \leq 3/2$, but getting the smaller such p_0 for a given $\{S_j\}_{j=1,2,3}$ is related to getting better sums difference estimates [5] as seen in the discrete analogues section below. The last two clams in the theorem give that for the rational commensurate case p cannot get below 1 and it can be arbitrarily close to 3/2.

In this summery we consider the one dimension case and it can be extend to higher dimensions [1].

17.3 Irrational case

In this case $T(f,g,h)(x) = \int_{|t|\leq 1} f(x+t)g(x-t)h(x-\theta t)dt$ where θ is irrational. Since θ is irrational we have [3] rational approximations such that $|\frac{1-\theta}{1+\theta} - \frac{p_n}{q_n}| \leq \frac{1}{q_n^2}$. Write N for q_n and set $\delta = CN^{-2}$ for sufficiently small C. Consider the sets

$F = \bigcup_{j=1}^{p} \{x : |x - jp^{-1}| < \delta\}$,
$G = \bigcup_{k=1}^{q} \{x : |x - kq^{-1}| < \delta\}$,
$H = \bigcup_{l=1}^{p+q} \{x : |x - ly| < C_2\delta\}$, where $y = (1-\theta)/2p$

Set f, g, h to be the characteristic functions of sets E, F, G respectively. Then each function has $\|.\|_p \sim N^{-1/p}$.

When $x + t \sim jp^{-1}$ and $x - t \sim kq^{-1}$, we have $x - \theta t \sim j\frac{1-\theta}{2p} + k\frac{1+\theta}{2q}$. Taking $y = (1-\theta)/2p \sim (1+\theta)/2q$ allows us to handle pq possible values of $x - \theta t$ by a function h with only $p + q$ bumps. We have

$$T(f, g, h)(x)/\|f\|_p\|g\|_p\|h\|_p \gtrsim N^{(\frac{3}{p}-2)} \tag{3}$$

for $x \in E_N$ where E_N is a subset of a bounded set with measure bounded below by constant. Having noticed that when $p < 3/2$ the exponent is positive, take $f = \sum_\nu c_\nu f_\nu$ and similarly for g and h. By choosing c_ν appropriately get $T(f, g, h)(x) = +\infty$ for a set with positive measure.

17.4 Rational case

In this case we prove boundedness of the operator (2). Here the integral is taken over $\mathbb{R}^d$ rather than $|t| \leq 1$. This operator (2) maps $L^2 \otimes L^2 \otimes L^1$ to L^1, which follows easily by changing variables and then applying Holder's inequality. Similarly we get the other two permutations; then using interpolation we get boundedness from $L^{3/2} \otimes L^{3/2} \otimes L^{3/2}$ to L^1.

By interpolation, obtaining a restricted weak type estimate $\lambda^{q_0}|E| \leq \{|A|.|B|.|C|\}^{\frac{q_0}{p_0}}$ is enough get the boundedness for $p < 3/2$. Here A, B, C are measurable subsets of $\mathbb{R}$ and $E = \{x : T(\chi_A, \chi_B, \chi_C)(x) > \lambda\}$. Further it is enough to have

$$\lambda^{1-\delta}|E| \leq |A|^{r_1}|B|^{r_2}|C|^s$$

for some $\delta \in (0, 1)$ and $0 < r_1, r_2, s$, since we can use estimates we get from boundedness from $L^2 \otimes L^2 \otimes L^1$ to L^1.

The rest of the proof uses a continuous version of the method used in [5].

17.5 Threshold exponent

Let $1 < p < 3/2$ and $\theta = r/s$ be a rational number. Consider the following sets,

$$A = \{\sum_{n=1}^{K} a_n(rs)^{-n} + z : a_n = 0, r, 2r, \ldots, (s-1)r; |z| \leq 2(rs)^{-K}\}$$

$$B = \{\sum_{n=1}^{K} b_n(rs)^{-n} + z : b_n = 0, s, 2s, \ldots, (r-1)s; |z| \leq 2(rs)^{-K}\}$$

This choice allows us to get a bound for the measure of $C = r^{-1}A + s^{-1}B$ that is the same size as A and B. i.e. $|A| \lesssim r^{-K}, |B| \lesssim s^{-K}$and $|C| \lesssim (r+s)^K(rs)^{-K}$. This gives

$$\frac{T(\chi_A, \chi_B, \chi_C)(x)}{|A|^{1/p}|B|^{1/p}|C|^{1/p}} \gtrsim r^{(-K+2K/p)}s^{(-K+2K/p)}(r+s)^{-K/p} \tag{4}$$

for x in a measurable subset of a bounded set with measure bounded below by a constant.

Now for $p < 3/2$ choose r, s big enough and close to each other so that the right hand side of (4) $\gtrsim (\alpha)^K$, where $\alpha > 1$. Then use the method used in the irrational case to conclude the proof.

Next fix r, s and choose p close to 1 so that the right hand side of (4) $\gtrsim (\beta)^K$, where $\beta > 1$. Again use the method used in the irrational case to conclude the proof.

17.6 Discrete analogoues

In the proof of the rational case we used a continuous analogue of [5]. In this section we look at the equivalent discrete versions of the main theorem.

Theorem 3. *Suppose that* $(m, k), (m', k') \in \mathbb{Z}^2$ *and no two vectors* $(1, 0), (0, 1), (m, k), (m', k')$ *are linearly dependent. Then there exist* $p < 3/2$ *and* $K < \infty$*, depending only on* m, k, m', k'*, such that for any trosion-free Abelian group* $\mathbb{G}$*, for any finite subsets* $A, B, C \subset \mathbb{G}$*, the multiplicity function*

$$\mu(x) = |\{(a, b, c) \in A \times B \times C : c = m'a + k'b, x = ma + kb\}|$$

satisfies for every $\lambda > 0$

$$\sum_{x \in \mathbb{G}} \mu(x)^q \leq K|A|^{p/q}|B|^{p/q}|C|^{p/q}$$

where $1 + q^{-1} = 3p^{-1}$

This easily follows from first reducing to $\mathbb{Z}$ and then using Theorem 2 for functions $f_A = \sum_{x \in A}[x-1, x+1]$. We have the following corollary which extends the sums difference in [5] to integer combinations as seen below.

Corollary 4. *There exists $\delta > 0$, depending on m,k, such that for any trosion-free Abelian group $\mathbb{G}$, any positive integer N, any finite subsets $A, B, C \subset \mathbb{G}$ such that $|A|, |B|, |C| \leq N$,*

$$|\{ma + kb : (a, b, m'a + k'b) \in A \times B \times C\}| \leq N^{2-\delta}.$$

References

[1] M. Christ, *On certain elementary trilinear operators,* Mathematical Research Letters 8, 4356 (2001).

[2] L. Grafakos and N. Kalton, *Some remarks on multilinear maps and interpolation,* Journal of Functional Analysis 199 (2003) 379385.

[3] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers, fifth edition,* the Clarendon Press, Oxford University Press, Oxford, 1985.

[4] S. Janson, *On interpolation of multilinear operators, Function spaces and applications* (Lund, 1986), 290302, Lecture Notes in Math. 1302, Springer, Berlin-New York, 1988.

[5] N. Katz and T. Tao, *Bounds on arithmetic projections, and applications to the Kakeya conjecture,* Math. Res. Lett. 6 (1999), 625630.

[6] C. Kenig and E. M Stein *Multilinear estimates and fractional integration,* Math. Res. Lett. 6 (1999), 115.

Prabath Silva, Indiana University Bloomington
email: **pssilva@indiana.edu**

18 Bilinear ergodic averages

after J. Bourgain in [1] and C. Demeter in [2]
A summary written by Betsy Stovall

Abstract

We discuss the almost sure convergence of certain bilinear averages associated to ergodic dynamical systems.

18.1 Introduction

Let T be an ergodic transformation acting on the probability space $(\Omega, \mathcal{B}, \mu)$. One of the most important theorems in ergodic theory states that, given any $h \in L^1(\Omega, \mu)$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} h(T^n \omega) = \int_\Omega f \, d\mu, \quad \text{almost surely in } \Omega. \tag{1}$$

In [1], Bourgain generalizes this theorem to cover bilinear averages by proving the following.

Theorem 1. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space acted on by the ergodic transformation T. Then for any $f, g \in L^\infty(\mu)$, the averages*

$$\frac{1}{N} \sum_{n=1}^{N} f(T^n \omega) g(T^{-n} \omega) \tag{2}$$

converge almost surely.

The methods used in [1] generalize to the case where T and T^{-1} in (2) are replaced by arbitrary powers of T.

In this summary, we will sketch the proof of Theorem 1, largely following the outline in [2].

18.2 Step one: Transfer to a problem on the integers

The bulk of the proof of Theorem 1 involves working with functions defined on the integers. In this section, we explain how that reduction is carried out.

To begin, it suffices to prove the theorem in the case when g is orthogonal to the eigenfunctions of the operator $g \mapsto g \circ T^{-1}$. Under this hypothesis, the uniform Wiener-Wintner theorem implies that

$$\lim_{N\to\infty} \sup_{\lambda\in\mathbb{T}} \Big|\frac{1}{N}\sum_{n=1}^{N} g(T^n\omega)e^{in\lambda}\Big| = 0, \quad \text{almost surely in } \Omega. \tag{3}$$

The ergodic theorem also holds, and so in proving Theorem 1, it suffices to consider those $\omega \in \Omega$ such that (1) holds for certain L^1 functions h and (3) also holds. With such a nice ω fixed, we consider the $\ell^\infty(\mathbb{Z})$ functions

$$f(n) := f(T^n\omega) \qquad g(n) := g(T^n\omega).$$

We now define some notation relevant to the finite problem. Integers $M_0 < N(M_0) \ll J$ will be given. For each $x \in [-J, J]$, we are also given an integer N_x, $M_0 \le N_x \le N(M_0)$. We note that

$$\frac{1}{2N_x}\sum_{[-N_x,N_x]} g(x-m)f(x+m) = \int_{\mathbb{T}} \hat{f}(\lambda)P_x(\lambda)e^{2ix\lambda}\,d\lambda,$$

where

$$P_x(\lambda) = \frac{1}{|I_x|}\sum_{n\in I_x} g(n)e^{-in\lambda}, \quad I_x = x + [-N_x, N_x].$$

We let

$$H_{M_0}(x) := \Big|\int_{\mathbb{T}} \hat{f}(\lambda)P_x(\lambda)e^{2ix\lambda}\,d\lambda\Big|,$$

and note that by our assumption on ω, for each $\delta_0 > 0$, the set

$$E_{M_0J\delta_0} := \{x \in [-J, J] : \sup_{\lambda\in\mathbb{T}} |P_x(\lambda)| < \delta_0\}$$

satisfies

$$\frac{1}{2J}|E_{M_0J\delta_0}| \to 1 \quad \text{as } M_0 \to \infty.$$

With the notation as above, it is then possible to prove that Theorem 1 follows from the following theorem (which is stated in [2]).

Theorem 2. *Let $M_0 < N(M_0) \ll J$ be integers, and let f, g be 1-bounded functions on the interval $[-2J, 2J] \subset \mathbb{Z}$. Assume that*

$$\frac{1}{2J}|E_{M_0J\delta_0}| \ge 1 - \delta_0 \tag{4}$$

Then we have that

$$\frac{1}{2J}\sum_{x\in E_{M_0J\delta_0}} H_{M_0}(x) \le \varepsilon(\delta_0), \tag{5}$$

where $\varepsilon(\delta_0)$, which depends only on δ_0, tends to zero with δ_0.

We remark here that once the transference has been carried out, further reducing to the consideration of dyadic integers adds little extra difficulty. Because of this we will henceforth assume that all integers are equal to powers of two.

The transference having been completed, the main part of the proof begins.

18.3 Step two: Reduction to certain cardinality bounds

In this step, we reduce the proof of (5) to a bound on the cardinality of certain subsets of the integers.

To begin, we dyadically decompose the range of the function P_x, and consider instead functions $P_{x\delta}$, which are supported where $\frac{\delta}{2} \le |P_x| \le 2\delta$ and which satisfy

$$\sum_{\delta \text{ dyadic}} P_{x\delta} = P_x.$$

A nice (and important) fact is that this decomposition can be done in such a way that $|P'_{x\delta}| \lesssim |P'_x|$. In addition, because we are only interested in integers $x \in E_{M_0J\delta_0}$, the sum may be taken over $\delta \lesssim \delta_0$.

The proof of Theorem 2 then reduces to proving that for each $\delta \lesssim \delta_0$,

$$\frac{1}{J}|\{x\in[-J,J] : O_\delta(x) := |\int_{\mathbb{T}} \hat{f}(\lambda)P_{x,\delta}(\lambda)e^{2ix\lambda}\,d\lambda| > (\log\frac{1}{\delta})^{-\varepsilon}\}| \lesssim \delta^{\varepsilon/2} \tag{6}$$

(to see the relevance, sum over dyadic $\delta \lesssim \delta_0$).

Of course, reducing to fixed size in this way will complicate the support of $P_{x\delta}$, but we will gloss over this issue in the summary.

18.4 Step three: Conditional proofs of the cardinality bounds

The rest of the proof breaks into two parts. In this step, it is shown that if the bound (6) holds when we exclude certain bad integers x, then the bound

(6) holds with no restrictions. In the next section, we will describe the final step, which is to show that the bounds do hold when we exclude certain bad integers.

The ultimate goal of this step is to prove the following proposition from [2].

Proposition 3. *Suppose that there exists $\varepsilon > 0$ such that whenever $I \subset [-J, J]$ is a dyadic interval and $F_I \subset \mathbb{T}$ is a Δ-separated set ($\Delta > 0$) with cardinality $\# F_I \lesssim \delta^{-2-3\varepsilon}$, we have the bound*

$$|\{x : I_x \subset I,\ N_x^{-1} \leq \frac{\Delta}{16},\ \|p_{x,\delta}\|_{L^2(\mathbb{T}\setminus(F_I+\omega_{x0}))} \leq \delta^{\varepsilon/2}|I_x|^{-1/2}, \tag{7}$$

$$O_\delta(x) \gtrsim (\log \delta^{-1})^{-2}\}| \lesssim \delta^{C\varepsilon}|I|, \tag{8}$$

where $\omega_{0x} = [-4(N_x)^{-1}, 4(N_x)^{-1}]$ under the identification of $\mathbb{T}$ with $[-1/2, 1/2]$. Then (6) holds.

We want to show that bounding the cardinality of the sets in the proposition lets us bound the cardinality of the set in (6). Thus the main thing is to prove that the restriction

$$\|p_{x,\delta}\|_{L^2(\mathbb{T}\setminus(F_I+\omega_{x0}))} \leq \delta^{\varepsilon/2}|I_x|^{-1/2} \tag{9}$$

does not exclude too many integers x, when we take the union over all of the possible sets F_I.

Why might such a thing be true? We'll concentrate on this in the case when the assumption of Δ-separation of the set F_I is dropped. Note that without the restriction to Δ-separated F_I, the hypothesis of the proposition is stronger (we have to prove that something holds for more sets F_I), and so the proposition is weaker and our job easier.

First, we discuss why it should be possible to find sets $F_I(x)$ such that

$$\|p_{x,\delta}\|_{L^2(\mathbb{T}\setminus(F_I(x)+\omega_{x0}))} \leq \delta^{\varepsilon/2}|I_x|^{-1/2}.$$

By the construction of the $p_{x,\delta}$ and our assumption that they are supported on the intervals I_x, we have that

$$\|p_{x,\delta}\|_{L^2(\mathbb{T})} \approx \sum_{\lambda\in\mathcal{F}} \|p_{x,\delta}\|_{L^2(\lambda+\omega_{x,0})},$$

where $\mathcal{F}$ is a finite set such that $|p_{x,\delta}| \sim \delta$ on much of $\lambda + \omega_{x,0}$ for $\lambda \in \mathcal{F}$. Thus, by removing such intervals from $\mathbb{T}$, we can decrease the L^2 norm of $p_{x,\delta}$.

In order to remove the dependence of the sets F_I on the integer x, something else is needed to remove bad points. This is where a time-frequency localization comes in. In this case, the important objects are tiles $I_s \times \omega_s$ where $I_s \subset \mathbb{Z}$ and $\omega_s \subset \mathbb{T}$ are dyadic intervals which are dual to each other in the sense that $\# I_s \sim |\omega_s|^{-1}$. These tiles have a natural partial ordering, under which $s \leq s'$ if $I_s \subset I_{s'}$ and $\omega_s \supset \omega_{s'}$. Thus s and s' are comparable if and only if $I_s \cap I_{s'}$ and $\omega_s \cap \omega_{s'}$ both have nonempty interior. The proof proceeds by initially working on maximal tiles (from some finite family), and refining as needed.

18.5 Step four: Bounding a certain maximal function

Our ultimate goal is to show that the hypothesis of Proposition 3 holds, i.e., to prove the bound (7) for $I \subset [-J, J]$ a dyadic interval and $F_I \subset \mathbb{T}$ a Δ-separated set with cardinality $\# F_I \lesssim \delta^{-2-3\varepsilon}$.

We let the elements of F_I be enumerated by λ_i, $1 \leq i \leq L$. Roughly, condition (9) means that we only have to examine the contribution to $O_\delta(x)$ coming from points λ near one of the λ_i. We have already mentioned that $p'_{x,\delta}$ is pretty small, and so for λ near λ_i, $p_{x,\delta}(\lambda_i)$ approximates $p_{x,\delta}(\lambda)$. The separation of the λ_i allows us to apply Bernstein's inequality, and so it is possible to show that

$$O_\delta(x) = |\sum_{\lambda_i \in F_I} \int_{\mathbb{T}} \hat{f}_I(\lambda) p_{x,\delta}(\lambda_i) e^{i\lambda_i x} \alpha_{\delta^{-\varepsilon} N_x}(\lambda - \lambda_i) e^{ix\lambda}\, d\lambda| + \text{small errors}. \tag{10}$$

Here α is a C^∞ cutoff function, $\alpha_R(\lambda) = \alpha(R\lambda)$, and f_I is the restriction of f to the interval I.

Via a bit of work, it is possible to reduce the bound for the main term in (10) to a bound for a certain maximal function, roughly

$$Mf(x) := \sup_{\Delta^{-1} \lesssim N \lesssim |I|} |\int_{\mathbb{T}} \hat{f}(\lambda) \sum_{l=1}^{L} e^{i\lambda_l x} \tilde{p}_{x,N,\delta}(\lambda_l) \alpha_N(\lambda - \lambda_l) e^{ix\lambda}\, d\lambda|,$$

where

$$\tilde{p}_{x,N,\delta}(\lambda) := \tilde{p}_{x,N}(\lambda)\chi_{|p_{x,N}|\sim\delta}(\lambda)$$
$$\tilde{p}_{x,N}(\lambda) := \frac{1}{N}\sum_{j\in D_N(x)} g(j)e^{-ij\lambda}.$$

Here, $D_N(x)$ is the dyadic interval of length N centered at x. In particular, the main theorem follows from the bound $\|Mf_I\|_{\ell^2(I)} \lesssim_\beta \delta^{1/3-\beta}L^\beta|I|$ (this is essentially the content of a theorem stated in [2]).

The proof of the bound for the maximal function can be simplified by expressing $\tilde{p}_{x,N,\delta}(\lambda_l) = w_{x,N,l}\sigma_{x,N,l}$, where w and σ are constant in x on dyadic intervals of length N (because D_N is constant on such intervals) and satisfy

$$\|w_{x,N,l}\|_{\ell^2(l)} = O(1) \qquad \|\sigma_{x,N,l}\|_{\ell^\infty(l)} \leq \delta.$$

Furthermore, it can be shown that the sets

$$\Lambda_x := \{(w_{x,N,l})_{1\leq l\leq L} : \Delta^{-1} \lesssim N \lesssim |I|\}$$

can be covered by a small number of $\ell^2([0,L])$ balls of radius τ for any $\tau > 0$ (the actual bound is quantitative in τ), though the proof of this fact is somewhat involved.

References

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

[2] Demeter, C., *Pointwise convergence of bilinear ergodic averages, Bourgain's argument revisited*, preprint.

Betsy Stovall, UC Berkeley and UCLA
email: **betsy@math.berkeley.edu**

19 On a Conjecture of E. M. Stein on the Hilbert Transform on Vector Fields

after M. Lacey and X. Li [7]
A summary of Part I written by Po-Lam Yung

Abstract

We describe a new kind of maximal function associated to a Lipschitz vector field on the plane $\mathbb{R}^2$, and show that it is of weak type (2,2) with a quantitative bound on its norm. We also describe a related conjecture on the L^p boundedness ($p < 2$) of a smaller maximal function, and verify it for vector fields that are real analytic.

19.1 Introduction

Let v be a Lipschitz unit vector field on the plane $\mathbb{R}^2$. The goal of the paper of Lacey and Li is to study the truncated Hilbert transforms associated to this vector field, namely

$$H_{v,\varepsilon} f(x) = \text{p.v.} \int_{-\varepsilon}^{\varepsilon} f(x - tv(x)) \frac{dt}{t}$$

where ε is a positive number. The authors attributed to E.M. Stein the following conjecture on $H_{v,\varepsilon}$:

Conjecture 1. *There is an absolute constant $K > 0$ such that if $\varepsilon = (K\|v\|_{Lip})^{-1}$, then $H_{v,\varepsilon}$ is of weak-type (2,2).*

To study questions related to this conjecture, Lacey and Li introduced a new kind of Kekeya maximal function associated to a Lipschitz unit vector field by taking the maximal averages of a function over a suitable collection of rectangles as follows. Given a rectangle R in the plane, let e be a unit vector parallel to the longer side of R, and we shall think of it as a point on the unit circle. Let $L(R)$ and $W(R)$ be the lengths of the longer and shorter sides of R. The interval of uncertainty $EX(R)$ was defined to be the subarc of the unit circle centered at e and of length $W(R)/L(R)$. The set $V(R)$ was defined to be the set of points x in R for which $v(x)$ lies in the interval of uncertainty $EX(R)$. The new Kekeya maximal function will be defined by taking maximal averages over rectangles R for which $V(R)$ occupies a fixed positive fraction of R (or in other words, where R 'aligns well' with the given vector field v), and where $L(R)$ is small! . More precisely, for $0 < \delta < 1$, the *Lipschitz Kekeya maximal function* associated to v was defined by

$$M_{v,\delta} f(x) = \sup_{\substack{R \ni x,\ |V(R)| \geq \delta |R| \\ L(R) \leq (100\|v\|_{\mathrm{Lip}})^{-1}}} \frac{1}{|R|} \int_R |f(y)| dy.$$

Here $|R|$ denote the Lebesgue measure of R.

One of the main results in their paper is the following:

Theorem 2. *If v is a Lipschitz unit vector field on $\mathbb{R}^2$ and $\delta > 0$, then $M_{v,\delta}$ is of weak type (2,2), with norm bounded by $C\delta^{-\frac{1}{2}}$.*

Lacey and Li also studied in this paper a variant of the Lipschitz Kekeya maximal function that averages only over rectangles of a given range of widths. More precisely, for $0 < \delta < 1$ and $0 < w < \frac{1}{100}\|v\|_{\mathrm{Lip}}$, they defined

$$M_{v,\delta,w} f(x) = \sup_{\substack{R \ni x,\ |V(R)| \geq \delta |R| \\ w \leq W(R) \leq 2w \\ L(R) \leq (100\|v\|_{\mathrm{Lip}})^{-1}}} \frac{1}{|R|} \int_R |f(y)| dy.$$

They conjectured that

Conjecture 3. *There exists $1 < p < 2$ and some finite N such that for all Lipschitz unit vector fields v in $\mathbb{R}^2$, all $0 < \delta < 1$ and all $0 < w < \frac{1}{100}\|v\|_{Lip}$, the maximal function $M_{v,\delta,w}$ is of weak type (p,p) with norm at most $C\delta^{-N}$.*

They verified the conjecture for a class of vector fields that includes the real analytic ones, and proved in particular the following theorem:

Theorem 4. *Let v be a real analytic unit vector field on $\mathbb{R}^2$. Suppose ε_0 is sufficiently small. Then for all $0 < \delta < 1$ and all $0 < w < \varepsilon_0$, $M_{v,\delta,w}$ is of weak type (1,1) with norm bounded by $C\delta^{-1}(1 + \log \delta^{-1})$.*

This concludes the summary of the first part of the paper. In the second part of the paper, Lacey and Li went on to prove the following theorem about the truncated Hilbert transforms associated to vector fields.

Theorem 5. *Assume v is a Lipschitz unit vector field on $\mathbb{R}^2$ for which the conclusion of the Conjecture 3 holds. Then $H_{v,\varepsilon} \circ P_N$ maps L^2 to L^2 with* *norm* $\lesssim 1$ *whenever* $N > \|v\|_{Lip}^{-1}$ *and* $\varepsilon = (K\|v\|_{Lip})^{-1}$*, where* P_N *is the Littlewood-Paley projection onto frequency* $|\xi| \simeq N$*. Furthermore, if in addition* $v \in C^{1+\eta}$ *for some* $\eta > 0$*, then* $H_{v,\varepsilon}$ *maps* L^2 *to* L^2*, with norm* $\lesssim (1 + \log \|v\|_{C^{1+\eta}})^2$*, whenever* $\varepsilon = (K\|v\|_{C^{1+\eta}})^{-1}$*.*

In particular these hold for all real analytic unit vector fields v. This will be the subject of the second talk on this paper by Kovac.

19.2 Backgrounds

The Hilbert transform associated to a unit vector field is a kind of Radon transform, and these have been extensively studied by various authors. Other versions of maximal functions, where the maximal averages are taken over very thin sets or over lower dimensional submanifolds, have also been investigated in the literature. Listed below are some of the relevant earlier papers.

1. Christ, Nagel, Wainger and Stein [3] proved the boundedness of a general class of maximal and singular Radon transforms on $L^p_{\text{loc}}(\mathbb{R}^n)$, $1 < p < \infty$. Note there they were working in the smooth setting integrating over general submanifolds (not just over straight line segments), and in this general formulation some kind of curvature conditions are necessary (contrary to Conjecture 1).

2. Strömberg [9] proved a quantitative bound on L^2 for maximal functions associated to rectangles pointing in N uniformly distributed directions in $\mathbb{R}^2$, while A. Córdoba and R. Fefferman [4] proved the boundedness of an analogous maximal function for lacunary directions on $L^2(\mathbb{R}^2)$. Their methods relied on a covering argument. Nagel, Stein and Wainger [8] later extended the latter result to $L^p(\mathbb{R}^2)$, $p > 1$, using Fourier analysis. See also Katz [5], where he proved a quantitative bound for the maximal function associated to an arbitrary set of N directions in $\mathbb{R}^2$.

3. In [9], Strömberg also obtained, as an easy consequence of the result of his stated above, quantitative bounds of the maximal function

$$M_\delta f(x) = \sup_{x \in R,\ |EX(R)| \geq \delta} \frac{1}{|R|} \int_R |f(y)| dy$$

on L^2 in dimension 2. See also the work of Bourgain [2], where he obtained quantitative bounds for an analogous Nikodym maximal function in dimensions $n \geq 3$ when $2 \leq p \leq p_0(n)$ for some $2 < p_0(n) < n$. The Kekeya conjecture is the conjecture that the Nikodym maximal function maps L^p to L^p on $\mathbb{R}^n$ with norm $\lesssim \delta^{-\varepsilon}$ when $p \geq n$, $n \geq 3$, and this remains an outstanding open problem in harmonic analysis.

In view of Strömberg's result, it is quite remarkable that the Lipschitz Kekeya maximal function is of weak-type (2,2) as in Theorem 2, since in the definition of the Lipschitz Kekeya maximal function rectangles of arbitarily small eccentricity is allowed (as long as they align substantially with the given vector field and are not too long).

4. Katz [6] proved that discretized maximal functions associated to a Lipschitz vector field are of weak-type (2,2) in $\mathbb{R}^2$, with a quantitative bound (depending on the number N of uniformly distributed discrete directions used to approximate the given Lipschitz vector field) that is better than the one in [5].

5. Bourgain [1] proved, among other things, the boundedness on L^2_{loc} of the following maximal function associated to real analytic vector fields v in $\mathbb{R}^2$:
$$\mathfrak{M}_{v,a} f(x) = \sup_{0<r\leq a} \frac{1}{r} \int_0^r |f(x+tv(x))| dt.$$
Here a is any sufficiently small parameter. Compare with Theorem 4.

It should be noted that the construction of the Besicovitch set in 2 dimensions shows that the truncated Hilbert transform $H_{v,\varepsilon}$ cannot be of weak-type (2,2) for vector fields that are only C^α for some $\alpha < 1$. Furthermore, if we had $\|H_{v,1}\|_{L^2 \to L^2} \lesssim 1$ for all C^2 vector fields v with $\|v\|_{C^2} \leq 1$, then it can be deduced from this that Carleson's maximal operator is bounded on $L^2(\mathbb{R})$.

19.3 Sketch of Proofs

There have been two main ways to prove boundedness of maximal operators. One is by using covering lemmas (as in Córdoba-Fefferman [4]), the other is by using Fourier analysis (to exploit the notions of curvature using oscillatory integrals, and orthogonality using square functions, as in Christ-Nagel-Stein-Wainger [3]). In this paper, Lacey and Li took the former approach.

To prove Theorem 2, the key covering lemma is the following:

Lemma 6. *Let $\mathcal{R}$ be any finite collection of rectangles satisfying $L(R) \leq (100\|v\|_{Lip})^{-1}$ and $V(R) \geq \delta|R|$ for all $R \in \mathcal{R}$. Then there is a subcollection $\mathcal{R}'$ of $\mathcal{R}$ such that*

$$\int (\sum_{R\in\mathcal{R}'} \chi_R)^2 \lesssim \delta^{-1} \sum_{R\in\mathcal{R}'} |R| \quad \textit{and} \quad \left|\bigcup_{R\in\mathcal{R}} R\right| \lesssim \sum_{R\in\mathcal{R}'} |R|.$$

The proof of this lemma is long. Let's just say in brief that the hardest part is in choosing $\mathcal{R}'$ correctly and in estimating

$$\int (\sum_{R\in\mathcal{R}'} \chi_R)^2 = \sum_{R\in\mathcal{R}'} |R| + \sum_{\rho,R\in\mathcal{R}',\ \rho\neq R} |\rho \cap R|.$$

The latter involved a case-by-case analysis of the areas of $\rho \cap R$ that appear in the second term, and a careful estimate of their overlaps. Two maximal functions were exploited in this connection: the ordinary maximal function associated to squares (using the key observation of Strömberg [9]) and a maximal function associated to a fixed set of uniformly distributed directions.

To prove Theorem 4, the following covering lemma was used instead.

Lemma 7. *Let v be a real-analytic vector field. Let $\mathcal{R}$ be a finite collection of rectangles satisfying $L(R) \leq (100\|v\|_{Lip})^{-1}$, $V(R) \geq \delta|R|$ and $w \leq W(R) \leq 2w$ for all $R \in \mathcal{R}$. Let $s \simeq \log\delta^{-1}$. Then we can partition $\mathcal{R}$ into s subcollections such that any 2 rectangles in the same subcollection, say R_1 and R_2 where $L(R_1) \geq L(R_2)$, satisfy either*

$$L(R_2) \geq \frac{1}{2}L(R_1) \quad \textit{or} \quad L(R_2) \leq 2^{-s}L(R_1).$$

Write this decomposition as $\mathcal{R} = \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_s$. Then among each $\mathcal{R}_j$, there is a subcollection $\mathcal{R}_j'$ such that at any point, the rectangles in $\mathcal{R}_j'$ overlap at most $\lesssim \delta^{-1}$ times, and

$$\left|\bigcup_{R\in\mathcal{R}_j} R\right| \lesssim \sum_{R\in\mathcal{R}_j'} |R|.$$

In fact Bourgain [1] proved that if a vector field v in $\mathbb{R}^2$ is real analytic, then writing $\omega_x(t) = |\det[v(x), v(x+tv(x))]|$, we have the following estimate: there exists $C, c > 0$ and a small $\varepsilon_0 > 0$ such that

$$\left|\{t \in [-\varepsilon,\varepsilon] \colon \omega_x(t) < \tau\|\omega_x\|_{L^\infty[-\varepsilon,\varepsilon]}\}\right| \leq C\tau^c\varepsilon$$

for all $x \in \mathbb{R}^2$, $\tau \in (0,1)$ and $\varepsilon \in (0, \varepsilon_0)$. It was the latter geometric fact that enters into the proof of Lemma 7, and Theorem 4 also holds for vector fields v that satisfy this condition.

References

[1] Bourgain, J., *A remark on the maximal function associated to an analytic vector field*, Analysis at Urbana, Vol. I (Urbana, IL, 1986–1987), 1989, 111–132.

[2] Bourgain, J., *Besicovitch type maximal operators and applications to Fourier analysis*, Geom. Funct. Anal. 1 (1991), No. 2, 147–187.

[3] Christ, M., Nagel, A., Wainger, S. and Stein, E.M., *Singular and maximal Radon transforms: analysis and geometry*, Ann. of Math. (2) 150 (1999), No. 2, 489–577.

[4] Córdoba, A. and Fefferman, R., *On differentiation of integrals*, Proc. Nat. Acad. Sci. U.S.A. 74 (1977), 2211–2213.

[5] Katz, N., *Maximal operators over arbitrary sets of directions*, Duke Math. J. 97, (1999) No. 1, 67–79.

[6] Katz, N., *A partial result on Lipschitz differentiation*, Harmonic analysis at Mount Holyoke (South Hadley, MA, 2001), Contemp. Math., vol. 320, Amer. Math. Soc. (2003), 217–224.

[7] Lacey, M. and Li, X., *On a Conjecture of E. M. Stein on the Hilbert Transform on Vector Fields*, to appear in Mem. Amer. Math. Soc. (http://arxiv.org/abs/0704.0808v3)

[8] Nagel, A., Wainger, S. and Stein, E.M., *Differentiation in lacunary directions*, Proc. Nat. Acad. Sci. U.S.A. 75 (1978), 1060–1062.

[9] Strömberg, J-O., *Maximal Functions Associated to Rectangles with Uniformly Distributed Directions*, Ann. of Math. (2), 107 (1978) No. 3, 399–402.

Po-Lam Yung, Princeton University
email: **pyung@math.princeton.edu**