Carleson theorems and Radon type behavior

Spring/Summer School, Bonn

May 26th - May 30th 2014

Organizers:

Lillian Pierce
University of Bonn, Germany

Christoph Thiele
University of Bonn, Germany

Po Lam Yung
University of Oxford, United Kingdom

*supported by Hausdorff Center for Mathematics and Hausdorff Research Institute for Mathematics, Bonn

Contents

1 A polynomial Carleson operator along the paraboloid

after L. Pierce and P. Yung [1]
A summary written by Theresa C. Anderson

Abstract

We prove an L^p bound for a polynomial Carleson operator integrated along a parabolloid. This integration introduces Radon-type behavior which leads to new innovation in the techniques and style of proof.

1.1 Introduction

Oscillatory integrals and exponential sums are at the heart of Fourier analysis, and recent activity in these areas has been indicative of the many open questions lying at this interface between harmonic analysis and number theory. This area has long been an interest of mine, so when I discovered that Pierce and Yung had taken the classical Carleson operator, important to many different areas of mathematics, and introduced a Radon-type behavior involving integration on a paraboloid, I was eager to learn their technique.

Both the history of this problem and the technical beauty of Pierce and Yung's approach is remarkable. We can begin with the important contribution of Carleson, which says that the operator $f \to \sup_\lambda |T_\lambda f|$ is bounded on $L^2(\mathbb{T})$ where

$$T_\lambda f(x) = p.v. \int_{\mathbb{T}} e^{ix \cdot y} \frac{dy}{y}.$$

This landmark result involved in the proof of pointwise almost-everywhere convergence of Fourier series resulted in many extensions and generalizations, including L^p bounds (Hunt), more general kernels (Sjölin), another proof by Fefferman, and connections to the bilinear Hilbert transform by Lacey and Thiele.

Next Stein and others asked about the boundedness of a Carleson-type operator with a polynomial phase, that is

$$T_\lambda f(x) = \int_{\mathbb{R}^n} f(x-y) e^{iP_\lambda(y)} K(y) dy$$

where

$$P_\lambda(y) = \sum_{1\le|a|\le d} \lambda_a y^a$$

is a polynomial (this is the standard multinomial notation).

One of the highlights of work on this question was the general result and new technique of Stein and Wainger [3], presented in another summary. We recall this here:

Theorem 1. *Let $P_\lambda(y)$ be a polynomial lacking linear terms. Then*

$$\|\sup_\lambda |T_\lambda f(x)|\|_{L^p(\mathbb{R}^{n+1})} \le A_p\|f\|_{L^p(\mathbb{R}^{n+1})}$$

for $1 < p < \infty$ where the suprenum is over all coefficients λ.

There are still open questions about generalizing this conjecture.

Seemingly inspired by the result of Stein and Wainger, Pierce and Yung add Radon-type behavior to T_λ by defining

$$T_\lambda f(x,t) = \int_{\mathbb{R}^n} f(x-y, t-|y|^2)e^{iP_\lambda(y)}K(y)dy,$$

an operator which integrates f along the paraboloid $(y, |y|^2) \subset \mathbb{R}^{n+1}$ against an oscillatory factor with a polynomial phase and a Calderön-Zygmund singular kernel K, a tempered distribution agreeing with a $\mathcal{C}^1$ function for all $x \neq 0$, satisfying for $0 \le |\alpha| \le 1$:

$$|\partial_x^\alpha K(x)| \le A|x|^{-n+|\alpha|}$$

and whose Fourier transform is in L^∞. Pierce and Yung's main result is that for a certain class of *allowable polynomials*, the operator $f \to \sup_\lambda |T_\lambda f|$ is bounded on L^p for all $1 < p < \infty$.

Definition 2. *Choose a degree $d \ge 2$ and let $p_m(y)$ be a fixed polynomial on $\mathbb{R}^n$,homogeneous of degree m. Additionally assume that $p_2(y) \neq C|y|^2$ for any constant $C \neq 0$. For coefficients λ_m, let*

$$P_\lambda(y) = \sum_{m=1}^{d} \lambda_m p_m(y).$$

This P_λ is called an allowable phase polynomial.

And now here is the statement of the theorem, in the $n = 2$ case (we will work with this case for simplicity):

Theorem 3. *Let T_λ and P_λ be as above. Then for all Schwatrz functions f and all $1 < p < \infty$ we have*

$$\| \sup_{\lambda} |T_\lambda f| \|_{L^p(\mathbb{R}^{n+1})} \leq A_p \|f\|_{L^p(\mathbb{R}^{n+1})} \tag{1}$$

where the suprenum is over all sets of coefficients λ.

Through some limiting arguments, we get the above bound for the polynomial Carleson operator in all dimensions. Note that the paraboloid was chosen since it is the simplest example of a hypersurface with nonvanishing Gaussian curvature at every point.

1.2 Method of proof and innovations

A real innovation in Pierce and Yung's volume of work is the treatment of the Radon-type behavior, which goes well beyond the techniques of Stein and Wainger. Like Stein and Wainger, the operator T_λ is split into two parts, T_λ^-, where the phase e^{iP} does not oscillate enough to contribute to the bound (this occurs for small y relative to the coefficients of P), and thus can be compared to a maximal truncated Radon transform, and T_λ^+, whose phase oscillates significantly (this is where y is large). Unlike Stein and Wainger, this second part will be bounded by a maximal oscillatory Radon transform I_a^λ, whose specific boundedness properties in the theorem below is an interesting result in itself, and whose proof constitutes the bulk of the paper.

Theorem 4. *Let $P_\lambda(y)$ be an allowable polynomial and define*

$$I_a^\lambda f(x,t) = \int_{\mathbb{R}^n} f(x-t, t-|y|^2) e^{iP_\lambda(y/a)} \frac{1}{a^n} \eta\,(y/a)\,dy$$

where the η_a are a supported in B_1 and have $\mathcal{C}^1$ norm bounded by L_0. Then there exists a $\delta > 0$ such that for any $f \in L^2(\mathbb{R}^{n+1})$ and any $r \geq 1$

$$\| \sup_{r \leq \|\lambda\| < 2r, k \in \mathbb{Z}} |I_{2^k}^\lambda f(x,t)| \|_{L^2(\mathbb{R}^{n+1})} \leq A L_0 r^{-\delta} \|f\|_{L^2(\mathbb{R}^{n+1})}.$$

The $\|\lambda\|$ is the sum of the absolute values of the nonconstant coefficients. Note the key feature of this theorem is the decay in r.

What Stein and Wainger use in their main theorem (Theorem 1) is a TT^* argument to an operator similar to I_a^λ and then van der Corput estimates to the kernel of TT^*, majorizing this by a maximal function known to be bounded. This approach fails for Pierce and Yung's operator; they must instead introduce a smoother version of I called

$$J_a^\lambda f(x,t) = \int_{\mathbb{R}^n} f(x-t, t-z) e^{iP_\lambda(y/a)} \frac{1}{a^n} \eta\left(y/a\right) \frac{1}{a^2} \zeta\left(z/a^2\right) dydz$$

where $\zeta \in \mathcal{C}_c^\infty(\mathbb{R})$ with integral 1, and then introduce a square function using I and J,

$$S_r(f)(x,t) = \left(\sum_{k \in \mathbb{Z}} \left(\sup_{r \le \|\lambda\| < 2r} |(I_{2^k}^\lambda - J_{2^k}^\lambda) f(x,t)| \right)^2 \right)^{1/2}$$

Proving the boundedness of the square function with the crucial decay in r is a large part of their paper, which uses the clever analogue of Lemma 4.1 of Stein and Wainger (Lemma 5 below), and whose proof is itself intricate and a mainstay of Pierce and Yung's technique of working with polynomials.

1.3 Sketch of parts of proof

In what remains of this summary, we state this lemma below, sketch ever-so-briefly how this is used in the square function boundedness proof, and indicate how the square function bounds give Theorem 1.2 for $p = 2$. The generalization to L^p is another technique that we will not have time to cover, but involves complex interpolation and is reminiscent of the argument in Stein and Wainger.

In the lemma below, it is crucial to capture the decay in r on $B_2 \times B_1$ given by the kernels $K_\#^{\nu,\mu}$ of TT^*, except over certain small exceptional sets. The lemma below allows use to still use a stopping time argument as we have eliminated extra variables that these small sets G and F depend on.

Proposition 5. *Let $P_\lambda(y)$ be an allowable polynomial and let μ, ν be two sets of fixed coefficients such that $r \le \|\nu\|, \|\mu\| \le 2r$ for a chosen $r \ge 1$. Define the kernel $K_\#$ as*

$$K_\#(u,\tau) = \int_{\mathbb{R}} e^{iP_\nu(u+z) - iP_\mu(u)} \Psi(u,z) d\sigma$$

where Ψ is a $\mathcal{C}^1$ function supported on $B_2 \times B_1 \subset \mathbb{R}^2 \times \mathbb{R}^2$ defined implicitly in terms of u, τ, σ by

$$\tau = \frac{u_1 z_1 + u_2 z_2}{|u|}, \sigma = \frac{-u_1 z_2 + u_2 z_1}{|u|}$$

(this is a rotation of u, z). Then there exists a small constant $\delta > 0$, a small set $G^\nu \subset B_2$ with $|G^\nu| \le r^{-\delta}$ and for each $u \notin G$, a small set $F_u^\nu \subset [-1, 1]$ such that $|F_u^\nu| \le r^{-\delta}$ and

$$|K_\#(u,\tau)| \le C(r^{-\delta}\chi_{B_2}(u)\chi_{B_1}(\tau) + \chi_{G^\nu}(u)\chi_{B_1}(\tau) + \chi_{{}^cG^\nu}(u)\chi_{F_u^\nu}(\tau))\|\Psi\|_{\mathcal{C}^1}.$$

A key remark is that the choice of δ and the exceptional sets is independent of u and Ψ. The proof is discussed in the second lecture in some sample cases.

We would like to show that the square function is bounded with decay in r, that is,

Theorem 6. *Under the usual assumptions (see Them 6.2 in [1]),*

$$\|S_r(f)\|_{L^p(\mathbb{R}^{n+1})} \le AL_0 r^{-\delta}\|f\|_{L^p(\mathbb{R}^{n+1})}$$

Pierce and Yung use Proposition 5 to prove key bounds for the square function and operator I with decay in r, which lead to Propositions 7.1 and 7.4 in [1], which we will not discuss here. Together these can be combined to prove Pierce and Yung's Theorem 6.3 (our Theorem 7):

Theorem 7. *There exist $\epsilon_0, \delta_0 > 0$ such that for any $F \in L^2(\mathbb{R}^n)$, any $j, k \in \mathbb{Z}$ and any $r \ge 1$,*

$$\|\sup_{\|\lambda\|\approx r} |(I_{2^k}^\lambda - J_{2^k}^\lambda)\Delta_j F|\|_{L^2} \le r^{-\delta_0} 2^{-\epsilon_0(j-k)}\|\eta_{2^k}\|_{\mathcal{C}^1}\|F\|_{L^2},$$

where the $\Delta_j F$ comes from a Littlewood-Paley decomposition of F.

We are interested in Theorem 6 since it is the main step in proving Theorem 4. Without going into full detail, we will say that to prove Theorem 4 we will use a pointwise bound of I, with a Littlewood-Paley decomposition of f $(L_N f)$. Roughly, with all suprenums over $\|\lambda\| \approx r$ and $k \in \mathbb{Z}$, we have

$$\sup |I_{2^k}^\lambda f| \le \sup |I_{2^k}^\lambda L_N f| + \sup |I_{2^k}^\lambda (f - L_N f)|.$$

We can bound the second term with decay in r, leaving us left to bound

$$\sup |I^{\lambda}_{2^k} L_N f| \leq \sup |J^{\lambda}_{2^k} L_N f| + S_r(f).$$

Again, we can treat the first term, so that the only term left to bound with good decay in r is the square function. The general technique of comparing an operator to a close one with a known bound, and then bounding the square function has been used before in the literature. See, for example, Stein, Chapter 11 [2].

Finally, Theorem 6 is proven using Theorem 7 in a straightforward argument, connecting everything together.

References

[1] Pierce, Lillian B. and Yung, Po-Lam. *A polynomial Carleson operator along the parabolloid*. Preprint, 2014;

[2] E.M. Stein, *Harmonic analysis: Real-variable methods, orthogonality, and oscillatory integrals*. Princeton University Press, Princeton NJ, 1993.

[3] Stein and Wainger. *Oscillatory Integrals Related to Carleson's Theorem.* Math. Res. Lett. 8 789-800 (2001).

THERESA C. ANDERSON, BROWN UNIVERSITY
email: **tcanderson@math.brown.edu**

2 L^p Estimates for the Hilbert Transform along a one-variable Vector Field

after M.Bateman and C. Thiele [2]
A summary written by Cristina Benea

Abstract

Boundedness of the Hilbert transform along a non-vanishing vector field was proved in [2], under the assumptions that the vector field only depends on one variable and $\frac{3}{2} < p < \infty$. Here we present the main ideas of the proof.

2.1 Introduction

The Hilbert transform along a non-vanishing, measurable vector field $v : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by

$$H_v f(x,y) := p.v. \int_{\mathbb{R}} \frac{f((x,y) - tv(x,y))}{t} dt.$$

For vector fields depending on one variable only, it was proved in [2] that H_v is a bounded operator on L^p for $p \in (\frac{3}{2}, \infty)$:

Theorem 1. *Suppose v is a non-vanishing measurable vector field such that for all $x, y \in \mathbb{R}$*

$$v(x,y) = v(x,0),$$

and suppose $p \in (\frac{3}{2}, \infty)$. Then

$$\|H_v f\|_p \lesssim \|f\|_p.$$

Furthermore, one can assume that the vector field is of the form $v(x,y) = (1, u(x))$, where $u(x)$ is a measurable function with

$$\|u\|_\infty \leq 10^{-2}. \tag{1}$$

While this is not obvious at first, it is a consequence of the following observations:

(i) The case of a constant vector field v follows from L^p estimates for the one-dimensional Hilbert transform.

(ii) Whenever T is a regular linear transformation of the plane,

$$(H_{T\circ v\circ T^{-1}}f)\circ T = H_v(f\circ T) \quad \text{and} \qquad \|H_v\|_{p\to p} = \|H_{T\circ v\circ T^{-1}}\|_{p\to p}.$$

(iii) The class of vector fields depending on the first variable is invariant under linear transformations which preserve the vertical direction. This group of symmetries is generated by transformations of the type

$$(x,y)\to(\lambda x,\lambda y), \quad (x,y)\to(x,\lambda y), \quad (x,y)\to(x,y+\lambda x).$$

Remark 2. *Due to an observation attributed to Coifman, the case $p=2$ of Theorem 1 is equivalent to the Carleson-Hunt theorem in L^2. This means that estimates are known for $p=2$, but on the other hand one should expect a problem as complex as the boundedness of the Carleson operator.*

2.2 Reduction to estimates for a single frequency band

Because of the assumption (1),

$$H_v(\mathbf{1}-P_c)f(\xi,\eta) = H_{(1,0)}(\mathbf{1}-P_c)f(\xi,\eta)$$

where P_c is the Fourier restriction operator to the cone $\{10|\xi|\leq|\eta|\}$. Hence it suffices to estimate H_vP_c. Also, let P_k be the Fourier projection operator to the horizontal pair of bands

$$B_k = \{(\xi,\eta)\in\mathbb{R}^2 : |\eta|\in[2^k,2^{k+\frac{1}{200}})\}.$$

Noting that $H_vP_k = P_kH_v$, and using Littlewood-Paley theory, L^p estimates for H_v follow from the vector-valued estimate

$$\|\left(\sum_k |H_kf_k|^2\right)^{1/2}\|_p \lesssim \|\left(\sum_k |f_k|^2\right)^{1/2}\|_p. \tag{2}$$

Here $H_k := P_kH_vP_c = P_kH_vP_cP_k$; the boundedness of H_k was proved in [1] and an equivalent, more transparent formulation of this result is

Theorem 3. *Assume that $1<p<\infty$ and that $\hat{f}(\xi,\eta)$ is supported in a horizontal pair of strips $A<|\eta|<2A$ for some $A>0$. Then*

$$\|H_vf\|_p \lesssim \|f\|_p.$$

The case $p = 2$ of (2) follows directly from the theorem above. The general estimate can be obtained using Marcinkiewicz interpolation for l^2 vector valued functions, once *restricted type estimates* are proved. This means that for any $G, H \subset \mathbb{R}^2$ and any sequence of functions $\{f_k\}_k$ with $\sum_k |f_k|^2 \leq \mathbf{1}_H$,

$$|\langle \left(\sum_k |H_k f_k|^2 \right)^{1/2}, \mathbf{1}_G \rangle| \lesssim |H|^{\frac{1}{p}} |G|^{\frac{1}{p'}}. \tag{3}$$

In fact, one proves estimates similar to generalized restricted type, by inductively constructing the exceptional sets. This is the subject of the lemma below:

Lemma 4. *Let $G', H' \subset [-N, N]^2$ be measurable and let $\frac{3}{2} < p < \infty$.*

If $p > 2$ and $10|G'| < |H'|$, then there exists a subset $H \subset H'$ depending only on p, G', and H' with $|H| \geq \frac{|H'|}{2}$ such that (3) holds with $G = G'$ and any sequence of functions $\{f_k\}_k$ with $\sum_k |f_k|^2 \leq \mathbf{1}_H$.

If $p < 2$ and $10|H'| < |G'|$, then there exists a subset depending only on p, G', and H' with $|G| \geq \frac{|G'|}{2}$ such that (3) holds with $H = H'$ and any sequence of functions $\{f_k\}_k$ with $\sum_k |f_k|^2 \leq \mathbf{1}_H$.

Using Hölder's inequality, it turns out that (3) is a consequnce of the estimate

$$\|H_{k,G,H} f\|_2 \lesssim \left(\frac{|G|}{|H|} \right)^{\frac{1}{2} - \frac{1}{p}} \|f\|_2 \tag{4}$$

where $H_{k,G,H} f = \mathbf{1}_G H_k(\mathbf{1}_H f)$.

While we know L^p estimates for H_k, for $1 < p < \infty$, we want L^2 estimates for the "localized" version, with operator norm $\lesssim \left(\frac{|G|}{|H|} \right)^{\frac{1}{2} - \frac{1}{p}}$. This is achieved by Marcinkiewicz interpolation, using estimates from [1] as well. Thus it suffices to prove the subsequent restricted type estimate:

Theorem 5. *Let $\frac{3}{2} < p < \infty$, and $G', H' \subset \mathbb{R}^2$ as in Lemma 4. Then there are sets G, H as in Lemma 4 such that for any measurable sets $E, F \subset \mathbb{R}^2$ and each k, we have*

$$\left| \langle H_{k,G,H} \mathbf{1}_F, \mathbf{1}_E \rangle \right| \lesssim \left(\frac{|G|}{|H|} \right)^{\frac{1}{2} - \frac{1}{p}} |F|^{1/2} |E|^{1/2}.$$

2.3 Construnction of the sets G and H

For this part, one needs the notion of *shifted dyadic grids*, and associated to parallelograms R with two vertical edges, the notions of *height* $H(R)$*, shadow* $I(R)$*, set of slopes* $U(R)$. Also set

$$E(R) := \{(x,y) \in R : u(x) \in U(R)\}$$

where $u : \mathbb{R} \to \mathbb{R}$ is a measurable function, corresponding to the slope function of the initial vector field.

1) construction of the set H:

   First, we construct exceptional sets H_1 and H_2 associated to shifted dyadic grids $\mathcal{I}_1$ and $\mathcal{I}_2$ respectively:

   $$H_i := \cup\{R \in \mathcal{R}_i : |E(R) \cap G'| \geq \delta |R|\}$$

   where $\delta = C_\alpha \left(\frac{|G'|}{|H'|}\right)^{1-\alpha}$. One can prove that $4|H_i| \leq |H'|$.

   Then setting $H := H' \setminus (H_1 \cup H_2)$, it will have all the expected properties.

2) construction of the set G:

   Define

   $$G_i = \bigcup_{\substack{k\in\mathbb{Z}\\ k<0}} \{R \in \mathcal{R}_i : \frac{|E(R)|}{|R|} \geq 2^k, \frac{|H' \cap R|}{|R|} \geq C_\epsilon 2^{-k(\frac{1}{2}+\epsilon)} \left(\frac{|H'|}{|G'|}\right)^{1/2} \}$$

   for some $\epsilon > 0$ to be determined later. Again, $G = G' \setminus (G_1 \cup G_2)$.

   Proving that $|G_i| \leq \frac{|G'|}{4}$ is quite demanding, but it can be done by carefully regrouping the collections of rectangles R. This is based on a stopping-time argument.

2.4 Proof of Theorem 5

At this point, it only remains to prove Theorem 5. In doing so, one has to understand the proof of boundedness of H_k in theorem 3, and find a way of localizing these estimates to now-fixed measurable sets G and H.

Standard in time-frequency analysis, the proof of the boundedness of H_k involves approximations of the bilinear form $\langle H_k \mathbf{1}_F, \mathbf{1}_E \rangle$ by linear combinations of a bounded number of model forms

$$\sum_{s \in \mathcal{U}_k} \langle \mathbf{C}_{s,k} \mathbf{1}_F, \mathbf{1}_E \rangle$$

where $\mathcal{U}_k$ is a set of parallelograms with vertical edges and constant height. This collection will be further decomposed into subcollections $\mathcal{U}_{\delta,\sigma}$, where we have control on "densities" and "sizes". Lastly, $\mathcal{U}_{\delta,\sigma}$ is written as a disjoint union of collections $\mathcal{T}_{\delta,\sigma}$ of *trees*. The trees are the very atoms of these type of decompositions.

It is at the level of the trees that one can refine the estimates for H_k, obtaining the desired estimates for $H_{k,G,H}$ after summing up the local estimates.

References

[1] Bateman, M., *Single annulus L^p estimates for Hilbert transforms along vector fields.* Rev. Mat. Iberoam. 29 no. 3, 1021 - 1069 (2013).

[2] Bateman, M. and Thiele, C. , *L^p estimates for the Hilbert transform along a one-variable vector field.* arxiv/1109.6396 (25 pp.).

Cristina Benea, Cornell University
email: **cmb387@cornell.edu**

3 Oscillatory integrals related to Carleson's theorem

after E. M. Stein and S. Wainger [3]
A summary written by José M. Conde-Alonso

Abstract

The authors of [3] prove an (almost) generalization of the famous Carleson theorem on oscillatory integrals of the second kind, hence initiating the study of variants of the Carleson operator. We summarize their results.

3.1 Introduction

Oscillatory integrals occur naturally in the study of the convergence of Fourier series (in fact, the Fourier transform is itself an oscillatory integral). The so-called oscillatory integrals of the second kind are defined as operators

$$T_\lambda f(x) = \int_{\mathbb{R}^n} e^{i\lambda\cdot\phi(x,y)} K(x,y) f(y) dy,$$

where λ is a parameter in $\mathbb{R}^n$ and K is some kernel which may be singular in the diagonal $\Delta = \{(x,y) \in \mathbb{R}^{2n} : \ x = y\}$. For us K will be a Calderón-Zygmund convolution kernel satisfying Lipschitz smoothness conditions. An important problem related to this kind of integrals is finding estimates for T_λ in $L^2(\mathbb{R}^n)$ that are uniform in λ. In this context, the starting point of the article is the following theorem of Sjölin:

Theorem 1 (Sjölin, [2])**.** *Let*

$$T_\lambda f(x) = \int_{\mathbb{R}^n} e^{i\lambda\cdot y} K(y) f(x-y) dy.$$

Then the mapping $f \mapsto \sup_{\lambda\in\mathbb{R}^n} |T_\lambda f(x)|$ is bounded on $L^2(\mathbb{R}^n)$.

Denote a polynomial of degree $\leq d$ without constant terms by

$$P_\lambda(x) = \sum_{1\leq|\alpha|\leq d} \lambda_\alpha x^\alpha,$$

and define $|\lambda| = \sum |\lambda_\alpha|$. The main result of the article is the validity of Theorem 1 when we replace the factor $y \cdot \lambda$ in the exponential by a polynomial $P_\lambda \in \mathcal{P}_d$, where $\mathcal{P}_d$ is the class of polynomials of degree $\leq d$ **without constant and linear terms**. In particular, the result reads as follows:

Theorem 2. *Denote*

$$T_{P_\lambda} f(x) = \int_{\mathbb{R}^n} e^{iP_\lambda(y)} K(y) f(x-y) dy.$$

Then,

$$\left\| \sup_{P_\lambda \in \mathcal{P}_d} |T_{P_\lambda} f(x)| \right\|_{L^2(\mathbb{R}^n)} \leq C \|f\|_{L^2(\mathbb{R}^n)}$$

Theorem 2 is proved using a maximal estimate which is interesting in its own right. In order to state it precisely, assume φ is some fixed C^1 function supported on the unit ball and denote, for $\lambda \in \mathbb{R}^n$ and $a \in \mathbb{R}$,

$$\Phi_a^\lambda(x) = a^{-n} e^{iP_\lambda(x/a)} \varphi(x/a).$$

Then, the maximal result reads as follows:

Theorem 3. *There exists some (small) positive number δ such that*

$$\left\| \sup_{|\lambda| \geq r,\, a>0} \left| (f * \Phi_a^\lambda)(x) \right| \right\|_{L^2(\mathbb{R}^n)} \leq C r^{-\delta} \|f\|_{L^2(\mathbb{R}^n)},\ r \geq 1.$$

It is worth noting that the results above do not strictly yield a generalization of Theorem 1. The reason of this is the fact that Theorem 3 does not hold for polynomials with linear terms, because in that case the decay on $r \leq |\lambda|$ is lost. Hence, the methods in the article are insufficient to achieve the full generalization. However, the problem has been totally solved by Victor Lie (see [1]).

3.2 Technical tools and methods

Let $B = B(0,1)$ be the unit ball. The first technical step in the proof of the results is the following elementary observation, which is valid for polynomials even with linear terms:

Lemma 4. *Let P_λ be a polynomial of degree $\leq d$. Then there exists some k, $1 \leq k \leq d$, and a unit vector v, so that*

$$\left|(v \cdot \nabla)^k P(x)\right| \geq C|\lambda|, \ \forall x \in B.$$

The previous result can be used to transfer to $\mathbb{R}^n$ two results that were previously known in dimension 1. They are the following:

- **A Van der Corput-like estimate:** if $\varphi \in C^1$ and $\Omega \subset B$ is any convex set, then

$$\left|\int_\Omega e^{iP_\lambda(x)}\varphi(x)dx\right| \leq C(n,d)|\lambda|^{-\frac{1}{d}} \sup_{x\in B}\left(|\varphi(x)| + |\nabla\varphi(x)\right).$$

- **A small set estimate:**

$$|\{x \in B :\ |P(x)| \leq \epsilon\}| \leq C(n,d)\epsilon^{\frac{1}{d}}|\lambda|^{-\frac{1}{d}}.$$

Both items are proved with the same technique: first, we make an orthogonal change of variables so that the distinguished vector v given by lemma 4 lies in the direction of the first coordinate axis. Then, for the first result, we are allowed to use the classical Van der Corput one dimensional result in the direction of v. Notice that we need the estimate of lemma 4, which is the reason why we perform the change of coordinates. Finally, the n-dimensional result is deduced by just integrating over the rest of the directions. We argue similarly for the second result.

Another necessary technical tool is the operator $\mathcal{M}_\epsilon$, which is defined as follows:

$$\mathcal{M}_\epsilon f(x) := \sup_{a>0,\ |E|\leq\epsilon} (a^{-n}\chi_E(\cdot/a) * |f|)(x).$$

$\mathcal{M}_\epsilon$ satisfies the following estimate:

$$\|\mathcal{M}_\epsilon f\|_{L^2(\mathbb{R}^n)} \leq C\epsilon^{\frac{1}{2}}\|f\|_{L^2(\mathbb{R}^n)}. \tag{1}$$

3.3 Proof of main results

The proof of Theorem 2 has the following scheme. First, by the assumptions on the kernel K, one can write

$$K = \sum_{j=-\infty}^{\infty} 2^{-nj}\varphi_j(2^{-j}x),$$

for some average zero C^1 functions φ_j which are supported in $\{1/4 \leq |x| \leq 1\}$. Then, we divide the sum above (depending on λ) according to j big or small and we write

$$T_{P_\lambda} = T^-_{P_\lambda} + T^+_{P_\lambda}.$$

The kernel of the operator $T^-_{P_\lambda}$ has compact support and is estimated by standard arguments. On the other hand, the bound for $T^+_{P_\lambda}$ follows from an inequality which is a consequence of Theorem 3, namely

$$\left\| \sup_{N(\lambda)\geq r,\, a>0} \left|(f * \Phi^\lambda_a)(x)\right| \right\|_{L^2(\mathbb{R}^n)} \leq Cr^{-\delta}\|f\|_{L^2(\mathbb{R}^n)},$$

where $N(\lambda) = \sum_{k=1}^{d}(\sum_{|\alpha|=k}|\lambda_\alpha|)^{\frac{1}{k}}$.

Finally, we sketch the proof of Theorem 3. The idea is to apply a variant of the $TT*$ method. To that end, we use the Van der Corput and the small set estimates to prove

$$\left|\left(\Phi^\nu_h * \tilde{\Phi}^\mu_1\right)(x)\right| \leq C\left(r^{-2\delta}\chi_{B(0,2)}(x) + \chi_{E_\mu}(x)\right), \tag{2}$$

where $E_\mu \subset B(0,2)$ satisfies $|E_\mu| \leq r^{-4\delta}$ and

$$\tilde{\Phi}^\lambda_a(x) := \overline{\Phi}^\lambda_a(-x) = a^{-n}e^{iP_\lambda(-x/a)}\overline{\varphi}(-x/a).$$

We define

$$Tf(x) := \int_{\mathbb{R}^n} \Phi^{\lambda(x)}_{a(x)}(y)f(x-y)dy.$$

for arbitrary functions $\lambda(x)$ and $a(x)$. To prove the result, it is enough to show that $\|T\|_{L^2\to L^2} \leq Cr^{-\delta}$ with C independent of $\lambda(x)$ and $a(x)$. First, assume that $r \leq \lambda(x) \leq 2r$, and denote the associated operator by T_r. Then

$$T_rT_r^*f(x) = \int \left(\Phi^\nu_{a_1} * \tilde{\Phi}^\mu_{a_2}\right)(x-y)f(y)dy,$$

with $\nu = \lambda(x)$, $\mu = \lambda(y)$, $a_1 = a(x)$, $a_2 = a(y)$. Applying (2) yields

$$\begin{aligned}
|\langle T_rT_r^*f, g\rangle| &= \int \left(\Phi^\nu_{a_1} * \tilde{\Phi}^\mu_{a_2}\right)(x-y)f(y)dy \\
&\leq Cr^{-\delta}\int_{\mathbb{R}^n} \left(\mathcal{M}f(x)|g(x)| + \mathcal{M}g(x)|f(x)|\right)dx \\
&+ C\int_{\mathbb{R}^n} \left(\mathcal{M}_\epsilon f(x)|g(x)| + \mathcal{M}_\epsilon g(x)|f(x)|\right)dx,
\end{aligned}$$

where $\mathcal{M}$ is the standard maximal function (which is L^2 bounded). Now we apply (1) (for $\epsilon = r^{-4\delta}$) and we get

$$|\langle T_r T_r^* f, g\rangle| \leq C r^{-2\delta} \|f\|_{L^2(\mathbb{R}^n)} \|g\|_{L^2(\mathbb{R}^n)}.$$

This gives $\|T_r T_r^*\| \leq C r^{-2\delta}$. Now, since $\|T_r T_r^*\| = \|T_r\|^2$, we obtain

$$\|T_r\| \leq C r^{-\delta},$$

as desired. Finally, we want to get rid of the assumption $|\lambda| \leq 2r$. To that end, compute

$$\begin{aligned}
\left\| \sup_{|\lambda|\geq r,\, a>0} \left|(f * \Phi_a^\lambda)(x)\right| \right\|_{L^2(\mathbb{R}^n)} &\leq \left\| \sum_{j=0}^{\infty} \sup_{2^j r \leq |\lambda| \leq 2^j r,\, a>0} \left|(f * \Phi_a^\lambda)(x)\right| \right\|_{L^2(\mathbb{R}^n)} \\
&\leq \left\| \sum_{j=0}^{\infty} T_{2^j r} f \right\|_{L^2(\mathbb{R}^n)} \\
&\leq \sum_{j=0}^{\infty} \|T_{2^j r} f\|_{L^2(\mathbb{R}^n)} \\
&\leq C r^{-\delta} \|f\|_{L^2(\mathbb{R}^n)} \sum_{j=0}^{\infty} 2^{-j\delta} \\
&\leq C r^{-\delta} \|f\|_{L^2(\mathbb{R}^n)}.
\end{aligned}$$

References

[1] Lie, V. *On Stein's conjecture on the polynomial Carleson operator.* http://arxiv.org/abs/0805.1580

[2] Sjölin, P. *Convergence almost everywhere of certain singular integral operators and multiple Fourier series.* Ark. Mat. **9** (1971), 65-90.

[3] Stein, E. and Wainger, S. *Oscillatory integrals related to Carleson's theorem.* Mathematical Research Letters **8** (2001), 789-800;

José M. Conde-Alonso, ICMAT-UAM
email: **jose.conde@icmat.es**

4 L^p theory for outer measures and two themes of Lennart Carleson united (Part 2)

after Y. Do and C. Thiele [1]
A summary written by Polona Durcik

Abstract

We discuss an application of outer L^p spaces in time-frequency analysis. Using a generalized Carleson embedding theorem we reprove bounds for the bilinear Hilbert transform.

4.1 Introduction

Let $\beta = (\beta_1, \beta_2, \beta_3)$ be a unit vector in $\mathbb{R}^3$ perpendicular to $(1,1,1)$ with pairwise distinct entries. For three Schwartz functions f_1, f_2, f_3 on $\mathbb{R}$ we define

$$\Lambda_\beta(f_1, f_2, f_3) := p.v. \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} \prod_{j=1}^{3} f_j(x - \beta_j t) dx \Big) \frac{dt}{t}$$

This family of trilinear forms is dual to a family of bilinear operators called the *bilinear Hilbert transform*.[1] In [2], Lacey and Thiele proved the following bounds.

Theorem 1. *There exists a constant C such that for all Schwartz functions f_1, f_2, f_3 the form Λ_β satisfies the estimate*

$$|\Lambda_\beta(f_1, f_2, f_3)| \le C \prod_{j=1}^{3} \|f_j\|_{p_j} \tag{1}$$

whenever the exponents p_j are such that $2 < p_1, p_2, p_3 < \infty$ and $\sum_{j=1}^{3} \frac{1}{p_j} = 1$.

Their proof employs techniques from time-frequency analysis. To prove (1) they pass through a discrete model sum

$$\Lambda(f_1, f_2, f_3) = \sum_{P \in \mathbf{P}} c_P \prod_{j=1}^{3} a_P(f_j)$$

[1]To obtain an explicit pairing of a bilinear operator in $f_j, j \neq i$, with the function f_i, one translates in the x variable to make one of the components β_i vanish. Then one interchanges the order of integrals.

where the summation index runs through a set of tiles in the phase plane. The sequence $(a_P(f_j))_{P\in\mathbf{P}}$ is of the form $a_P(f_j) = \langle f_j, \phi_P\rangle$ for the L^1 normalized wave packets

$$\phi_P(x) = 2^{-k}\phi(2^{-k}x - n)e^{2\pi i 2^{-k}xl}$$

where ϕ is a suitable Schwartz function and $k, l, n \in \mathbb{Z}$ parametrize $\mathbf{P}$.

The new observation of Do and Thiele is that the bound on Λ is in this case a Hölder inequality with respect to an outer measure on $\mathbf{P}$

$$|\Lambda(f_1, f_2, f_3)| \leq C \sup_{P\in\mathbf{P}} |c_P| \prod_{j=1}^{3} \|a_P(f_j)\|_{L^p_j(\mathbf{P},\dots)}$$

where $\dots$ stands for an explicit outer measure structure. For the rest of the proof one has to establish estimates of the form

$$\|a_P(f_j)\|_{L^p_j(\mathbf{P},\dots)} \leq \|f_j\|_{p_j}$$

for each j separately, which we call *generalized Carleson embedding theorems.*

The authors of [1] develop a theory of outer L^p spaces and employ these ideas to reprove the bounds for the bilinear Hilbert transform. However, they do not pass through a discrete model form but rather work with an outer measure space on a continuum. This avoids the usual technicalities in the discretization process.

4.2 Generalized tents and Carleson embedding

Let us start by discussing a generalized Carleson embedding theorem. For the definition of outer measures, sizes, outer L^p spaces and interpolation theorems consult [1] or the summary by Y. Ou.

Consider the space $X := \mathbb{R} \times \mathbb{R} \times (0, \infty)$. Let $0 < |\alpha| \leq 1$ and $|\beta| \leq 0.9$ be two real parameters. For $(x, \xi, s) \in X$ we define a tent

$$T_{\alpha,\beta}(x, \xi, s) := \{(y, \eta, t) \in X : t \leq s, |y - x| \leq s - t, |\alpha(\eta - \xi) + \beta t^{-1}| \leq t^{-1}\}$$

Note that this is a generalization of the classical tent from Example 3, [1]. The projection of $T_{\alpha,\beta}$ onto the first and the last variable is exactly a classical tent. In addition, the generalized tent involves a frequency variable. The collection of all tents $\mathbf{E}$ generates an outer measure if we define

$$\sigma(T_{\alpha,\beta}(x, \xi, s)) = s$$

For $0 < b < 1$ and a Borel measurable function F on X we define a size S by

$$S^b(F)(T_{\alpha,\beta}(x,\xi,s)) :=$$
$$(s^{-1}\int_{T_{\alpha,\beta}(x,\xi,s)\setminus T^b(x,\xi,s)} |F(y,\eta,t)|dyd\eta dt)^{1/2} + \sup_{(y,\eta,t)\in T_{\alpha,\beta}(x,\xi,s)} |F(y,\eta,t)|$$

where $T^b(x,\xi,s)$ is another tent

$$T^b(x,\xi,s) := \{(y,\eta,t) \in X : t \le s, |y-x| \le s-t, |\eta-\xi| \le bt^{-1}\}$$

The following is a version of the Carleson embedding theorem in the setting of generalized tents. It is the main ingredient in the proof of Theorem 1.

Theorem 2. *Let α, β be as above and $0 < b \le 2^{-8}$. Let ϕ be a Schwartz function such that $\widehat{\phi}$ is supported in $(-2^{-8}b, 2^{-8}b)$ and let $2 \le p \le \infty$. For $f \in L^p(\mathbb{R})$ define the function F on X by*

$$F(y,\eta,t) := \int_{\mathbb{R}} f(x)e^{i\eta(y-x)}t^{-1}\phi(t^{-1}(y-x))dx$$

Then there is a constant $C = C(\alpha,\beta,b,\phi,p)$ such that if $p > 2$,

$$\|F\|_{L^p(X,\sigma,S^b)} \le C\|f\|_p$$

and if $p = 2$,

$$\|F\|_{L^{2,\infty}(X,\sigma,S^b)} \le C\|f\|_2$$

The proof follows by Marcinkiewicz interpolation, Proposition 3.5 [1], between the endpoints $p = 2$ and $p = \infty$. As it turns out it suffices to work with a discrete variant of the theorem, considering only tents $T(x,\xi,\eta)$ with tips (x,ξ,s) of the form

$$x = 2^{k-4}n, \xi = 2^{-k-8}bl, s = 2^k$$

for some integers k, n, l.

The endpoint $p = \infty$ is easier to estimate. We estimate the L^∞ and the L^2 portion of the size S^b separately, which is done by analyzing contributions of $\langle f, \phi_{y,t}\rangle$ for suitable wave packets $\phi_{y,t}$. In the case of the weak type $(2,2)$ bound we use an iterative procedure to select two collections of tents, one of the larger L^∞ portion of the size and one of the larger L^2 portion of the size. Then we carefully estimate terms of the form $\langle f, \phi_{y,t}\rangle$ in relation to our collections of tents and derive the required size estimates.

4.3 Boundedness of the bilinear Hilbert transform

Now we are ready to reduce the basic estimates for the bilinear Hilbert transform to the generalized Carleson embedding theorem.

Let α be a unit vector perpendicular to $(1,1,1)$ and β. To prove Theorem 1 it suffices to show that there is a constant C depending only on p_1, p_2, p_3 and ϕ such that

$$\left| \int_0^\infty \int_{\mathbb{R}} \int_{\mathbb{R}} \prod_{j=1}^{3} F_j(y, \alpha_j \eta + \beta_j t^{-1}, t) d\eta dy dt \right| \leq C \|f_j\|_{L^{p_j}(\mathbb{R})} \tag{2}$$

Here $F_j(y, \eta, t) := \int_{\mathbb{R}} f_j(x) e^{i\eta(y-x)} t^{-1} \phi(t^{-1}(y-x)) dx$ and ϕ is a real valued function such that $\widehat{\phi}$ is non negative, non vanishing at 0 and supported in $[-\epsilon, \epsilon]$ for suitably small ϵ.

To deduce (1) from (2) we transform the left hand side of (2) into

$$\int_0^\infty \int_{\mathbb{R}} g(v) t^{-2} e^{-it^{-1}v} \psi(t^{-1}v) dv dt \tag{3}$$

for $g(v) := \int_{\mathbb{R}} \prod_{j=1}^{3} f_j(u - \beta_j v) du$ and $\psi(w) := \int_{\mathbb{R}} \prod_{j=1}^{3} \phi((z - \beta_j w)) dz$. Using Plancherel one can see that (3) is a non-zero multiple of $\int_{-\infty}^{0} \widehat{g}(\zeta) d\zeta$, which turns (3) into a nontrivial linear combination of

$$g(0) = \int_{\mathbb{R}} \prod_{j=1}^{3} f_j(u)$$

and

$$p.v. \int_{\mathbb{R}} g(t) \frac{dt}{t} = p.v. \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} f_1(u - \beta_1 t) f_2(u - \beta_2 t) f_1(u - \beta_3 t) du \Big) \frac{dt}{t}$$

The bound for the former follows by Hölder's inequality, while the bound for the latter follows from (2). This is the estimate we are looking for.

To prove (2) we consider the space $X = \mathbb{R} \times \mathbb{R} \times (0, \infty)$ and the outer measure generated by the collection of all tents $T(x, \xi, s) := T_{1,0}(x, \xi, s)$ where $(x, \xi, s) \in X$. Using Proposition 3.6 from [1] we can estimate the left hand side of (2) by

$$C \|G_1 G_2 G_3\|_{L^1(x, \sigma, S)} \tag{4}$$

where $G_j(y,\eta,t) := F_j(y,\alpha_j\eta+\beta_j t^{-1},t)$. The size S is defined as

$$S(G)(T(x,\xi,\eta)) = s^{-1}\int_{T(x,\xi,\eta)} |G(y,\eta,t)|dyd\eta dt$$

Now we would like to apply the outer Hölder inequality, Proposition 3.4 [1]. This requires to define three further sizes. For $j=1,2,3$ we define

$$S_j(G)(T) := (s^{-1}\int_{T\setminus T^{(j)}} |G(y,\eta,t)|^2 dyd\eta dt)^{1/2} + \sup_{(y,\eta,t)\in T} |G(y,\eta,t)|$$

where $T^{(j)}$ is the region

$$\{(y,\eta,t)\in X : t<s, |y-x|<s-t, |\alpha_j^{-1}(\eta-\xi)+\alpha_j^{-1}\beta_j t^{-1}|\leq bt^{-1}\}$$

With some work one can show that for any $T\in\mathbf{E}$ holds

$$S(G_1G_2G_3)(T)\leq 4\prod_{k=1}^{3} S_k(G_k)(T)$$

Using the outer Hölder inequality we then bound (4) by

$$C\prod_{j=1}^{3}\|G_j\|_{L^{p_j}(X,\sigma,S_j)}$$

It remains to show that

$$\|G_j\|_{L^{p_j}(X,\sigma,S_j)}\leq C\|f_j\|_{p_j}$$

This follows from the generalized Carleson embedding after a reparametrization of the space X under the homeomorphism

$$\Phi_j : X\to X,\ (y,\eta,t)\mapsto (y,\alpha_j\eta+\beta_j t^{-1},t)$$

References

[1] Do, Y. and Thiele, C., *L^p theory for outer measures and two themes of Lennart Carleson united,* preprint, available at arXiv:1309.0945v2 [math.CA].

[2] Lacey, M. and Thiele, C., *L^p estimates on the bilinear Hilbert transform,* Proc. Nat. Acad. Sci. U.S.A. 94. no. 1, (1997), 33-35.

Polona Durcik, Universität Bonn
email: **durcik@math.uni-bonn.de**

5 Singular and Maximal Radon Transform: Analysis and Geometry - Part I

after M. Christ, A. Nagel, E. Stein, S. Wainger [2].
A summary written by Shaoming Guo

5.1 Main Objects and Examples

The targets are the following singular integral and maximal operator

$$Tf(x) := \psi(x) \int_{|t| \le \epsilon_0} f(\gamma(x,t))K(t)dt, \tag{1}$$

and

$$Mf(x) := \sup_{0<\epsilon\le\epsilon_0} \frac{1}{2\epsilon}\psi(x) \int_{|t|\le\epsilon} |f(\gamma(x,t))|dt, \tag{2}$$

where $K(t)$ is a Calderon-Zygmund kernel and γ is a C^∞ function defined in a neighbourhood of the origin in $\mathbf{R}^n \times \mathbf{R}$, taking values in $\mathbf{R}^n$, satisfying

$$\gamma(x,0) = x \tag{3}$$

and for any fixed t small,

$$\gamma(\cdot,t) \text{ is a local diffeomorphism.} \tag{4}$$

Here the funciton ψ, which a C^∞ cut-off function with support near $0 \in \mathbf{R}^n$, plays the role of localization.

Example 5.1. *On the plane $\mathbf{R}^2$, for a function $h : \mathbf{R} \to \mathbf{R}$, we define*

$$Tf(x_1,x_2) = \int_{\mathbf{R}} f(x_1 - t, x_2 - h(t))dt/t, \tag{5}$$

which is the Hilbert transform along curve $(t, h(t))$. This example appeared in the study of the singular integrals associated to the heat equation.

Example 5.2.

$$Tf(x) := \int_{\mathbf{R}} f(x_1 - t, x_2 - tx_1)dt/t, \tag{6}$$

which is related to the Hilbert transform on the Heisenberg group.

Example 5.3. *For any measurable function $u : \mathbf{R} \to \mathbf{R}$, define*

$$Hf(x) := \int_{\mathbf{R}} f(x_1 - t, x_2 - tu(x_1))dt/t, \tag{7}$$

which is the Hilbert transform along the one-variable vector field $v(x_1, x_2) = (1, u(x_1))$ and is the content of the papers that Kevin Hughes and Cristina Benea will present. (the function γ in this example is not smooth.)

Example 5.4. *Take a small positive number $\epsilon_0 \ll 1$, let $u : \mathbf{R} \to \mathbf{R}$ be a measurable function with $\|u\|_\infty \leq \epsilon_0$, let $h : \mathbf{R}^2 \to \mathbf{R}$ be a Lipschitz function with $\|\nabla h - (1,0)\|_\infty \leq \epsilon_0$. For the vector field $v(x_1, x_2) = (1, u(h(x_1, x_2)))$, the associated Hilbert transform is defined as*

$$Hf(x) := \int_{\mathbf{R}} f(x_1 - t, x_2 - t \cdot u(h(x_1, x_2)))dt/t. \tag{8}$$

Example 5.5. *The following is the maximal operator along a general planar vector field $v : \mathbf{R}^2 \to S^1$ with cut-off ϵ_0:*

$$M_{v,\epsilon_0} f(x) := \sup_{0<\epsilon\leq\epsilon_0} \frac{1}{2\epsilon} \int_{-\epsilon}^{\epsilon} |f(x + tv(x))|dt. \tag{9}$$

Example 5.6. *The Hilbert transform along a planar vector field $v : \mathbf{R}^2 \to S^1$ with cut-off ϵ_0:*

$$H_{v,\epsilon_0} f(x) := \int_{-\epsilon_0}^{\epsilon_0} f(x + tv(x))dt/t. \tag{10}$$

Most of these specific examples have been well understood, except the following two long standing conjectures (which are not the main concern of the paper we are reading):

Conjecture (Zygmund/Stein)**.** *If we assume that the vector field v is Lipschitz, then the associated maximal operator M_{v,ϵ_0} and Hilbert transform H_{v,ϵ_0} with $\epsilon_0 := \kappa/\|v\|_{Lip}$ for some universal constant $\kappa > 0$, is of weak type $(2,2)$.*

The assumption of the result that I am presenting here is in terms of the underlying geometry, instead of the (optimal) regularity of the function γ.

5.2 Main Result

Theorem 5.7 (Main Theorem)**.** *Suppose that the function γ satisfies either (C_M) or (C_g) or (C_J) (all these three are equivalent), then the singular Radon transform T defined as in (1) and the maximal operator M defined as in (2), are bounded in L^p for all $p \in (1, \infty)$.*

5.2.1 The curvature condition (C_M)

Definition 5.8. *We say that γ satisfies the curvature condition (C_M) at x_0 if there exists no C^∞ submanifold of $\mathbf{R}^n$, of positive codimension, that is invariant under γ to infinite order at x_0.*

Definition 5.9. *A submanifold $M \subset \mathbf{R}^n$ is locally invariant under γ at x_0 if there exists a neighbourhood V of $(x_0, 0)$ in $M \times \mathbf{R}$ such that $\gamma(x,t) \in M$ for every $(x,t) \in V$.*

Definition 5.10. *A submanifold M of $\mathbf{R}^n$ containing x_0 is invariant under γ to infinite order at x_0 if for all $(x,t) \in M \times \mathbf{R}$ sufficiently close to $(x_0, 0)$,*

$$\textit{distance } (\gamma(x,t), M) = O(\textit{distance } (x, x_0) + |t|)^N \textit{ as } x \to x_0 \textit{ and } t \to 0, \tag{11}$$

for every positive integer N.

Example 5.11. *Consider $\gamma(x,t) = (x_1 + t, x_2 + x_2 t)$. The line $\{(x_1, x_2) : x_2 = 0\}$ is an invariant submanifold. Hence it does not satisfy the curvature condition (C_M).*

Example 5.12. *Consider $\gamma(x,t) = (x_1 + t, x_2 + e^{-1/t^2})$. The line $\{(x_1, x_2) : x_2 = 0\}$ is not locally invariant, but it is locally variant to infinite order near $(0,0)$. This can be seen from the fact that the function e^{-1/t^2} vanishes to infinite order near $t = 0$. (This is an example that can be avoided by assuming γ to be analytic, see [1] and [3].)*

Remark 5.13. *Although the above example does not satisfy the curvature condition (C_M), the associated Hilbert transform and maximal operator are still bounded!*

Example 5.14. *Consider $\gamma(x,t) = (x_1 + t, x_2 + x_1 t)$, for which there exists no locally invariant submanifold.*

5.2.2 The curvature condition (C_g)

Theorem 5.15 (Taylor expansion; Theorem 8.5 in [2])**.** *Let γ be any C^∞ mapping from a neighbourhood of $(x_0, 0) \in \mathbf{R}^n \times \mathbf{R}$ to $\mathbf{R}^n$, satisfying $\gamma(x, 0) = x$. Then there exists a unique collection $\{X_\alpha : \alpha \in \mathbf{N}_+\}$ of C^∞ vector fields, all defined in some common neighbourhood U of x_0, such that*

$$\gamma(x,t) = \exp\Big(\sum_{0<\alpha<N} t^\alpha X_\alpha/\alpha!\Big)(x) + O(|t|^N) \tag{12}$$

for each positive integer N, for all $x \in U$, as $|t| \to 0$.

Definition 5.16. *We say that γ satisfies the curvature condtion (C_g) at x_0 if the vector fields $\{X_\alpha\}_{\alpha \in \mathbf{N}_+}$ together with all their commutators span the tangent space to $\mathbf{R}^n$ at x_0.*

Example 5.17. *Still consider $\gamma(x,t) = (x_1 + t, x_2 + h(t))$. In this case*

$$X_1 = \frac{\partial(t, h(t))}{\partial t}\bigg|_{t=0} = (1, h'(0)), \tag{13}$$

and for all $\alpha \geq 2$

$$X_\alpha = \frac{\partial^\alpha(t, h(t))}{\partial t^\alpha}\bigg|_{t=0} = (0, h^{(\alpha)}(0)). \tag{14}$$

Notice that all these vector fields $\{X_\alpha\}$ are actually constant, hence all their commutators vanish. Then the curvature condition (C_g) will be satisfied iff not all $h^{(\alpha)}(0)$ for $\alpha \geq 2$ vanish. (compare with $h(t) = e^{-1/t^2}$.)

5.2.3 The curvature condition (C_J)

Notations: set $\Gamma^1(x,t) = \gamma(x,t)$ and for $2 \leq j \leq n$ (where n is the dimension)

$$\Gamma^j(x, t^1, ..., t^j) = \gamma(\Gamma^{j-1}(x, t^1, ..., t^{j-1}), t^j). \tag{15}$$

Among all these we single out the n-th iterate

$$\Gamma(x,\tau) := \Gamma^n(x,\tau), \tag{16}$$

for $\tau = (\tau_1, \tau_2, ..., \tau_n) \in \mathbf{R}^n$. Set

$$J(x,\tau) = \det\left(\frac{\partial\Gamma(x,\tau)}{\partial(\tau_1, \tau_2, ..., \tau_n)}\right) \tag{17}$$

Definition 5.18. *We say that γ satisfies the curvature condition (C_J) at x_0 if there exists a multi-index β such that*

$$\partial_\tau^\beta J(x,\tau)\Big|_{\tau=0} \neq 0. \tag{18}$$

5.3 Smoothing Properties

Trivial example that does not satisfy the curvature condtion: we consider the case $\gamma(x,t) = (x_1 + t, x_2)$. For the associated Hilbert transform

$$Hf(x) := \int_{\mathbf{R}} f(x_1 + t, x_2)dt/t, \tag{19}$$

it is bounded by the one dimensional Hilbert transform and Fubini's theorem. But it is not difficult to see that γ does not satisfy the curvature condition. For example, any horizontal line is an invariant submanifold.

Smoothing property that the trivial example does not possess: this time instead of a singular kernel, we take $K \in C^\infty(\mathbf{R}^n \times \mathbf{R})$ and consider the operator

$$Tf(x) := \psi(x) \int f(\gamma(x,t))K(x,t)dt. \tag{20}$$

Theorem 5.19. *If γ satisfies any of the curvature conditions at x_0 then there exists $s > 0$ and a neighbourhood U of $(x_0, 0)$ such that for every $K \in C^\infty$ supported in U, the operator T defined as in* (20) *maps L^2 to H^s.*

Theorem 5.20. *If γ satisfies any of the curvature conditions at x_0 then there exists a neighbourhood U of $(x_0, 0)$ such that for each $p \in (1, \infty)$ there exists an exponent $q > p$ such that for any $K \in L^\infty(\mathbf{R}^n \times \mathbf{R})$ supported in U, the operator defined as in* (20) *maps $L^p(\mathbf{R}^n)$ to $L^q(\mathbf{R}^n)$.*

Remark 5.21. *The curvature condition and the smoothing property are if and only if.*

References

[1] J. Bourgain: A remark on the maximal function associated to an analytic vector field. Analysis at Urbana. Vol. I (Urbana, IL, 1986-1987), 111-132.

[2] M, Christ, A. Nagel, E. Stein and S. Wainger: Singular and maximal Radon transforms: analysis and geometry. Ann. of Math. (2) 150 (1999), no. 2, 489-577.

[3] E. Stein and B. Street: Multi-parameter singular Radon transforms III: Real analytic surfaces. Adv. Math. 229 (2012), no. 4, 2210-2238.

Shaoming Guo, Universität Bonn
email: **shaoming@math.uni-bonn.de**

6 Single Annulus L^p Estimates for Hilbert Transforms along Vector Fields

after Michael Bateman [2]
A summary written by Kevin Hughes

6.1 Introduction

We assume that v is a nonvanishing planar vector field $v : \mathbb{R}^2 \to \mathbb{R}^2 \setminus \{0\}$ and define the Hilbert transform along v by

$$H_v f(x) := p.v. \int \frac{f(x - tv(x))}{t} \, dt. \tag{1}$$

We will consider vector fields that depend on only the first variable; that is, $v(x_1, x_2) = v(x_1)$. The point in considering this class of vector fields is that the class of vector vields is now more symmetric, and these symmetries interact nicely with those of the Hilbert transform.

Using the dilation invariance of dt/t, we normalize the vector field so that $v(x) = (1, u(x_1))$. Thus u is the slope of v. We assume that the slope is bounded by 1; that is, $|u(x) \leq 1$ for all $x \in \mathbb{R}^2$. With this normalization, the multiplier $\mu(\xi_1, \xi_2)$ of the Hilbert transform along v becomes $sgn(\xi_1 + \xi_2 u(x_1))$ where sgn denotes the signum function. We can now see that if our function is frequency-supported in the cone given by a 45 degree angle of the x-axis, then $sgn(\xi_1 + \xi_2 u(x_1)) = sgn(\xi_1)$ which is the multiplier for the constant vector field $v(x) \equiv (1, 0)$. $H_{(1,0)}$ is essentially the familiar 1-dimensional Hilbert transform and so we have boundedness on $L^p(\mathbb{R}^2)$ for all $1 < p < \infty$. This allows us to restrict the frequency support of our function to the cone given by a 45 degree angle around the y-axis.

We are interested in estimates where the frequency is supported in an annulus (centered at the origin). Instead of an annulus, we will prove estimates for trapezoids; the combinatorics for the part of an annulus supported in a cone and trapezoids are very similar, but the trapezoids is slightly simpler. Let $\tau(W)$ be the trapezoid determined by the corners $(-1/W, 1/W)$, $(1/W, 1/W)$, $(2/W, 2/W)$ and $(-2/W, 2/W)$ for a fixed $W > 0$. This allows us to define the frequency projection $\widehat{\Pi_\tau f}(\xi) := \mathbf{1}_\tau(\xi) \cdot \widehat{f}(\xi)$ for a reasonable function $f : \mathbb{R}^2 \to \mathbb{C}$. Throughout $\widehat{f}$ is the Fourier transform of f on $\mathbb{R}^2$.

Motivated by his previous work on related maximal functions in [1, 3], Bateman builds on the ideas and machinery of Lacey–Li in [4, 6] (which are in turn related to Lacey–Thiele's proof of Carleson's theorem in [5]) to prove the following result in [2].

Theorem 1 (Bateman)**.** *For a vector field v depending on the first variable and with slope bounded by 1, we have*

$$||H_v(\Pi_\tau f)||_p \lesssim_p ||\Pi_\tau f||_p \tag{2}$$

for all $1 < p < \infty$*.*

From now on, fix $W > 0$ and assume that f is frequency-supported on trapezoid $\tau := \tau(W)$; that is, $supp(\widehat{f}) \subset \tau$.

6.2 Reducing to the model operator

Since the proof of Theorem 1 is based on Lacey–Thiele proof of Carleson's theorem in [5], we will require a good understanding of phase space. In order to avoid technicalities, we will pretend that we can sharply localize our operators and functions in phase space – this is impossible due to the *Uncertainty Principle*. Hopefully, this simplification will provide some good intuition, even though our "lemmas" may not be correct as stated. For correct statements and proofs, we refer to the original paper [2] which treats the *Schwartz tails* that arise from the Uncertainty Principle.

We proceed in several steps. First, we replace the multiplier $\mu(\xi) = sgn(\xi_1 + \xi_2 u(x_1))$ with the multiplier $\mathbf{1}_{(0,\infty)}(\xi_1 + \xi_2 u(x_1))$. This is possible since $sgn + Id = 2 \cdot \mathbf{1}_{(0,\infty)}$. Next, we dyadically decompose our new kernel; let $\psi_k(t) := \mathbf{1}_{\{2^k \leq t < 2^{k+1}\}}(t)$ so that our Hilbert transform along v is now replaced by $\sum_{k\in\mathbb{Z}} \int \check{\psi}_k(t) \cdot f(x - tv(x))\, dt =: \sum_{k\in\mathbb{Z}} H_v^k$.

For $l \in \mathbb{N}$, let $\mathcal{D}_l$ be the dyadic intervals in $[-1, 1]$ and $\mathcal{D}$ be $\cup_{l\in\mathbb{N}}\mathcal{D}_l$. If $\omega \in \mathcal{D}_l$, then ω has a left-half $Lt(\omega)$ and a right-half $Rt(\omega)$. We define $\widehat{\varphi_\omega}(\xi) := \mathbf{1}_{[1,2]}(\xi_2) \cdot \mathbf{1}_{Rt(\omega)}(\frac{\xi_1}{\xi_2})$ supported in the set $\{\xi : \xi_1/\xi_2 \in Rt(\omega),\ \xi_2 \in [1, 2]\} \subset \mathbb{R}^2$ which looks like a slanted trapezoid. In fact, a simple linear transformation sends this trapezoid to one of the form such as τ above. Now we wish to decompose $\mathbb{R}^2$ into tiles. Fix $\omega \in \mathcal{D}$, let $c(\omega)$ be the center of ω and partition $\mathbb{R}^2$ into a collection of parallelograms $\mathcal{U}_\omega$ such that each parallelogram has the following properties:

- the short side is parallel to the y-axis and has width W,

- the projection onto the x-axis of the long side has length $W/|\omega|$, and
- the long side has slope θ where $\tan\theta \approx c(\omega)$.

We refer to a parallelogram $s \in \mathcal{U}_\omega$ as a *tile* and let $\mathcal{U} := \cup_{\omega\in\mathcal{D}} \mathcal{U}_\omega$ be the collection of all possible tiles. By ω_s, we mean the associated ω such that $s \in \mathcal{U}_\omega$. Note that $[\widehat{\varphi_\omega}]^2 = \widehat{\varphi_\omega}$ (therefore $\varphi_\omega * \varphi_\omega = \varphi_\omega$) which allows us to define an L^2-normalized *wave packet* for a tile s as $\varphi_s(y) = |s|^{1/2}\varphi_\omega(y - c(s))$ where $c(s)$ is the center of the tile s. Two packets $\varphi_{s_1}, \varphi_{s_2}$ are orthogonal if $Rt(\omega_{s_1})$ and $Rt(\omega_{s_2})$ are disjoint. The importance of these decompositions is captured in the following lemma, which can be proved by a straightforward computation.

Lemma 2.

$$f * \varphi_\omega(x) = \lim_{N\to\infty} \frac{1}{4N^2} \int_{[-N,N]^2} \sum_{s\in\mathcal{U}} \langle f, \varphi_s(y + \cdot)\rangle \varphi_s(y + x)\, dy. \tag{3}$$

Applying the Hilbert transform, this leads us to define our model operator as

$$\mathbf{C}f := \sum_{s\in\mathcal{S}} \langle f, \varphi_s\rangle \phi_s \tag{4}$$

where $\phi_s := \int \check{\psi}_k(t)\varphi_s(x_1 - t, x_2 - tv(x))\, dt$ and $\mathcal{S}$ is any fixed finite subset of tiles in $\mathcal{U}$. Note that $\phi_s(x) = 0$ unless $u(x) \in Rt(\omega_s)$. Also keep in mind that our model operator depends on $\mathcal{S}$ even though we are suppressing this dependence in our notation.

With these reductions, we now want to prove the following restricted weak-type estimate for the model operator.

Theorem 3 (Bateman)**.** *Let $1 < p < \infty$. For any subsets E, F of $\mathbb{R}^2$, let E and F respectively denote their characteristic functions. Then for any finite subset $\mathcal{S} \in \mathcal{U}$ we have the bound:*

$$|\langle \mathbf{C}E, F\rangle| \lesssim |E|^{1-\frac{1}{p}}|F|^{\frac{1}{p}} \tag{5}$$

with implied constants independent of $\mathcal{S}$, E and F.

By restricted weak-type interpolation, Theorem 3 implies Theorem 1. From now on E and F are fixed subsets of $\mathbb{R}^2$.

6.3 The Tree Lemma and the Organizational Lemma

Now that we have reduced matters to our model operator $\mathbf{C}$, we wish to follow the Lacey–Thiele paradigm by organizing our collection $\mathcal{S}$ of tiles into collections of trees for which we have a favorable estimates. To do this there is a partial ordering on the tiles given by $s_1 \leq s_2$ if $s_1 \subset 10s_2$ and $\omega_{s_2} \subset \omega_{s_1}$ where $10s_2$ is the parallelogram with the same center as s_2 but dilated by a factor of 10. This tells us that if $s_1 \cap s_2 \neq \emptyset$ and $\omega_{s_2} \subset \omega_{s_1}$, then $s_1 \leq s_2$. We can now define a *tree* as a collection of tiles T such that there exists a *top* of the tree, $top(T) \in \mathcal{U}$ with $s \leq top(T)$ for all $s \in T$.

Now that we have defined our tiles and trees. We need to understand how large they can be in terms of their *density* and *size*. For any tile s write E_s for the set $\{x \in E : u(x) \in \omega_s\} = E \cap u^{-1}(\omega_s)$, and define the density of a tile as

$$dense(s) := \frac{1}{|s|} \int_{E_s} \mathbf{1}_s = \frac{|E_s \cap s|}{|s|} \,, \tag{6}$$

the upper density of a tile as

$$\overline{dense}(s) := \sup_{s' > s} dense(s') \,, \tag{7}$$

and the size of a set of tiles as

$$size(\mathcal{S}) := \sup_{right-trees T \subset \mathcal{S}} \left(\frac{1}{top(T)} \sum_{s \in T} |\langle F, \varphi_s \rangle|^2 \right)^{1/2} . \tag{8}$$

Furthermore, we extend the definition of upper density to a collection of tiles $\mathcal{S}$, e.g. a tree by letting $\overline{dense}(\mathcal{S}) = \sup_{s \in \mathcal{S}} dense(s)$ be the supremum over all densities of tiles in the collection. The following Tree Lemma bounds our operator in terms of properties of the tree.

Tree Lemma. *If T be a tree with top, $top(T)$, upper density bounded by δ and size bounded by σ, then*

$$\sum_{s \in T} |\langle F, \varphi_s \rangle \langle E, \phi_s \rangle| \lesssim \delta\sigma |top(T)|. \tag{9}$$

Note that the dependence on E and F are in the quantities δ and σ respectively.

The Tree Lemma is proved by decomposing our tiles, and consequently operator into pieces that are close to the top of the tree and pieces that are far; the main contribution comes from pieces that are close to the top. To be more precise, fix a tree T with top $top(T)$ that has slope 0 (the long side is parallel to the x-axis). There is a partition of $\mathbb{R}^2$ into rectangles $\mathcal{P}$ such that the y-axis is partitioned into intervals of width proportional to the width of the top and the x-direction projection is a dyadic interval not containing (the x-projection) of any tile in T; choose the maximal dyadic interval so that this is true. For a rectangle $P \in \mathcal{P}$, split the tree into two sets of tiles depending on whether the x-projection of the tile is larger or smaller than the x-projection of P; we refer to these as the small tiles and large tiles respectively. To handle the small tiles we use the localization of our wave packets. The large tiles are more intricate. The important points for the large tiles is that there cannot be too many of them and their supports cannot be too large. With a precise formulation of this, it is easy to show that right-trees have the right bound by exploiting orthogonality between wave packets. The left-trees are harder. To handle them, we approximate by a (partitioned) flat Hilbert transform applied to φ_s (giving us another version of ϕ_s) and use orthogonality between these to show that it is bounded.

In order to effectively use the Tree Lemma, we need to sort our tiles into trees efficiently. This is accomplished by the Organizational Lemma below. Before we state the Organizational Lemma, we briefly sketch how to sort our tiles. By a greedy algorithm, we can partition our collection of tiles $\mathcal{S}$ into a doubly-dyadic family of trees $T_{\delta,\sigma}$ (δ and σ are dyadic parameters). To do this first partition $\mathcal{S}$ into collections $\mathcal{S}_\delta := \{s \in \mathcal{S} : \overline{dense s} \in (\delta/2, \delta]\}$. Sort a fixed $\mathcal{S}_\delta$, by choosing the left-tree T with $size(T) \geq \sigma/10$, $top(T) \in T$ and $c(\omega_{top(T)})$ most clockwise (with smallest slope) and putting this into the subcollection $\mathcal{S}_{\delta,\sigma}$; repeat until there are no such left-trees left. Then another greedy algorithm (this time choosing trees by a tile s with maximal length and considering left-trees with top s) shows that $size(T) < \sigma/2$. Now repeat replacing σ with $\sigma/2$. This process terminates for both δ and σ since our collection $\mathcal{S}$ is finite.

Organizational Lemma. *By the above process, we can decompose any finite collection of tiles* $\mathcal{S}$ *into* $\cup_{\delta,\sigma}\mathcal{S}_{\delta,\sigma}$ *such that* $\mathcal{S}_{\delta,\sigma}$ *has* $\sigma/2 < size(\mathcal{S}_{\delta,\sigma}) \leq \sigma$ *and*

$\delta/2 < \overline{dense}(\mathcal{S}_{\delta,\sigma}) \leq \delta$. *Furthermore, we have the following estimates:*

$$\text{Orthogonality Estimate:} \qquad \sum_{s\in\mathcal{S}_{\delta,\sigma}} |top(T)| \lesssim \sigma^{-2}|F| \tag{10}$$

$$\text{Density Estimate:} \qquad \sum_{s\in\mathcal{S}_{\delta,\sigma}} |top(T)| \lesssim \delta^{-1}|E| \tag{11}$$

$$\text{Maximal Estimate:} \qquad \sum_{s\in\mathcal{S}_{\delta,\sigma}} |top(T)| \lesssim \delta^{-1}\sigma^{-(1+\epsilon)}|F|^{1-\epsilon}|E|^{\epsilon} \tag{12}$$

Now Theorem 3 follows from balancing the estimates of the Organization Lemma with the Tree Lemma. From now on fix $\delta, \sigma \lesssim 1$ as dyadic values.

We briefly indicate why the Density Estimate is true; the Orthogonality Estimate and Maximal Estimate are too intricate to describe here. Many of the technicalities are in the *Schwartz tails* that arise from applying a smoothed localization to our wave packets. These are handled by a *localized Bessel inequality* – see Lemma 36 in [2]. The Density Estimate follows in a few steps. First we find a parallelogram associated to each tree. These parallelograms are incomparable between trees, then we cover them by a subcollection whose tops are disjoint. Finally there is a maximal theorem to bound the parallelograms that appear. We make this precise in the following statements.

Proposition 4. *For each tree T in $\mathcal{S}_{\delta,\sigma}$, there exist a parallelogram R_T such that $top(T) \leq R$ and $dense(R) \geq \delta$. Furthermore, the R_T are pairwise incomparable under $\leq$ as T varies.*

Let $\mathcal{R}$ be the collection of R_T arising in the above proposition, and for any $R \in \mathcal{R}$, let $\mathcal{T}_R := \{T \in \mathcal{S}_{\delta,\sigma} : R_T = R\}$. Then $\mathcal{S}_{\delta,\sigma} = \cup_{R\in\mathcal{R}}\mathcal{T}_R$. The next proposition essentially says that we can refine our collection in the above proposition so that the tops of the trees are disjoint as well.

Proposition 5. *For each $R \in \mathcal{R}$, there exists a subcollection of trees $\overline{\mathcal{T}_R} \subset \mathcal{T}_R$ such that the tops are pairwise disjoint and*

$$\sum_{T\in\mathcal{T}_R} |top(T)| \lesssim \sum_{T\in\overline{\mathcal{T}_R}} |top(T)|. \tag{13}$$

Now we have $\sum_{T\in\mathcal{S}_{\delta,\sigma}} |top(T)| \lesssim \sum_{R\in\mathcal{R}} \sum_{T\in\overline{\mathcal{T}_R}} |top(T)| \lesssim \sum_{R\in\mathcal{R}} |R|$; the last inequality follows since the tops are now pairwise disjoint, but $top(T) \leq R$ for $T \in \mathcal{T}_R$, we have $\sum_{T\in\overline{\mathcal{T}_R}} |top(T)| \lesssim |R|$. The following lemma compares the last sum to the bound we want.

Lemma 6. *If $\mathcal{R}$ is a collection of pairwise incomparable parallelograms under $\leq$, each with the same width and density at least δ, then*

$$\sum_{R\in\mathcal{R}} |R| \lesssim \delta^{-1}|E|. \tag{14}$$

Using our sharply localized definition of density, we easily see that if $x \in E_{R_1} \cap E_{R_2}$, then $R_1 \leq R_2$ or $R_2 \leq R_1$. Therefore since our collection of parallelograms is incomparable, the sets $E_R = \{u^{-1}(\omega_R) \cap E\}$ are disjoint. The bound follows.

References

[1] Bateman, Michael; L^p estimates for maximal averages along one-variable vector fields in $\mathbf{R^2}$, Proc. Amer. Math. Soc., 2009, 137, 955–963, 3, PAMYAR, 10.1090/S0002-9939-08-09583-X, http://dx.doi.org/10.1090/S0002-9939-08-09583-X

[2] Bateman, Michael; Single annulus \$L\^p\$ estimates for Hilbert transforms along vector fields, ArXiv e-prints, 2011.

[3] Bateman, Michael; Maximal averages along a planar vector field depending on one variable, ArXiv e-prints, 2011.

[4] Lacey, Michael and Li, Xiaochun; On a conjecture of E. M. Stein on the Hilbert transform on vector fields, Mem. Amer. Math. Soc., 2010.

[5] Lacey, Michael and Thiele, Christoph; A proof of boundedness of the Carleson operator, Math. Res. Lett., 2000.

[6] Lacey, Michael T. and Li, Xiaochun; Maximal theorems for the directional Hilbert transform on the plane, Trans. Amer. Math. Soc., 2006.

Kevin Hughes, University of Edinburgh
email: **k.hughes@ed.ac.uk**

7 Uniform bounds for a Walsh model of the bilinear Hilbert transform

after R. Oberlin and C. Thiele [6]
A summary written by Luis Daniel López-Sánchez

Abstract

We study the L^p boundedness behaviour of Walsh analogues of bilinear Hilbert transforms at the known region of exponents and beyond. The main tool at our disposal for exponents close to 1 is a multi-frequency Calderón-Zygmund decomposition.

7.1 Introduction

In the Fourier setting, a *bilinear Hilbert transform* is a bilinear singular integral operator of the form

$$\mathrm{BHT}_b(f_1, f_2) = \text{p.v.} \int_{\mathbb{R}} f_1(x - b_1 t) f_2(x - b_2 t) \, \frac{dt}{t}.$$

Dual to the family of bilinear Hilbert transforms are the trilinear forms

$$\Lambda_\beta(f_1, f_2, f_3) = \int_{\mathbb{R}} \text{p.v.} \int_{\mathbb{R}} f_1(x - \beta_1 t) f_2(x - \beta_2 t) f_3(x - \beta_3 t) \, \frac{dt}{t} \, dx,$$

with parameters β and b related by $\beta_1 - \beta_3 = b_1$ and $\beta_2 - \beta_3 = b_2$. By scaling and translation invariance the parameter β is restricted to be a unit vector orthogonal to $(1, 1, 1)$. If any two components of β are equal, the form is reduced to a composition of a pointwise multiplier with the (dual of the) classical Hilbert transform. Thus, the boundedness properties of this reduced form, called *degenerate*, are provided by the classical Calderón-Zygmund theory.

A priori L^p estimates for the *non-degenerate* case where first given in the breakthrough papers [3] and [4]. Part of interest in these estimates lies in their method of proof, as it is closely related to the techniques first developed by Carleson and Fefferman for the proof of the pointwise convergence of Fourier articles [1, 2]. Roughly speaking, the analysis consists of replacing functions which have perfect time and scale localization—wavelets—by wave packets, which have good time, scale and frequency localization.

7.2 Walsh models

Somewhat naïvely speaking, the dyadic analogue of non-negative integer powers of simple sine and cosine functions is the *Walsh system*, recursively given by

$$\begin{aligned} W_0(x) &= 1_{[0,1)}(x) \\ W_{2m}(x) &= W_m(2x) + W_m(2x-1) \\ W_{2m+1}(x) &= W_m(2x) - W_m(2x-1), \end{aligned}$$

where $1_{[0,1)}$ denotes the characteristic function of the unit interval. The Walsh system constitutes an orthonormal basis of $L^2[0,1)$. Thus the expansion of a given function in the Walsh basis is the analogue of its Fourier series.

The *Walsh phase plane* is the closed first quadrant $\mathbb{R}_+ \times \mathbb{R}_+$ of the plane. A dyadic rectangle $p \subset \mathbb{R}_+ \times \mathbb{R}_+$ is a rectangle of the form

$$p = I_p \times \omega_p = [2^k n, 2^k(n+1)) \times [2^{-l}m, 2^{-l}(m+1)) \in \mathscr{D} \times \mathscr{D}, \tag{1}$$

where $\mathscr{D}$ stands for the standard dyadic grid on $\mathbb{R}_+$, hence having scale parameters $k, l \in \mathbb{Z}$ and time and frequency parameters $n, m \in \mathbb{N}$. A *tile* is a dyadic rectangle of area 1. Let $\mathbf{t}$ denote the set of all tiles in $\mathbb{R}_+ \times \mathbb{R}_+$. Given a tile $p \in \mathbf{t}$ with parameters k, n, m determined by (1), the associated *Walsh wave packet* is the function

$$w_p(x) = 2^{-k/2} W_m(2^{-k}x - n),$$

which is supported on I_p, normalized in L^2 and has constant modulus $2^{-k/2}$. Clearly, for $p \in \mathbf{t}$ with frequency parameter $m = 0$, the associated Walsh wave packet is $w_p = \chi_{I_p}$ the L_2 normalized characteristic function of I_p. Further, if $m = 1$ then $w_p = h_{I_p}$, the Haar function associated to I_p. For any pair $p, q \in \mathbf{t}$ the corresponding Walsh wave packets satisfy the fundamental localization property

$$\langle w_p, w_q \rangle = \sqrt{|p \cap q|}. \tag{2}$$

By this relation, if a set $S \subset \mathbb{R}_+ \times \mathbb{R}_+$ can be decomposed as a disjoint union of tiles $\mathbf{p} \subset \mathbf{t}$, the *phase space projection* associated to S

$$\Pi_S f = \sum_{p \in \mathbf{p}} \langle f, w_p \rangle w_p,$$

is independent of the particular tiling $\mathbf{p}$ of S. A *bitile* is a dyadic rectangle $P = I \times \omega$ of area 2. Call $\mathbf{B}$ the set of all bitiles in the Walsh phase plane. For

$I \in \mathscr{D}$ we write I_- and I_+ to denote the left and right parts of I respectively. Note that each bitile $P = I \times \omega$ can be splitted into upper tile $P_u = I \times \omega_+$ and lower tile $P_d = I \times \omega_-$ (*frequency brothers*), or alternatively into left $P_l = I_- \times \omega$ and right $P_r = I_+ \times \omega$ tiles (*time brothers*). The associated Walsh wave packet are related via

$$w_{P_u} = 2^{-1/2}(w_{P_l} - w_{P_l}), \qquad\qquad w_{P_d} = 2^{-1/2}(w_{P_l} + w_{P_l}).$$

This relationship can be recursively applied to obtain Walsh wave packets starting from the frequency origin.

The first Walsh model of a non-degenerate bilinear Hilbert transform, known as the *quartile operator*, was introduced in [7]. The study of the boundedness behaviour of the quartile operator has been instrumental not only for the proof of pointwise convergence for the Walsh-Fourier series, but also provided the essential machinery to deal with the Fourier analogue, since technical difficulties such as the lack of a localization relation (2) for disjoint tiles are avoided. Ever since, related families of discrete models of near degenerate cases have been introduced to address uniform estimates. Here we will be concerned with the family of quartile forms with parameter $L \geq 2$ introduced in [6] as

$$\Lambda_L(f_1, f_2, f_3) = \int_{\mathbb{R}} \sum_{P \in \mathbf{B}} \tilde{w}_{P_d}(x) \Pi_{P_u} f_1(x) \prod_{j=2}^{3} \Pi_{2^L P_d} f_j(x)\, dx, \tag{3}$$

where $\tilde{w}_p$ denotes the L^∞ normalized wave packet $|I_p|^{1/2} w_p$ and $2^L S$ denotes the frequency-dilated set $\{(x, 2^L \xi) : (x, \xi) \in S\}$.

To avoid technicalities we will restrict ourselves to the bitiles contained the strip $\mathbb{R}_+ \times [0, 2^N)$ for some large $N \in \mathbb{N}$. This in turn restricts the finest time scale to 2^{-N}, making all test functions (to wit, f_1) to be constant on dyadic intervals of that size. We are now in position to state the main results of [6].

Theorem 1. *For any Hölder triple of exponents (p_1, p_2, p_3) with $1 < p_j < \infty$, $j = 1, 2, 3$, we have the a priori strong type estimate*

$$|\Lambda_L(f_1, f_2, f_3)| \leq C_{p_1, p_2, p_3} \prod_{j=1}^{3} \|f_j\|_{p_j},$$

with constant C_{p_1, p_2, p_3} uniform on L and N.

The condition to have a Hölder triple of exponents (p_1, p_2, p_3), namely that $\sum_j \alpha_j = 1$ with $\alpha_j = 1/p_j$, is necessary by the dilation symmetry of Λ_L when defined in the whole phase space. Thus, while fixing N breaks the dilation symmetry of the form Λ_L, uniform estimates on N make the family of all Λ_L for all such N to retain the dilation symmetry and thus the condition on the exponents prevails.

Generalized restricted type inequalities are available for extended ranges of exponents under the extra assumption that $|f_j| \leq 1_{E_j}$ a.e. for some measurable set $E_j \subset \mathbb{R}_+$. Namely, for a fixed exceptional index j such that $\alpha_j < 0$ restricted estimates for Λ_L hold if $|f_j| \leq 1_{E'_j}$ for some *major* subset $E'_j \subset E_j$ that depends on E_1, E_2 and E_3. Where it is said that $E' \subset E$ is major if it is such that $|E'| \geq \frac{1}{2}|E|$. A model result is the following.

Theorem 2. *Let $0 < \alpha_1, \alpha_3 < 1$ and $-\frac{1}{2} < \alpha_2 \leq 0$ such that $\sum_j \alpha_j = 1$. For any $E_j \subset \mathbb{R}_+$, $j = 1, 2, 3$, such that $|E_2|$ is maximal among the $|E_j|$, there is a major subset $E'_2 \subset E_2$ such that for any f_j such that $|f_j| \leq 1_{E_j}$ for $j \neq 2$ and $|f_2| \leq 1_{E'_2}$, we have the generalized restricted weak type estimate*

$$|\Lambda_L(f_1, f_2, f_3)| \leq C_{\alpha_1, \alpha_2, \alpha_3} \prod_{j=1}^{3} |E_j|^{\alpha_j}, \tag{4}$$

with constant $C_{\alpha_1, \alpha_2, \alpha_3}$ uniform on L and N.

7.3 The strategy

The full range of triples $(\alpha_1, \alpha_2, \alpha_3)$, $\sum_j \alpha_j = 1$, for which estimates of the form (4) hold is depicted as the unshaded area in Figure 1. Namely the convex hull of the open triangles a_1, a_2 and a_3. The open triangle c represents the case $2 \leq p_j < \infty$ and the convex hull of the open triangles b_1, b_2 and b_3 determines the reflexive Banach triangle $1 < p_j < \infty$.

The strategy will be the following. Theorem 1 will be proved first in the open triangle c. This will be done first by proving certain restricted weak type estimates and then strong type estimates will be obtained by extending the support of the f_j. Restricted weak type estimates will the be proven in the open triangle $b_3 \cup d_{23}$ and use multilinear Marcinkiewicz interpolation [8] to obtain strong type estimates in the open triangle b_3 and generalized restricted weak type estimates in the open triangle d_{23}. Symmetric arguments can then be applied to get Theorem 1 and Theorem 2 in their full extent.

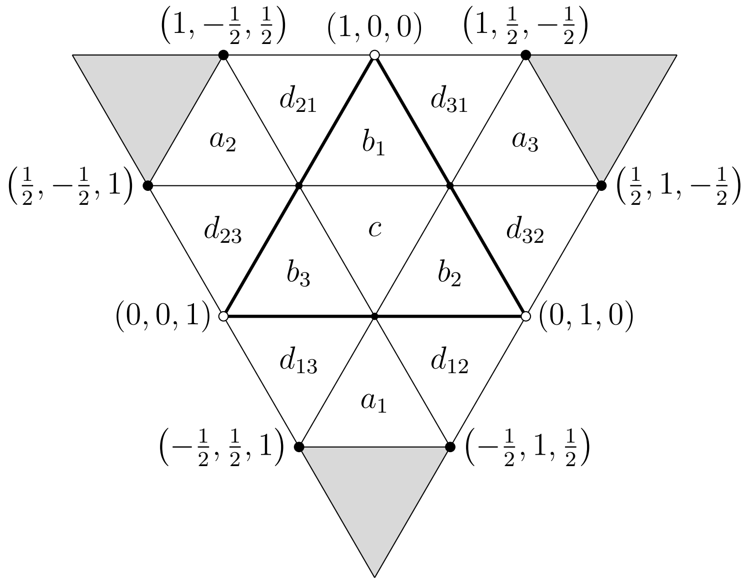

Figure 1: Region of exponents $\alpha_j = 1/p_j$

The technique that will be used to prove generalized restricted weak type estimates in the region $b_1 \cup d_{23}$ is the multi-frequency Calderón-Zygmund decomposition obtained in [5], which we state below.

Theorem 3. *Let $\xi_1 < \ldots < \xi_N \in \mathbb{R}$ for $N \geq 1$ and let $f \in L^1$ and $\lambda > 0$. Then f can be decomposed as*

$$f = g + \sum_{I \in \mathbf{I}} b_I$$

for a disjoint family of maximal dyadic intervals $\mathbf{I}$ such that

$$\sum_{I \in \mathbf{I}} |I| \leq CN^{1/2} \frac{1}{\lambda} \|f\|_1, \qquad \|f \cdot 1_I\|_1 \leq CN^{-1/2} \lambda |I|$$

for every $I \in \mathbf{I}$ and

$$\|g\|_2^2 \leq CN^{1/2} \lambda \|f\|_1, \quad \|f \cdot 1_I - b_I\|_2 \leq C\lambda |I|^{1/2}, \quad \int b_I(x) e^{i\xi_j x}\, dx = 0;$$

for each $I \in \mathbf{I}$ and $1 \leq j \leq N$. Further, $\operatorname{supp} b_I = 3I$*, the interval with the same center as I and sidelenght $\ell(3I) = 3\ell(I)$.*

References

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

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

[3] Lacey, M. and Thiele, C., *L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$.* Ann. of Math. (2) **146** (1997), no. 3, 693–724.

[4] Lacey, M. and Thiele, C., *On Calderón's conjecture.* Ann. of Math. (2) **149** (1999), no. 2, 475–496.

[5] Nazarov, F. and Oberlin, R. and Thiele, C., *A Calderón-Zygmund decomposition for multiple frequencies and an application to an extension of a lemma of Bourgain.* Math. Res. Lett. **17** (2010), no. 3, 529–545.

[6] Oberlin, R. and Thiele, C., *New Uniform Bounds for a Walsh Model of the Bilinear Hilbert Transform.* Indiana Univ. Math. **60** (2011), no. 5, 1693–1712.

[7] Thiele, C., *Time-frequency analysis in the discrete phase plane.* PhD thesis, Yale University, 1995.

[8] Thiele, C., *Wave Packet Analysis.* CMBS Regional Conference Series in Mathematics, no. 105, Published for the Conference Board of the Mathematical Sciences, Washington DC, 2006.

Luis Daniel López-Sánchez, ICMAT
email: **luisd.lopez@icmat.es**

8 Weak-type $(1, 1)$ bounds for oscillatory singular integrals with rational phases

after M. Folch-Gabayet and J. Wright [1]
A summary written by Prince Romeo Mensah

Abstract

This paper considers singular integral operators on $\mathbb{R}$ with an oscillatory factor that has a rational phase $R(x) = P(x)/Q(x)$. Relying only on the degrees of P and Q, the paper derives weak-type $(1, 1)$ bounds for such operators and establishes conditions for which these operators map the Hardy norm H^1 into itself or into L^1.

8.1 Introduction

The focus here will be the study of the oscillatory integral operators given by 1

$$Tf(x) = p.v \int_{\mathbb{R}} \frac{e^{iR(y)}}{y} f(x-y)dy. \tag{1}$$

The primary goal will be to consider the class of rational functions that combines previously known properties for such maps and give a uniform bound on L^1. The main result is given by the following theorem.

Theorem 1. *Let $R(x) = P(x)/Q(x)$ be a rational function with coefficients in $\mathbb{R}$ and consider the associated operator T given in 1. Then T is weak-type $(1, 1)$ and with bounds depending only on the degrees of P and Q, that is,*

$$\alpha \left|\{x \in \mathbb{R} : |Tf(x)| \geq \alpha\}\right| \leq C\|f\|_{L^1(\mathbb{R})},$$

where the constant C depends only on the degrees of P and Q and in particular, C may be taken to be independent of the coefficients.

Determining this theorem will stem from observing how this polynomial behaves around a bounded number of dyadic intervals, after $P(x)$ and $Q(x)$ have been decomposed into its linear factors. The consequence of this will be the reduction to already know singular integral operators such as that given by C. Fefferman and the Calderón-Zygmund operator given certain conditions.

Notation: Give two positive quantities, A and B. We write $A \lesssim B$ or $A = O(B)$ to denote the estimate $A \leq CB$ where C depends only on the degrees of P and Q. Then use $A \sim B$ to denote the estimates $A \lesssim B \lesssim A$.

Given these notations, the proof essentially relies on the following 3 lemmas:

Lemma 2. *Let $P(t) = a\prod_{j=1}^{d}(t - z_j) = \sum_{k=0}^{d} p_k t^k$ be a polynomial of degree d whose roots are ordered so that $|z_1| \leq \ldots \leq |z_d|$. For each $A > 0$, we define the following intervals (possibly empty) on $\mathbb{R}^+$: for $1 \leq j \leq d-1$, we set $G_j = G_j(A) := [A|z_j|, A^{-1}|z_{j+1}|]$ and for $j = d$, we set $G_d := [A|z_d|, \infty)$. Furthermore if $z_1 \neq 0$, we set $G_0 = G_0(A) = [0, A^{-1}|z_1|]$.*

Then there exists a constant $C = C(d) > 0$ such that for any $A \geq C(d)$ and $0 \leq j \leq d$ with G_j nonempty,

1. *$|P(t)| \sim |p_j||t|^j$ for $|t| \in G_j$ and*
2. *$|p_j| \sim |a|\prod_{l=j+1}^{d}|z_l|$ in particular $p_j \neq 0$*

Lemma 3. *Let $R = P/Q$ be a rational function and G a gap as described above. Then for any integer $n \geq 0, A \geq C_n$ can be chosen large enough so that on G, if $j \geq k$,*

$$R^{(n)}(t) = R(t)\left[\sum_{k+1 \leq l_1 \neq \ldots \neq l_n \leq j} \prod_{m=1}^{n} \frac{1}{t - z_{l_m}} + E_n(t)\right]$$

where $|(d/dt)^r E_n(t)| \lesssim C_{n,r} A^{-1}|t|^{-n-r}$ on G for all $r \geq 0$.

Lemma 4. *Let $R = P/Q$ and G be as in lemma 3 but where now $j < k$. For any integer $n \geq 1, A \geq C_n$ can be chosen large enough so that on G,*

$$R^{(n)}(t) = R(t)\Bigg[(-1)^n \sum_{m=1}^{n} \sum_{|\alpha|=n, l(\alpha)=m} d(\alpha) \sum_{j+1 \leq l_1, \ldots, l_m \leq k} \frac{1}{(t - w_{l_1})^{\alpha_1}} \cdots \frac{1}{(t - w_{l_m})^{\alpha_m}} + F_n(t)\Bigg]$$

where $|(d/dt)^r F_n(t)| \lesssim C_{n,r} A^{-1} |t|^{-n-r}$ *for all* $r \geq 0$. *Here* $\{d(\alpha)\}$ *are combinatorial numbers defined on strictly positive multi-indices* α *such that the sums*

$$c_m(n) = \sum_{|\alpha|=n, l(\alpha)=m} d(\alpha)$$

are the Sterling numbers of the second kind; i.e., $\{c_m(n)\}_{m=1}^n$ *are the coefficients of the polynomial*

$$x(x+1)\dots(x+n-1) = \sum_{m=1}^{n} c_m(n) x^m$$

Now depending on the degrees of P and Q, we may state some results on the classical Hardy space $H^1(\mathbb{R})$ and also give a necessary and sufficient condition for this operator to map $H^1(\mathbb{R})$ into the scaled $L^1(\mathbb{R})$ space. This is given by the following theorem:

Theorem 5. *Let* $R(x) = P(x)/Q(x)$ *be a rational function in* $\mathbb{R}$ *and let the degrees of* P *and* Q *be* d *and* e *respectively. Then given the operator* T *in 1, we have that:*

1. *if* $d \neq e+1$, *then* $T : H^1 \to H^1(\mathbb{R})$.
2. *if* $d = e+1$, *then* $T : H^1 \to L^{1,q}(\mathbb{R})$ *if and only if* $q = \infty$.

Establishing Theorem 2 will essentially invole splitting the operator T into three parts $T = T_1 + T_2 + T_3$ where

$$T_j(f(x)) := \int_{\mathbb{R}} f(x-t)\psi_j(t) \frac{e^{iR(t)}}{t} dt \qquad \text{for every } j = 1,2,3$$

such that $\sum_{j=1}^{3} \psi_j(t) = 1$ for all $t \in \mathbb{R}$ and where $\psi_2(t)$ vanishes for $|t|$ 'small' and 'large'. T_2 will therefore map $H^1(\mathbb{R})$ into itself therefore requiring the prove of Theorem 2 to rely only on T_1 and T_3

8.2 Remark

For the fixed Calderon-Zygmund kernel, $K(y)$ on $\mathbb{R}^n$, the Carleson operator as defined by Sjölin[2] is given by:

$$Cf(x) := \sup_{\lambda \in \mathbb{R}^n} \left| p.v. \int_{\mathbb{R}^n} f(x-y) K(y) e^{i\lambda . y} dy \right|.$$

Stein-Wainger however initiated the approach where $\lambda.y$ is replaced by a real polynomial of some degree that vanishes at the origin to some order.

This paper therefore opens the question of whether this approach could be pushed to cover the Carleson's theorem stated above for some rational phase.

References

[1] Folch-Gabayet, M. and Wright, J., *Weak-Type* $(1,1)$ *Bounds For Oscillatory Singular Integrals With Rational Phases.* Mathematics Subject Classification. **42B15** (1991).

[2] Yung, P-L., *Carleson's Theorem of Radon type with polynomial phases.* http://people.maths.ox.ac.uk/yung/Carleson_talk.pdf

Prince Romeo Mensah, Universität Bonn
email: **prmens@uni-bonn.de**

9 L^p theory for outer measures

after Y. Do and C. Thiele [2]
A summary written by Yumeng Ou

Abstract

We develop a theory of L^p spaces based on outer measures, which includes as a special case the classical L^p theory on Euclidean spaces. As an application, we rephrase several classical results concerning Carleson embedding, paraproducts and the $T(1)$ theorem in the language of outer measure spaces.

9.1 Introduction

Carleson measures and time-frequency analysis have been developed as indispensable tools for applications in a great number of problems in singular integral theory and related areas. It has been found later that these two theories are in fact closely related, for instance see [1]. The present paper offers a unifying language for both Carleson measures and time-frequency analysis by developing a natural L^p theory for outer measures. A novelty is that instead of passing through a discrete model form, one can work with an outer measure space on a continuum, which avoids the usual technicalities in the discretization process.

An outer measure on a set X is a monotone and subadditive function on the collection of subsets of X with values in the extended nonnegative real numbers, and with the value 0 attained by the empty set. The lack of additivity for disjoint sets prevents us from expecting a useful linear theory of integrals with respect to outer measures. A good replacement is a quasi sub-linear theory which leads directly to quasi norms rather than integrals. Moreover, instead of basing on the outer measure of super level sets $\{x : f(x) > \lambda\}$ for a function f, we develop the L^p theory using a more subtly defined quantity, which involves a pre-defined averages over the generating sets of the outer measure.

This paper is organized as follows. First, we introduce the necessary quantities and develop the L^p theory, which is a subject not only useful in our intended applications but also of independent interest. Second, we describe how the theory can be used in the context of Carleson measures, where we rephrase several classical results concerning Carleson embedding, paraproducts and a $T(1)$ theorem.

9.2 Outer measure spaces, sizes

In order to describe an abstract outer measure, one can first specify concretely a pre-measure on a small collection of subsets, and then pass abstractly to the outer measure through a covering process, which is described as follows.

Proposition 1. *Let X be a set and $\mathbf{E}$ a collection of subsets of X. Let σ be a function from $\mathbf{E}$ to $[0,\infty)$. Define for an arbitrary subset E of X*

$$\mu(E) := \inf_{\mathbf{E}'} \sum_{E' \in \mathbf{E}'} \sigma(E') \ ,$$

where the infimum is taken over all subcollections $\mathbf{E}'$ of $\mathbf{E}$ which cover the set E, that is whose union contains E. Then μ is an outer measure.

Classically, for $1 \le p < \infty$, an L^p norm of a nonnegative function f is defined via

$$(\int_0 p\lambda^{p-1}\mu(\{x \in X : f(x) > \lambda\})\, d\lambda)^{1/p} \ .$$

However, rather than regarding f as a pointwise assignment, we build the L^p theory on outer measure spaces via averages (”sizes”) over generating sets. In the absence of measurability we will require the averages to be merely sub-linear or even quasi sub-linear. For simplicity, we assume that X is a metric space and every set of the collection $\mathbf{E}$ is Borel.

Definition 2. *Let X be a metric space. Let σ be a function on a collection $\mathbf{E}$ of Borel subsets of X and let μ be the outer measure generated by σ. A size is a map*

$$S : \mathcal{B}(X) \to [0,\infty]^{\mathbf{E}}$$

satisfying for every $f, g \in \mathcal{B}(X)$ and every $E \in \mathbf{E}$ the following properties:

1. *Monotonicity: if $|f| \le |g|$, then $S(f)(E) \le S(g)(E)$.*
2. *Scaling: $S(\lambda f)(E) = |\lambda| S(f)(E)$ for every $\lambda \in \mathbb{C}$.*
3. *Quasi-subadditivity:*

$$S(f+g)(E) \le C[S(f)(E) + S(g)(E)] \qquad (1)$$

for some constant C depending only on S but not on f, g, E.

For example, let $X = \mathbb{R}^m$ and $\mathbf{E}$ be the set of all dyadic cubes. For each dyadic cube Q with side length 2^k, set $\sigma(Q) = 2^{mk}$, then σ generates the classical Lebesgue outer measure on $\mathbb{R}^m$. We also define for every $f \in \mathcal{B}(X)$ and every cube Q the size

$$S(f)(Q) = \mu(Q)^{-1} \int_Q |f(x)|\, dx \ .$$

Another example which we will explore in detail in the last section is outer measure generated by tents. Let $X = \mathbb{R} \times (0, \infty)$ be the open upper half plane and $\mathbf{E}$ be the set of tents of the form

$$T(x,s) = \{(y,t) \in \mathbb{R} \times (0,\infty) : t < s, |x-y| < s-t\}, \quad (x,s) \in \mathbb{R} \times (0,\infty).$$

Define $\sigma(T(x,s)) = s$, $S_\infty(F)(T(x,s)) := \sup_{(y,t)\in T(x,s)} |F(y,t)|$ and for $1 \le p < \infty$ the sizes

$$S_p(F)(T(x,s)) = \left(s^{-1} \int_{T(x,s)} |F(y,t)|^p \, dy \, \frac{dt}{t} \right)^{1/p} \ .$$

Now there is only the most subtle piece left for the development of outer L^p theory, which is the super level measure. Given an outer measure space, denoted by the triple (X, σ, S). Let $f \in \mathcal{B}(X)$ and $\lambda > 0$, define $\mu(S(f) > \lambda)$ to be the infimum of all values $\mu(F)$ where F runs through all Borel subset of X which satisfy $\mathrm{outsup}_{\mathrm{X}\setminus\mathrm{F}}\mathrm{S(f)} \le \lambda$, where in general, the outer essential supremum of $f \in \mathcal{B}(X)$ on F is defined as

$$\mathrm{outsup_F S(f)} := \sup_{\mathrm{E}\in\mathbf{E}} \mathrm{S(f1_F)(E)}.$$

Note that in general, $\mu(S(f) > \lambda)$ is not the outer measure of the Borel set where $|f|$ is larger than λ, even though in the first example of Lebesgue measure we've mentioned above it is indeed the case. And it can be easily verified that the super level measure has properties such as monotonicity, scaling invariance and quasi-subadditivity.

9.3 Outer L^p spaces

Now we define the strong and weak outer L^p spaces in a standard fashion. Let $f \in \mathcal{B}(X)$ and $0 < p < \infty$, define

$$\|f\|_{L^\infty(X,\sigma,S)} = \|f\|_{L^{\infty,\infty}(X,\sigma,S)} := \mathrm{outsup_X S(f)} = \sup_{\mathrm{E}\in\mathbf{E}} \mathrm{S(f)(E)} \ ,$$

$$\|f\|_{L^p(X,\sigma,S)} := \left(\int_0^\infty p\lambda^{p-1}\mu(S(f) > \lambda)\,d\lambda\right)^{1/p} ,$$

$$\|f\|_{L^{p,\infty}(X,\sigma,S)} := \left(\sup_\lambda \lambda^p \mu(S(f) > \lambda)\right)^{1/p} .$$

As in the classical case we trivially have $\|f\|_{L^{p,\infty}(X,\sigma,S)} \le \|f\|_{L^p(X,\sigma,S)}$. And the outer L^p quasi norm satisfies monotonicity and quasi-subadditivity due to the corresponding properties of super level measure. Many of the classical results for L^p theory have their counterparts in outer L^p theory as well, among which we have the following Hölder inequality and Marcinkiewicz interpolation, whose proofs are in the similar fashion as the classical ones but adapted to the outer measure setting.

Proposition 3 (Hölder inequality)**.** *Let X be a metric space, $\mathbf{E}, \mathbf{E}_1, \mathbf{E}_2$ be three collections of Borel subsets, and $\sigma, \sigma_1, \sigma_2$ be three functions on these collections generating outer measures μ, μ_1, μ_2 on X. Assume $\mu \le \mu_j$ for $j = 1, 2$, and S, S_1, S_2 are three respective sizes such that for any $E \in \mathbf{E}$ there exist $E_1 \in \mathbf{E}_1$ and $E_2 \in \mathbf{E}_2$ such that for all $f_1, f_2 \in \mathcal{B}(X)$ we have*

$$S(f_1 f_2)(E) \le S_1(f_1)(E_1) S_2(f_2)(E_2) .$$

Let $p, p_1, p_2 \in (0, \infty]$ such that $1/p = 1/p_1 + 1/p_2$, then

$$\|f_1 f_2\|_{L^p(X,\sigma,S)} \le 2\|f_1\|_{L^{p_1}(X,\sigma_1,S_1)} \|f_2\|_{L^{p_2}(X,\sigma_2,S_2)} .$$

Proposition 4 (Marcinkiewicz interpolation)**.** *Given outer measure space (X, σ, S) and $1 \le p_1 < p_2 \le \infty$. Let T be an operator that maps $L^{p_1}(Y, \nu)$ and $L^{p_2}(Y, \nu)$ to the space of Borel functions on X, such that for any $f, g \in L^{p_1}(Y, \nu) + L^{p_2}(Y, \nu)$ and $\lambda \ge 0$ we have $|T(\lambda f)| = |\lambda T(f)|$, $|T(f + g)| \le C(|T(f)| + |T(g)|)$ and the weak boundedness:*

$$\|T(f)\|_{L^{p_1,\infty}(X,\sigma,S)} \le A_1 \|f\|_{L^{p_1}(Y,\nu)} ,$$

$$\|T(f)\|_{L^{p_2,\infty}(X,\sigma,S)} \le A_2 \|f\|_{L^{p_2}(Y,\nu)} .$$

Then we also have

$$\|T(f)\|_{L^p(X,\sigma,S)} \le A_1^{\theta_1} A_2^{\theta_2} C_{p_1,p_2,p} \|f\|_{L^p(Y,\nu)} ,$$

where $p_1 < p < p_2$ and θ_1, θ_2 are such that $\theta_1 + \theta_2 = 1$ and $1/p = \theta_1/p_1 + \theta_2/p_2$.

Before we move on to the application, the following proposition, being a simple variant of a classical fact about measures, acts as a bridge connecting the measure on X and our outer L^p space.

Proposition 5. *Given an outer measure space (X, σ, S) and assume that about every point in X there is an open ball for which there exists $E \in \mathbf{E}$ which contains the ball. Let ν be a positive Borel measure on X and assume that $\forall f \in \mathcal{B}(X)$, $\forall E \in \mathbf{E}$ we have $\int_E |f|\, d\nu \leq CS(f)(E)\mu(E)$. Then, $\forall f \in \mathcal{B}(X)$ with finite $\|f\|_{L^\infty(X,\sigma,S)}$ we have*

$$|\int_X f\, d\nu| \leq C\|f\|_{L^1(X,\sigma,S)} \ .$$

9.4 Applications

In this section we will work with the example of outer measure generated by tents and rephrase in the language of outer measure spaces several classical results. Let ϕ be a smooth function supported in $[-1, 1]$, for any locally integrable function f define $F_\phi(f)(y, t) := f * \phi_t(y)$ where $\phi_t(y) = 1/t\phi(y/t)$. Then the map F_ϕ is reminiscent of Carleson embeddings.

Theorem 6 (Carleson embedding theorem)**.** *Let $1 < p \leq \infty$. We have for ϕ as above*

$$\|F_\phi(f)\|_{L^p(X,\sigma,S_\infty)} \leq C_{p,\phi}\|f\|_p \ .$$

If in addition $\int \phi = 0$, then

$$\|F_\phi(f)\|_{L^p(X,\sigma,S_2)} \leq C_{p,\phi}\|f\|_p \ .$$

The proof of the theorem involves Marcinkiewicz interpolation for both parts, where in the second part one apply Calderón's reproducing formula to prove the $L^\infty \to L^\infty$ boundedness, and use Calderón-Zygmund decomposition for the demonstration of the weak type bound at $p = 1$. This theorem can be used to prove the following classical results.

9.4.1 Paraproducts

A classical paraproduct after pairing with a third function can be viewed as a trilinear form of the type

$$\Lambda(f_1, f_2, f_3) = \int_{\mathbb{R}\times(0,\infty)} \prod_{j=1}^{3} F_{\phi_j}(f_j)(x, t)\, dx\frac{dt}{t}$$

with compactly supported smooth functions $\{\phi_j\}$ where two of them, say ϕ_1, ϕ_2 have vanishing integrals. An application of Proposition 5 implies $|\Lambda(f_1, f_2, f_3)| \le C \| \prod_{j=1}^3 F_{\phi_j}(f_j)\|_{L^1(X,\sigma,S_1)}$. Let $1 < p_1, p_2, p_3 \le \infty$ such that $1 = 1/p_1 + 1/p_2 + 1/p_3$, one then have according to the Hölder inequality and Theorem 6 that $|\Lambda(f_1, f_2, f_3)| \le C\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_3}$, the classical estimates.

9.4.2 A simplified $T(1)$ theorem

Let ϕ be some nonzero smooth function supported in $[-1, 1]$ with $\int \phi = 0$ and define $\phi_{x,s}(y) = 1/s\phi((y - x)/s)$, $\forall(x, s) \in \mathbb{R} \times (0, \infty)$. We have

Theorem 7 ($T(1)$ theorem)**.** *Let T be a bounded linear operator in $L^2(\mathbb{R})$ such that for all x, y, s, t*

$$| \langle T(\phi_{x,s}), \phi_{y,t}\rangle | \le \frac{\min(t, s)}{\max(t, s, |y - x|)^2} \ . \tag{2}$$

Then we have bounds $\|T\|_{L^p\to L^p} \le C_p$ for $1 < p < \infty$, where C_p depends only on ϕ, p and in particular not on T.

The proof involves Calderón's reproducing formula, Hölder inequality, Proposition 5, and as a key element, Theorem 6 as well as its modified version, which we introduce as a corollary at the end, whose proof can be obtained by formulation of a modified outer measures and sizes, together with a pullback result of outer measures.

Corollary 8. *Let $1 < p \le \infty$, $-1 \le \alpha \le 1$, $0 < \beta \le 1$, and assume $\int \phi = 0$. Define $F_{\alpha,\beta,\phi}(f)(y, t) = F_\phi(f)(y + \alpha t, \beta t)$. Then there exists $\epsilon > 0$ such that*

$$\|F_{\alpha,\beta,\phi}(f)\|_{L^p(X,\sigma,S_2)} \le C_{p,\phi}\beta^{-1/p+\epsilon}\|f\|_p.$$

References

[1] P. Auscher, S. Hofmann, C. Muscalu, T. Tao, C. Thiele, *Carleson measures, trees, extrapolation, and $T(b)$ theorems.* Publ. Mat. 46 (2002), 257-325.

[2] Y. Do, C. Thiele, *L^p theory for outer measures and two themes of Lennart Carleson united.* preprint, arXiv: 1309.0945v2 (2014);

Yumeng Ou, Brown University
email: **yumeng_ou@brown.edu**

10 A proof of boundedness of the Carleson operator

after Michael Lacey and Christoph Thiele [5]
A summary written by Guillermo Rey

Abstract

We give a short summary of the proof of the weak-type (2,2) boundedness of the Carleson operator in [5].

10.1 Introduction

It was conjectured by Luzin in 1915 that the Fourier series of every L^2 function converges pointwise almost everywhere to the original function. Kolmogorov showed in 1923, by constructing a counterexample, that the result cannot be true if the function is just in L^1. The conjecture was finally settled in 1966 by L. Carleson [1] (later extended by Hunt in [3] to the whole range $1 < p < \infty$) with a very technical argument which decomposes the function in a very precise way. C. Fefferman gave an alternate proof in 1973 [2] where now the *operator* is decomposed into simpler "almost orthogonal" operators.

Here we give a summary of the proof in [5]. After standard reductions and passing to a dyadic model, the proof is based on a decomposition of dyadic trees into unions of "simpler" trees.

We will use the essentially the same notation as in the paper:

$$\begin{aligned}
\operatorname{Trans}_y f(x) &= f(x-y) \\
\operatorname{Mod}_\eta f(x) &= e^{2\pi i \eta x} f(x) \\
\operatorname{Dil}^p_\lambda f(x) &= \lambda^{-1/p} f(x/\lambda).
\end{aligned}$$

We note that $\operatorname{Dil}^p_\lambda$ is the dilation that preserves the $L^p(\mathbb{R})$ norm. Also, if $\mathcal{F}$ denotes the Fourier transform:

$$\mathcal{F} f(\xi) = \int_{\mathbb{R}} e^{-2\pi i x \xi} f(x)\, dx,$$

then we have the following relations:

$$\mathcal{F} \operatorname{Trans}_y = \operatorname{Mod}_{-y} \mathcal{F}, \quad \mathcal{F} \operatorname{Mod}_\eta = \operatorname{Trans}_\eta \mathcal{F} \quad \text{and} \quad \mathcal{F} \operatorname{Dil}^p_\lambda = \operatorname{Dil}^{p'}_{\lambda^{-1}} \mathcal{F},$$

where p' is the Hölder conjugate exponent.

The constant ν will be some large integer which will be fixed for the rest of the summary. We define

$$w(x) = (1 + |x|)^{-\nu} \quad \text{and} \quad w_I(x) = \operatorname{Trans}_{c(I)} \operatorname{Dil}^1_{|I|} w(x),$$

where $c(I)$ denotes the center of the interval I and $|I|$ its length. Observe that $\|w\|_{L^1} = \|w_I\|_{L^1} < \infty$ for any I.

We will use a function ϕ which is chosen from the Schwartz class so that $\widehat{\phi}$ is real, non-negative, supported in $[-0.1, 0.1]$ and equal to 1 on $[-0.09, 0.09]$. If P is a rectangle of area 1 with $P = I_P \times \omega_P$, we define

$$\phi_{1P} = \operatorname{Mod}_{c(\omega_{1P})} \operatorname{Trans}_{c(I_P)} \operatorname{Dil}^2_{|I_P|} \phi,$$

where ω_{1P} and ω_{2P} are the lower and upper halves of ω_P.

We will let $\mathbb{P}$ denote the collection of all rectangles $I \times \omega$ of area 1 and with I and ω dyadic intervals, elements of $\mathbb{P}$ will be called tiles. We define a partial order on $\mathbb{P}$ by setting $P < P'$ if $I_P \subseteq I_{P'}$ and $\omega_{P'} \subseteq \omega_P$. A set of tiles T is called a tree if there is a tile P_T (called the top) such that $P < P_T$ for all $P \in T$. A tree is called a j-tree if $\omega_{jP_T} \subseteq \omega_{jP}$ for all $P \in T$.

10.1.1 Setting

The goal of the article is to show that we have

$$\lim_{N \to \infty} \int_{-\infty}^{N} e^{2\pi i x \xi} \widehat{f}(\xi)\, d\xi \to f(x)$$

almost everywhere for $f \in L^2(\mathbb{R})$.

By Stein's maximal principle, this is equivalent to showing that the Carleson operator

$$\mathcal{C}f(x) = \sup_{N \in \mathbb{Z}} \left| \int_{-\infty}^{N} e^{2\pi i x \xi} \widehat{f}(\xi)\, d\xi \right|$$

is of weak-type $(2, 2)$.

Let $\theta \in \mathbb{R}$, then there is a unique (up to a multiplicative constant) non-zero linear operator T on $L^2(\mathbb{R})$ which is bounded, commutes with translations and $\operatorname{Dil}^2_\lambda \operatorname{Mod}_{(\lambda-1)\theta}$ for all $\lambda > 0$, and is identically 0 for all functions whose Fourier support lies in (θ, ∞).

Indeed, any bounded linear operator on L^2 which commutes with translations must be a Fourier multiplier:

$$\widehat{Tf}(\xi) = m(\xi)\widehat{f}(\xi).$$

If we now impose the relation $\mathrm{Dil}^2_\lambda \mathrm{Mod}_{(\lambda-1)\theta} T = T \mathrm{Dil}^2_\lambda \mathrm{Mod}_{(\lambda-1)\theta}$ for all $\lambda > 0$, we arrive at

$$m(\xi)\widehat{f}(\xi\lambda - (\lambda-1)\theta) = m(\xi\lambda - (\lambda-1)\theta)\widehat{f}(\xi\lambda - (\lambda-1)\theta).$$

Now let $\xi, \eta < \theta$ be two different numbers, then we can find a $\lambda \in (0,1]$ such that $\eta = \xi\lambda - (\lambda-1)\theta$ and a function f such that $\widehat{f}(\eta) \neq 0$. This now implies

$$m(\xi) = m(\eta) \quad \forall \eta, \xi < \theta.$$

The condition that Tf should vanish whenever the Fourier support of f lies in (θ, ∞) translates to the requirement of $m(\xi) = 0$ for all $\xi > \theta$, so we have determined T up to a multiplicative constant as desired.

The authors now take a further step towards the final form on which they will primarily work. The reduce to studying just the operators

$$A_\theta f(x) = \sum_{P \in \mathbb{P}} \mathbb{1}_{\omega_{2P}}(\theta)\langle f, \phi_{1P}\rangle \phi_{1P}.$$

To this end, they write a certain average over all translations, dilations and modulations of these operators and then show that this average a linear bounded operator on L^2 which commutes with translations and $\mathrm{Dil}^2_\lambda \mathrm{Mod}_{(\lambda-1)\theta}$. Indeed, define $\Pi_\theta f(x)$ pointwise by the following limit:

$$\lim_{n\to\infty} \frac{1}{|F_n|} \int_{F_n} \mathrm{Mod}_{-\eta} \mathrm{Trans}_{-y} \mathrm{Dil}^2_{2^{-k}} A_{2^{-k}(\theta+\eta)} \mathrm{Dil}^2_{2^k} \mathrm{Trans}_y \mathrm{Mod}_\eta f \, dy \, d\eta \, d\kappa,$$

where $F_n = [-n,n] \times [-n,n] \times [0,1]$. It is routine to check (using the properties of A) that this operator is a well defined bounded linear operator which commutes with translations and $\mathrm{Dil}^2_\lambda \mathrm{Mod}_{(\lambda-1)\theta}$, and vanishes whenever the Fourier support of f lies in (θ, ∞). Hence by the observation above:

$$\widehat{\Pi_\theta f}(\xi) = c_\theta \mathbb{1}_{(-\infty,\theta)}(\xi)\widehat{f}(\xi),$$

which implies

$$\mathcal{C}f(x) = \sup_{N \in \mathbb{Z}} \frac{1}{|c_N|} |\Pi_N f(x)|.$$

One can check that with this definition of Π_N, the constant c_N is independent of N, so we arrive at

$$\mathcal{C}f(x) = C \sup_{N\in\mathbb{Z}} |\Pi_N f(x)|$$

for some $C > 0$.

By the convexity of $L^{2,\infty}$ we see that it will suffice to show that

$$\| \sup_N |A_N f| \|_{L^{2,\infty}} \lesssim \|f\|_{L^2}.$$

By linearizing the operator, we see that it would be enough to show that

$$\|Q_N f\|_{L^{2,\infty}} \lesssim \|f\|_{L^2}$$

uniformly over all measurable functions N, where

$$Q_N f(x) = \Pi_{N(x)} f(x).$$

By duality and the triangle inequality, this would follow from the estimate

$$\sum_{P\in\mathbb{P}} |\langle f, \phi_{1P}\rangle \langle (\mathbb{1}_{\omega_{2P}} \cdot \phi_{1P}, \mathbb{1}_E \circ N)\phi_{1P}\rangle| \lesssim \|f\|_{L^2} |E|^{1/2}. \tag{1}$$

10.2 Main ingredients of the proof

Having reduced to showing (1), the rest of the proof consists on showing how to decompose a collection of tiles into a union of tress which satisfy certain "mass" and "energy" properties.

More precisely, for a collection $\mathbf{P}$ of tiles, we define

$$\operatorname{mass}(\mathbf{P}) = \sup_{P\in\mathbf{P}} \sup_{P'\in\mathbb{P}:\, P<P'} \int_{E_{P'}} w_{P'}(x)\, dx$$

and

$$\operatorname{energy}(\mathbf{P}) = \sup\Big\{ \Big(|I_T|^{-2} \sum_{P\in T} |\langle f, \phi_{1P}\rangle|^2 \Big)^{1/2} : \text{ over all trees } T \subseteq \mathbf{P} \Big\}.$$

As in [5], we are using the notation

$$E_P = E \cap \{x : N(x) \in \omega_P\}, \quad E_{2P} = E \cap \{x : N(x) \in \omega_{2P}\}.$$

With this notation (1) becomes

$$\sum_{P\in\mathbb{P}} |\langle f,\phi_{1P}\rangle\langle(\mathbb{1}_{\omega_{2P}}\cdot\phi_{1P},\mathbb{1}_{E_{2P}}\circ N)\phi_{1P}\rangle| \lesssim \|f\|_{L^2}|E|^{1/2}. \tag{2}$$

By the invariants of the problem, we can assume $\|f\|_{L^2} = 1$ and tha t$|E| \leq 1$. Under these conditions it can be proved, see [5] that, for any tree T, we have

$$\sum_{P\in T} |\langle f,\phi_{1P}\rangle\langle\phi_{1P},\mathbb{1}_{E_{2P}}\rangle| \lesssim \mathrm{energy}(T)\,\mathrm{mass}(T)|I_T|.$$

Finally, the authors in [5] prove an elegant decomposition of collections of tiles into unions of tress with controlled energy and mass. In particular they prove that any finite collection of tiles $\mathbf{P}$ can be decomposed into a union of sets $\mathbf{P}_n$, where n runs through a set of integers, such that

$$\mathrm{mass}(\mathbf{P}_n) \leq 2^{2n}, \quad \mathrm{energy}(\mathbf{P}_n) \leq 2^n$$

and $\mathbf{P}_n$ is a union of a set of trees $\mathbf{T}_n$ with

$$\sum_{T\in\mathbf{T}_n} |I_T| \lesssim 2^{-2n}.$$

Taking this decomposition for granted, we can finish the proof:

$$\begin{aligned}
\sum_{P\in\mathbb{P}} |\langle f,\phi_{1P}\rangle\langle(\mathbb{1}_{\omega_{2P}}\cdot\phi_{1P},\mathbb{1}_{E_{2P}}\circ N)\phi_{1P}\rangle| &\leq \sum_n \sum_{T\in\mathbf{T}_n} \sum_{P\in T} |\langle f,\phi_{1P}\rangle\langle\phi_{1P},\mathbb{1}_{E_{2P}}\rangle| \\
&\lesssim \sum_n \sum_{T\in\mathbf{T}_n} \mathrm{energy}(T)\,\mathrm{mass}(T)|I_T| \\
&\leq \sum_n \sum_{T\in\mathbf{T}_n} 2^n \min(\mathrm{mass}(T),C)|I_T|,
\end{aligned}$$

since the mass of any collection of tiles must be bounded by a universal constant (since $w_{P'}$ in the definition of mass is just a translation and a dilation of w which preserves the L^1 norm). Therefore, we may continue:

$$\begin{aligned}
\sum_n \sum_{T\in\mathbf{T}_n} 2^n \min(\mathrm{mass}(T),C)|I_T| &\leq \sum_n 2^n \min(\mathrm{mass}(T),C) \sum_{T\in\mathbf{T}_n} |I_T| \\
&\lesssim \sum_n 2^n \min(2^{2n},C)2^{-2n} \\
&\lesssim 1.
\end{aligned}$$

The proof of the decomposition resembles a "Calderón-Zygmund" decomposition, we will give details about this in the summer school.

References

[1] L. Carleson. On convergence and growth of partial sums of Fourier series. *Acta Math.*, 116:135–157, 1966.

[2] C. Fefferman. Pointwise convergence of Fourier series. *Ann. of Math. (2)*, 98:551–571, 1973.

[3] R. Hunt. On the convergence of Fourier series: Orthogonal Expansions and their Continuous Analogues Proc. Conf., Edwardsville, Ill., 1967), 1968

[4] M. Lacey. Carleson's theorem: proof, complements, variations. arXiv: math/0307008

[5] M. Lacey and C. Thiele. A proof of boundedness of the Carleson operator. *Math. Res. Lett.*, 7(4):361–370, 2000.

Guillermo Rey, MSU
email: **reyguill@math.msu.edu**

11 Singular and maximal Radon transforms: analysis and geometry. Part 2

after M. Christ, A. Nagel, E. Stein and S. Wainger [1]
A summary written by Joris Roos

Abstract

The authors of [1] show L^p boundedness of a class of singular Radon transforms and their corresponding maximal operators under some curvature assumption. This condition can be formulated in essentially three different ways which are all equivalent.

11.1 Introduction

Singular Radon transforms are operators of the type

$$T(f)(x) = \int_{M_x} f(y) K_x(y) d\sigma_x(y)$$

where $x \in \mathbf{R}^n$, $(M_x)_x$ is a family of k-dimensional submanifolds varying in some sense "smoothly" with x, $(K_x)_x$ is a family of Calderón-Zygmund kernels and $(\sigma_x)_x$ surface measures, where it should be imagined that also the maps $x \mapsto K_x$, $x \mapsto \sigma_x$ are smooth in some way. The integral is made sense of as usual by taking suitable truncations.

The corresponding maximal operator is

$$M(f)(x) = \sup_{r>0} \left| \int_{M_x \cap B(x,r)} f(y) d\sigma_x(y) \right|$$

In [1] the authors take M_x to be parametrized by a smooth map $\gamma : \mathbf{R}^n \times U \to \mathbf{R}^n$ with $U \subset \mathbf{R}^k$ a neighborhood of the origin and $\gamma(x, 0) \equiv x$. We set

$$M_x = \{\gamma(x,t) \,:\, t \in U\}$$

Depending on context it is also useful to look at γ as a family $(\gamma_t)_t$ of local diffeomorphisms of $\mathbb{R}^n$ which are given by $\gamma_t(x) = \gamma(x,t)$.

Choose K to be a standard Calderón-Zygmund kernel on $\mathbf{R}^k$ and ψ a smooth cut-off function as well as a small positive constant a. Then the singular Radon transform from above takes the form

$$T(f)(x) = \psi(x)\, p.v. \int_{|t|\leq a} f(\gamma(x,t))K(t)dt \tag{1}$$

The a and ψ serve the purpose of localizing to a small neighborhood of $(x_0, 0) \in \mathbf{R}^n \times \mathbf{R}^k$ for some x_0 in the support of ψ. The maximal operator in this context is

$$M(f)(x) = \sup_{0<r<a} r^{-k} \left| \psi(x) \int_{|t|\leq r} f(\gamma(x,t))dt \right| \tag{2}$$

The question is now whether T and M define bounded operators $L^p(\mathbf{R^n}) \to L^p(\mathbf{R}^n)$ for $1 < p \leq \infty$.

The main result of the paper [1] is that this is true assuming a certain curvature condition $(\mathcal{C})$ on the family $(M_x)_x$, which will be formulated in three equivalent forms below, each being useful in different parts of the proof. The equivalence of these curvature conditions is by itself also a central result of the paper and makes up the geometric part.

Theorem 1. *Assuming the curvature condition* $(\mathcal{C})$*, then* T*,* M *from* (1), (2) *define bounded operators from* $L^p(\mathbb{R}^n)$ *to* $L^p(\mathbb{R}^n)$ *for all* $1 < p \leq \infty$*.*

Counterexamples show that this theorem is false without any curvature assumptions.

11.2 Curvature conditions

We now state the curvature conditions. They are denoted $(\mathcal{C}_M), (\mathcal{C}_\mathfrak{g}), (\mathcal{C}_J)$ as they involve a submanifold M, a Lie algebra $\mathfrak{g}$ and a Jacobian J, respectively.

The first curvature condition $(\mathcal{C}_M)$ uses the concept of an invariant submanifold M.

Definition 2. *A submanifold* $M \subset \mathbf{R}^n$ *is* locally invariant under γ at x_0 *if there exists a neighborhood* V *of* $(x_0, 0)$ *in* $M \times \mathbf{R}^k$ *such that* $\gamma(x,t) \in M$ *for all* $(x,t) \in V$*.*

This is the notion of invariance that we have should keep in mind but for technical reasons we need the following slightly weaker definition.

Definition 3. *A submanifold $M \subset \mathbf{R}^n$ with $x_0 \in M$ is called* invariant under γ to infinite order at x_0 *if there exists a neighborhood V of $(x_0, 0)$ in $M \times \mathbf{R}^k$ such that*

$$\operatorname{dist}(\gamma(x,t), M) = O(\operatorname{dist}(x, x_0) + |t|)^N$$

as $x \to x_0$ and $t \to 0$ for all $N \in \mathbf{N}$ and $(x,t) \in V$.

Local invariance and invariance to infinite order are equivalent if γ and M are real analytic. We now turn to the curvature condition $(\mathcal{C}_M)$.

Definition 4. *γ is said to satisfy the curvature condition $(\mathcal{C}_M)$ at x_0 if there exists no smooth submanifold $M \subset \mathbb{R}^n$ such that M has positive codimension and is invariant under γ to infinite order at x_0.*

The next curvature condition is formulated in terms of vector fields arising from a sort of noncommutative Taylor expansion of γ. Namely, with γ as above, there exists a unique family $\{X_\alpha : \alpha \in \mathbf{N}^n \backslash \{0\}\}$ of smooth vector fields defined in a common neighborhood U of x_0 such that

$$\gamma(x,t) = \exp\left(\sum_{0<|\alpha|<N} \frac{t^\alpha X_\alpha}{\alpha!}\right)(x) + O(|t|^N) \tag{3}$$

Definition 5. *Let S be the smallest set such that $X_\alpha \in S$ for all α and if $Y, Y' \in S$ then also $[Y, Y'] \in S$. S is called the set of* iterated commutators *of $\{X_\alpha\}$. We say that γ satisfies curvature condition $(\mathcal{C}_\mathfrak{g})$ at x_0 if S is a generating system for the tangent space to $\mathbf{R}^n$ at x_0.*

The last curvature condition is stated using the Jacobian of certain iterates of the map $t \mapsto \gamma(x,t)$. We define $\Gamma^1(x,t) = \gamma(x,t)$ and

$$\Gamma^j(x, t^1, \ldots, t^j) = \gamma(\Gamma^{j-1}(x, t^1, \ldots, t^{j-1}), t^j) \tag{4}$$

for $j \in \{2, \ldots, n\}$ and $(t^1, \ldots, t^j) \in \mathbf{R}^{kj}$ in a small neighborhood of 0. Now set

$$\Gamma(x, \tau) = \Gamma^n(x, \tau) \tag{5}$$

where $\tau \in \mathbf{R}^{kn}$ and we fix an arbitrary ordering of the real coordinates $\tau = (\tau_1, \ldots, \tau_{kn}) \in \mathbf{R}^{kn}$.

Now for each choice of indices $\xi = (\xi_1, \dots, \xi_n)$ from $\{1, \dots, kn\}$ we define the $J_\xi(x, \tau)$ to be the determinant of the corresponding $n \times n$-submatrix of the Jacobian of Γ. That is,

$$J_\xi(x, \tau) = \det \left(\frac{\partial \Gamma(x, \tau)}{\partial(\tau_{\xi_1}, \dots, \tau_{\xi_n})} \right) \tag{6}$$

Definition 6. *We say that γ satisfies curvature condition $(\mathcal{C}_J)$ at x_0 if there exist $\xi = (\xi_1, \dots, \xi_n) \subset \{1, \dots, kn\}$ and a multi-index $\beta \in \mathbf{N}_0^{kn}$ such that*

$$\delta_\tau^\beta J_\xi(x_0, \tau) \mid_{\tau=0} \neq 0$$

The core geometric statement of the paper is the following.

Theorem 7. *The curvature conditions $(\mathcal{C}_M), (\mathcal{C}_\mathfrak{g}), (\mathcal{C}_J)$ are equivalent.*

The proof will be detailed in the talks.

References

[1] Christ, M. and Nagel, A. and Stein, E. and Wainger S., *Singular and maximal Radon transforms: analysis and geometry.* Ann. of Math. (2) **150** (1999), no. 2, 489–577.

Joris Roos, Bonn University
email: **jroos@math.uni-bonn.de**

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

after V. Lie [1]
A summary written by Gennady Uraltsev

Abstract

[1] provides a proof of weak L^2 boundedness of the Carleson operator with both linear and quadratic modulation terms:

$$\mathcal{C}f(x) = \sup_{\substack{P\in\mathbb{R}[y]\\ \deg P\leq 2}} \left| \int_{\mathbb{T}} e^{iP(x)} f(x-y) e^{-iP(x-y)} \frac{\mathrm{d}y}{y} \right|.$$

The proof is similar to [3] and it extends the time-frequency tile approach to be able to study the quadratic Carleson operator in terms of its vaster space of symmetries.

12.1 Introduction

The original formulation of Carleson's theorem states that given any L^2 function f on the interval $\mathbb{T} = [-\pi;\pi)$, its Fourier series converges to f a.e. The proof of the theorem essentially follows from the fact that the maximal-type operator

$$\mathcal{C}f(x) = \sup_{N\in\mathbb{Z}} \left| \sum_{k=-\infty}^{N} \hat{f}(k) e^{ikx} \right| \qquad \text{with } \hat{f}(k) = \int_{\mathbb{T}} f(x) e^{-ikx}\,\mathrm{d}x$$

is weakly bounded on $L^2(\mathbb{T})$.

More generally $\mathcal{C}$ (or a modified version of it) is a maximal operator defined by the Hilbert transform on $\mathbb{T}$ and by a group of symmetries (unitary operators) on $L^2(\mathbb{T})$ given by $M_c f(x) = e^{icx} f(x)$:

$$\mathcal{C}f(x) = \sup_{c\in\mathbb{R}} |M_c \mathcal{H} M_c^* f(x)| = \sup_{c\in\mathbb{R}} \left| \int_{\mathbb{T}} f(x-y) \frac{e^{icy}}{y}\,\mathrm{d}y \right| \tag{1}$$

with $\mathcal{H}$ being the Hilbert transform. Many proofs of the boundedness of such an operator have been obtained e.g. in [3] and [2]. An important question is what happens if one substitutes the linear phase in (1) by a polynomial one:

$$f \longmapsto \sup_{\substack{P\in\mathbb{R}[y]\\ \deg P\leq d}} \left| \int_{\mathbb{T}} f(x-y)e^{iP(x-y)}\frac{\mathrm{d}y}{y} \right|. \tag{2}$$

Abstractly this is equivalent to extending the group of symmetries by adding polynomial modulations and to considering the associated maximal operator

$$\mathcal{C}_d f(x) = \sup_{\mathbf{c}\in\mathbb{R}^d} |U_{\mathbf{c}}\mathcal{H}U_{\mathbf{c}}^{*} f(x)| \tag{3}$$

where $\{U_{\mathbf{c}}\}_{\mathbf{c}\in\mathbb{R}^d}$ are a group of unitary operators defined by $U_{\mathbf{c}}f(x) = e^{iP_{\mathbf{c}}(x)}f(x)$ with $P_{\mathbf{c}}(x) = \sum_{k=1}^{d} c_k x^k$. Herein we give an overview of the proof in [1] of the following theorem.

Theorem 1. *Let $d = 2$. The quadratic Carleson operator on $\mathbb{T} = [-\pi;\pi)$ satisfies*

$$\|C_2 f(x)\|_{L^p(\mathbb{T})} \leq C_p \|f\|_{L^2(\mathbb{T})} \tag{4}$$

for all $p \in [1, 2)$.

An abstract factorization principle due to Nikishin and Stein would allows us to conclude that $\mathcal{C}_2$ is weakly bounded on L^2.

12.2 An outline of the proof

The proof of Theorem 1 is of time-frequency nature and basically consists of the following procedure.

1. We begin by studying how the symmetries of the problem act on the time-frequency space. This corresponds on establishing what are the "tiles" in our problem.

2. We find a discretization of the Hilbert transform and of the Carleson operator that allows us to express its action in a way closely related to the our language of "tiles". We try to identify groups of tiles which are closely related with respect to the action of the Hilbert transform. These will be the so-called "trees".

3. We separate the tiles into a series of classes that have uniform bounds on how much they are involved with the action of the discretized operator. Each of these levels get separated into trees and boundedness estimates are done on each of these sets and then summed up to obtain the bound on the operator.

12.3 Symmetries and the the time-frequency picture

Our problem is essentially defined by its symmetries: the modulation symmetries, together with the translation symmetry (the Hilbert transform is translation invariant) and the dilation symmetry (the Hilbert kernel is homogeneous of degree -1) [2]. The latter two define the Hilbert transform up to a linear combination with the identity operator while the former intervene in the definition of the Carleson maximal operator.

Definition 2 (Symmetries and commutation relations)**.** *Suppose we are working on $L^2(\mathbb{R})$. We can define the following symmetries*

$$\begin{aligned} T_y f(x) &= f(x-y) \quad & M_c f(x) &= e^{icx} f(x) \\ Q_b f(x) &= e^{ibx^2} f(x) \quad & D_t f(x) &= t^{-\frac{1}{2}} f\left(\frac{x}{t}\right) \end{aligned} \tag{5}$$

The generated group of symmetries acts faithfully on the time-frequency plane $\mathbb{R}^2$ by the relation $(b,c,y,t) \circlearrowleft (x;\xi) = (t(x+y); t^{-1}(\xi+c+2bx))$ and thus also on the tile $[0,1)\times[0,1) \subset \mathbb{R}^2$. The set of images of this tile is the set of parallelograms of area 1 with two sides parallel to the frequency axis (see Figure 2). For some test function $\phi \in S(\mathbb{R})$ define the wavelets $\phi_{(b,c,y,t)} = Q_b M_c T_y D_t \phi$. The scalar product $\left\langle \phi_{(b,c,y,t)};\ \phi_{(b',c',y',t')} \right\rangle$ is going to be small if the tiles associated to (b,c,y,t) and (b',c',y',t') are disjoint and far away.

[2] Even though dilations are not naturally defined on a periodic interval $\mathbb{T}$ and in any case they are not a group of unitary transformations on $L^2(\mathbb{T})$ they still help in the understanding of our operator by analogy to the case of the real line

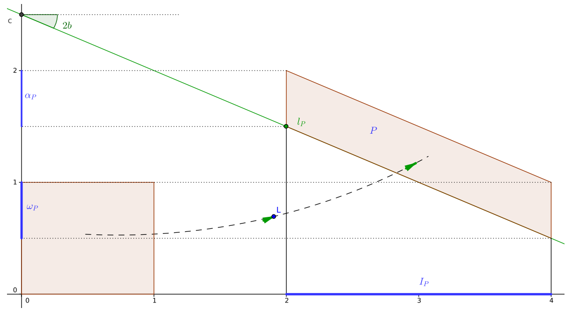

Figure 2: A tile

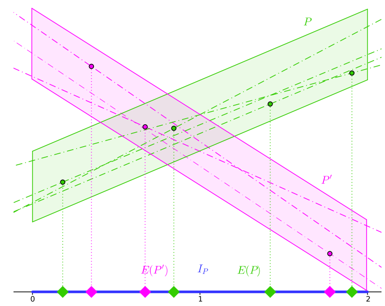

Figure 3: Tiles and lines

12.4 Discretization of the operator

To prove Theorem 1 it is sufficient to obtain uniform boundedness for any choice of measurable function $b, c : \mathbb{T} \to \mathbb{R}$ of the linearized operator

$$\begin{aligned} \mathcal{C}_{2,b,c} f(x) = M_{c(x)} Q_{b(x)} \mathcal{H} Q^*_{b(x)} M^*_{c(x)} f(x) = \\ \int e^{i\left(c_x y + 2b_x x y - i b_x y^2\right)} f(x-y) \frac{\mathrm{d}y}{y} = \int e^{i\left(l_x(x) y - i b_x y^2\right)} f(x-y) \frac{\mathrm{d}y}{y}. \end{aligned} \tag{6}$$

where we denote by l_x the line $l_x(z) = c_x + 2b_x z$. We will denote $\mathcal{C}_{2,b,c}$ simply by $\mathcal{C}$.

Definition 3. *Identifying any tile P with the triplet of intervals $(I_P, \alpha_P, \omega_P)$ as in Figure 2, let*

$$\mathbb{P} = \left\{ \textit{tile } P \subset \mathbb{T} \times \mathbb{R} \mid I_P, \alpha_P, \omega_P \textit{ dyadic intervals, } |I_P|^{-1} = |\alpha_P| = |\omega_P| = 2^{-k} \right\}.$$

Furthermore we say that a line $l_x(z) = c_x + 2b_x z$ satisfies $l_x \in P$ if $x \in I_p$ and $(z, l_x(z)) \in P$ for all $z \in P$. Also set $E(P) = \{x \in I_P \mid l_x \in P\}$ (see Figure 3).

Two tiles $P \neq P' \in \mathbb{P}$ with the same spatial interval $I_P = I_{P'}$ cannot contain a line in common i.e. if $|I_P| = |I_{P'}|$ then $E(P)$ and $E(P')$ are disjoint. Setting $\frac{1}{y} = \sum_{k \in \mathbb{Z}} \psi_{2^k}(y) = \sum_{k \in \mathbb{Z}} 2^{-\frac{k}{2}} \psi\left(\frac{y}{2^k}\right)$ for some zero-mean, C^∞_c function ψ we have

$$\mathcal{C} f(x) = \sum_{k \in \mathbb{Z}} \sum_{\substack{P \in \mathbb{P} \\ |I_P| = 2^k \\ l_x \in P}} \left(\int_{\mathbb{T}} e^{i\left(l_x(x) y - i b_x y^2\right)} f(x-y) \psi_{2^k}(y) \, \mathrm{d}y \right) = \sum_{P \in \mathbb{P}} T_P f(x)$$

where

$$
\begin{aligned}
T_P f(x) &= \left(\int_{\mathbb{T}} e^{i\left(l_x(x)y - ib_x y^2\right)} f(x-y)\psi_{|I_P|}(y)\,\mathrm{d}y \right) \not\Vdash_{E(P)}(x) \\
T_P^* f(x) &= \int_{\mathbb{T}} e^{i\left(l_{(x-y)}(x-y)y - ib_{(x-y)}y^2\right)} \left(f \not\Vdash_{E(P)}\right)(x-y)\ \ \psi_{|I_P|}(y)\,\mathrm{d}y.
\end{aligned}
\tag{7}
$$

In the proof we are allowed to "rarefy" somewhat the operator (the space of tiles) by sacrificing a constant in our bound. Under these assumptions it can be seen that the operators T_P and T_P^* have time-frequency localization controlled by the tile P

Definition 4. *Given two tiles $P, P' \in \mathbb{P}$, suppose that $|I_{P'}| \leq |I_P|$ then we define their phase distance as*

$$
\Delta(P,P') = \inf_{\substack{l \in P \\ l' \in P'}} \sup_{x \in I_{P'}} \frac{|l(x) - l'(x)|}{|I_{P'}|^{-1}} \qquad \langle \Delta(P,P') \rangle = 1 + \Delta(P,P') \tag{8}
$$

Lemma 5 (Time-frequency localization)**.** *Given a tile P let $I_P^* = 2I_P \pm \frac{9}{2}|I_P|$. Then $\operatorname{spt} T_P f \subset I_P$ and $\operatorname{spt} T_P^* f \subset I_P^*$. Furthermore given two tiles $P_1, P_2 \in \mathbb{P}$*

$$
\begin{aligned}
\left| \int_{\mathbb{T}} \tilde{1}_{I_{1,2}} T_{P_1}^* f \overline{T_{P_2}^* g} \right| &\leq C_\epsilon \langle \Delta(P_1,P_2) \rangle^{-\left(\frac{1}{2}-\epsilon\right)} \min\left(|I_{P_1}|; |I_{P_2}|\right) \frac{\int_{E(P_1)} |f|}{|I_{P_1}|} \frac{\int_{E(P_2)} |g|}{|I_{P_2}|} \\
\left| \int_{\mathbb{T} \setminus I_{1,2}} T_{P_1}^* f \overline{T_{P_2}^* g} \right| &\leq C_N \langle \Delta(P_1,P_2) \rangle^{-N} \min\left(|I_{P_1}|; |I_{P_2}|\right) \frac{\int_{E(P_1)} |f|}{|I_{P_1}|} \frac{\int_{E(P_2)} |g|}{|I_{P_2}|}
\end{aligned}
$$

for arbitrarily large N and some small $\epsilon > 0$. Where $I_{1,2}$ is a small (ϵ)-critical intersection interval for the pair P_1 and P_2 centered at the intersection of the lines of P_1 and P_2 and of length $|I_{1,2}| = \langle \Delta(P_1,P_2) \rangle^{-\left(\frac{1}{2}-\epsilon\right)} \min\left(|I_{P_1}|; |I_{P_2}|\right)$.

This lemma is proved using standard stationary phase integral techniques and illustrates that operators associated to tiles that are far away or with high incidence angle interact weakly and the interaction is concentrated along a critical intersection interval.

12.5 Geometry and combinatorics

It is necessary to group the tiles in $\mathbb{P}$ into sets that behave uniformly with respect to different conjugated versions of the Hilbert transform. These objects will be trees and then groups of trees: forests. The main difficulty of the proof of Theorem 1 consists in controlling interaction between trees in a forest. This is the most technically involved part of the article that solves this problem by cutting away small enough spatial intervals around the area where the time-frequency supports of the trees intersect in a manner somewhat similar to the one suggested by Lemma 5.

Definition 6 (Relations between tiles)**.** *Let P and P' be two tiles. We say $P \leq P'$ if $I_P \subseteq I_{P'}$ and there exists a line l such that $l \in P$ and $l \in P'$*

Definition 7 (Mass)**.** *The mass is a measure of auto-interaction of tiles*

$$A(P) = \sup_{\substack{P' \in \mathbb{P} \\ I_{P'} \supseteq I_P}} \frac{|E(P')|}{|I_{P'}|} \langle \Delta(P, P') \rangle^{-N} \tag{9}$$

The proof of Theorem 1 is based on subdividing all tiles into groups with uniformly controlled mass. A careful choice of constants for each mass level $\delta_n \approx 2^{-n}$ yields a factor of δ^η in bounds on operators associated to tiles, trees, and forests. That allows the estimates to be finally summed up.

Apart from technical details of the definition, a tree groups together strongly interacting tiles which, together, determine an operator that behaves like a truncated Hilbert transform. This statement would actually be precise if the linearization functions b, c were constant. Otherwise one must account for small deviations in the phase factor.

Definition 8 (Forest)**.** *A set of tiles $\mathcal{P} \subset \mathbb{P}$ that is a union family of trees $\mathcal{P} = \bigcup_j \mathcal{P}_j$ with tops $\{\tilde{P}_j\}_j$ is a (δ, K)-forest (with parameter $K > 0$) for some $0 < \delta \leq 1$ if it satisfies the following conditions*

1. *Uniform mass bounds on the tiles of the trees i.e. for all tiles of all trees $P \in \mathcal{P}$ we have $A(P) < \delta$.*

2. *Trees have controlled tile overlap i.e. if $P \in \mathcal{P}_j$ then $2P \not\leq \tilde{P}_k$ for $k \neq j$.*

3. *Trees have controlled spatial overlap i.e. $\#\{\mathcal{P}_j \mid x \in I_{\tilde{P}_j}\} < K\delta^{-2}$ for all point x.*

Proposition 9 (Forest estimate)**.** *A (δ, K)-forest $\mathcal{P} = \{\mathcal{P}_j\}$ satisfies $\|\sum_j T_{\mathcal{P}_j} f\|_{L^2(F^c)} \leq C\delta^{\tilde{\eta}} \log K \|f\|_2$ outside an exceptional set F with $|F| \leq C\delta^{50} K^{-1}$, for some constant $\eta \in (0, \frac{1}{2})$.*

The statement of Theorem 1 follows optimizing the choice of the constant K for each super-level set of $\mathcal{C}$ and integrating.

References

[1] Lie, V. *The (Weak-L 2) Boundedness of the Quadratic Carleson Operator* Geometric and Functional Analysis, 19 (2009), no. 2, pp. 457–497, Springer;

[2] Lacey, M. and Thiele, C. *A proof of boundedness of the Carleson operator* Mathematical Research Letters, 7 (2000), no. 4, pp. 361–370, International Press;

[3] Fefferman, C. *Pointwise convergence of Fourier series* Annals of Mathematics, 1973, pp. 551–571.

Gennady Uraltsev, Universität Bonn
email: **guraltse@math.uni-bonn.de**

13 On maximal ergodic theorem for certain subsets of the integers

after J. Bourgain [1]
A summary written by Bartosz Trojan

Abstract

We prove L^2-boundedness of a maximal function for averages along the squares $(n^2 : n \in \mathbb{N})$.

13.1 Introduction

Let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space with an invertible measure preserving transformation $T : X \to X$. We consider the averages along the squares

$$A_N f(x) = N^{-1} \sum_{n=1}^{N} f(T^{n^2} x)$$

for $f \in L^2(X, \mu)$. We prove

Theorem 1 ([1, Theorem 1]). *There is $C > 0$ such that for any $f \in L^2(X, \mu)$*

$$\big\| \sup_{N \in \mathbb{N}} |A_N f| \big\|_{L^2} \leq C \|f\|_{L^2}.$$

As an application we show

Theorem 2 ([1, Theorem 5]). *For any $f \in L^2(X, \mu)$ there is $f^* \in L^2(X, \mu)$ such that*

$$\lim_{N \to \infty} A_N f(x) = f^*(x)$$

μ-almost everywhere on X.

13.2 Maximal function

In view of Calderón's transference principle we reduce proving Theorem 1 to a model dynamical system $\mathbb{Z}$ with the counting measure and the bilateral shift operator. For $f \in \ell^2(\mathbb{Z})$ we set

$$M_N f(x) = N^{-1} \sum_{n=1}^{N} f(x - n^2).$$

We are going to prove

Theorem 3. *There is $C > 0$ such that for any $f \in \ell^2(\mathbb{Z})$*

$$\big\| \sup_{N \in \mathbb{N}} \big| M_N f \big| \big\|_{\ell^2} \leq C \|f\|_{\ell^2}.$$

Let us observe that we may assume $f \geq 0$ and restrict the supremum to the set

$$\mathcal{D} = \big\{ \lfloor \tau^n \rfloor : n \in \mathbb{N} \big\}$$

for any $\tau \in (1, 2]$. For each $j \in \mathbb{N}$ we have

$$M_{2^j} f = \mathcal{F}^{-1}\big(m_j \hat{f} \big)$$

where

$$m_j(\xi) = \tau^{-j} \sum_{n=1}^{\tau^j} e^{2\pi i \xi n^2}.$$

Let

$$\Phi_j(\xi) = \tau^{-j} \int_0^{\tau^j} e^{2\pi i \xi x^2} dx.$$

Then integration by parts implies

$$|\Phi_j(\xi)| \lesssim \min\{|\xi \tau^{2j}|, |\xi \tau^{2j}|^{-1/2}\}. \tag{1}$$

In what follows we also need a *Gaussian sums* defined for any rational number $a/q \in \mathbb{Q}$ by

$$G(a/q) = q^{-1} \sum_{r=1}^{q} e^{2\pi i (a/q) r^2}.$$

Let us recall that

$$|G(a/q)| \lesssim q^{-1/2}. \tag{2}$$

To better understand the multiplier m_j we apply Hardy and Littlewood circle method. First, using mean value theorem we show

Proposition 4. *Let $1 \leq q \leq \tau^{j\alpha}$, $\gcd(a, q) = 1$. There is $C > 0$ such that if*

$$\Big| \xi - \frac{a}{q} \Big| \leq \tau^{-2j(1-\alpha)}$$

then

$$\big| m_j(\xi) - G(a/q) \Phi_j(\xi - a/q) \big| \leq C \tau^{-j/2}$$

provided $\alpha < 1/10$.

We fix $\rho \in (0,1)$. Let $\eta : \mathbb{R} \to \mathbb{R}$ be a smooth function such that $0 \le \eta \le 1$ and

$$\eta(\xi) = \begin{cases} 1 & \text{for } |\xi| \le 1/4, \\ 0 & \text{for } |\xi| \ge 1/2. \end{cases}$$

We may assume η is a convolution of two smooth nonnegative functions with compact supports contained in $[-1/2, 1/2]$. For any $s \in \mathbb{N} \cup \{0\}$ we define a multiplier

$$\nu_j^s(\xi) = \sum_{a/q \in \mathscr{R}_s} G(a/q) \Phi_j(\xi - a/q) \eta_s(\xi - a/q)$$

where $\eta_s(\xi) = \eta\big(2^{2^{\rho(s+1)}}\xi\big)$ and

$$\mathscr{R}_s = \big\{a/q \in \mathbb{Q} : 2^s \le q < 2^{s+1}, 1 \le a \le q, \text{ and } \gcd(a,q) = 1\big\}.$$

Let $\nu_j = \sum_{s \ge 0} \nu_j^s$. We need one more tool (see [3]).

Lemma 5 (Weyl's inequality)**.** *Let $\xi \in \mathbb{T}$ and $a, q \ge 1$, $\gcd(a,q) = 1$ such that*

$$\Big| \xi - \frac{a}{q} \Big| \le \frac{1}{q^2}.$$

Then

$$\Big| \sum_{n=0}^{N} e^{2\pi i \xi n^2} \Big| \lesssim N^{-1} q^{1/2} + \big(N \log q\big)^{1/2} + \big(q \log q\big)^{1/2}.$$

Based on Weyl's inequality, Proposition 4 and estimates (1) and (2) we show

Proposition 6. *There is $C > 0$ such that for all $\xi \in \mathbb{T}$*

$$\big| m_j(\xi) - \nu_j(\xi) \big| \le C j^{-1/(2\rho)}.$$

Now, let us observe that we may replace the multiplier m_j by ν_j. Indeed, by Plancherel's theorem

$$\begin{aligned} \big\| \sup_{j \ge 1} \big| \mathcal{F}^{-1}\big(m_j \hat{f}\big) - \mathcal{F}^{-1}\big(\nu_j \hat{f}\big) \big| \big\|_{\ell^2}^2 &= \big\| \sup_{j \ge 1} \big| \mathcal{F}^{-1}\big(m_j \hat{f}\big) - \mathcal{F}^{-1}\big(\nu_j \hat{f}\big) \big|^2 \big\|_{\ell^1} \\ &\le \sum_{j \ge 1} \big\| m_j - \nu_j \big\|_{L^\infty}^2 \|f\|_{\ell^2}^2 \le \sum_{j \ge 1} j^{-1/\rho} \|f\|_{\ell^2}^2. \end{aligned}$$

Therefore it is enough to prove

Theorem 7. *There are $C > 0$ and $\delta' > 0$ such that for any $f \in \ell^2(\mathbb{Z})$*

$$\big\| \sup_{j \geq 1} \big| \mathcal{F}^{-1}\big(\nu_j^s \hat{f}\big) \big| \big\|_{\ell^2} \leq C 2^{-\delta' s} \|f\|_{\ell^2}.$$

Proof. For any $0 < \delta < \rho$ we split the set $\{2^s, \ldots, 2^{s+1} - 1\}$ into $2^{s(1-\delta)}$ subsets Λ_k each of which contains at most $2^{s\delta}$ elements. Then, by change of variables

$$\mathcal{F}^{-1}\big(\nu_j^s \hat{f}\big)(x) = \sum_{1 \leq k \leq 2^{s(1-\delta)}} \mathcal{F}^{-1}\big(\Phi_j \eta_s F_k(\cdot; x)\big)(x)$$

where

$$F_k(\xi; x) = \sum_{q \in \Lambda_k} \sum_{\substack{1 \leq a \leq q \\ \gcd(a,q)=1}} G(a/q) e^{2\pi i (a/q) x} \hat{f}(\xi + a/q).$$

If Q_k denotes the common multiplicity of $q \in \Lambda_k$ then

$$\begin{aligned}
\big\| \sup_{j \geq 1} &\big| \mathcal{F}^{-1}\big(\Phi_j \eta_s F_k(\cdot; x)\big)(x) \big| \big\|_{\ell^2(x)}^2 \\
&= \sum_{l=1}^{Q_k} \big\| \sup_{j \geq 1} \big| \mathcal{F}^{-1}\big(\Phi_j \eta_s F_k(\cdot; l)\big)(Q_k x + l) \big| \big\|_{\ell^2(x)}^2.
\end{aligned}$$

Next, we show that there is $C > 0$ such that for any $f \in \ell^2(\mathbb{Z})$ and $l \in \{1, \ldots, Q_k\}$

$$\big\| \sup_{j \geq 1} \big| \mathcal{F}^{-1}\big(\Phi_j \eta_s \hat{f}\big)(Q_k x + l) \big| \big\|_{\ell^2(x)} \leq C \big\| \mathcal{F}^{-1}\big(\eta_s \hat{f}\big)(Q_k x + l) \big\|_{\ell^2(x)}. \tag{3}$$

Let us denote the left-hand-side of (3) by J_l. Since $\eta_s = \eta_s \eta_{s-1}$ we get

$$\sum_{l=1}^{Q_k} J_l^2 = \big\| \sup_{j \geq 1} \big| \mathcal{F}^{-1}\big(\Phi_j \eta_s \hat{f}\big) \big| \big\|_{\ell^2}^2 \lesssim \|\mathcal{F}^{-1}\big(\eta_s \hat{f}\big)\|_{\ell^2}^2. \tag{4}$$

We need the following two Lemmas.

Lemma 8 ([2, Lemma 1])**.** *There is $C > 0$ such that for any $s \geq 1$ and $u \in \mathbb{R}$*

$$\Big\| \int_{\mathbb{T}} e^{-2\pi i \xi x} \eta_s(\xi) d\xi \Big\|_{\ell^1(x)} \leq C,$$

$$\Big\| \int_{\mathbb{T}} e^{-2\pi i \xi x} \big(1 - e^{2\pi i \xi u}\big) \eta_s(\xi) d\xi \Big\|_{\ell^1(x)} \leq C |u| 2^{-2^{\rho s}}.$$

Lemma 9 ([2, Lemma 2]). *For any $s \geq 1$ and $l \in \{1, \ldots, Q_k\}$*

$$\big\|\mathcal{F}^{-1}\big(\eta_s \hat{f}\big)(Q_k x + l)\big\|_{\ell^2}(x) \simeq Q_k^{-1/2}\big\|\mathcal{F}^{-1}\big(\eta_s \hat{f}\big)\big\|_{\ell^2}.$$

Now, using Lemma 8 and estimate $Q_k \leq 2^{s2^{\delta s}}$ we get

$$|J_l - J_{l'}| \leq C Q_k 2^{-2^{\rho s}} \big\|\mathcal{F}^{-1}\big(\eta_s \hat{f}\big)\big\|_{\ell^2} \lesssim \big\|\mathcal{F}^{-1}\big(\eta_s \hat{f}\big)\big\|_{\ell^2},$$

thus by (4)

$$Q_k J_l^2 \lesssim \big\|\mathcal{F}^{-1}\big(\eta_s \hat{f}\big)\big\|_{\ell^2}^2$$

what together with Lemma 9 implies (3). Finally, by Plancherel's theorem, estimate (2) and disjointness of supports of $\eta_s(\cdot - a/q)$'s while a/q varies over $\mathscr{R}_s$ we get

$$\begin{aligned} \Big\|\sup_{j \geq 1}\big|\mathcal{F}^{-1}\big(\Phi_j \eta_s F_k(\cdot; x)\big)(x)\big|\Big\|_{\ell^2(x)}^2 &\lesssim \sum_{l=1}^{Q_k} \big\|\mathcal{F}^{-1}\big(\eta_s F_k(\cdot; l)\big)(Q_k x + l)\big\|_{\ell^2(x)}^2 \\ &= \Big\|\sum_{q \in \Lambda_k} \sum_{a \in A_q} G(a/q) \mathcal{F}^{-1}\big(\eta_s(\cdot - a/q)\hat{f}\big)\Big\|_{\ell^2}^2 \lesssim 2^{-s}\|f\|_{\ell^2}^2. \end{aligned}$$

Summing up over $1 \leq k \leq 2^{s(1-\delta)}$ we obtain

$$\Big\|\sup_{j \geq 1}\big|\mathcal{F}^{-1}\big(\nu_j^s \hat{f}\big)\big|\Big\|_{\ell^2} \lesssim 2^{s(1/2-\delta)}\|f\|_{\ell^2}.$$

Therefore, it is enough to take $1/2 < \delta < \rho < 1$ to finish the proof. □

13.3 Pointwise convergence

Suppose for some bounded function $f \in L^2(X, \mu)$ a sequence $(A_N f : N \in \mathbb{N})$ does not converge μ-almost everywhere. Then there is $\epsilon > 0$ such that

$$\mu\Big\{x \in X : \limsup_{M,N \to \infty}\big|A_N f(x) - A_M f(x)\big| > 4\epsilon\Big\} > 4\epsilon.$$

Now, we construct a strictly increasing sequence of integers $(k_j : j \in \mathbb{N})$ such that for each j

$$\mu\Big\{x \in X : \sup_{N_j \leq N \leq N_{j+1}}\big|A_N f(x) - A_{N_j} f(x)\big| > \epsilon\Big\} > \epsilon$$

where $N_j = \tau^{k_j}$ and $\tau = 1 + \epsilon/(2\|f\|_{L^\infty})$. If $\tau^k \leq N < \tau^{k+1}$ then

$$|A_N f(x) - A_{\tau^k} f(x)| \leq (\tau - 1)\|f\|_{L^\infty} \leq \epsilon/2.$$

In particular,

$$\mu\big\{x \in X : \sup_{\tau^k \in \mathcal{D}_j} \big|A_{\tau^k} f(x) - A_{N_j} f(x)\big| > \epsilon/2\big\} > \epsilon$$

where $\mathcal{D}_j = \mathcal{D} \cap (N_j, N_{j+1}]$. Using Calderón transference principle it is enough to prove

Theorem 10. *For each $J \in \mathbb{N}$ there is $C_J > 0$ such that*

$$\sum_{j=0}^{J} \Big\| \sup_{\tau^k \in \mathcal{D}_j} \big|M_{\tau^k} f(x) - M_{N_j} f(x)\big| \Big\|_{\ell^2}^2 \leq C_J \big\|f\big\|_{\ell^2}^2$$

and $\lim_{J \to \infty} C_J / J = 0$.

References

[1] Bourgain, J., *On the maximal egodic theorem for certain subsets of the integers*, Isr. J. Math., Vol. 61, no. 1 (1988), 39–72

[2] Mirek, M. and Trojan B., *Cotlar's ergodic theorem along the prime numbers*, preprint, 2014

[3] Sarközy, A., *On difference sets of sequences of integers, I*, Acta Math. Acad. Sci. Hung., 31, (1978), 125–149

Trojan Bartosz, University of Wroclaw
email: **trojan@math.uni.wroc.pl**

14 Extention of a multi-frequency maximal inequality of Bourgain

after C.Thiele, F.Nazarov and R.Oberlin [1]
A summary written by Ioann Vasilyev

Abstract

We give a proof of a strenghtened version of Bourgain's multi-frequency maximal inequality. Proof contains one nice version of Calderon-Zygmund decomposition.

14.1 Introduction

In our talk we will present a variant of a Calderon-Zygmund decomposition where the mean zero condition is replaced by a collection of conditions for a number of frequencies. Precisely the following theorem holds.

Theorem 1. *There exists a universal constant C, such that the following holds. Let $\xi_1 < \cdots < \xi_N$ be any real numbers for some $N \geq 1$. Let $f \in L^1(\mathbb{R})$ and let $\lambda > 0$. Then there exists a decomposition*

$$f = g + \sum_{I \in \mathbf{I}} b_I$$

for some disjoint collection $\mathbf{I}$ of intervals, for which

$$\sum_{I \in \mathbf{I}} |I| \leq C N^{1/2} ||f||_1 \lambda^{-1}.$$

Also for any $I \in \mathbf{I}$?? $1 \leq j \leq N$ the following holds:

$$||g||_2^2 \leq C ||f||_1 N^{1/2} \lambda$$
$$||f_I||_1 \leq C |I| \lambda$$
$$||f_I - b_I||_2 \leq C |I|^{1/2} \lambda N^{1/2}$$
$$\int b_I(x) e^{i\xi_j x} dx = 0,$$

where f_I - is a product of f with characteristic function of an interval I. Finally the support of b_I is $3I$ (e.g. the interval with the same center, as I and with 3 times the length.)

Theorem 1 will help us to prove an extention of multi-frequency maximal inequality of Bourgain.

Let us introduce some notaions. For each dyadic interval

$$\omega = [2^k n, 2^k(n+1))$$

with $k, n \in \mathbb{Z}$ let ϕ_ω be a Schwartz function whose Fourier transform $\hat{\phi}_\omega$ is supported on ω. Let $\xi_1 < \cdots < \xi_N$ be real numbers and denote by X the set $\{\xi_1, \ldots, \xi_N\}$. We are interested in bounds for the vector valued operator

$$\Delta_k[f] = \sum_{\substack{|w|=2^k \\ \omega \cap X \neq \emptyset}} f * \phi_\omega$$

whose vector components are parameterized by the integer k.

Definition 2. *For $1 < r < \infty$, define the* **r - variation seminorms** *of a sequence g_k by*

$$||g_k||_{\tilde{V}_k^r} := \sup_{M, k_0 < \ldots k_M} \left(\sum_{j=1}^{M} |g_{k_j} - g_{k_{j-1}}|^r \right)^{1/r}$$

where the supremum is over all strictly increasing finite sequences k_j of arbitrary finite length $M+1$ and define the **variation norm**

$$||g_k||_{V_k^r} := \sup_k |g_k| + ||g_k||_{\tilde{V}_k^r}.$$

When $r = \infty$, we replace the sum by a supremum in the usual manner.

It was proven in [3] that for $r > 0$ we have

$$||\Delta_k[f](x)||_{L_x^2(L_k^\infty)} \leq (1+\log(N)) N^{\frac{1}{2}-\frac{1}{r}} \left(D_1 + \sup_{j=1,\ldots,N} || \sum_{|\omega|=2^k} \hat{\phi}_\omega(\xi_j) ||_{V_k^r} \right) ||f||_{L^2}$$

with the convention

$$D_M := \sup_{\omega, x} |\omega|^M |\hat{\phi}_\omega^{(M)}(x)|$$

for any integer $M \geq 0$ where the supremum is over all dyadic intervals ω, real numbers x, and where $\hat{\phi}_\omega^{(M)}$ is the M-th derivative of $\hat{\phi}_\omega$. This is a weighted version of the above mentioned bound of Bourgain???s originating in [2]. Our aim is to strengthen this result in two directions. First, we would like to replace L^2 by L^p for $1 \leq p < 2$. Second we would like to replace the L^∞_k norm by the stronger q- variation norm. Specifically, we will show:

Theorem 3. *Suppose* $1 < p \leq 2 < r < q$. *Then there exists an* M *depending only on* q *and* r *such that*

$$||\Delta_k[f](x)||_{L^p_x(V^q_k)} \leq$$
$$C_{p,q,r}(1+\log(N))N^{(\frac{1}{2}-\frac{1}{r})\frac{q}{q-2}+\frac{1}{p}-\frac{1}{2}}\left(D_M + \sup_{j=1,\dots,N}||\sum_{|\omega|=2^k}\hat{\phi}_\omega(\xi_j)||_{V^r_k}\right)||f||_{L^p}. \tag{1}$$

Theorem 3 is the main result of this talk.

Now we provide a short proof of a nice Calderon-Zygmund decomposition(Theorem 1).

14.2 Proof of Theorem 1. (e.g. of a variant of a Calderon-Zygmund decomposition)

Proof. Let $f \in L^1(\mathbb{R})$. Consider the set

$$E = \{x : \mathcal{M}[f](x) > \lambda N^{-1/2}\}$$

where $\mathcal{M}$ is the Hardy-Littlewood maximal operator. By the Hardy-Littlewood maximal theorem we have

$$|E| \leq \frac{CN^{1/2}||f||_1}{\lambda}.$$

Let $\mathbf{I}$ be the collection of maximal dyadic interval contained in E such that $6I$ is also contained in E. Clearly the collection $\mathbf{I}$ covers E and the collection of intervals $3I$ has bounded overlap.

Consider the finite dimensional subspace

$$A := \text{span}\{e^{i\xi_j x} : 1 \leq j \leq N\}$$

of the Hilbert space $L^2(3I)$. For each element v in this space let us prove the estimate

$$||v||_{L^\infty(I)} \leq N^{1/2}|I|^{-1/2}||v||_{L^2(3I)}.$$

Let $v_1, \ldots, v_N$ be an ortonormal basis of A, considered as subspace of $L^2(I)$. Since

$$\int_I \sum_{j=1}^N |v_j(x)|^2 dx = N,$$

there exists a point $x_0 \in I$ such that

$$|I| \sum_{j=1}^N |v_j(x_0)|^2 \leq N.$$

Hence for every element v in A,

$$|v(x_0)| \leq \sum_{j=1}^N |\langle v, v_j\rangle v_j(x_0)| \leq$$

$$\leq (\sum_{j=1}^N |\langle v, v_j\rangle|^2)^{1/2} (\sum_{j=1}^N |v_j(x_0)|^2)^{1/2} \leq N^{1/2}|I|^{-1/2}||v||_{L^2(I)}.$$

To estimate v at general point $x_1 \in I$ we apply this estimate to

$$\tilde{v}(x) = v(x - x_0 + x_1)$$

Which is also in A, and thus we get the required inequality.
Now we see that the mapping

$$v \to v \cdot f_I$$

is a linear bounded functional on the subspace A of $L_2(3I)$ with norm bounded by $N^{1/2}|I|^{-1/2}||f_I||_1$. That is so because

$$\left|\int v \cdot f_I\right| \leq ||f_I||_1 \cdot ||v||_{L^\infty(I)} \leq ||v||_{L^2(3I)} ||f_I||_1 N^{1/2}|I|^{-1/2}$$

Now by the Riesz representation theorem, there is an element g_I in A such that

$$\int f_I(y) e^{2\pi i \xi_j y} dy = \int_{3I} g_I(y) e^{2\pi i \xi_j y}$$

and such that

$$||g_I||_{L^2(3I)} \leq N^{1/2}|I|^{-1/2}||f_I||_1$$

We extend g_I to a function on all $\mathbb{R}$ by setting it equal to 0 outside $3I$.

For each $I \in \mathbf{I}$ consider the restriction f_I of f to I and observe that by looking at the maximal function on $12I$ we have

$$||f||_1 \leq 24|I|\lambda N^{-1/2}$$

Define

$$b_I = f_I - b_I,$$
$$b = \sum_{\mathbf{I}} b_I,$$
$$g = f - b.$$

Function b is supported on E. The functions g_I have bounded overlap, hence

$$||g||_2^2 \leq \int\limits_{E^c} |f(x)|^2 dx + \int (\sum_{\mathbf{I}} g_I(x))^2 dx$$

$$\leq \int\limits_{E^c} |f(x)|\lambda N^{-1/2} dx + C \sum_{\mathbf{I}} \int g_I^2(x) dx$$

$$\leq ||f||_1 \lambda N^{-1/2} + C \sum_{\mathbf{I}} \frac{N}{|I|} |I|^2 \lambda^2 \frac{1}{N} \leq ||f||_1 \lambda N^{-1/2} + C\lambda^2 |E| \leq$$

$$\leq C||f||_1 \lambda N^{-1/2}$$

□

References

[1] F.Nazarov, R.Oberlin, C.Thiele, *A Calderon Zygmund decomposition for multiple frequencies and an application to an extension of a lemma of Bourgain.* Math. Res. Lett. **17** (2010), no. 3, 529–545;

[2] Jean Bourgain, *On the maximal ergodic theorem for certain subsets of the integers,* Inst. Hautes Études Sci. Publ. Math. (1989), no. 69, 5–45.

[3] Ciprian Demeter, Michael T. Lacey, Terence Tao, and Christoph Thiele, *Breaking the duality in the return times theorem,* Duke Math. J. **143** (2008), no. 2, 281–355.

Ioann Vasilyev, SPbSU
email: **milavas@mail.ru**

15 Cotlar's ergodic theorem along the prime numbers

after M. Mirek and B. Trojan [1]
A summary written by Michal Warchalski

Abstract

We prove Cotlar's ergodic theorem modeled on the set of primes.

15.1 Introduction

We consider a dynamical system $(X, \mathcal{B}, \mu, S)$ on a measure space X endowed with a σ-algebra $\mathcal{B}$, a σ-finite measure μ and an invertible measure preserving transformation $S\colon X \to X$. The almost everywhere convergence of the ergodic truncated Hilbert transform

$$\lim_{N\to\infty} \sum_{1\le|n|\le N} \frac{f(S^n x)}{n}$$

for $f \in L^r(\mu)$, $1 \le r < \infty$ was proven by Cotlar[5] in 1955. Motivated by this result we show the corresponding theorem with the natural numbers replaced by the set of the prime numbers.

Theorem 1. *For a given dynamical system $(X, \mathcal{B}, \mu, S)$ the almost everywhere convergence of the ergodic truncated Hilbert transform along $\mathbb{P}$*

$$\lim_{N\to\infty} \sum_{p\in\pm\mathbb{P}_N} \frac{f(S^p x)}{p} \log|p|$$

holds for all $f \in L^r(\mu)$ with $1 < r < \infty$.

We obtain the aforementioned theorem using transference principle showing first the respective theorem for the set of integers and the counting measure. Let $K \in C^1(\mathbb{R} \setminus \{0\})$ be a Calderón-Zygmund kernel satisfying for $|x| \ge 1$

$$|x||K(x)| + |x|^2|K'(x)| \le 1$$

as well as having a cancellation property

$$\sup_{\lambda\ge1} \left| \int_{1\le|x|\le\lambda} K(x)dx \right| \le 1.$$

Define Tf for a function $f\colon \mathbb{Z} \to \mathbb{C}$ as

$$Tf(n) = \sum_{p\in\mathbb{P}} f(n-p)K(p)\log|p|.$$

We also introduce the truncation T_N and the maximal function T^*

$$T_N f(n) = \sum_{p\in\mathbb{P}_N} f(n-p)K(p)\log|p|$$

$$T^* f(n) = \sup_{N\in\mathbb{N}} |T_N f(n)|.$$

Now we can state the announced result.

Theorem 2. *The maximal function $T^*f(n)$ is bounded on $\ell^r(\mathbb{Z})$ for any $1 < r < \infty$. Moreover, the pointwise limit*

$$\lim_{N\to\infty} T_N f(n)$$

exists and coincides with the Hilbert transform Tf which is also bounded on ℓ^r for any $1 < r < \infty$.

We follow Bourgain's approach from the sequence of his papers [2], [3], [4] for $r = 2$ using Hardy and Littlewood circle method. However for $r \neq 2$, we show two crucial lemmas which simplify Bourgain's arguments making them more elementary, what we will see in the talk *On the maximal ergodic theorem for certain subsets of the integers.*

15.2 ℓ^2 theory

We fix $\tau \in (1, 2]$ and consider localizations of the kernel K.

$$K_j(x) = K(x)\mathbb{1}_{|x|\in(\tau^j,\tau^{j+1}]}$$

Thus now we are dealing with a sequence of multipliers m_j given by

$$m_j(\xi) = \sum_{p\in\pm\mathbb{P}} e^{2\pi i\xi p} K_j(p)\log|p|.$$

As mentioned before in the ℓ^2 case we explore Hardy and Littlewood circle method. For any $\alpha > 0$ to be properly chosen later and fixed $j \in \mathbb{N}$ we define major arcs as

$$\mathcal{M}_j = \bigcup_{1\le q\le j^\alpha} \bigcup_{a\in A_q} \mathcal{M}_j(a/q)$$

where

$$\mathcal{M}_j(a/q) = \{\xi \in [0,1] \colon |\xi - a/q| \leq \tau^{-j} j^{\alpha}\}$$

and A_q is the set of natural numbers smaller than q and relatively prime with q. Minor arcs are just ξ's, such that $\xi \in [0,1] \setminus \mathcal{M}_j$. The idea is that on major arcs, using certain disjointness properties, we can approximate m_j's well by easier controllable functions ν_j such that

$$\nu_j(\xi) = \sum_{a/q \in \mathbb{Q},\, (a,q)=1} \frac{\mu(q)}{\varphi(q)} \Phi_j(\xi - a/q) \eta_{\lfloor \log q / \log 2 \rfloor}(\xi - a/q)$$

where μ denotes Möbius function, φ Euler's totient function, Φ_j is the Fourier transform of K_j and $(\eta_n : n \in \mathbb{N})$ is a sequence of smooth cutoff functions. On the other hand on minor arcs both m_j and ν_j are small(for m_j's it is shown applying Vinogradov's theorem).

15.3 Oscillatory norm estimate

Let H_N denote the truncated Hilbert transform along primes

$$H_N f(n) = \sum_{p \in \mathbb{P}_N} \frac{f(n-p)}{p} \log |p|$$

Our key to show pointwise convergence in the general setting $(X, \mathcal{B}, \mu, S)$ is to show a type of oscillatory norm estimate, which we first prove on $\mathbb{Z}$ and then transfer to X. Let $\tau \in (1,2]$ be fixed, $\Lambda := \{\tau^n \colon n \in \mathbb{N}\}$, $(k_j \colon j \in \mathbb{N})$ be a strictly increasing sequence of the natural numbers, $N_j := \tau^{k_j}$ and $\Lambda_j := \Lambda \cap (N_j, N_{j+1}]$. We have

Theorem 3. *For every $J \in \mathbb{N}$ there is C_J such that*

$$\sum_{j=0}^{J} \| \sup_{\tau^k \in \Lambda_j} |H_{\tau^k} f - H_{N_j} f| \|_{\ell^2}^2 \leq C_J \|f\|_{\ell^2}^2$$

and $\lim_{J \to \infty} C_J / J = 0$.

15.4 Obtaining Theorem 1 from Theorem 2

Here we outline the transfer of the result from Theorem 2 to Theorem 1. Recall, we consider the counterpart of the truncated Hilbert transform along the prime numbers for a general dynamical system $(X, \mathcal{B}, \mu, S)$ given by

$$\mathcal{H}_N f(x) = \sum_{p \in \mathbb{P}_N} \frac{f(S^{-p}x)}{p} \log |p|.$$

Let k_j, N_j, Λ_j be as before. First, we have the corresponding oscillatory norm type estimate for truncations $\mathcal{H}_N$. This is essentially our transference principle for the oscillatory norm.

Proposition 4. *For each $J \in \mathbb{N}$ there is C_J such that*

$$\sum_{j=0}^{J} \| \sup_{N \in \Lambda_j} |\mathcal{H}_N - \mathcal{H}_{N_j}| \|_{L^2(\mu)}^2 \leq C_J \|f\|_{L^2(\mu)}^2$$

and $\lim_{J \to \infty} C_J / J = 0$.

The proof of the Theorem 1 can now be obtained.

Theorem 5. *Let $f \in L^r(\mu)$, $1 < r < \infty$. For μ-almost every $x \in X$*

$$\lim_{N \to \infty} \mathcal{H}_N f(x) = \mathcal{H} f(x)$$

and $\mathcal{H}$ is bounded on $L^r(\mu)$.

Proof. Let $f \in L^2(\mu)$, since the maximal function $\mathcal{H}^*$ is bounded on $L^2(\mu)$ we may assume f is bounded by 1. Suppose $(\mathcal{H}_N f \colon N \in \mathbb{N})$ does not converge μ-almost everywhere. Then there is $\epsilon > 0$ such that

$$\mu\{x \in X \colon \limsup_{M,N \to \infty} |\mathcal{H}_N f(x) - \mathcal{H}_M f(x)| > 4\epsilon\} > 4\epsilon.$$

Now one can find a strictly increasing sequence of integers $(k_j \colon j \in \mathbb{N})$ such that for each $j \in \mathbb{N}$

$$\mu\{x \in X \colon \sup_{N_j \leq N \leq N_{j+1}} |\mathcal{H}_N f(x) - \mathcal{H}_{N_j} f(x)| > \epsilon\} > \epsilon,$$

where we put $N_j = \tau^{k_j}$ and $\tau = 1 + \epsilon/4$. If $\tau^k \leq N < \tau^{k+1}$ then setting $P_k = \mathbb{P} \cap (\tau^k, \tau^{k+1}]$ we get

$$|\mathcal{H}_N f(x) - \mathcal{H}_{\tau^k} f(x)| \leq \tau^{-k} \sum_{p \in P_k} \log p.$$

By Siegel-Walfisz theorem we get

$$\sum_{p \in \mathbb{P}_N} \log p = N + \mathcal{O}(N(\log N)^{-1})$$

thus there is $C > 0$ such that

$$\left| \tau^{-k} \sum_{p \in P_k} \log p - \tau + 1 \right| \leq C k^{-1} (\log \tau)^{-1}.$$

Hence, whenever $k \geq 4C\epsilon^{-1}(\log \tau)^{-1}$ we have

$$|\mathcal{H}_N f(x) - \mathcal{H}_{\tau^k} f(x)| \leq \epsilon/2.$$

In particular, we conclude

$$\mu\{x \in X \colon \sup_{\tau^k \in \Lambda_j} |\mathcal{H}_{\tau^k} f(x) - \mathcal{H}_{N_j} f(x)| > \epsilon/2\} > \epsilon$$

for each $k_j \geq 4C\epsilon^{-1}(\log \tau)^{-1}$ which contradicts to Proposition 4. Indeed,

$$\epsilon^3 \lesssim \frac{1}{J - J_0} \sum_{j=0}^{J} \| \sup_{\tau^k \in \Lambda_j} |\mathcal{H}_{\tau^k} f - \mathcal{H}_{N_j} f| \|_{L^2(\mu)}^2 \leq \frac{C_J}{J - J_0} \|f\|_{L^2(\mu)}^2$$

where $J_0 = \min\{j \in \mathbb{N} \colon k_j \geq 4C\epsilon^{-1}(\log \tau)^{-1}\}$. Now, the standard density argument implies pointwise convergence for each $f \in L^r(\mu)$ where $r > 1$, and the proof of the theorem is completed. $\square$

References

[1] Mirek M., Trojan B. *Cotlar's ergodic theorem along the prime numbers* arXiv:1311.7572

[2] Bourgain J. *On the maximal ergodic theorem for certain subsets of the integers* Israel J. Math. **61** (1988), 39-72

[3] Bourgain J. *On the pointwise ergodic theorem on L^p for arithmetic sets* Israel J. Math. **61** (1988), 73-84

[4] Bourgain J. *Pointwise ergodic theorems for arithmetics sets. With appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein.* Publ. Math.-Paris (1989), no. 69, 545.

[5] Cotlar M. *A unified theory of Hilbert transforms and ergodic theorems* Rev. Mat. Cuyana **1** (1955), no. 2, 105-167

Michal Warchalski, Universität Bonn
email: **waral91@gmail.com**

16 Pointwise ergodic theorems for arithmetic sets

after J. Bourgain [Bou89]
A summary written by Pavel Zorin-Kranich

Abstract

Let X be a σ-finite measure space equipped with a measure-preserving $\mathbb{Z}$-action. We will describe Bourgain's 1989 approach to the study of the asymptotic behavior of the ergodic averages $K_N * f$, $f \in L^2(X)$, where

$$K_N = \frac{1}{N} \sum_{n=1}^{N} \delta_{n^d}, \quad d \geq 1.$$

The best-known result regarding the ergodic averages $K_N * f$ is the *Hardy–Littlewood maximal inequality*, which says

$$\| \sup_N |K_N * f| \|_2 \lesssim \|f\|_2 \tag{1}$$

in the case $d = 1$, $X = \mathbb{Z}$. In view of Calderón's transference principle [Cal68], the maximal inequality (1) for an arbitrary measure-preserving action $\mathbb{Z} \curvearrowright X$ is a formal consequence of the case $X = \mathbb{Z}$. A closely related result is the *pointwise ergodic theorem (PET)* which says that $K_N * f$ converges almost surely as $N \to \infty$. Unlike the maximal inequality, the PET is mostly interesting for finite measure spaces X, since for instance on the integers $K_N * f \to 0$ pointwise as $N \to \infty$. However, pointwise convergence is not the right property to look at, because it is not *local* in the sense required for the transference principle to work.

Properties that *can* be transferred from the integers to measure-preserving actions and imply pointwise convergence include finite q-variation and sublinearly growing oscillation (defined in §16.1 and §16.3, respectively). In this summary we outline Bourgain's 1989 proofs [Bou89] of the maximal inequality and an oscillation inequality on ℓ^2 for the kernels K_N with $d > 1$. The first ingredient is number-theoretic and regards the approximation of the Fourier multiplier $\hat{K}_N$ by more analytically tractable functions. Since a very similar estimate is also used in the earlier paper [Bou88b], which will be discussed in a different talk, we only recall the end result:

$$\|\hat{K}_N - \hat{L}_N\|_\infty \lesssim N^{-\epsilon}, \tag{2}$$

where

$$\hat{L}_N = \sum_{s \geq 0} \hat{L}_{s,N}, \quad \hat{L}_{s,N}(\alpha) = \sum_{\theta \in R_s} S(\theta) v(N^d(\alpha - \theta)) \phi(10^s(\alpha - \theta)), \tag{3}$$

v is the Fourier transform of a compactly supported finite measure with $|v(\xi)| \lesssim |\xi|^{-b}$, ϕ is a compactly supported smooth function, and R_s is the set of rationals of height in $[2^s, 2^{s+1})$ in the interval $[0,1)$. Whereas the L^2 theory in [Bou88b] uses the special form of the coefficients $S(\theta)$ (which are complete exponential sums) to recombine the corresponding terms in a somewhat cancellative way, the L^2 theory in [Bou89] uses only the size estimate $|S(\theta)| \lesssim 2^{-s\epsilon}$ for $\theta \in R_s$.

We observe that the error terms $N^{-\epsilon}$ are not summable for $N \in \mathbb{N}$. This necessitates the restriction of N's to a lacunary sequence $Z_\epsilon = \{\lfloor (1+\epsilon)^k \rfloor, k \in \mathbb{N}\}$. It is immediately clear that this does not affect the maximal inequality since there is no cancellation in the kernels K_N. The situation for the oscillation inequality is different, and we cannot expect to obtain oscillation estimates along $\mathbb{N}$ from oscillation estimates along Z_ϵ (without controlling the constants as $\epsilon \to 1$). However, for the end goal of obtaining a pointwise ergodic theorem this turns out to be immaterial: by the transfer principle the oscillation inequality will yield the PET on $L^2(X)$ along the sequence Z_ϵ for every $\epsilon > 1$. For bounded functions this implies the PET along $\mathbb{N}$. This can be in turn used as a dense subclass result for the maximal inequality, and leads to the PET on $L^2(X)$ along $\mathbb{N}$.

16.1 A variational inequality

In this section we present a result which implies the pointwise ergodic theorem for $d = 1$ and also plays a crucial role in the arguments involved in the case $d > 1$. The *r-variation norm* of a function f_t is defined by

$$\|f_t\|_{V_t^r} = \sup_{t_1 < \cdots < t_T} (\sum_j |f_{t_j} - f_{t_{j-1}}|^r)^{1/r},$$

where the supremum is taken over all finite increasing sequences of arguments. It is obvious that any function with finite r-variation norm, $r < \infty$, converges as $t \to \infty$. The fundamental result on variation is the Lépingle inequality for martingales.

Theorem 1. *Let $r > 2$ and let $(f_t)_t$ be a martingale in L^2 on some measure space. Then*

$$\|\|f_t\|_{V_t^r}\|_{L^2} \lesssim \frac{r}{r-2}\|f_\infty\|_{L^2}.$$

For a modern sketch of proof see [JSW08]. The fact that the constant grows as $\frac{r}{r-2}$ for $r \to 2$ is due to Bourgain and is important for the maximal inequality in §16.2, although this will not be apparent from this summary.

The Lépingle inequality has the following consequence, which can be thought of as a joint quantitative extension of the Hardy–Littlewood maximal inequality and the Lebesgue differentiation theorem.

Theorem 2. *Let φ be a finite measure on $\mathbb{R}$ such that $|\hat{\varphi}(\xi)| \lesssim |\xi|^{-b}$. Then, for $r > 2$,*

$$\|\|f * \varphi_t\|_{V_{t>0}^r}\|_{L^2(\mathbb{R})} \lesssim \frac{r}{r-2}\|f\|_{L^2(\mathbb{R})},$$

where φ_t denotes the t-dilate of φ.

Sketch of proof. By a density argument we may restrict t to an arithmetic progression. We compare the convolution $f * \varphi_t$ with $P_t f$, where P_t is the Poisson semigroup whose Fourier multiplier is $e^{-t|\xi|}$. It follows that the operator $f \mapsto f * \varphi_t - P_t f$ is a Fourier multiplier that decays both at 0 and at ∞. For dyadic t its contribution can be estimated by a square function and the contribution of t's between dyadic values can be estimated by interpolation between V^1 and $V^\infty = \ell^\infty$ estimates, see [Bou89, Lemma 3.28] for details.

It therefore remains to show

$$\|\|P_t f\|_{V_{t\in\mathbb{N}}^r}\|_{L^2(\mathbb{R})} \lesssim \frac{r}{r-2}\|f\|_{L^2(\mathbb{R})}.$$

To this end one can use Rota's dilation theorem [Ste70, §IV.4]. By Rota's theorem there exists a measure space $\tilde{X}$ with a measure-preserving projection $\pi : \tilde{X} \to \mathbb{R}$ such that $P_t f = \pi_* E_t(f \circ \pi)$, where π_* denotes conditional expectation onto $\mathbb{R}$ under π and $(E_t)_t$ is a decreasing sequence of conditional expectations. Since variational estimates are clearly preserved under conditional expectation, the claim follows from Theorem 1. □

The case $d = 1$ of (3) is special because $S(\theta) = 0$ for $\theta \neq 0$ (which follows from the proof of (2)). Therefore, considering L_N as a function on $\mathbb{R}$ rather than on the torus, Theorem 2 provides a variational estimate for L_N. Standard arguments (see e.g. [Bou89, Lemma 4.4] or [ZK14, Lemma 4.1]) allow one to transfer this estimate to the integers. This allows us to consider the functions $L_{s,N}$ as being defined on $\mathbb{R}$ from now on.

16.2 A maximal inequality

One thing that one might want to try for $d > 1$ is applying Theorem 2 to the multipliers (3) in the crudest way possible. Firstly, the functions $\phi(10^s(\cdot - \theta))$ have disjoint supports for distinct $\theta \in R_s$ provided that the support of ϕ is small enough. By Theorem 2 we can estimate

$$\|\|\mathscr{F}^{-1}(\hat{f}S(\theta)v(N^d(\cdot-\theta))\phi(10^s(\cdot-\theta)))\|_{V^r_{N>0}}\|_{L^2(\mathbb{R})} \\ \lesssim_r |S(\theta)|\|\mathscr{F}^{-1}(\hat{f}\phi(10^s(\cdot-\theta))\|_{L^2(\mathbb{R})}.$$

Since the functions on the right-hand side have disjoint Fourier support, the ℓ^2 norm of their norms is bounded by $\|f\|_{L^2(\mathbb{R})}$. Unfortunately, this only gives the bound

$$\|\|L_{s,N} * f\|_{V^r_{N>0}}\|_{L^2(\mathbb{R})} \lesssim \|S(\theta)\|_{\ell^2_{\theta\in R_s}}\|f\|_{L^2(\mathbb{R})} \approx 2^{s(1/2-\epsilon)}\|f\|_{L^2(\mathbb{R})}$$

for some small $\epsilon > 0$, and this is not summable in s.

The large loss comes from the fact that we have considered each frequency $\theta \in R_s$ separately. The main part of Bourgain's proof is a multi-frequency argument which gives the bound

$$\|\|L_{s,N} * f\|_{\ell^\infty_{N\in Z_\epsilon}}\|_{L^2(\mathbb{R})} \lesssim \max_{\theta\in R_s}|S(\theta)|(1+\log|R_s|)^2\|f\|_{L^2(\mathbb{R})} \tag{4}$$

using only the fact that that the functions $\phi(10^s(\cdot - \theta))$ have disjoint support. In particular, it does not make any reference to rationality of $\theta \in R_s$ or specific structure of $S(\theta)$'s. These bounds can be summed in s using the size bound on the $S(\theta)$'s, yielding the maximal inequality (1).

For space reasons it is not feasible to give a faithful outline of Bourgain's argument here, so we refer to the original article [Bou89] and a modernized version [ZK14], the latter also giving a variational estimate in the variable N. Here we assume the maximal inequality (4) as a black box and outline the proof of the oscillation inequality.

16.3 The oscillation inequality

Let $N_1 < N_2 < \dots$ be an increasing sequence in Z_ϵ with $N_{j+1} \geq 2N_j$ and let

$$\mathcal{M}_j f = \sup_{N\in I_j}|f * (K_N - K_{N_j})|, \quad I_j = [N_j, N_{j+1}] \cap Z_\epsilon.$$

Theorem 3. *For some $\delta > 0$ and all J we have*

$$\| \| \mathcal{M}_j f \|_{\ell^2} \|_{\ell^2_{j \le J}} \lesssim J^{1/2-\delta} \|f\|_2.$$

Let us briefly explain why this oscillation inequality (transferred to X) implies the pointwise ergodic theorem on $L^2(X)$ along Z_ϵ. If the pointwise ergodic theorem fails, then for a positive measure subset of X we can find an $\epsilon > 0$ such that for any N_j there exists $N \in Z_\epsilon$ such that $|K_{N_j} * f - K_N * f| > \epsilon$. A standard measure theory argument shows that we may assume $N < N_{j+1}$ for some N_{j+1} depending on N_j and also that a single $\epsilon > 0$ works on a set of positive measure. Hence $\|\mathcal{M}_j f\|_{\ell^2}$ is bounded below by a constant for all j, which is a contradiction.

Sketch of proof of Theorem 3. By a square function argument using (2) we may replace K_N by L_N. Let s_0 be chosen later and split

$$\begin{aligned} &\| \sup_{N \in I_j} |(L_N - L_{N_j}) * f| \|_2 \\ &\qquad \le \sum_{s \le s_0} \| \sup_{N \in I_j} |(L_{s,N} - L_{s,N_j}) * f| \|_2 + 2 \sum_{s > s_0} \| \sup_{N \in Z_\epsilon} |L_{s,N} * f| \|_2. \end{aligned} \tag{5}$$

Using the maximal inequality (4) and the bound for $S(\theta)$, the second term of (5) can be estimated by $2^{-s_0 \delta} \|f\|$. In the first term of (5) we separate the frequencies and use the crude bound $|S(\theta)| \le 1$ to estimate

$$\begin{aligned} \| \sup_{N \in I_j} |(L_{s,N} - L_{s,N_j}) * f| \|_2 &\le \sum_{\theta \in R_s} \| \sup_{N \in I_j} |\mathscr{F}^{-1}(v(N^d(\cdot - \theta)) \phi(10^s(\cdot - \theta)) \hat{f})| \|_2 \\ &= \sum_{\theta \in R_s} \| \sup_{N \in I_j} |\mathscr{F}^{-1}(v(N^d \cdot) \hat{f}_\theta)| \|_2, \end{aligned}$$

where $\hat{f}_\theta = \phi(10^s \cdot) \hat{f}(\cdot + \theta)$ We estimate the ℓ^2_j norm of each summand changing the order of summation and using Hölder's inequality:

$$\begin{aligned} \| \| \sup_{N \in I_j} |\mathscr{F}^{-1}(v(N^d \cdot) \hat{f}_\theta)| \|_2 \|_{\ell^2_{j \le J}} &= \| (\sum_{j=1}^{J} \sup_{N \in I_j} |(\check{v}_{N^d} - \check{v}_{N_j^d}) * f_\theta|^2)^{1/2} \|_2 \\ &\le \| J^{\delta/2(1+\delta)} (\sum_{j=1}^{J} \sup_{N \in I_j} |(\check{v}_{N^d} - \check{v}_{N_j^d}) * f_\theta|^{2+2\delta})^{1/2(1+\delta)} \|_2 \\ &\le J^{\delta/2(1+\delta)} \| \| \check{v}_t * f_\theta \|_{V_t^{2+2\delta}} \|_{L^2} \lesssim J^{\delta/2(1+\delta)} \|f_\theta\|_{L^2}, \end{aligned}$$

where the last inequality is given by Theorem 2. Since the functions $\hat{f}_\theta$ have disjoint support, this gives the bound $\sum_{s\leq s_0} J^{\delta/2(1+\delta)}|R_s|^{1/2}\|f\|_2 \lesssim 2^{s_0}J^{\delta}\|f\|_2$ for the $\ell^2_{j\leq J}$ norm of the first term in (5). The $\ell^2_{j\leq J}$ norm of the second term can be estimated by $J^{1/2}2^{-s_0\delta}\|f\|_2$. Choosing s_0 such that $2^{s_0} \approx J^{1/3}$, say, we obtain the conclusion of the theorem. □

References

[Bou88a] J. Bourgain. An approach to pointwise ergodic theorems. In *Geometric aspects of functional analysis (1986/87)*, volume 1317 of *Lecture Notes in Math.*, pages 204–223. Springer, Berlin, 1988.

[Bou88b] J. Bourgain. On the maximal ergodic theorem for certain subsets of the integers. *Israel J. Math.*, 61(1):39–72, 1988.

[Bou89] Jean Bourgain. Pointwise ergodic theorems for arithmetic sets. *Inst. Hautes Études Sci. Publ. Math.*, (69):5–45, 1989.

[Cal68] A.-P. Calderón. Ergodic theory and translation-invariant operators. *Proc. Nat. Acad. Sci. U.S.A.*, 59:349–353, 1968.

[JSW08] Roger L. Jones, Andreas Seeger, and James Wright. Strong variational and jump inequalities in harmonic analysis. *Trans. Amer. Math. Soc.*, 360(12):6711–6742, 2008.

[Ste70] Elias M. Stein. *Topics in harmonic analysis related to the Littlewood–Paley theory.* Annals of Mathematics Studies, No. 63. Princeton University Press, Princeton, N.J., 1970.

[ZK14] Pavel Zorin-Kranich. Variation estimates for averages along primes and polynomials, 2014. Preprint, `arXiv:1403.4085`.

Pavel Zorin-Kranich, Hebrew University
email: **pzorin@math.huji.ac.il**