POINTWISE CONVERGENCE OF FOURIER SERIES

PO-LAM YUNG

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, and $\mathcal{K}(y) = p.v.\frac{1}{y}$ on $[-1/2, 1/2)$ and extend it periodically so that it becomes a distribution on $\mathbb{T}$. Let $N\colon \mathbb{T} \to 2\pi\mathbb{N}$ be a measurable function. For $f \in C^\infty(\mathbb{T})$, let

$$Tf(x) = \int_{\mathbb{T}} f(x-y)\mathcal{K}(y)e^{iN(x)y}dy \quad \text{for } x \in \mathbb{T}.$$

We want to show that $T\colon L^2(\mathbb{T}) \to L^{2-\varepsilon}(\mathbb{T})$ for any $\varepsilon > 0$. Below we give a white lie proof of this fact, following Fefferman [2] (our exposition also draws heavily on Demeter's points of view in [1]).

First decompose $\mathcal{K}(y) = \sum_{k=0}^\infty \psi_k(y)$ where each $\psi_k(y)$ is an odd C^∞ function supported on $|y| \simeq 2^{-k}$. It follows that

$$Tf(x) = \sum_{k=0}^{\infty} \int_{\mathbb{T}} f(x-y)\psi_k(y)e^{iN(x)y}dy. \tag{1}$$

If $I \subset \mathbb{T}$ is a dyadic interval and $\omega \subset [1,\infty)$ is a dyadic interval of length $1/|I|$, then the pair (I,ω) is called a tile. (Let's take all dyadic intervals to be half-open and half-closed, so that all dyadic intervals contain the left endpoint but not the right endpoint; then the collection of all dyadic intervals of the same length scale form a partition of the full space.) If p is a tile, we usually denote by I_p and ω_p the time and frequency intervals such that $p = (I_p, \omega_p)$. The set of all tiles is denoted by $\mathfrak{P}$. It should be thought of as a 3-parameter family, indexed by the length of the time interval, the position of the time interval and the position of the frequency interval.

We denote by E_p the set $N^{-1}(\omega_p) \cap I_p$. Note that for every $x \in \mathbb{T}$ and every $k \geq 0$, there exists one and only one tile p with $|I_p| = 2^{-k}$ such that $x \in I_p$ and $N(x) \in \omega_p$. In other words, there exists a unique tile p with $|I_p| = 2^{-k}$ such that $x \in E_p$. Hence from (1), we have

$$Tf = \sum_{p \in \mathfrak{P}} T_p f$$

where

$$T_p f(x) := \chi_{E_p}(x) \int_{\mathbb{T}} f(x-y)\psi_k(y)e^{iN(x)y}dy.$$

Define a tentative mass of a tile p by

$$A_0(p) = \frac{|E_p|}{|I_p|}.$$

Since

$$|T_p f(x)| \leq \chi_{E_p}(x)\frac{1}{|I_p|}\int_{|y|\leq|I_p|}|f(x-y)|dy \leq \chi_{E_p}(x)\frac{1}{|I_p|^{1/2}}\|f\|_{L^2},$$

it follows that

$$\|T_p\|_{L^2\to L^2} \leq A_0(p)^{1/2}.$$

We upgrade this to a *single tree estimate*. First we introduce a partial ordering on the set of all tiles. If p_1, p_2 are tiles, then we say p_2 is an ancestor of p_1, written $p_1 < p_2$, if $I_{p_1} \subseteq I_{p_2}$ and $\omega_{p_1} \supseteq \omega_{p_2}$. This is a partial ordering on $\mathfrak{P}$; indeed two tiles are either disjoint, or comparable under this partial ordering. A finite collection $\mathfrak{p}$ of tiles is said to be convex if $p, p' \in \mathfrak{p}$ and $p < p'' < p$ implies $p'' \in \mathfrak{p}$. If p_0 is a tile, then a tree with top p_0 is a convex collection $\mathfrak{p}$ of tiles, such that $p < p_0$ for all $p \in \mathfrak{p}$ (note that we do not require p_0 to be in $\mathfrak{p}$). We write $T_{\mathfrak{p}}$ for $\sum_{p\in\mathfrak{p}} T_p$ whenever $\mathfrak{p}$ is a finite collection of tiles.

Proposition 1. *Let $\delta \in (0,1)$. If $\mathfrak{p}$ is a tree such that $A_0(p) \leq \delta$ for all $p \in \mathfrak{p}$, then*

$$\|T_{\mathfrak{p}}\|_{L^2\to L^2} \lesssim \delta^{1/2}.$$

The idea is that $T_{\mathfrak{p}}$ is then like a maximally truncated Hilbert transform localized to a set of measure $\lesssim \delta$ in $\mathbb{T}$. More precisely, let $\mathfrak{p}$ be a tree with top p_0. For every dyadic interval $I \subset \mathbb{T}$, let p_I be the unique tile with time interval I such that ω_{p_I} contains ω_{p_0}. Since $\mathfrak{p}$ is finite, we may partition $\mathbb{T}$ into a disjoint union of dyadic intervals $I_1, I_2, \dots$, such that each I_i is a maximal dyadic interval satisfying the following condition:

$$\tilde{I}_i \subseteq I_p \text{ for all } p \in \mathfrak{p} \text{ with } I_p \cap I_i \neq \emptyset.$$

Here $\tilde{I}_i$ is the 'parent' of I_i, namely the dyadic interval that has length $2|I_i|$ that contains I_i. Then every $\tilde{I}_i = I_{p_i}$ for some unique $p_i \in \mathfrak{p}$, and if we let $E_i := E(p_i) \cap I_i$, then $|E_i| \leq A_0(p_i)|I_{p_i}| \leq 2\delta|I_i|$. Also, for every i, we have $E(p) \cap I_i \subseteq E_i$ for every $p \in \mathfrak{p}$ (this is trivial if $E(p)$ doesn't intersect I_i; on the other hand, if $E(p)$ intersects I_i, then $I_{p_i} = \tilde{I}_i \subseteq I_p$, so $\omega_p \subseteq \omega_{p_i}$, which implies $E(p) \cap I_i \subseteq N^{-1}(\omega_{p_i}) \cap I_i = E_i$). Now define a maximal function

$$\mathcal{M}F(x) = \begin{cases} \sup_{I \supseteq I_i} \frac{1}{|I|} \int_I |F(y)| dy & \text{if } x \in E_i \text{ for some } j \\ 0 & \text{otherwise.} \end{cases}$$

Then one can show that $\mathcal{M} \colon L^2 \to L^2$ with $\|\mathcal{M}\|_{L^2 \to L^2} \lesssim \delta^{1/2}$, and for every $x \in \mathbb{T}$, we have

$$T_{\mathfrak{p}} f(x) \lesssim \mathcal{M}(f * \mathcal{K})(x) + \mathcal{M}f(x).$$

Since $\|f * \mathcal{K}\|_{L^2} \lesssim \|f\|_{L^2}$, we obtain the conclusion of Proposition 1.

(Indeed the above is adapted from Demeter's survey [1], and is not exactly Fefferman's proof. In Fefferman's paper, the mass was defined so that the mass of a tile p is at least the supremum of the A_0 mass of all ancestors of p. We may assume, without loss of generality, that N is bounded. In that case, if $|I|$ is sufficiently small, then the mass of p_I is 1. Hence we may partition $\mathbb{T}$ into a disjoint union of dyadic intervals $I_1, I_2, \dots$, such that each I_i is a maximal dyadic interval with the mass of p_{I_i} being $> \delta$. Then because we are using Fefferman's definition of mass, where the mass was defined so that the mass of a tile p is at least the supremum of the A_0 mass of all ancestors of p, we still have $\tilde{I}_i \subseteq I_p$ for all $p \in \mathfrak{p}$ with $I_p \cap I_i \neq \emptyset$. Also, if we let $E_i := E(p_{\tilde{I}_i}) \cap I_i$, then for every i, we still have $|E_i| \leq 2\delta|I_i|$, and $E(p) \cap I_i \subseteq E_i$ for all $p \in \mathfrak{p}$. Thus we can proceed as above, and finish the proof of Proposition 1.)

To proceed further, note that if two tiles p_1 and p_2 are disjoint, then at least one of the following happens: $I_{p_1} \cap I_{p_2} = \emptyset$ or $\omega_{p_1} \cap \omega_{p_2} = \emptyset$. In either case, $T_{p_1} f(x) T_{p_2} g(x) = 0$ for all $f, g \in C^\infty(\mathbb{T})$ and all $x \in \mathbb{T}$; in particular,

$$T_{p_1}^* T_{p_2} = 0.$$

We will pretend that we also have
White lie:

$$T_{p_1} T_{p_2}^* = 0$$

whenever p_1, p_2 are disjoint tiles; this is based on the heuristic that $T_p f$ is morally supported on p for all tiles p and all $f \in C^\infty(\mathbb{T})$.

Proposition 2. *Let $\mathfrak{p}$ be a convex collection of tiles. Suppose for any $p \in \mathfrak{p}$, and for any two ancestors p_1, p_2 of p with $p_1, p_2 \in \mathfrak{p}$, we have p_1 comparable to p_2. Then*
(a) $\mathfrak{p}$ can be organized into a union of trees, so that any two tiles from two different trees are disjoint.
(b) Hence if $A_0(p) \leq \delta$ for every $p \in \mathfrak{p}$, then by Proposition 1 and Cotlar-Stein, we have

$$\|T_{\mathfrak{p}}\|_{L^2 \to L^2} \lesssim \delta^{1/2}.$$

A collection of tiles $\mathfrak{p}$ satisfying the conditions of Proposition 2 is called a Fefferman forest.

From now on let K be a large positive integer to be determined.

Proposition 3. *Let $\mathfrak{p}$ be a convex collection of tiles, such that $A_0(p) \leq \delta$ for every $p \in \mathfrak{p}$. Then there exists a convex collection $\mathfrak{p}' \subset \mathfrak{p}$ such that the following holds:*
(a) $A_0(p') \leq \delta/2$ for all $p' \in \mathfrak{p}'$;
(b) $\mathfrak{p} \setminus \mathfrak{p}'$ can be organized into a union of trees $\mathfrak{p}_1, \mathfrak{p}_2, \dots$ with tops $p_1, p_2, \dots$, so that

$$\left\| \sum_j \chi_{I_{p_j}}(x) \right\|_{L^1(\mathbb{T})} \lesssim \delta^{-1}.$$

In particular, there exists a (small) exceptional set $F_\delta \subset \mathbb{T}$, such that $|F_\delta| \leq \delta^{99}/K$, and such that for $x \notin F_\delta$, the number of p_j's with $x \in I_{p_j}$ is at most $K\delta^{-100}$.

Indeed, if $A_0(p) \leq \delta/2$ for all $p \in \mathfrak{p}$, then just set $\mathfrak{p}' = \mathfrak{p}$. If not, then we pick a maximal $p_1 \in \mathfrak{p}$ with $A_0(p_1) \in (\delta/2, \delta]$, and let $\mathfrak{p}_1$ be the set of all $p \in \mathfrak{p}$ with $p < p_1$. Clearly $\mathfrak{p}_1$ is a tree with top p_1. Assume that we have chosen trees $\mathfrak{p}_1, \dots, \mathfrak{p}_\ell$. Then $\mathfrak{p} \setminus (\mathfrak{p}_1 \cup \dots \cup \mathfrak{p}_\ell)$ is a convex collection of tiles, and if $A_0(p) \leq \delta/2$ for all $p \in \mathfrak{p} \setminus (\mathfrak{p}_1 \cup \dots \cup \mathfrak{p}_\ell)$, just set $\mathfrak{p}' = \mathfrak{p} \setminus (\mathfrak{p}_1 \cup \dots \cup \mathfrak{p}_\ell)$; if not, we pick a maximal $p_{\ell+1} \in \mathfrak{p} \setminus (\mathfrak{p}_1 \cup \dots \cup \mathfrak{p}_\ell)$ such that $A_0(p_{\ell+1}) \in (\delta/2, \delta]$, and let $\mathfrak{p}_{\ell+1}$ be the set of all $p \in \mathfrak{p} \setminus (\mathfrak{p}_1 \cup \dots \cup \mathfrak{p}_\ell)$ with $p < p_{\ell+1}$. This process terminate (since $\mathfrak{p}$ is finite), and it remains to compute $\left\| \sum_j \chi_{I_{p_j}}(x) \right\|_{L^1(\mathbb{T})}$. But this norm is equal to

$$\sum_j |I_{p_j}| \leq \frac{2}{\delta} \sum_j |E_{p_j}| \leq \frac{2}{\delta};$$

indeed $E_{p_1}, E_{p_2}, \dots$ are disjoint subsets of $\mathbb{T}$, since the $p_1, p_2, \dots$ were chosen to be incomparable. This proves Proposition 3.

We remark that a small modification of the above proof shows that the dyadic BMO norm of $\sum_j \chi_{I_{p_j}}$ is also bounded by δ^{-1}. The John-Nirenberg inequality then gives a better control of the size of the exceptional set, and this is useful when one wants to prove pointwise a.e. convergence of Fourier series for functions in $L^p(\mathbb{T})$, $p \in (1,2)$.

Proposition 4. *Let $\mathfrak{p}$ be as in Proposition 3. Let $\mathfrak{p}'$ and F_δ be as in the conclusion of Proposition 3. Let $\mathfrak{p}''$ be the collection of tiles $p \in \mathfrak{p}$ such that $I_p \subset F_\delta$. Then $\mathfrak{p} \setminus (\mathfrak{p}' \cup \mathfrak{p}'')$ can be reorganized into a union of M Fefferman forests $\mathfrak{p}^{(1)}, \dots \mathfrak{p}^{(M)}$ with $M \lesssim \log(K\delta^{-100})$. Hence*

$$\|T_{\mathfrak{p} \setminus \mathfrak{p}'}\|_{L^2(\mathbb{T} \setminus F_\delta)} \lesssim \log(K\delta^{-100}) \delta^{1/2} \|f\|_{L^2}.$$

Indeed, let M be the smallest integer such that $2^M > K\delta^{-100}$; in particular, $M \lesssim \log(K\delta^{-100})$. Recall that we had a list of trees $\mathfrak{p}_1, \mathfrak{p}_2, \dots$ with tops $p_1, p_2, \dots$ from Proposition 3. For $m = 1, 2, \dots, M$, let $\mathfrak{p}^{(m)}$ be the set of all $p \in \mathfrak{p} \setminus (\mathfrak{p}' \cup \mathfrak{p}'')$ such that

$$2^{m-1} \leq \text{the number of tree tops } p_j\text{'s that are ancestors of } p < 2^m.$$

Then $\mathfrak{p} \setminus (\mathfrak{p}' \cup \mathfrak{p}'')$ is the disjoint union of $\mathfrak{p}^{(1)}, \dots, \mathfrak{p}^{(M)}$, because each $p \in \mathfrak{p} \setminus \mathfrak{p}'$ belongs to at least one tree (hence at least one of the tree tops is an ancestor of p), and if $p \in \mathfrak{p} \setminus (\mathfrak{p}' \cup \mathfrak{p}'')$ has at least $2^M > \log(K\delta^{-100})$ ancestors that are tree tops, then I_p is contained in I_{p_j} for more than $\log(K\delta^{-100})$ tree tops p_j, contradicting Proposition 3.

Furthermore, we check that each $\mathfrak{p}^{(m)}$ is a Fefferman forest: the key is to show that if $p \in \mathfrak{p}^{(m)}$ and $p', p'' \in \mathfrak{p}^{(m)}$ are both ancestors of p, then p' and p'' are comparable. Assume not. Then since $p', p'' \in \mathfrak{p}^{(m)}$, there are at least 2^{m-1} tree tops that are ancestors of p', and at least 2^{m-1} tree tops that are ancestors of p''. Furthermore, these two sets of tree tops are disjoint, since if p', p'' are incomparable ancestors of p, then p' and p'' cannot have ancestors in common. Hence there are at least $2^{m-1} + 2^{m-1} = 2^m$ tree tops that are ancestors of p, contradicting that $p \in \mathfrak{p}^{(m)}$. It remains to apply Proposition 2 to estimate $\|T_{\mathfrak{p}^{(m)}}\|_{L^2 \to L^2}$ for each forest $\mathfrak{p}^{(m)}$. Since

$$T_{\mathfrak{p} \setminus \mathfrak{p}'} = \sum_{m=1}^{M} T_{\mathfrak{p}^{(m)}} \quad \text{on } \mathbb{T} \setminus F_\delta,$$

it follows that $\|T_{\mathfrak{p} \setminus \mathfrak{p}'} f\|_{L^2(\mathbb{T} \setminus F_\delta)} \lesssim \log(K\delta^{-100}) \delta^{1/2} \|f\|_{L^2}$.

By successively applying Propositions 3 and 4 (with $\delta = 1, 2^{-1}, 2^{-2}, \dots$ and $\mathfrak{p}$ replaced by $\mathfrak{p}'$ from the previous step each time), we obtain the following proposition:

Proposition 5. *Let $\mathfrak{p}$ be a convex collection of tiles. Then there exists a (small) exceptional set $F \subset \mathbb{T}$ such that $|F| \leq 1/K$, with*

$$\|T_{\mathfrak{p}} f\|_{L^2(\mathbb{T} \setminus F)} \lesssim \log K \|f\|_{L^2}.$$

Hence if $\|f\|_{L^2} = 1$, then for any $\alpha > 1$, the set $\{x \in \mathbb{T} \colon |T_{\mathfrak{p}} f(x)| > \alpha\}$ has measure

$$\lesssim \frac{(\log K)^2}{\alpha^2} + \frac{1}{K}.$$

This is true for all $K \geq 1$, so letting $K = \alpha^2$ we see that

$$|\{x \in \mathbb{T} \colon |T_{\mathfrak{p}} f(x)| > \alpha\}| \lesssim_\varepsilon \frac{1}{\alpha^{2-\varepsilon}}$$

for any $\varepsilon > 0$. It follows that

$$\|T_{\mathfrak{p}} f\|_{L^{2-\varepsilon}(\mathbb{T})} \lesssim_\varepsilon 1$$

for all $\varepsilon > 0$, which shows that $\|T_{\mathfrak{p}}\|_{L^2 \to L^{2-\varepsilon}} \lesssim_\varepsilon 1$.

We remark that in order to deal with the white lie, we need a more complicated definition of mass (namely $A(p)$ in Fefferman's paper), that takes into account tiles whose frequency intervals are further away. In that case, we can only say that $\|T_{p_1} T_{p_2}^*\|_{L^2 \to L^2}$ is small whenever $A(p_1)$ and $A(p_2)$ are small (say $\leq \delta$), and $\omega_{p_1}, \omega_{p_2}$ are far apart (at least distance δ^{-1} when measured in the correct length scale). This led Fefferman to introduce the notion of separated trees; see Lemma 4 in his paper.

Next, let's have a row of trees where the tops have disjoint time intervals. At some point one needs to study the orthogonality between separated rows. In order for this to work, we need to localize in frequency for each tree, and this corresponds to a blurring of the support for each tree on the time side. In order for this blurring not to ruin things, Fefferman introduced the notion of normal trees, and required that all trees in a row to be normal. See Lemma 5 in his paper. It is for this reason that Fefferman had to remove an exceptional set in the conclusion of his main lemma; indeed, to create normal trees, on top of removing tiles whose time interval lives inside the exceptional set, Fefferman considered what he called central dyadic frequency intervals, and invoked a convenient inclusion property for these central dyadic frequency intervals.

Another complication that arise because we don't have perfect frequency localization is that for any union of trees $\mathfrak{p}$, in order to control $T_{\mathfrak{p}}$, we need to control the number of rows present in $\mathfrak{p}$. This is why in Fefferman's main lemma, he needs to assume that each point in $\mathbb{T}$ belongs to no more than a certain number of time intervals of the tops.

Finally, although Fefferman used a more complicated definition of mass, he would still use the exact same tree selection algorithm we used in the proof of our Proposition 3. In particular, one should still consider maximal elements whose A_0 mass is big (rather than the A mass), and remove all tiles in the remaining collection that are $<$ than this maximal element. In order to guarantee that one would remove at least one tile from the collection unless the A mass of all remaining tiles are small, one needs to reconcile between the A mass and the A_0 mass again. This is done again via the convenient inclusion property for central dyadic frequency intervals. See the first page of Section 7 in Fefferman's paper.

For reference, our Proposition 1 is basically Fefferman's Lemma 3. Our Proposition 2 is basically Fefferman's Corollary to his Main Lemma. Our Propositions 3,4 and 5 are in Section 7 of Fefferman's paper.

References

[1] Ciprian Demeter, *A guide to Carleson's theorem*, Rocky Mountain J. Math. **45** (2015), no. 1, 169–212.

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

[3] Charles L. Fefferman, *Erratum: "Pointwise convergence of Fourier series" [Ann. of Math. (2)* ***98*** *(1973), no. 3, 551–571; MR0340926 (49 #5676)]*, Ann. of Math. (2) **146** (1997), no. 1, 239.