Harmonic Analysis Working Group :-

Bourgain's Circular Maximal Theorem

Lecture 2

Recall from last time

$$Mf(x) := \sup_{1 \le r \le 2} |Af(x,r)| = \sup_{1 \le r \le 2} |f * \sigma_r(x)|$$

is the local spherical maximal function. Using geometric arguments, we proved the L^2 estimate

$$\|Mf\|_{L^2(\mathbb{R}^d)} \lesssim \|f\|_{L^2(\mathbb{R}^d)}$$

for $d \geq 3$.

Stein's theorem asserts that for $d \geq 3$ we have

$$\|Mf\|_{L^p(\mathbb{R}^d)} \lesssim \|f\|_{L^p(\mathbb{R}^d)} \quad \text{for } p > \frac{d}{d-1}.$$

It seems tricky to prove this using geometric tools alone. (I am not aware of a purely geometric proof of Stein's theorem; it would be interesting to give one. One can easily reduce the problem to understanding intersections between d spherical annuli - if you are interested, ask me about this). To prove Stein's theorem in the full range, we will introduce tools from Fourier analysis.

Littlewood - Paley decomposition :-

We perform a standard Littlewood-Paley decomposition of the operator $Af(\cdot, r)$. Fix $\eta \in C^\infty(\mathbb{R})$ even such that

$$\eta(r) = 1 \text{ if } |r| \le 1 \ ; \quad \eta(r) = 0 \text{ if } |r| \geq 2.$$

Define $\beta(r) := \eta(r) - \eta(2r)$ so that

$$\operatorname{supp} \beta \subseteq [-2,2] \setminus [-1/2, 1/2] \quad \text{and}$$

$$1 = \eta(r) + \sum_{j=1}^{\infty} \beta(2^{-j} r) \quad \text{for } r \neq 0.$$

By the Fourier inversion formula, we have

$$Af(x,r) = f * \sigma_r(x) = \int_{\hat{\mathbb{R}}^d} e^{2\pi i x\cdot\xi} \, \hat{\sigma}_r(\xi)\, \hat{f}(\xi)\, d\xi .$$

We decompose in the frequency space by writing

$$Af(x,r) = \sum_{j=0}^{\infty} A^j f(x,r)$$

where

$$A^0 f(x,r) := \int_{\hat{\mathbb{R}}^d} e^{2\pi i x\cdot\xi}\, \eta(|\xi|)\, \hat{\sigma}_r(\xi)\, \hat{f}(\xi)\, d\xi$$

and

$$A^j f(x,r) := \int_{\hat{\mathbb{R}}^d} e^{2\pi i x\cdot\xi}\, \beta(2^{-j}|\xi|)\, \hat{\sigma}_r(\xi)\, \hat{f}(\xi)\, d\xi$$

for $j \geq 1$.

Alternatively, we can express

$$A^j f(x,r) = f * (\varphi_j * \sigma_r)(x)$$

where
$$\varphi_0(x) = \check{\eta}(x)$$
$$\varphi_j(x) = 2^{-dj}\, \check{\beta}(2^{-j}x) \qquad \text{for } j \geq 1$$

(Here we are identifying η, β (functions on $\mathbb{R}$) with the radial functions $\eta(|\cdot|)$, $\beta(|\cdot|)$ on $\mathbb{R}^d$).

To prove an L^p bound

$$\|Mf\|_{L^p(\mathbb{R}^d)} \lesssim \|f\|_{L^p(\mathbb{R}^d)}$$

it suffices to show a sequence of L^p bounds

$$\|M^j f\|_{L^p(\mathbb{R}^d)} \lesssim 2^{-j\varepsilon(p)} \|f\|_{L^p(\mathbb{R}^d)} \qquad (1)$$

for some $\varepsilon(p) > 0$, where the M^j are the frequency localised maximal operators

$$M^j f(x) := \sup_{1 \leq r \leq 2} |A^j f(x,r)|.$$

Henceforth, we will focus on proving bounds of the form (1). Noticer that the dependence on the frequency 2^j in the operator norm is crucial – we need the geometric decay $2^{-j\varepsilon(p)}$ in order to sum the frequency localised pieces.

Remark:- Via simple Littlewood - Paley arguments, the estimate (1) also implies global maximal estimates (where the supremum is taken over the unrestricted range $r > 0$) for $p \geq 2$.

Low frequency contribution

The low frequency term M^0 is easy to estimate and in fact satisfies

$$\|M^0 f\|_{L^p(\mathbb{R}^d)} \lesssim \|f\|_{L^p(\mathbb{R}^d)} \qquad (2)$$

for all $1 \leq p \leq \infty$ in all dimensions $d \geq 2$. This is because the convolution kernel

$$K_r^0 := \varphi_0 * \sigma_r$$

is dominated by some non-negative radially non-increasing function $\mathcal{K}^0$ (for $r = 1$, say) such that $\|\mathcal{K}^0\|_1 \lesssim 1$. By standard results, it follows that

$$M^0 f(x) \lesssim M_{HL} f(x)$$

where M_{HL} is the Hardy-Littlewood maximal operator, and so (2) follows by the Hardy-Littlewood maximal theorem.

This can be understood at a heuristic level via the uncertainty principle: by frequency localising to scale 1, we blur the circles on the physical side at scale 1. In particular, we replace $C(x,r)$ with its 1-neighbourhood. This fattened set essentially looks like a ball, so we can compare M^0 with M_{HL}.

The same argument can be applied at all frequencies 2^j :– we can dominate

$$K_r^j := \varphi_j * \sigma_r$$

by some radially non-increasing $\mathcal{K}^j$ when $r=1$. However, the norms $\|\mathcal{K}^j\|_1$ will blow up badly in 2^j, and so we certainly can't expect to establish (1) this way!

Discretisation, linearisation and duality.

We will now apply the same reductions to the frequency localised operators M^j as those used in lecture 1.

By the uncertainty principle, we expect the kernels $K_r^j = \varphi_j * \sigma_j$ to satisfy

$|K_r^j|$ is essentially constant at scale 2^{-j},

where 2^{-j} is of course the reciprocal scale to the frequency decomposition. It is therefore natural to break the spatial domain into cubes of side-length 2^{-j}.

We will again work with local estimates, since the original operator M is local (the M^j will only be "pseudo-local" because of Schwartz tails arising from the frequency decomposition, but one could still pass from local to global for the M^j if one wished).

Let $\bar{Q}_0$ be the cube of side-length $1/10n$ centred at 0 and let $\mathcal{Q}_j$ denote a decomposition of $\bar{Q}_0$ into dyadic cubes of side-length 2^{-j}.

Writing $\|M^j f\|_{L^p(\bar{Q}_0)} = \left(\sum_{Q_j \in \mathcal{Q}_j} \|M^j f\|^p_{L^p(Q_j)} \right)^{1/p}$,

we see that

$$\|M^j f\|^p_{L^p(\bar{Q}_0)} \lesssim \sum_{Q_j \in \mathcal{Q}_j} \delta^d \cdot \left| \int_{\mathbb{R}^d} f(y)\, \varphi_j * \sigma_{r_{Q_j}}(x_{Q_j} - y)\, dy \right|^p$$

where $\delta := 2^{-j}$ and

where $(x_{Q_j}, r_{Q_j}) \in Q_j \times [1,2]$ for all $Q_j \in \mathcal{Q}$ denotes a choice of centre and radius.

To clean things up, we introduce some notation :-

Let $\tilde{\chi}_{C^\delta}(y) := 2^{-j} \cdot \varphi_j * \sigma_r(x-y)$

for $C = C(x,r)$ a sphere centred at x of radius r.

By the uncertainty principle we should think of $\tilde{\chi}_{C^\delta}$ as being supported on C^δ. Crucially, $\tilde{\chi}_{C^\delta}$ also carries some oscillation, and so we can roughly think of

Heuristic :- $$\tilde{\chi}_{C^\delta}(y) = e^{2\pi i \xi_{C^\delta} \cdot y} \chi_{C^\delta}(y) \qquad (3)$$

The oscillation is a crucial feature of the problem. (3) can be thought of as an analogue of the wave packet decomposition from Fourier restriction theory.

From this definition, we can write

$$\| M^j f \|_{L^p(Q_0)} \lesssim \delta^{d/p - 1} \Big(\sum_{C \in \mathcal{C}} \Big| \int_{\mathbb{R}^d} f(y) \tilde{\chi}_{C^\delta}(y)\, dy \Big|^p \Big)^{1/p}$$

where $\mathcal{C}$ is a collection of spheres centred at δ-separated points in Q_0 with radii lying in $[1,2]$.

We can now argue via duality exactly as in Lecture 1 to deduce the following :-

Lemma (Duality) :- To prove the maximal bound (1) for some fixed p, it suffices to show

$$\Big\| \sum_{C \in \mathcal{C}} a_C \tilde{\chi}_{C^\delta} \Big\|_{L^{p'}(\mathbb{R}^d)} \lesssim \delta^{1 - d/p + \varepsilon(p)} \Big(\sum_{C \in \mathcal{C}} |a_C|^{p'} \Big)^{1/p'} \qquad (4)$$

for all $\mathcal{C}$ as above, $\delta := 2^{-j}$ and all sequences $(a_C)_{C\in\mathcal{C}}$.

Note, in contrast with the geometric argument, we must now show additional gain in the form of the $\delta^{\varepsilon(p)}$ factor, which will arrive from the oscillation.

It turns out that this set up is much easier to work with and, in particular, we will be able to adapt the arguments from the previous ~~lecture~~ lecture to the frequency-localised setting to prove Stein's theorem in the full range $p > \frac{d}{d-1}$ (for $d \geq 3$).

The idea is as follows:-

- We will show (4) holds for $p = 2$ with a very good choice of exponent $\varepsilon(p) > 0$. In fact, we have

Lemma (Frequency localised L^2-bound): For all $d \geq 2$, we have

$$\Big\| \sum_{C\in\mathcal{C}} a_C \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)} \lesssim \delta^{1-d/2+\varepsilon_d(2)} \Big(\sum_{C\in\mathcal{C}} |a_C|^2\Big)^{1/2} \qquad (5)$$

where $\varepsilon_d(2) := \dfrac{d-2}{2}$

Note that (5) also holds for $d = 2$ (the corresponding lemma in Lecture 1 failed for $d=2$). However, when $d = 2$ the exponent

$\varepsilon_2(2) = 0$

so we cannot sum together the frequency localised pieces. Nevertheless (5) will play an important role in the proof of Bourgain's theorem.

From (5), we immediately recover the L^2 bound for M from Lecture 1. However, we can leverage the fact that the exponent $\varepsilon_d(2)$ is more than we need just to prove the L^2 bound (ie any $\varepsilon_p(2) > 0$ would have sufficed for the L^2 bound) to go beyond L^2. In particular, we will interpolate (5) against a trivial L^∞ estimate.

Lemma (Frequency localised L^∞ bound). For all $d \geq 2$, we have

$$\Big\| \sum_{C \in \mathcal{C}} a_C \tilde{\chi}_{C^\delta} \Big\|_{L^\infty(\mathbb{R}^d)} \lesssim \delta^{1-d+\varepsilon_d(1)} \sup_{C \in \mathcal{C}} |a_C|$$

where $\varepsilon_d(1) := -1$

Proof :- We simply bound

$$\Big\| \sum_{C \in \mathcal{C}} a_C \tilde{\chi}_{C^\delta} \Big\|_\infty \leq \# \mathcal{C} \cdot \sup_{C \in \mathcal{C}} \| \tilde{\chi}_{C^\delta} \|_\infty \sup_{C \in \mathcal{C}} |a_C|.$$

Since the δ-separation of the centres ensures that $\# \mathcal{C} \lesssim \delta^{-d}$, it suffices to show

$$\| \tilde{\chi}_{C^\delta} \|_\infty \lesssim 1 \quad \text{for all} \quad C \in \mathcal{C}.$$

(note this is consistent with the heuristic (3)).

In view of the definition of $\tilde{\chi}_{C^\delta}$, the problem boils down to showing

$$\sup_{x \in \mathbb{R}^d,\, 1 \leq r \leq 2} \int_{S^{d-1}} | \varphi(\delta^{-1}(x - ry)) | \, d\sigma \lesssim \delta^{d-1}.$$

If we ignore Schwartz tails, then the function φ can be thought of as concentrated on $B(0,1)$. The above integral is then bounded by the maximum area of intersection between rS^{d-1} and a ball $B(x,\delta)$, which clearly gives the desired bound. It is not difficult to make this rigorous. $\square$

Corollary: (Frequency localised L^p bound): For all $d \geq 2$, we have

$$\Big\| \sum_{C \in \mathcal{C}} a_C \tilde{\chi}_{C^\delta} \Big\|_{L^{p'}(\mathbb{R}^d)} \lesssim \delta^{1 - d/p + \varepsilon_d(p)} \Big(\sum_{C \in \mathcal{C}} |a_C|^{p'} \Big)^{1/p'}$$

for all $1 \leq p \leq 2$, where

$$\boxed{\varepsilon_d(p) := d - 1 - d/p}.$$

Proof: This follows by interpolating the previous lemmas. □

If $d \geq 3$, then $\varepsilon_d(p) > 0$ for $p > \frac{d}{d-1}$ and so the corollary implies Stein's maximal theorem.

For $d = 2$, we still have $\varepsilon_2(p) > 0$ for $p > 2$, but the problem is that this falls strictly outside the admissible range of exponents $1 \leq p \leq 2$ for the corollary.

Nevertheless, the corollary is very close to proving the circular maximal theorem, in the sense that the $\varepsilon_2(2) = 0$ result "just fails" to prove the summability of the frequency localised pieces for $p > 2$.

L^2 theory:-

The above argument reduces the entirity of Stein's maximal theorem to proving an L^2 bound for the localised pieces. To prove this L^2 bound, we argue as in Lecture 1, first writing

$$\Big\| \sum_{C \in \mathcal{C}} a_C \tilde{\chi}_{C^\delta} \Big\|^2_{L^2(\mathbb{R}^d)} = \sum_{C \in \mathcal{C}} |a_C|^2 \| \tilde{\chi}_{C^\delta} \|^2_{L^2(\mathbb{R}^d)}$$

$$+ \sum_{\substack{C_1, C_2 \in \mathcal{C} \\ C_1 \neq C_2}} a_{C_1} \overline{a_{C_2}} \langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle .$$

So now we wish to bound the inner products $|\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle|$. In the geometric argument in Lecture 1, this corresponded to estimating the volumes $|C_1^\delta \cap C_2^\delta|$ but here cancellation will play a central rôle.

Lemma (Diagonal estimate):–

$$\| \tilde{\chi}_{C^\delta} \|_{L^2(\mathbb{R}^d)} \lesssim \delta^{1/2}$$

This lemma relies on the decay properties of $\hat{\sigma}$. In particular, since σ is a radial measure,

$$\hat{\sigma}(\xi) = B_d(|\xi|) \quad \text{is radial,}$$

for some function $B_d : [0,\infty) \to \mathbb{C}$. Moreover, we have the standard van der Corput estimate

$$|B_d(r)| \lesssim (1+r)^{-(d-1)/2}. \qquad (6)$$

Proof (of Lemma):– By Plancherel,

$$\| \tilde{\chi}_{C^\delta} \|^2_{L^2(\mathbb{R}^d)} = \delta^2 \int_{\hat{\mathbb{R}}^d} |\hat{\sigma}_r(\xi)|^2 \, |\hat{\varphi}_j(\xi)|^2 \, d\xi$$

$$= \delta^2 \int_{\hat{\mathbb{R}}^d} |\hat{\sigma}(-\xi)|^2 \, |\mathcal{B}(\delta|\xi|)|^2 \, d\xi$$

$$\lesssim \delta^2 \int_{\delta^{-1}/2 \le |\xi| \le 2\delta^{-1}} \left[(1+|\xi|)^{-(d-1)/2} \right]^2 d\xi$$

$$\sim \delta^2 \cdot \delta^{2(d-1)/2} \, \delta^{-d} = \delta \qquad \square$$

We can use a variant of this argument to bound the off-diagonal terms. It turns out that the Fourier decay condition (6) alone is not sufficient for our purposes, however, and a more sophisticated bound is needed. Nevertheless, we give the off-diagonal generalisation of the above lemma as a step towards understanding what follows.

Lemma (Off-diagonal estimate : Take 1).

$$|\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle| \lesssim \delta \, (1 + \delta^{-1} \mathrm{dist}(C_1, C_2))^{-\frac{d-1}{2}} \qquad (7)$$

Here, as in Lecture 1, $\mathrm{dist}(C_1, C_2)$ denotes the distance between the centres of C_1 and C_2.

Proof :- Let $C_\ell = C(x_\ell, r_\ell)$ so that

$$\tilde{\chi}_{C_\ell^\delta}(y) = \delta \cdot \varphi_j * \sigma_{r_\ell}(x_\ell - y).$$

Taking Fourier transforms, we have

$$(\tilde{\chi}_{C_\ell^\delta})^\wedge(\xi) = \delta \cdot e^{-2\pi i x_\ell \cdot \xi} \, \hat{\varphi}_j(\xi) \, \hat{\sigma}_{r_\ell}(\xi)$$

$$= \delta \cdot e^{-2\pi i x_\ell \cdot \xi} \, \beta(\delta|\xi|) \, B_d(r_\ell |\xi|)$$

Applying Plancherel,

$$\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle = \delta^2 \int_{\hat{\mathbb{R}}^d} B_d(r_1|\xi|) \overline{B_d(r_2|\xi|)} \, |\beta(\delta|\xi|)|^2 e^{-2\pi i (x_1 - x_2)\cdot\xi} \, d\xi$$

Since many of the functions in the integrand are radial, it makes sense to apply polar coordinates. This gives

$$\delta^2 \int_0^\infty B_d(r_1 \rho) \overline{B_d(r_2 \rho)} \, |\beta(\delta\rho)|^2 \int_{S^{d-1}} e^{-2\pi i \rho (x_1 - x_2)\cdot\theta} \, d\sigma(\theta) \, \rho^{d-1} d\rho$$

$$= \delta^2 \int_0^\infty B_d(r_1 \rho) \overline{B_d(r_2 \rho)} \, B_d(\rho|x_1 - x_2|) \, |\beta(\delta\rho)|^2 \rho^{d-1} d\rho.$$

We now estimate as in the diagonal case. This gives

$$|\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle| \lesssim \delta^2 \cdot \int_{\delta^{-1}/2}^{2\delta^{-1}} \rho^{-(d-1)/2} \rho^{-(d-1)/2} (1 + \rho|x_1 - x_2|)^{-\frac{d-1}{2}} \rho^{d-1} d\rho$$

$$\sim \delta \cdot (1 + \delta^{-1}|x_1 - x_2|)^{-(d-1)/2},$$

as required. $\square$

The diagonal estimate is in fact better than what we need. The off-diagonal estimate is, however, too weak. To improve it, we take into account not only the decay of the Fourier transform $\widehat{\sigma}$, but also its oscillation. In particular, we have the classical stationary phase formula

$$\widehat{\sigma}(\xi) = \sum_{\pm} e^{\pm 2\pi i |\xi|} a_{\pm}(|\xi|)$$

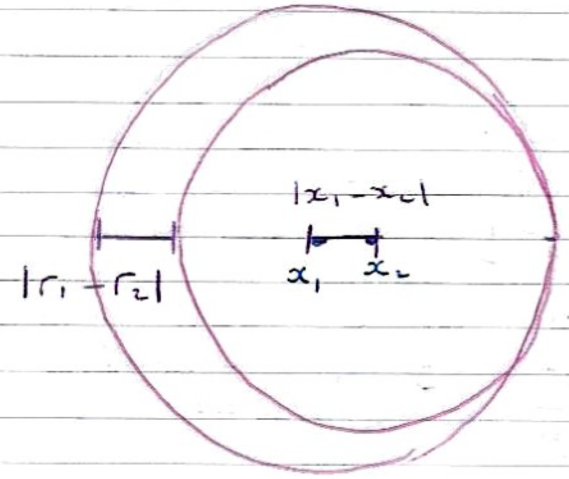
where $|a_{\pm}(r)| \lesssim (1+r)^{-\frac{d-1}{2}}$. It will be crucial to exploit cancellation arising from the oscillatory terms $e^{\pm 2\pi i |\xi|}$.

The refined version of the off-diagonal estimate involves the quantity

$$\Delta(C_1, C_2) := \big| |x_1 - x_2| - |r_1 - r_2| \big| \qquad C_\ell = C(x_\ell, r_\ell)$$

which measures tangency between two spheres. In particular, note C_1 and C_2 are interior tangent if and only if $\Delta(C_1, C_2) = 0$, ie.

$$|x_1 - x_2| = |r_1 - r_2| :-$$



Interior tangent spheres.

Lemma (Off-diagonal estimates: Take 2) For all $N \in \mathbb{N}$,

$$|\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle| \lesssim_N \delta \cdot (1 + \delta^{-1} \mathrm{dist}(C_1, C_2))^{-\frac{d-1}{2}} (1 + \delta^{-1} \Delta(C_1, C_2))^{-N} \tag{8}$$

Thus, the terms $\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle$ contribute very little to the sum unless $\Delta(C_1, C_2) \lesssim \delta$.

We postpone the proof of the off-diagonal estimate until later in the lecture. For now we show how it can be used to conclude the proof of the frequency localised L^2 bound.

Proof (of Freq. localised L^2 lemma) :-

We first note that $\mathrm{dist}(C_1, C_2) = \Delta(C_1, C_2)$ if $r_1 = r_2$. Thus, if $r_1 = r_2$, then (8) implies

$$|\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle| \lesssim \delta \left(1 + \delta^{-1} \mathrm{dist}(C_1, C_2)\right)^{-(d+1)},$$

which improves over (7). This continues to hold if $|r_1 - r_2| \lesssim \delta$.

Recall, we wish to show

$$\Big\| \sum_{C \in \mathcal{C}} a_C \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)} \lesssim \Big(\sum_{C \in \mathcal{C}} |a_C|^2 \Big)^{1/2}. \qquad (9)$$

To make things simpler, we will assume $a_C = 1$ for all $C \in \mathcal{C}$ so that the bound becomes

$$\Big\| \sum_{C \in \mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)} \lesssim [\#\mathcal{C}]^{1/2}. \qquad (10)$$

We can in fact prove (9) using the same methods, but this involves a few extra applications of Cauchy-Schwarz: see Lecture 1.

We first split up the function $\sum_{C \in \mathcal{C}} \tilde{\chi}_{C^\delta}$ according to the size of the radii. We let $(I_k)_{k=1}^{\overline{K}}$ be a partition of $[1,2]$ into $\overline{K} \sim \delta^{-1}$ intervals of length $\sim \delta$. For each $1 \leq k \leq \overline{K}$ we let $\mathcal{C}_k := \{ C \in \mathcal{C} : \mathrm{rad}(C) \in I_k \}$ where $\mathrm{rad}(C)$ denotes the radius of C.

By Cauchy-Schwarz,

$$\Big\| \sum_{C \in \mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)} \leq \Big\| \sum_{k=1}^{\overline{K}} \Big| \sum_{C \in \mathcal{C}_k} \tilde{\chi}_{C^\delta} \Big| \Big\|_{L^2(\mathbb{R}^d)}$$

$$\lesssim \delta^{-1/2} \left(\sum_{k=1}^{K} \Big\| \sum_{C \in \mathcal{C}_k} \tilde{\chi}_{C^\delta} \Big\|^2_{L^2(\mathbb{R}^d)} \right)^{1/2} \qquad (11).$$

We now fix k and write

$$\Big\| \sum_{C \in \mathcal{C}_k} \tilde{\chi}_{C^\delta} \Big\|_2^2 = \sum_{C_k \in \mathcal{C}_k} \| \tilde{\chi}_{C^\delta} \|_2^2 + \sum_{\substack{C_1, C_2 \in \mathcal{C}_k \\ C_1 \neq C_2}} \langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle$$

$$\lesssim \delta \cdot \# \mathcal{C}_k + \delta \sum_{\ell=0}^{\lceil \log_2 \delta^{-1} \rceil} \sum_{C_1 \in \mathcal{C}_k} \sum_{\substack{C_2 \in \mathcal{C}_k \\ \mathrm{dist}(C_1, C_2) \sim 2^\ell \delta}} (1 + \delta^{-1} \mathrm{dist}(C_1, C_2))^{-(d+1)}$$

$$\lesssim \delta \cdot \# \mathcal{C}_k + \delta \sum_{\ell=0}^{\lceil \log_2 \delta^{-1} \rceil} \sum_{C_1 \in \mathcal{C}_k} \# \{ C_2 \in \mathcal{C}_k : \mathrm{dist}(C_1, C_2) \sim 2^\ell \delta \} \, 2^{-(d+1)\ell}$$

$$\lesssim \delta \, \# \mathcal{C}_k \left(1 + \sum_{\ell=0}^{\lceil \log_2 \delta^{-1} \rceil} 2^{-\ell} \right) \lesssim \delta \, \# \mathcal{C}_k .$$

Thus, $\Big\| \sum_{C \in \mathcal{C}_k} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)} \lesssim \delta^{1/2} [\# \mathcal{C}_k]^{1/2}$

and combining this with (11) gives

$$\Big\| \sum_{C \in \mathcal{C}_k} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)} \lesssim \left(\sum_{k=1}^{K} \# \mathcal{C}_k \right)^{1/2} = [\# \mathcal{C}]^{1/2},$$

which is precisely the desired bound (10). $\square$.

Weak orthogonality: the proof of the refined off-diagonal estimate.

Proof:- Without loss of generality, assume $r_1 \geq r_2$. Arguing as before, we can write

$$\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle = \qquad (12)$$

$$\delta^2 \int_0^\infty B_d(r_1 \rho) \overline{B_d(r_2 \rho)} \, B_d(\rho |x_1 - x_2|) \, |\beta(\delta \rho)|^2 \rho^{d-1} \, d\rho .$$

Rather than apply the van der Corput bound

$$|B_d(r)| \lesssim (1 + r)^{-\frac{d-1}{2}}$$

we apply the asymptotic expansion

$$B_d(r) = \sum_{m=0}^{M} \Big(\sum_{\pm} c^{\pm}_{m,d}\, e^{\pm 2\pi i r} \Big) r^{-m-\frac{d-1}{2}} + r^{-M-\frac{d-1}{2}} \mathcal{E}_{M,d}(r)$$

where $\left| \frac{\partial^k}{\partial r^k} \mathcal{E}_{M,d}(r) \right| \lesssim_{M,d,k} 1$ for all $k \in \mathbb{N}_0$.

We first consider the expression resulting from the leading term of this expansion. More precisely, if we make the substitutions

$$B_d(r_1 e) \longrightarrow c^{+}_{0,d}\, e^{2\pi i r_1 e} (r_1 e)^{-\frac{d-1}{2}}$$

$$B_d(r_2 e) \longrightarrow c^{+}_{0,d}\, e^{2\pi i r_2 e} (r_2 e)^{-\frac{d-1}{2}}$$

$$B_d(e|x_1 - x_2|) \longrightarrow c^{-}_{0,d}\, e^{-2\pi i e|x_1 - x_2|} (e|x_1 - x_2|)^{-\frac{d-1}{2}}$$

in (12), then we obtain an expression of the form

$$C(r_1, r_2) \cdot \delta^2 \int_0^\infty e^{2\pi i (r_1 - r_2 - |x_1 - x_2|)e} |\mathcal{B}(\delta^{+1} e)|^2 e^{\frac{d-1}{2}} de \cdot |x_1 - x_2|^{-\frac{d-1}{2}}$$

$$= C(r_1, r_2) \cdot \delta \cdot (\delta^{-1}|x_1 - x_2|)^{-\frac{d-1}{2}} . \qquad (13)$$

$$(\mathcal{F}^{-1}\alpha)(\delta^{-1}(r_1 - r_2 - |x_1 - x_2|))$$

where $\alpha(e) := |\mathcal{B}_+(e)|^2 \cdot e^{\frac{d-1}{2}}$ is a Schwartz function.

Thus, by rapid decay of $(\mathcal{F}^{-1}\alpha)$, this term is dominated by

$$C_N\, \delta\, (\delta^{-1} |x_1 - x_2|)^{-\frac{d-1}{2}} (\delta^{-1} \Delta(C_1, C_2))^{-N} \qquad (14)$$

for all $N \in \mathbb{N}_0$.

There are other contributions to the first order term, coming from different choices of sign in the phases:

$$e^{\pm 2\pi i r_1 \ell}, \quad e^{\pm 2\pi i r_2 \ell}, \quad e^{\pm 2\pi i \ell |x_1 - x_2|}.$$

These different choices lead to similar expressions to (13) but with $(\mathcal{F}^{-1}\alpha)(\delta^{-1}\Delta(C_1, C_2))$ replaced with

$$(\mathcal{F}^{-1}\alpha)(\delta^{-1}(-r_1 + r_2 - |x_1 - x_2|)) \qquad (15a),$$

$$(\mathcal{F}^{-1}\alpha)(\delta^{-1}(r_1 + r_2 + |x_1 - x_2|)) \qquad (15b),$$

$$(\mathcal{F}^{-1}\alpha)(\delta^{-1}(r_1 + r_2 - |x_1 - x_2|)) \qquad (15c);$$

there are also 4 other terms, but these correspond to reflecting the arguments in the terms considered above and are therefore treated identically.

For (14a), we have

$$\begin{aligned} |-r_1 + r_2 - |x_1 - x_2|| &= r_1 - r_2 + |x_1 - x_2| \\ &\geq ||r_1 - r_2| - |x_1 - x_1|| \\ &= \Delta(C_1, C_2) \end{aligned}$$

and so the corresponding term can be bonded by (14) as before.

For (15b) and (15c), we have

$$r_1 + r_2 + |x_1 - x_2| \gtrsim 1,$$

$$r_1 + r_2 - |x_1 - x_2| \gtrsim 1,$$

since $r_1, r_2 \in [1, 2]$ and $|x_1 - x_2| \in [0, 1]$, by our hypothesis on ℓ. Thus (15b) and (15c) are both $O_N(\delta^{100dN})$ and therefore lead to negligible contributions.

The same analysis can be applied to other terms in the asymptotic expansion. Finally, by choosing M sufficiently large, we can ensure the contribution arising from the error term is, say,

$$O_N(\delta^{100dN}) \quad \text{and therefore negligible.}$$

From all this, we see

$$|\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle| \lesssim_N \delta \cdot (\delta^{-1} \operatorname{dist}(C_1, C_2))^{-\frac{d-1}{2}} \cdot (\delta^{-1} \Delta(C_1, C_2))^{-N}.$$

Combining this with the Cauchy-Schwarz bound $|\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle| \lesssim \delta$, we conclude

$$|\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle| \lesssim_N \delta \, (1 + \delta^{-1} \operatorname{dist}(C_1, C_2))^{-\frac{d-1}{2}} \cdot (1 + \delta^{-1} \Delta(C_1, C_2))^{-N},$$

as required. $\square$