IMPROVED DECOUPLING FOR THE PARABOLA

LARRY GUTH, DOMINIQUE MALDAGUE, AND HONG WANG

Abstract. We prove an \((L^2, L^6)\) decoupling inequality for the parabola with constant \((\log R)^c\). In the appendix, we present an application to the sixth-order correlation of the integer solutions to \(x^2 + y^2 = m\).

Contents

1. Introduction and main results 1
   1.1. Technical issues 3
   2. Intuitive explanation of the argument 9
      2.1. Technical issues 9
   3. Proof of Theorem 1.2 – the broad, well-spaced case 12
      3.1. Statement of Proposition 3.4 12
      3.2. Auxiliary functions 13
      3.3. Pruning wave packets 17
      3.4. High/low lemmas for \(g_k\) 21
      3.5. The sets \(\Omega_k\) 22
      3.6. Proof of Proposition 3.4 24
   4. Proof of Theorem 1.2 – the general case 28
      4.1. Removing the broad hypothesis 28
      4.2. Removing the well-spaced hypothesis 33
      4.3. Proof of Theorem 1.2 35
   5. Showing \(\text{Dec}_6(R) \lesssim (\log R)^c\) from Theorem 1.2 35
      5.1. Wave packet decomposition and pigeonholing 35
      5.2. Proof of Theorem 1.1 37
   Appendix 39
   References 40

1. Introduction and main results

Let \(f : \mathbb{R}^n \to \mathbb{C}\) be in the Schwartz class \(S\) with Fourier support contained in \(\mathcal{N}_{R^{-1}}(\mathbb{P}^{n-1})\), the \(R^{-1}\) neighborhood of \(\mathbb{P}^{n-1} := \{(\xi, |\xi|^2), |\xi| \leq 1, \xi \in \mathbb{R}^{n-1}\}\). Let \(\{\theta\}\) be a tiling of \(\mathcal{N}_{R^{-1}}(\mathbb{P}^{n-1})\) by approximately \(R^{-1/2} \times \cdots \times R^{-1/2} \times R^{-1}\) rectangular boxes \(\theta\) and define \(f_\theta = (\hat{f} \chi_\theta)^\vee\).

Let \(D_{n,p}(R)\) denote the smallest constant such that

\[
\|f\|_{L^p(\mathbb{R}^n)} \leq D_{n,p}(R) \left( \sum_\theta \|f_\theta\|_{L^p(\mathbb{R}^n)}^2 \right)^{1/2},
\]

for any \(f \in S\) with \(\text{supp} \hat{f} \subset \mathcal{N}_{R^{-1}}(\mathbb{P}^{n-1})\).

Date: September 18, 2020.
A trivial estimate using Cauchy-Schwarz and the triangle inequality yields $D_{n,p}(R) \leq R^{(n-1)/2}$. And we have $D_{n,p}(R) \geq 1$ by taking $f = f_\theta$. Bourgain and Demeter [3] proved that for $2 \leq p \leq \frac{2(n+1)}{n-1}$, $D_{n,p}(R) \leq C_\epsilon R^\epsilon$ for any small $\epsilon > 0$. Such estimates are possible due to the curvature of $\mathbb{P}^{n-1}$ and are sharp up to $R^\epsilon$-loss. The estimates have many applications in harmonic analysis, PDE and number theory. It was conjectured that $D_{n,p} \leq C_p$ for $1 \leq p < \frac{2(n+1)}{n-1}$.

In this paper, we focus on the case $n = 2$ and write $D_p(R) = D_{n,p}(R)$. At the end point $p = 6$, Bourgain proved in [2] that $D_6(R) \geq (\log R)^{1/6}$. Based on the Bourgain-Demeter decoupling, Zane Li [7] proved that $D_6(R) \leq \exp(O(\frac{\log R}{\log \log R}))$. Then, by adapting ideas from efficient congruencing, he proved [9] that $D_6(R) \leq \exp(O(\frac{\log R}{\log \log R}))$. This was the best previous bound for $D_6(R)$. In this paper, we prove

**Theorem 1.1.** $D_6(R) \lesssim (\log R)^c$ for an absolute constant $c$.

Theorem 1.1 is a corollary of our main theorem, which estimates the $L^6$-norm of $f$ on a subset of $\mathbb{R}^2$.

**Theorem 1.2.** There exists $c > 0$ such that the following holds. If $f \in S$ has Fourier support contained in $N_{R^{-1}}(\mathbb{P}^1)$ and $Q_R \subset \mathbb{R}^2$ is any cube of sidelength $R$, then

$$
\|f\|_{L^6(Q_R)}^6 \leq (\log R)^c \left(\sum_\theta \|f_\theta\|_{L^\infty(\mathbb{R}^2)}^2\right)^2 \sum_\theta \|f_\theta\|_{L^2(\mathbb{R}^2)}^2 \quad \forall R \geq 2.
$$

The proof of Theorem 1.2 is related to an incidence estimate between points and rectangles used in [2] and [5]. These arguments are based on the following idea. We consider the square function $g = \sum_\theta |f_\theta|^2$, and we divide it into a high-frequency part and a low frequency part. For the high frequency part, the different terms $|f_\theta|^2$ are essentially orthogonal, and this gives a powerful tool when the high frequency part of $g$ dominates. When the low frequency part of $g$ dominates, we try to reduce the whole problem to a similar problem at a coarser scale.

We use these tools to give a different proof of decoupling for the parabola. Compared to the two previous proofs (by Bourgain-Demeter [3] and Li [9], our proof leans less heavily on induction on scales, and we think this is the main reason it gives a stronger estimate. In order to obtain the $(\log R)^c$ bound, we also need to deal carefully with a number of technical difficulties. These include a wave packet decomposition using Gaussian partitioning of unity, carefully modifying the function at each scale and reducing to a well-spaced frequency case.

We give an intuitive explanation of the argument in Section 1.

The bound $(\log R)^c$ is useful compared to $R^\epsilon$ in some diophantine equation problems. Let $\Lambda_m = \{(x, y) \in \mathbb{Z}^2, x^2 + y^2 = m\}$. In [1], Bombieri and Bourgain studied the number of solutions of the system $\lambda_1 + \lambda_2 + \lambda_3 = \lambda_4 + \lambda_5 + \lambda_6$ with $\lambda_j \in \Lambda_m$. In [8], Li and Bourgain applied decoupling to this problem. They were able to prove a very strong bound for the number of solutions provided that $\Lambda_m$ is very large. Using our stronger estimate for $D_6(R)$, we can extend their bound to a wider range of $\Lambda_m$. We present this application in the appendix.

Another corollary of Theorem 1.1 concerns the discrete Fourier restriction on $\{(n, n^2), n \in \mathbb{Z}\}$.

**Corollary 1.3.** Let $K_p(N)$ denote the smallest constant such that for any $\{a_n\}_{|n| \leq N}$,

$$
\| \sum_{|n| \leq N} a_n e^{2\pi i (nx + n^2 t)} \|_{L^p(\mathbb{T}^2)} \lesssim K_p(N) (\sum_{|n| \leq N} |a_n|^2)^{1/2}.
$$
Then $K_6(N) \lesssim (\log N)^C$.

Bourgain showed in [2] ((2.51), Proposition 2.36) that $c(\log N)^{1/6} \leq K_6(N) \leq \exp(c(\log N)^{1/6})$. He also asked whether $K_p(N)$ is bounded independent of $N$ for each $p < 6$.

2. Intuitive explanation of the argument

In this section, we outline the main ideas of the proof. For simplicity, we suppress some minor technical details, but at the end we will discuss the most important technical issues that come up.

Let $\text{Dec}(R)$ be the optimal constant in the decoupling inequality

$$\|f\|_{L^6(Q_R)} \leq \text{Dec}(R) \left( \sum_\theta \|f_\theta\|_{L^6(\omega_R)}^2 \right)^{1/2}.$$  

Here the $\theta$ denote $\sim R^{-1/2} \times R^{-1}$ approximate rectangles which partition an $R^{-1}$ neighborhood of $\mathbb{P}^1$ and $f_\theta = f \chi_\theta$ where $f$ is a Schwartz function. The original arguments of Bourgain and Demeter to prove that $\text{Dec}(R) \leq C_\varepsilon R^\varepsilon$ involve analysis of $f_\tau$ where $\tau$ is a rectangle in a neighborhood of $\mathbb{P}^1$ and $\tau$ is at various scales between 1 and the final scale $R$. This is also true for the proof of Theorem 1.2 which involves analysis of $f$ at $\sim \log R$ many scales. Use the notations $A \lesssim B$ and $A \approx B$ to mean $A \leq CB$ and $A \leq (\log R)^C B$, respectively, for some absolute constant $C$.

By a standard pigeonholing argument (see [3]), the decoupling inequality above follows from the estimate

$$\alpha^6|\{x \in Q_R : |f| \sim \alpha\}| \lesssim \left( \sum_\theta \|f_\theta\|_{L^\infty(Q_R)}^2 \right)^2 \sum_\theta \|f_\theta\|_{L^2(Q_R)}^2$$

where we may assume that for each $\theta$,

$$\|f_\theta\|_{L^\infty(Q_R)} \sim 1 \text{ or } f_\theta = 0$$

and that $\|f_\theta\|_{L^p(Q_R)}$ are comparable for all non-zero $f_\theta$ and all $2 \leq p \leq 6$. Note that inequality (2.2) is also (roughly) the statement of Theorem 1.2. In this section, we are suppressing the weight functions localized to $Q_R$ which are present in the $L^2$-norms on the right hand side above.

Recall the reverse square function estimate for $L^4$, which says that

$$\alpha^4|\{x \in Q_R : |f| \sim \alpha\}| \lesssim \int_{Q_R} \left( \sum_\theta |f_\theta|^2 \right)^2.$$  

If we use Minkowski’s integral inequality to interchange the $\ell^2$ and $L^4$ norms, then we immediately get $L^4$-decoupling

$$\alpha^4|\{x \in Q_R : |f| \sim \alpha\}| \lesssim \left( \sum_\theta \left( \int_{Q_R} |f_\theta|^4 \right)^{1/2} \right)^2.$$  

The heart of our argument involves analyzing special cases where we can upgrade (2.3) into something that implies $L^6$-decoupling. We will describe the simplest special case of the argument now.

Special case: High frequency dominance. Consider the square function that appears on the right hand side of (2.3), $\sum_\theta |f_\theta|^2$. Each summand $|f_\theta|^2 = f_\theta \overline{f_\theta}$ has Fourier support in $\theta - \theta$, which looks like a copy of the $R^{-1/2} \times R^{-1}$ rectangle $\theta$ that is dilated by a factor
of 2 and translated to the origin. Let $\eta$ be a smooth approximation of the characteristic function of a ball of radius $R^{-1/2}/\log R$. Define the low frequency part as

$$\left( \sum_\theta |f_\theta|^2 \right)_\ell := \sum_\theta |f_\theta|^2 * \tilde{\eta}$$

and the high part by

$$\left( \sum_\theta |f_\theta|^2 \right)_h := \sum_\theta |f_\theta|^2 - \left( \sum_\theta |f_\theta|^2 \right)_\ell.$$

In this special case, we assume that

$$\int_{Q_R} \left| \left( \sum_\theta |f_\theta|^2 \right)_h \right|^2 \lesssim \int_{Q_R} \left| \left( \sum_\theta |f_\theta|^2 \right)_h \right|^2.$$

We would like to use Plancherel’s theorem to analyze the right hand side, but we are integrating over $Q_R$ instead of $\mathbb{R}^2$. This is solved by using a weight function $w_{Q_R}$ to approximate $\chi_{Q_R}$ that has the property that $\hat{w}_{Q_R}$ is supported in a ball of radius $R^{-1}$. This blurs the support properties of the Fourier transform of $\left( \sum_\theta |f_\theta|^2 \right)_h$ by a factor of $R^{-1}$, but does not alter the main properties. In particular, $\left( \sum_\theta |f_\theta|^2 \right)_h$ has Fourier transform supported on $\cup (\theta - \theta)$ intersected with the compliment of the ball centered at the origin of radius $R^{-1/2}/\log R$. Since the $\theta$ cover a small $R^{-1}$-neighborhood of $\mathbb{P}^1$ (which curves), the rectangles $\theta - \theta$ are oriented at angles which are pairwise $\gtrsim R^{-1/2}$-separated. The overlap of these rectangles outside of $B(R^{-1/2}/\log R)$ is $\sim \log R$. Thus, by Cauchy-Schwarz,

$$\int_{Q_R} \left| \left( \sum_\theta |f_\theta|^2 \right)_h \right|^2 \lesssim \log R \int_{Q_R} \sum_\theta |f_\theta|^4$$

where technically the auxiliary function $\eta$ associated to the definition of the high and low frequency parts is absorbed into a weight function discussed in more detail below.

To summarize, we have so far that

$$\alpha^4 | \{ x \in Q_R : |f| \sim \alpha \} | \lesssim \int_{Q_R} \sum_\theta |f_\theta|^4.$$

This means that we have upgraded our $(\ell^2, L^4)$-decoupling result (2.2) into an $(\ell^4, L^4)$-decoupling result. Note that for some $x \in Q_R$, $\alpha \sim |f(x)| \leq \sum_\theta \|f_\theta\|_{L^\infty(Q_R)} \sim \sum_\theta \|f_\theta\|_{L^\infty(Q_R)}^2$ and each $\|f_\theta\|_{L^\infty} \lesssim 1$. Thus,

$$\alpha^4 | \{ x \in Q_R : |f| \sim \alpha \} | \lesssim \alpha^{-2} \left( \sum_\theta \|f_\theta\|_{L^\infty(Q_R)}^2 \right)^2 \sum_\theta \int_{Q_R} |f_\theta|^2$$

which is the $L^6$-decoupling result we were aiming for and concludes the special case.

Suppose that we are not in the high frequency dominating case above, but that we have a high frequency dominance at a different scale $\tilde{R}$:

$$\int_{Q_{\tilde{R}}} \left( \sum_\tau |f_\tau|^2 \right)^2 \lesssim \int_{Q_{\tilde{R}}} \left| \left( \sum_\tau |f_\tau|^2 \right)_h \right|^2.$$

where the $\tau$ are $\tilde{R}^{-1/2} \times \tilde{R}^{-1}$ rectangles covering a $\tilde{R}^{-1}$-neighborhood of $\mathbb{P}^1$ and the high part is with respect to this new scale. If we repeat the above argument, we obtain the
\((\ell^4, L^4)\) result
\[
\alpha^4 \{x \in Q_R : |f| \sim \alpha \} \lesssim \int_{Q_R} \sum_{\tau} |f_\tau|^4.
\]
If we try to relate the right hand side to a sum of \(L^2\)-norms, then
\[
\alpha^4 \{x \in Q_R : |f| \sim \alpha \} \lesssim (\max_{\tau} \|f_\tau\|_\infty^2) \sum_{\tau} \int_{Q_R} |f_\tau|^2.
\]
By \(L^2\)-orthogonality, this is equivalent to
\[
\alpha^4 \{x \in Q_R : |f| \sim \alpha \} \lesssim (\max_{\tau} \|f_\tau\|_\infty^2) \sum_{\theta} \int_{Q_R} |f_\theta|^2.
\]

The issue now is that in the special case above, we had good control over \(\|f_\theta\|_\infty\) given in (2.2), but we don't have any corresponding estimate for \(\|f_\tau\|_\infty\). A key part of the proof is a pruning process for the wave packets of \(f_\tau\) which will allow us to control \(\|f_\tau\|_\infty\).

Our argument will involve many scales, and so we introduce a sequence of intermediate scales and high-low decompositions for each scale. Denote the intermediate scales by
\[
1 < R_1 < \cdots < R_k < R_{k+1} < \cdots < R_N = R.
\]
We will use scales which have the property that
\[
\frac{R_{k+1}}{R_k} \sim (\log R)^c.
\]
Let \(\{\tau_k\}\) denote \(R_k^{-1/2} \times R_k^{-1}\) rectangles which partition an \(R_k^{-1}\) neighborhood of \(\mathbb{P}^1\). Note that for each \(k = 1, \ldots, N\),
\[
f = \sum_{\tau_k} f_{\tau_k}.
\]
We analyze the square functions
\[
g_k = \sum_{\tau_k} |f_{\tau_k}|^2.
\]
Intuitively since the first scale \(R_1 \approx 1\), we have by Cauchy-Schwarz that
\[
|f| \lesssim g_1^{1/2},
\]
and we should think of \(g_1\) as being close to \(|f|^2\). On the other hand, \(g_N\) is our original square function \(\sum_{\theta} |f_\theta|^2\).

We define a high-low decomposition for \(g_k\), building on [4] and [5]. As in the special case above, observe that the Fourier transform of \(g_k\) is
\[
\hat{g}_k = \sum_{\tau_k} |f_{\tau_k}|^2 = \sum_{\tau_k} \hat{f}_{\tau_k} \ast \hat{f}_{\tau_k}.
\]
By definition, \(\hat{f}_{\tau_k}\) has support on \(\tau_k\) and \(\hat{f}_{\tau_k}\) has support on \(-\tau_k\). Thus
\[
\text{supp} |f_{\tau_k}|^2 \subset \tau_k - \tau_k
\]
and \(\tau_k - \tau_k\) is the same as \(\tau_k\) translated to the origin and dilated by a factor of 2. Since the \(\{\tau_k\}\) formed a partition of the neighborhood of the parabola, the Fourier support of \(g_k\) is a union of \(\sim R_k^{-1/2} \times R_k^{-1}\) rectangles centered at the origin oriented at \(\sim R_k^{-1/2}\)-separated angles. Note that the intersection of all of these tubes is an \(R_k^{-1}\)-ball centered at the origin,
and outside of some neighborhood of the origin, the tubes look more disjoint (or at least finitely overlapping). This is the setting for a high-low frequency decomposition.

We separate out a low-frequency part of \( g_k \) and a high frequency part of \( g_k \).

**Definition 2.1.** Let \( \eta_k(\xi) \) be a bump function associated to a ball of radius \( \rho_k \), where \( \rho_k = (\log R)^{-c} R_k^{-1/2} \) with \( c \) in (2.5). We have \( \eta_k(\xi) = 1 \) on \( B_{\rho_k} \) and \( \eta_k(\xi) = 0 \) outside of \( 2B_{\rho_k} \).

Define the low frequency part of \( g_k \) by

\[
\hat{g}_{k,\ell} = \eta_k \hat{g}_k
\]

Let the high part of \( g_k \) be equal to

\[
g_{k,h} = g_k - g_{k,\ell}.
\]

The following lemmas describe the good features of the low-frequency part and the high-frequency part of \( g_k \).

**Lemma 2.2** ("Low lemma"). Given \( \rho_k = (\log R)^{-c} R_k^{-1/2} \leq R_{k+1}^{-1/2} \),

\[
|g_{k,\ell}(x)| \leq C_{\text{low}} g_{k+1}(x).
\]

**Sketch.** First write

\[
g_{k,\ell}(x) = \sum_{\tau_k} \int |f_{\tau_k}|^2(y) \eta_k(x - y)dy.
\]

Each \( \tau_k \) is a union of \( R_k^{-1/2} \times R_k^{-1} \) rectangles, so we write \( f_{\tau_k} = \sum_{\tau_{k+1} \subseteq \tau_k} f_{\tau_{k+1}} \). Morally, the \( f_{\tau_{k+1}} \) are orthogonal in \( L^2 \) on balls of radius \( \gtrsim R_{k+1}^{1/2} \). Since \( |\eta_k| \) is locally constant on balls of radius \( \sim \rho_k^{-1} \gtrsim R_{k+1}^{1/2} \), we invoke the local orthogonality of the \( f_{\tau_{k+1}} \) to conclude

\[
|g_{k,\ell}(x)| \lesssim \sum_{\tau_k} \sum_{\tau_{k+1} \subseteq \tau_k} \int |f_{\tau_{k+1}}|^2(y) |\eta_k|(x - y)dy.
\]

Since \( g_{k+1} \) is locally constant on balls of radius \( \sim R_{k+1}^{1/2} \), we may also update the low lemma above to say

**Corollary 2.3** (White lie! "Low corollary").

\[
g_{k,\ell}(x) \leq C_{\text{low}} g_{k+1}(x).
\]

**Lemma 2.4** ("High lemma"). For any \( B_{R_k} \) ball,

\[
\int_{B_{R_k}} |g_{k,h}|^2 \lesssim \rho_k^{-1} R_k^{-1/2} \sum_{\tau_k} \int |f_{\tau_k}|^4 \omega_{R_{k+1}}
\]

where \( 1_{B_{R_k}} \leq \omega_{R_{k+1}} \) and \( \omega_{R_{k+1}} \) is supported in a ball of radius \( 2R_{k+1}^{-1} \).

**Remark.** Since \( \rho_k = (\log R)^{-c} R_k^{-1/2} \), and so the factor \( \rho_k^{-1} R_k^{-1/2} \) is bounded by \( (\log R)^c \).
For $k$ that we can choose later. This leads to a partition of the domain into the following sets:

\[ \text{Define } \rho_k \text{ as most } \| \omega \|_{k+1}, \| \omega \|_k \text{ or for some } \]

\[ \| \| \|_k \|_k \|_k \|_k. \]

The Fourier transform of $g_{k,h}$ is supported outside of a ball centered at the origin of radius $\rho_k$. The support of $\hat{g}_{h,k} \ast \hat{\omega}_{R_{k+1}}$ is outside of a ball of radius $\rho_k - R_{k+1}^{-1} \sim \rho_k$. This means that the previous displayed math equals

\[ \int_{B_{R_{k+1}}} |g_{k,h}|^2 \leq \int |g_{k,h}|^2 \omega_{R_{k+1}} \]

\[ = \int \hat{g}_{k,h} (\hat{\omega}_{k,h} \ast \hat{\omega}_{R_{k+1}}). \]

On the region $|\xi| \geq \rho_k$, the support of $|\hat{f}_{\tau_k}|^2$ intersects the support of $|\hat{f}_{\tau'_k}|^2 \ast \hat{\omega}_{R_{k+1}}$ for at most $\rho_k^{-1} R_{k+1}^{-1/2}$ many $\tau'_k$. It follows from Plancherel’s theorem and Cauchy-Schwarz that

\[ \sum_{\tau_k} \sum_{\tau'_k} \int_{|\xi| \geq \rho_k} |\hat{f}_{\tau_k}|^2 (|\hat{f}_{\tau'_k}|^2 \ast \hat{\omega}_{R_{k+1}}) \lesssim \rho_k^{-1} R_{k+1}^{-1/2} \sum_{\tau_k} |\hat{f}_{\tau_k}|^4 \omega_{R_{k+1}}. \]

The “low lemma” tells us that

\[ g_k = g_{k,\ell} + g_{k,h} \leq C_{\text{low}} g_{k+1} + |g_{k,h}|. \]

Therefore, either $g_k(x) \leq A |g_{k,h}|$ or $g_k(x) \leq A^{1/k} C_{\text{low}} g_{k+1}(x)$. Here $A$ is a parameter that we can choose later. This leads to a partition of the domain into the following sets:

\[ Q_R = \Omega \cup \Omega_1 \cup \cdots \cup \Omega_{N-1}. \]

Define

\[ \Omega_{N-1} = \{ x \in Q_R : g_{N-1}(x) \leq A |g_{N-1,h}(x)| \}. \]

For $k = 1, \ldots, N - 2$, define

\[ \Omega_k = \{ x \in Q_R \setminus (\Omega_{k+1} \cup \cdots \cup \Omega_{N-1}) : g_k(x) \leq A |g_{k,h}(x)| \} \]

\[ \subseteq \{ x \in Q_R : g_k(x) \leq A |g_{k,h}(x)|, g_{k,\ell} \leq A^{1/k} C_{\text{low}} g_{k+1} \text{ for } k + 1 \leq \ell \leq N - 1 \}. \]

The set $L$ is defined as

\[ Q_R \setminus (\Omega_1 \cup \cdots \cup \Omega_{N-1}) \subseteq \{ x \in Q_R : g_{\ell}(x) \leq A^{1/k} C_{\text{low}} g_{k+1} \text{ for } 1 \leq \ell \leq N - 1 \}. \]

Since we have partitioned $Q_R$ into $\sim \log R$ many sets, it suffices to consider the cases

\[ \| f \|_{L^6(Q_R)} \lesssim (\log R) \| f \|_{L^6(L)} \]

or for some $k$,\n
\[ \| f \|_{L^6(Q_R)} \lesssim (\log R) \| f \|_{L^6(\Omega_k)}. \]

The case where $L$ dominates, which means (2.7) holds, is simple because for $x \in L$, $|f(x)|^2 \lesssim g_1(x) \lesssim g_N(x) = \sum_{\theta} |f_{\theta}(x)|^2$, and so

\[ \int_L |f|^6 \lesssim \int_L \left( \sum_{\theta} |f_{\theta}|^2 \right)^3 \lesssim \left( \sum_{\theta} \| f_{\theta} \|_{L^6(Q_R)}^2 \right)^2 \sum_{\theta} \| f_{\theta} \|_{L^2(Q_R)}^2 \]
which implies the conclusion of Theorem \ref{thm:main}

More quantitatively, for \( x \in L \), we have the bound

\[
|f(x)|^2 \lesssim g_1 \leq \left( \frac{A}{A-1} C_{\text{low}} \right)^N \sum_{\theta} |f_{\theta}(x)|^2.
\]

This ultimately gives us a bound for \( \text{Dec}(R) \) of the form \( \left( \frac{A}{A-1} C_{\text{low}} \right)^N \). Recall that \( N \sim \log R/\log \log R \). We can choose \( A \) to be large to control the contribution of \( \frac{A}{A-1} \). But if \( C_{\text{low}} \) is a constant bigger than 1, then \( C_{\text{low}}^N \) will be much larger than \( (\log R)^C \) (although still smaller than \( C_{\text{low}}^N \)). We will have to work more carefully in the low lemma to make \( C_{\text{low}} \) very close to 1. We will return to discuss this more below.

But first we discuss the case where one of the \( \Omega_k \) dominates. In this case we begin by applying a broad/narrow analysis. The narrow case is handled by an induction on scales argument. For the broad case, we consider the set

\[
U := \{ x \in Q_R : |f(x)| \sim \alpha, \quad |f(x)| \leq \max_{\tau_1, \tau_1^* \text{ non-adjacent}} |f_{\tau_1} f_{\tau_1^*}|^{1/2}(x) \}
\]

where \( \tau_1 \) and \( \tau_1^* \) are \( R_1^{-1} \times R_1^{-1/2} \) rectangles.

The \( L^\infty \) norm of \( g_N \) plays an important role in our argument, so we give it a name:

\[
r = \|g_N\|_{L^\infty(Q_R)}.
\]

Since \( r = \| \sum_{\theta} |f_{\theta}|^2 \|_{L^\infty(Q_R)} \leq \sum_{\theta} \| f_{\theta} \|^2_{L^\infty(Q_R)} \), the main estimate (2.1) follows from the bound

\[
(2.9) \quad \alpha^6 |U \cap \Omega_k| \lesssim r^2 \sum_{\theta} \| f_{\theta} \|^2_{L^2(Q_R)}.
\]

Focusing on the broad case allows us to use bilinear restriction, which leads to the following bound:

\[
\alpha^4 |U \cap \Omega_k| \lesssim \int_{\Omega_k} g_k^2.
\]

From here, use that \( g_k \lesssim |g_{k,h}| \) on \( \Omega_k \) and proceed as in the special case above, using the “high lemma”, Lemma \ref{lem:high}, to obtain

\[
(2.10) \quad \alpha^4 |U \cap \Omega_k| \lesssim \max_{\tau_h} \| f_{\tau_h} \|^2_{L^\infty(Q_R)} \sum_{\theta} \| f_{\theta} \|^2_{L^2(Q_R)}.
\]

Recall that we must modify the function \( f \) to get a good bound for the \( \| f_{\tau_h} \|_{L^\infty(Q_R)} \). Here is the idea of modifying \( f \). If \( x \in U \), then we know that \( |f(x)| \sim \alpha \). By the definition of \( \Omega_k \), we know that for \( x \in \Omega_k \),

\[
\sum_{\tau_h} |f_{\tau_h}(x)|^2 \leq K r,
\]

where \( K = (\log R)^c \left( \frac{A}{A-1} C_{\text{low}} \right)^{N-k-1} \). Even though \( K \) might be as large as \( C_{\text{low}}^N \), this is accounted for in the proof with a more careful definition of \( \Omega_k \). Inequality (2.10) implies that for \( x \in U \cap \Omega_k \), the \( f_{\tau_h} \) with \( |f_{\tau_h}(x)| > 100 K r/\alpha \) make a small contribution to \( f(x) \).

More precisely, if \( x \in U \cap \Omega_k \), then
We define
\begin{equation}
\lambda = 100 Kr/a.
\end{equation}
Roughly speaking, the parts of $f_\tau$ with norm bigger than $\lambda$ don’t make a significant contribution to $f$ on the set $U \cap \Omega_k$. To take advantage of this observation, we divide each $f_\tau$ into wave packets, and then prune the wave packets with amplitude bigger than $\lambda$.

The pruning process goes roughly as follows (but this account is a little oversimplified). First we expand $f_\tau$ into wave packets,
\begin{equation}
f_\tau = \sum_T \psi_T f_\tau.
\end{equation}
Here $T$ denotes a translate of the dual convex $\tau_k^*$ (see Definition 3.11), and the sum is over a collection of translates that tile the plane. The function $\psi_T$ is a smooth approximation of the characteristic function of $T$, and the $\psi_T$ form a partition of unity. Each $\psi_T f_\tau$ is called a wave packet, and it has Fourier support essentially contained in $\tau_k$. We define $\tilde{f}_\tau$ be the result of pruning the high amplitude wave packets from $f_\tau$:
\begin{equation}
\tilde{f}_\tau = \sum_{T: ||\psi_T f_\tau||_\infty \leq 1} \psi_T f_\tau.
\end{equation}
The Fourier support of $\tilde{f}_\tau$ is still essentially contained in $\tau_k$. Suppose for a moment that $\psi_T$ was just $\chi_T$, the characteristic function of $T$. Then because of our pruning, $||\tilde{f}_\tau||_\infty \leq \lambda$. Next we define $f_k = \sum_{\tau_k} \tilde{f}_\tau$.

To analyze $|U \cap \Omega_k|$, we use the argument above with $f_k$ in place of $f$ and $\tilde{f}_\tau$ in place of $f_\tau$. There are the key features of $f_k$ that makes this possible:

- The function $f_k$ is close to $f$ on $U \cap \Omega_k$. If $\psi_T$ was just $\chi_T$, then the analysis in (2.11) would show that for $x \in U \cap \Omega_k$, $|f(x) - f_k(x)| \leq \frac{1}{100}$. We will ultimately define $f_k$ in a slightly more complicated way, and we will prove this bound for $|f(x) - f_k(x)|$.
- We now have the bound $||\tilde{f}_\tau||_\infty \leq \lambda \approx r/a$.
- The function $f_k$ has Fourier support properties similar to those of $f$ so that we can run the argument above. For instance, the Fourier support of $\tilde{f}_\tau$ is essentially contained in $\tau_k$.

When we run the argument above with $f_k$ in place of $f$, and then plug in the bound $||\tilde{f}_\tau||_\infty \lesssim r/a$, we get the estimate
\begin{equation}
\alpha^4 |U \cap \Omega_k| \lesssim (r/a)^2 \int_{Q_R} \sum_\theta |f_\theta|^2.
\end{equation}
This is our desired estimate (2.9).

2.1. Technical issues. There are two main sources of technical difficulties that come up in implementing the sketch above. One has to do with proving the low lemma with very sharp control. The other has to do with pruning wave packets, which we have to do at many different scales. To make the argument work rigorously, $g_k$ and $f_k$ both have to be defined in a more complex way than they were above.
In order to obtain the bound $\text{Dec}(R) \leq (\log R)^C$, we also have to make many small strategic decisions in our argument. The guiding principle behind all these decisions is the following: to obtain a $(\log R)^C$ bound with an argument involving $N \sim \log R/\log \log R$ steps, we must carefully control any step which involves $N$ iterations. On the other hand, for steps that are only iterated $O(1)$ times, we can prove lossier bounds.

A good example is the low lemma. If we are not very careful with how we formulate the “low lemma”, we will get a bound for $\text{Dec}(R)$ which is much larger than $(\log R)^C$. As we discussed above, if we prove the low lemma in the form $|g_{k,\ell}(x)| \leq C_{\text{low}} g_{k+1}(x)$, then we will get a bound for $\text{Dec}(R)$ which is at least as big as $C_{\text{low}}^N$. To get our desired bound for $\text{Dec}(R)$, we need $C_{\text{low}}$ to be almost 1.

Above we gave a non-rigorous sketch of the low lemma. To get some perspective, let us now rigorously prove a version of the low lemma to get a perspective on $C_{\text{low}}$.

**Lemma 2.5** (Baby low lemma). Let $\eta_k(\xi)$ be a bump function defined as in Definition 2.1. Then

$$
|g_{k,\ell}| = |\sum_{\tau_k} |f_{\tau_k}|^2 * \tilde{\eta}_k| \leq 2 \sum_{\tau_{k+1}} |f_{\tau_{k+1}}|^2 * |\tilde{\eta}_k|.
$$

*Proof.* We write $\sum_{\tau_k} |f_{\tau_k}|^2 * \tilde{\eta}_k(x)$ using Fourier inversion.

$$
\sum_{\tau_k} |f_{\tau_k}|^2 * \tilde{\eta}_k(x) = \sum_{\tau_k} \int \hat{f}_{\tau_k} * \hat{f}_{\tau_k} e^{2\pi i \xi \cdot x} \eta_k(\xi)d\xi.
$$

Now $f_{\tau_k} = \sum_{\tau_{k+1} \subset \tau_k} f_{\tau_{k+1}}$, so we can expand out the last expression:

$$
= \sum_{\tau_k} \sum_{\tau_{k+1} \subset \tau_k} \int \hat{f}_{\tau_{k+1}} * \hat{f}_{\tau_{k+1}'} e^{2\pi i \xi \cdot x} \eta_k(\xi)d\xi.
$$

Now the point is that most of the integrals in the sum above vanish. The convolution $\hat{f}_{\tau_{k+1}} * \hat{f}_{\tau_{k+1}'}$ is supported in $\tau_{k+1} - \tau_{k+1}'$, and $\eta_k$ is supported in a ball of radius $2\rho_k \leq R_{k+1}^{-1/2}$. Now each rectangle $\tau_{k+1}$ has dimensions $R_{k+1}^{-1} \times R_{k+1}^{1/2}$. So $\tau_{k+1} - \tau_{k+1}'$ intersects the support of $\eta_k$ only if $\tau_{k+1}'$ is equal to or adjacent to $\tau_{k+1}$. We keep only these terms in the sum to get

$$
\sum_{\tau_k} |f_{\tau_k}|^2 * \tilde{\eta}_k = \sum_{\tau_{k+1}, \tau_{k+1}' \text{equal or adjacent}} (f_{\tau_{k+1}} \overline{f_{\tau_{k+1}'}}) * \tilde{\eta}_k.
$$

For the cross terms, we note that

$$
|(f_{\tau_{k+1}} \overline{f_{\tau_{k+1}'}}) * \tilde{\eta}_k| \leq (|f_{\tau_{k+1}}| |f_{\tau_{k+1}'}|) * |\tilde{\eta}_k| \leq \left(\frac{1}{2} |f_{\tau_{k+1}}|^2 + \frac{1}{2} |f_{\tau_{k+1}'}|^2\right) * |\tilde{\eta}_k|.
$$

Finally, grouping all the terms gives the desired bound:

$$
|\sum_{\tau_k} |f_{\tau_k}|^2 * \tilde{\eta}_k| \leq 2 \sum_{\tau_{k+1}} |f_{\tau_{k+1}}|^2 * |\tilde{\eta}_k|.
$$

□

There are a couple of issues with this bound. One issue is that we have an unwanted factor of 2 on the right-hand side. A second issue is that we have a convolution on the right-hand side. If we take $g_k = \sum_{\tau_k} |f_{\tau_k}|^2$, then we have $|g_{k,\ell}| \leq 2g_{k+1} * |\tilde{\eta}_k|$. 

To deal with the factor of 2, we consider a special case when the Fourier support of \( f \) has a helpful spacing condition. Let \( \Theta = \{ \theta \} \) be a collection of \( \sim R^{-1/2} \times R^{-1} \) rectangles contained in an \( R^{-1} \)-neighborhood of \( \mathbb{P}^1 \). The collection \( \Theta \) has the spacing property at scale \( R_k \) if there exists a collection of \( \sim R_k^{-1/2} \times R_k^{-1} \)-rectangles \( \tau_k \) which cover \( \cup_{\theta \in \Theta} \theta \) and such that
\[
\text{dist}(\tau_k, \tau'_k) \geq (\log R)^{-1} R_k^{-1/2}
\]
whenever \( \tau_k \) and \( \tau'_k \) are distinct. If \( \Theta \) has the spacing property at scales \( R_1, \ldots, R_N-1 \), then say \( \Theta \) is well-spaced. A well-spaced collection of rectangles \( \theta \) can include most of the rectangles needed to cover the parabola, and we will be able to reduce our theorem for a general \( f \) to the case that the Fourier support of \( f \) is well-spaced. The spacing condition helps us because whenever \( \tau_{k+1}, \tau'_{k+1} \) are distinct, \( \tau_{k+1} - \tau'_{k+1} \) is supported outside the ball of radius \( (\log R)^{-1} R_k^{-1/2} \). Now we choose \( \rho_k \leq (\log R)^{-1} R_k^{-1/2} \), and we see that all the cross terms in Lemma 2.5 vanish. This gets rid of the factor of 2. Assuming that \( f \) obeys the spacing condition, we conclude that
\[
\sum_{\tau_k} |f_{\tau_k}|^2 * \tilde{\eta}_k | \leq \sum_{\tau_{k+1}} |f_{\tau_{k+1}}|^2 * |\tilde{\eta}_k|.
\]

Next we discuss the \( *|\tilde{\eta}_k| \) on the right-hand side. In order to deal with this factor, we define \( g_k \) in a more complicated way. Our definition of \( g_k \) has the following form:

\[
g_k := \sum_{\tau_k} |f_{k+1, \tau_k}|^2 * \varphi_{\tilde{T}_{\tau_k}}.
\]

Here \( f_{k+1, \tau_k} \) is given by pruning high amplitude wave packets from \( f_{\tau_k} \), and we will discuss it more below. The function \( \varphi_{\tilde{T}_{\tau_k}} \) is roughly \( \frac{1}{|\tau_k|} \chi_{\tau_k^*} \). It’s a bit bigger than this, so a more accurate model is
\[
\varphi_{\tilde{T}_{\tau_k}} \sim \frac{(\log R)^c}{|\tau_k^*|} \chi(\log R)^c \tau_k^*.
\]

Let’s see why this extra convolution helps us. In the well-spaced case, the argument above shows that
\[
|g_k * \tilde{\eta}_k| \leq \sum_{\tau_{k+1}} |f_{k+1, \tau_{k+1}}|^2 * \varphi_{\tilde{T}_{\tau_k}} * |\tilde{\eta}_k|.
\]

On the right-hand side, \( \tau_k \) denotes the parent of \( \tau_{k+1} \). We choose the functions \( \varphi_{\tilde{T}_{\tau_k}} \) so that for any \( \tau_{k+1} \subset \tau_k \),

\[
\varphi_{\tilde{T}_{\tau_k}} * |\tilde{\eta}_k| \leq \varphi_{\tilde{T}_{\tau_{k+1}}}.
\]

With this choice, the right-hand side of (2.16) is bounded by \( \sum_{\tau_{k+1}} |f_{k+1, \tau_{k+1}}|^2 * \varphi_{\tilde{T}_{\tau_{k+1}}} \leq g_{k+1} \). So with this definition, we get \( |g_k, \ell| \leq g_{k+1} \). (We will prove this in Lemma 3.24)

Redefining \( g_k \) in this way makes the statement of the low lemma very clean. It does have a cost though. We have to make sure that the contribution of \( \varphi_{\tilde{T}_{\tau_k}} \) is not too big. To control their size, we have to choose \( \eta_k \) carefully, and the key bound is the \( \|\tilde{\eta}_k\|_{L^1} \leq 1 + C/\log R \), which is proved in Lemma 3.3.
Finally, let us briefly discuss pruning wave packets. Our argument involves many different scales and we have to prune wave packets at all of the scales. We can define \( f_N \) to be our initial function \( f \). We decompose \( f_N \) into wave packets by combining (2.6) and (2.13)

\[
f_N = \sum_\theta \sum_T \psi_T f_{N,\theta}.
\]

Then we remove the wave packets with amplitude bigger than \( \lambda = 100Kr/a \). The resulting function is called \( f_{N-1} \):

\[
f_{N-1} = \sum_{\theta, T : \|\psi_T f_{N,\theta}\|_\infty \leq \lambda} \psi_T f_{N,\theta}.
\]

Next, we decompose \( f_{N-1} \) into wave packets at the next scale:

\[
f_{N-1} = \sum_{\tau_{N-1}} \sum_{T_{N-1}} \psi_{T_{N-1}} f_{N-1,\tau_{N-1}}.
\]

Here \( \tau_{N-1} \) is a rectangle of dimensions \( R_{N-1}^{1/2} \times R_{N-1}^{-1} \), and \( T_{N-1} \) is roughly a translate of \( \tau_{N-1} \). We remove the wave packets with amplitude bigger than \( \lambda \) and call the resulting function \( f_{N-2} \). This iterative pruning is necessary to make our argument work, but it also makes it fairly complex. In particular, since the pruning has \( N \) steps, we have to be very careful with all the estimates related to the pruning process. For example we have to define the smooth cutoff functions \( \psi_T \) carefully.

3. Proof of Theorem 1.2 – the broad, well-spaced case

The argument outlined in the above intuition section leads to the \((\log R)^c\) upper bound in Theorem 1.2 for functions which satisfy two extra properties. The function \( f \) being broad allows us to bound an \( L^4 \) norm of \( f \) by an \( L^2 \) norm of a square function \( g_k \). The property that \( f \) is well-spaced allows us to replace Lemma 2.5 with

\[
|g_{k,\ell}| \leq \sum_{\tau_{k+1}} |f_{\tau_{k+1}}|^2 \ast |\tilde{g}_k|
\]

(so we have no accumulated constant after iterating the inequality \( \lesssim \log R \) times). Theorem 1.2 in the special case of broad, well-spaced functions \( f \) is called Proposition 3.4 which we prove in this section. In §4 we remove the assumptions on \( f \).

3.1. Statement of Proposition 3.4. Let \( f \in S \) have Fourier support in \( N_{R^{-1}}(\mathbb{P}^1) \). \( \theta \) is always an approximate \( R^{-1/2} \times R^{-1} \) rectangle in a neighborhood of \( \mathbb{P}^1 \). The property that \( f \) is well-spaced allows us to replace Lemma 2.5 with

\[
|g_{k,\ell}| \leq \sum_{\tau_{k+1}} |f_{\tau_{k+1}}|^2 \ast |\tilde{g}_k|
\]

(\( \sim \) here, we mean within a factor of 2).

Our argument involves a sequence of scales \( R_k \) defined as follows:
Definition 3.2. For $k \in \mathbb{N}$, let $R_k = (\log R)^{12k}$. We analyze scale $R_1, \ldots, R_N$ where $R_N = R$. This means that $N = \frac{\log R}{12 \log \log R}$.

Definition 3.3 (Spacing property). Let $\Theta = \{\theta\}$ be a collection of $\sim R^{-1/2} \times R^{-1}$ rectangles contained in an $R^{-1}$-neighborhood of $\mathbb{P}^1$. The collection $\Theta$ has the spacing property at scale $R_k$ if there exists a collection of $\sim R_k^{-1/2} \times R_k^{-1}$-rectangles $\tau_k$ which cover $\bigcup_{\theta \in \Theta} \theta$ and such that

$$\text{dist}(\tau_k, \tau_k') \geq \frac{1}{2} R_k^{1/2}$$

whenever $\tau_k$ and $\tau_k'$ are distinct. If $\Theta$ has the spacing property at scales $R_1, \ldots, R_N-1$, then say $\Theta$ is well-spaced. A function $f \in \mathcal{S}$ is well-spaced if $f$ is supported in $\bigcup_{\theta \in \Theta} \theta$ for some well-spaced $\Theta$.

Proposition 3.4. There exist $c \in (0, \infty)$ such that for all well-spaced $f \in \mathcal{S}$,

$$\alpha^6 |U_\alpha \cap Q_R| \leq (\log R)^c \left( \sum_{\theta} ||f_\theta||_{L_\infty(\mathbb{R}^2)}^2 \right)^2 \sum_{\theta} ||f_\theta||_{L_2(\mathbb{R}^2)}^2.$$

Lemma 3.5. For any $p \geq 1$, $||f_\theta||_{L_\infty(\mathbb{R}^2)} \lesssim ||f_\theta||_{L_p(\mathbb{R}^2)}$.

Proof. Since the Fourier transform of $f_\theta$ is supported on $2\theta$, we can choose a smooth cutoff function $\phi_\theta$ such that $\phi_\theta = 1$ on $2\theta$ and $\phi_\theta = 0$ outside of $3\theta$. Now define $\psi_\theta(x) = \max_{y \in x + B(0,2)} |\hat{\phi}_\theta|$, then for any unit ball $B$

$$||f_\theta||_{L_\infty(B)} \leq \min_{x \in B} |f_\theta| * \psi_\theta(x) \leq ||f_\theta| * \psi_\theta||_{L_p(B)} \lesssim ||f_\theta||_{L_p(\mathbb{R}^2)}.$$

Note 3.6. For the remainder of $\S3$, assume that we have replaced $f$ by a constant multiple of $\mathcal{S}$ so that $\max_{\theta} ||f_\theta||_{L_\infty(\mathbb{R}^2)} = 1$. Note that this means that $\alpha$ is replaced by $c\alpha$ and $r$ is replaced by $c^2 r$. The purpose of this assumption is to simplify the error terms which are often written as negative powers of $R$. Note for example that by Lemma 3.5

$$R^{-50} \ll \frac{r^2}{\alpha^2} \max_{\theta} ||f_\theta||_{L_\infty(\mathbb{R}^2)} \lesssim \frac{r^2}{\alpha^2} \sum_{\theta} ||f_\theta||_{L_2(\mathbb{R}^2)}^2,$$

because $r \geq 1$ and $\alpha \leq R^{1/2}$.

3.2. Auxiliary functions. There are two places described in $\S2$ that involve auxiliary bump functions, Definition 2.11 and (2.13), which we analyze carefully in this section. To formally carry out the pruning process from $f$ to $f_k$, we define $\varphi_{\tilde{T}_k}$, and to define the high/low decomposition of $g_k$, define $\tilde{\eta}_k$. We control the $L^1$ norms of these functions in Lemma 3.9 and Lemma 3.13. This is important for achieving the $(\log R)^c$ upper bound in Proposition 3.4.

Notation 3.7. Let $\delta = \frac{1}{\log R}$. 

Definition 3.8. Let $\xi \in \mathbb{R}$ and $G_0(\xi) = e^{-(\log R)^{1/2}\xi^2}$. Define the Gaussian-like function

$$G(\xi) = G_0(\xi)\chi_{[-1, 1] \setminus [-\delta, \delta]}(\xi) + G_0(\delta)\chi_{[-\delta, \delta]}(\xi) - G_0(1)\chi_{[-1, 1]}(\xi).$$

Define $\eta : \mathbb{R}^2 \to \mathbb{R}$ by

$$\eta(\xi_1, \xi_2) = G(0)^{-2}G(\xi_1)G(\xi_2).$$

For each $k$, define

$$(3.1) \quad \eta_k(\xi) = \eta(4R_{k+2}^{1/2}\xi).$$

Note that $\eta_k(\xi) = 1$ for all $|\xi| \leq \frac{1}{4\log R} R_{k+2}^{-1/2}$ and $\eta_k$ is supported in $|\xi| \leq \frac{1}{4} R_{k+2}^{-1/2}$.

Lemma 3.9. Let $R > 0$ be larger than a certain absolute constant. Then for each $k$,

$$\|\eta_k\|_1 \leq 1 + \frac{c}{\log R}.$$

Proof. It suffices to show the claim for $\eta$ since $\eta_k$ is equal to $\eta$ composed with an affine transformation. Recall the definition of $\eta$

$$\eta(\xi_1, \xi_2) = G(0)^{-2}G(\xi_1)G(\xi_2).$$

Since $G(0)^{-1} = (e^{3/2} - e^{-3/2})^{-1} \leq (1 + \frac{1}{\log R})$ and Fourier transforms of products factor, it suffices to show that

$$\|\tilde{G}\|_1 \leq 1 + \frac{c}{\log R}.$$

Recall the definition of $G$: $G_0(\xi) = e^{-(\log R)^{1/2}\xi^2}$ and

$$G(\xi) = G_0(\xi)\chi_{[-1, 1] \setminus [-\delta, \delta]}(\xi) + G_0(\delta)\chi_{[-\delta, \delta]}(\xi) - G_0(1)\chi_{[-1, 1]}(\xi).$$

Also define the functions

$$G_1(\xi) = (G_0(\xi) - G_0(\delta))\chi_{[-\delta, \delta]}(\xi), \quad G_2(\xi) = G_0(\xi)\chi_{[-1, 1]}(\xi) + G_0(1)\chi_{[-1, 1]}(\xi)$$

and note that

$$G(\xi) = G_0(\xi) - G_1(\xi) - G_2(\xi).$$

Since $\tilde{G}_0 \geq 0$, $\|\tilde{G}_0\|_1 = \|G_0\|_\infty = 1$, which means it suffices to show that

$$\|\tilde{G}_1\|_1 \lesssim \frac{1}{\log R}$$

and

$$\|\tilde{G}_2\|_1 \lesssim \frac{1}{\log R}.$$

Observe that $G_1$ and $G_2$ are continuous, $L^1$ functions (though not differentiable at a few points). $G_1$ is Riemann-integrable, so for each $x \neq 0$, we can use integration by parts to compute $x\tilde{G}_1(x)$ as the inverse Fourier transform of a function. For $G_2$, the same is true after a limiting argument to approximate $G_2$ by Riemann-integrable functions.
First, we have
\[ \|\tilde{G}_1\|_1 = \left\|(1 + x^2)^{-1/2}(1 + x^2)^{1/2}\tilde{G}_1\right\|_1 \]
\[ \lesssim \left\|(1 + x^2)^{1/2}\tilde{G}_1\right\|_2 \]
\[ \lesssim \|G_1\|_2 + \|x\tilde{G}_1\|_2 \]
\[ \lesssim (\delta |1 - G_0(\delta)|^2)^{1/2} + (\log R)^{1/2} \left(\int_{[-\delta,\delta]} |\xi|e^{-\left((\log R)^{3/2}\right)^2}d\xi\right)^{1/2} \]
\[ \lesssim \delta^{1/2}(\log R)^{1/2}\delta^2 + (\log R)^{1/2}\delta^{3/2} \sim \frac{1}{\log R}. \]
Similarly,
\[ \|\tilde{G}_2\|_1 \lesssim \|G_2\|_2 + \|x\tilde{G}_2\|_2 \]
\[ \lesssim G_0(1) + (\log R)^{1/2} \left(\int_{[-1,1]} |\xi|e^{-\left((\log R)^{1/2}\right)^2}d\xi\right)^{1/2} \]
\[ \lesssim e^{-\left((\log R)^{1/2}\right)} + (\log R)^{1/8} \left(\int_{[-\delta^{-1/4},\delta^{-1/4}]} e^{-\xi^2/2}d\xi\right)^{1/2} \ll \frac{1}{\log R}. \]

\[ \square \]

**Definition 3.10.** Let \( \rho : \mathbb{R}^2 \to \mathbb{R} \) be equal \( \rho(\xi) = \eta(\delta \xi) \) where \( \eta \) is defined in Definition 3.8 and \( \delta = \frac{1}{\log R} \). For each scale \( R_k \) and each \( \sim R_k^{-1/2} \times R_k^{-1} \)-rectangle \( \tau_k \), let
\[ \rho_{\tau_k} = \rho \circ \ell_{\tau_k} \]
where \( \ell_{\tau_k} \) is an affine transformation mapping the smallest ellipse containing \( 2\tau_k \) to \( \mathbb{R} \).

**Definition 3.11.** If \( \tau \) is a symmetric convex set with center \( C(\tau) \), then the dual of \( \tau \) is defined as
\[ \tau^* = \{x : |x \cdot (y - C(\tau))| \leq 1, \forall y \in \tau\}. \]

**Definition 3.12.** For each \( R_1^{-1/2} \times R_1^{-1} \)-rectangle \( \tau_1 \), let
\[ \varphi_{\tilde{T}_{\tau_1}}(x) = \sup_{y \in x + (\log R)^{1/4} \tilde{T}_{\tau_1}} |\tilde{\rho}_{\tau_1}(y)|. \]
For each \( 2 \leq k \leq N \) and each \( R_k^{-1/2} \times R_k^{-1} \)-rectangle \( \tau_k \), define \( \varphi_{\tilde{T}_{\tau_k}} \) inductively by
\[ \varphi_{\tilde{T}_{\tau_k}}(x) = \max\left(\sup_{y \in x + (\log R)^{1/4} \tilde{T}_{\tau_k}} |\tilde{\rho}_{\tau_k}(y)|, \varphi_{\tilde{T}_{\tau_{k-1}}} \star |\tilde{\eta}_{k-1}|(x)\right), \]
where \( \tau_{k-1} \) is the \( R_{k-1}^{-1/2} \times R_{k-1}^{-1} \)-rectangle containing \( \tau_k \).

**Lemma 3.13.** For each \( k \) and \( \tau_k \),
\[ \|\varphi_{\tilde{T}_{\tau_k}}\|_1 \lesssim (\log R)^{\frac{3}{2}}. \]
The implicit constant is uniform in \( k \) and \( \tau_k \).
Proof. If \( k = 1 \), then \( \| \varphi_{\tilde{T}_1} \|_1 \lesssim (\log R)^c \) is clear from the definition. Recall the definition of \( \varphi_{\tilde{T}_k} \) for \( k \geq 2 \) to be
\[
\varphi_{\tilde{T}_k}(x) = \max \left( \sup_{y \in x + (\log R)^{11} \tau_k^*} |\tilde{\rho}_{\tau_k}(y)|, \varphi_{\tilde{T}_{k-1}} * |\tilde{\eta}_{k-1}|(x) \right),
\]
where \( \tau_{k-1} \) is the \( R_{k-1}^{-1/2} R_{k-1}^{-1} \)-rectangle containing \( \tau_k \). Let \( A_k \) be the set on which
\[
\sup_{y \in x + (\log R)^{11} \tau_k^*} |\tilde{\rho}_{\tau_k}(y)| \geq \varphi_{\tilde{T}_{k-1}} * |\tilde{\eta}_{k-1}|(x)
\]
and let \( B_k \) be the complement of \( A_k \). Note that
\[
\| \varphi_{\tilde{T}_k} \|_1 = \sup_{y \in + (\log R)^{11} \tau_k^*} |\tilde{\rho}_{\tau_k}(y)||L^1(A_k) + \| \varphi_{\tilde{T}_{k-1}} * |\tilde{\eta}_{k-1}| ||L^1(B_k)
\]
If
\[
\sup_{y \in + (\log R)^{11} \tau_k^*} |\tilde{\rho}_{\tau_k}(y)||L^1(A_k) \geq \delta \| \varphi_{\tilde{T}_{k-1}} * |\tilde{\eta}_{k-1}| ||L^1(B_k)
\]
then
\[
\| \varphi_{\tilde{T}_k} \|_1 \leq 2^\delta^{-1} \sup_{y \in + (\log R)^{11} \tau_k^*} |\tilde{\rho}_{\tau_k}(y)||L^1(\mathbb{R}^2)
\]
\[
= 2^\delta^{-1} \sup_{y \in + B(0,(\log R)^{11})} |\tilde{\rho}(y)||L^1 \lesssim (\log R)^c
\]
since \( |\tilde{\rho}| \) is bounded by \( (\log R)^c \) and is rapidly decaying outside a ball of radius \( (\log R)^c \).

If not, then
\[
\| \varphi_{\tilde{T}_k} \|_1 \leq (1 + \delta) \| \varphi_{\tilde{T}_{k-1}} * |\tilde{\eta}_{k-1}| ||L^1(B_k).
\]
By Young’s convolution inequality and Lemma 3.9
\[
\| \varphi_{\tilde{T}_k} \|_1 \leq (1 + \delta)^2 \| \varphi_{\tilde{T}_{k-1}} \|_1.
\]
If we iterate this argument for \( \leq N \leq \frac{\log R}{\log \log R} \) times, then we also obtain the desired conclusion.

\[\square\]

Definition 3.14 (Gaussian partition of unity). First define \( \psi_0 \) by
\[
\psi_0(x) = c \int_Q g(x - y)dy
\]
where \( Q \) is the unit cube \( Q = [-1/2,1/2]^2 \), \( g(x) = e^{-|x|^2} \), and \( c = (\int g)^{-1} \). Note that for every \( x \)
\[
\sum_{n \in \mathbb{Z}^2} \psi_0(x - n) = \sum_{n \in \mathbb{Z}^2} \int_{n+Q} g(x - y)dy = 1.
\]
Let \( T \) be any rectangle in \( \mathbb{R}^2 \). Let \( A : \mathbb{R}^2 \to \mathbb{R}^2 \) be an affine transformation mapping \( T \) to \( Q \). Define the Gaussian bump function adapted to \( T \) \( \psi_T \) by
\[
\psi_T(x) = c|T|^{-1} \int_T g(A(x - y))dy.
\]
(There are several different affine transformations \( A \) taking \( T \) to \( Q \), but they all give the same function \( \psi_T \) because the Gaussian \( g \) and the area form are invariant under the affine automorphisms of \( Q \).)
If $T$ is a set of congruent rectangles $T$ tiling the plane, then
\[ \sum_{T \in T} \psi_T(x) = 1, \]
and so the Gaussians $\{\psi_T\}_{T \in T}$ form a partition of unity.

The Fourier transform of $\psi_T$ is
\[ \widehat{\psi}_T(\xi) = c|T|^{-1} \int_T |A|^{-1} g((A^{-1})^t \xi) e^{-2\pi i A^{-1} y} dy \]
and satisfies
\[ |\widehat{\psi}_T(\xi)| \leq |T| e^{-\|(A^{-1})^t \xi\|^2}. \]

If $x \notin 100\sqrt{\log RT}$, then
\[ |\psi_T(x)| \leq c e^{-\inf_{y \in T} |x-y|^2} \leq R^{-100}. \]

If $\xi \notin 100\sqrt{\log RT^*}$, then
\[ |\widehat{\psi}_T(\xi)| \leq c|T|R^{-10^4}. \]

In the rest of the paper, a function being essentially supported in $S$ means that $|\cdot| \leq R^{-1000}$ off of $S$ and the function rapidly decays away from $S$. A weight function $w_S$ is localized to $S$ means that $w_S \sim 1$ on $S$ and $|w_S| \leq R^{-1000}$ off of $\log RS$.

### 3.3. Pruning wave packets

We define pruned versions of the function $f$ and the intermediate square functions $\sum_{\tau_k}|f_{\tau_k}|^2$. The pruning process depends on the parameter $\alpha > 0$ which measures the bilinearized version of $f$, and a new parameter $r > 0$ (defined below) related to the final square function $\sum_{\theta}|f_{\theta}|^2$.

**Definition 3.15.** Define $g_N$ as
\[ g_N = \sum_{\theta} |f_\theta|^2 \star \varphi_{T\theta}. \]
where $\varphi_{T\theta}$ is defined in Definition 3.12 for $\theta = \tau_N$.

**Definition 3.16.** Define $r > 0$ to be
\[ r = \|g_N\|_{L^\infty(\mathbb{R}^2)}. \]
Note that $r$ satisfies
\[ r \lesssim (\log R)^m \sum_{\theta} \|f_\theta\|_{L^\infty(\mathbb{R}^2)} \]
for the constant $\tilde{c}$ in Lemma 3.13.

**Notation 3.17.** The parameter $\lambda$ measures the ratio between $r$ and $\alpha$:
\[ \lambda = (\log R)^m \frac{r}{\alpha}, \]
where the exponent $m$ is sufficiently large as required by the proof of Lemma 3.22 and Lemma 3.27 and Proposition 3.4.
Definition 3.18. For each rectangle $\tau_k$, we write $T_{\tau_k}$ to denote a translate of $(\log R)^9 \tau_k^*$. We let $\mathcal{T}_{\tau_k}$ be a tiling of the plane by rectangles $T_{\tau_k}$. In Definition 3.14, we defined a Gaussian partition of unity associated to such a tiling:

$$\sum_{T_{\tau_k} \in \mathcal{T}_{\tau_k}} \psi_{T_{\tau_k}}(x) = 1.$$ 

Definition 3.19 (Defining $f_k$ with respect to $\lambda$). Define $f_N$ by

$$f_N = \sum_\theta \sum_{T_{\tau_k} \in \mathcal{T}_{\tau_k}} \psi_{T_{\tau_k}} f_{\theta}.$$ 

Let $\mathcal{T}_{\theta,\lambda} = \{T_{\theta} \in \mathcal{T}_{\theta} : \|\psi_{T_{\theta}} f_{\theta}\|_\infty \leq \lambda\}$. Define

$$f_{N,\theta} = \sum_{T_{\theta} \in \mathcal{T}_{\theta,\lambda}} \psi_{T_{\theta}} f_{\theta}$$

and note that $\hat{f}_{N,\theta}$ is essentially supported in $(1 + (\log R)^{-8})\theta$. Also define

$$f_{N,\tau_{N-1}} = \sum_{\theta \subset \tau_{N-1}} f_{N,\theta}.$$ 

Now we define $f_k, f_k, \tau_k$, and $f_k, \tau_{k-1}$, starting with $k = N - 1$ and going down to $k = 1$.

$$f_k = \sum_{\tau_k} \sum_{T_{\tau_k} \in \mathcal{T}_{\tau_k,\lambda}} \psi_{T_{\tau_k}} f_{k+1,\tau_k}$$

where $\mathcal{T}_{\tau_k,\lambda} = \{T_{\tau_k} \in \mathcal{T}_{\tau_k} : \|\psi_{T_{\tau_k}} f_{k+1,\tau_k}\|_\infty \leq \lambda\}$. Similarly we define

(3.5) $$f_{k,\tau_k} = \sum_{T_{\tau_k} \in \mathcal{T}_{\tau_k,\lambda}} \psi_{T_{\tau_k}} f_{k+1,\tau_k}$$

(3.6) $$f_{k,\tau_{k-1}} = \sum_{\tau_k \subset \tau_{k-1}} f_{k,\tau_k}.$$ 

Lemma 3.20 (Properties of $f_k$).

(1) $f_k = \sum_{\tau_k} f_{k,\tau_k}$. 

(2) $|f_{k,\tau_k}(x)| \leq |f_{k+1,\tau_k}(x)|$. 

(3) $\|f_{k,\tau_k}\|_{L^\infty} \leq C(\log R)^2 \lambda + R^{-1000}$. 

(4) ess supp$f_{k,\tau_k} \subset (1 + (\log R)^{-8}) \cdot \tau_k$. 

(5) ess supp$f_{k,\tau_{k-1}} \subset (1 + (\log R)^{-10}) \tau_{k-1}$. 

Proof. The first property follows straight from the definition. The second property follows because $\sum_{T_{\tau_k} \in \mathcal{T}_{\tau_k}} \psi_{T_{\tau_k}}$ is a partition of unity, and

$$f_{k,\tau_k} = \sum_{T_{\tau_k} \in \mathcal{T}_{\tau_k,\lambda} \subset \mathcal{T}_{\tau_k}} \psi_{T_{\tau_k}} f_{k+1,\tau_k}.$$ 

Now consider the $L^\infty$ bound in the third property. We write

$$f_{k,\tau_k}(x) = \sum_{T_{\tau_k} \in \mathcal{T}_{\tau_k,\lambda}, x \in (\log R)T_{\tau_k}} \psi_{T_{\tau_k}} f_{k+1,\tau_k} + \sum_{T_{\tau_k} \in \mathcal{T}_{\tau_k,\lambda}, x \notin (\log R)T_{\tau_k}} \psi_{T_{\tau_k}} f_{k+1,\tau_k}.$$
The first sum has at most $C(\log R)^2$ terms, and each term has norm bounded by $\lambda$ by the definition of $T_{r_k,\lambda}$. By the normalization in Note 3.6 it follows easily that

$$|f_{k+1,r_k}(x)| \leq R^{100}.$$  

But if $x \notin (\log R)T_{r_k}$, then $\psi_{T_{r_k}}(x) \leq e^{-(\log R)^2} \leq R^{-2000}$. Moreover, as $T_{r_k}$ gets further away from $x$, $\psi_{T_{r_k}}(x)$ is rapidly decaying. Therefore, the second sum has norm at most $R^{-1000}$.

The fourth and fifth properties depend on the essential Fourier support of $\psi_{T_{r_k}}$ (and on similar trivial bounds as (3.7)). Recall from Definition 3.18 that $T_{r_k}$ is a translate of $(\log R)^9 r_k$. Because of this factor $(\log R)^9$, the essential Fourier support of $\psi_{T_{r_k}}$ is contained in $100\sqrt{\log R}(\log R)^{-9}r_k$ (see Definition 3.14).

Initiate a 2-step induction with base case $k = N$: $f_{N,\theta}$ has essential Fourier support in $(1 + (\log R)^{-8})\theta$ because of the above definition. Then

$$f_{N,\tau_{N-1}} = \sum_{\theta \subset \tau_{N-1}} f_{N,\theta}$$

has essential Fourier support in $\bigcup_{\theta \subset \tau_{N-1}} (1 + (\log R)^{-8})\theta$, which is contained in $(1 + (\log R)^{-10})\tau_{N-1}$. Since each $\psi_{T_{\tau_{N-1}}}$ has essential Fourier support in $100\sqrt{\log R}(\log R)^{-9}\tau_{N-1}$,

$$f_{N-1,\tau_{N-1}} = \sum_{T_{\tau_{N-1}} \in \tau_{N-1}} \psi_{T_{\tau_{N-1}}} f_{N,\tau_{N-1}}$$

has essential Fourier support in

$$(100\sqrt{\log R}(\log R)^{-9} + 1 + (\log R)^{-10})\tau_{N-1} \subset (1 + (\log R)^{-8})\tau_{N-1}.$$  

Iterating this reasoning until $k = 1$ gives (4) and (5).

**Definition 3.21** (Definition of $g_k$). For $k = 1, \ldots, N - 1$, define

$$g_k := \sum_{\tau_k} |f_{k+1,\tau_k}|^2 \ast \varphi_{T_{\tau_k}}$$

where $\varphi_{T_{\tau_k}}$ is specified in Definition 3.12.

The following lemma shows that the difference between the $k$th and $(k + 1)$st versions of $f_{\tau}$ is controlled by $\lambda^{-1}g_k$. We eventually apply this lemma for $x \in \Omega_k$, defined in Definition 3.26 where we know that $g_k \sim r$. We will see that on this set, the differences between the different versions of $f_{\tau}$ are negligible.

**Lemma 3.22.** Suppose $\tau$ is a $\rho^{-1/2} \times \rho^{-1}$ rectangle in a $\rho^{-1}$-neighborhood of $\mathbb{P}^1$, at any scale $1 \leq \rho \leq R$. For $k = 1, \ldots, N - 1$, if $R_k \geq \rho$,

$$|\sum_{\tau_k \subset \tau} f_{k+1,\tau_k}(x) - \sum_{\tau_k \subset \tau} f_{k,\tau_k}(x)| \leq (\log R)^5 \lambda^{-1}g_k(x) + R^{-1000}.$$  

**Proof.** By 3.25,

$$\sum_{\tau_k \subset \tau} f_{k+1,\tau_k}(x) - \sum_{\tau_k \subset \tau} f_{k,\tau_k}(x) = \sum_{\tau_k \subset \tau} \sum_{T \in T_{\tau_k}} \psi_{T}(x)f_{k+1,\tau_k}(x).$$

$$\|\psi_{T}f_{k+1,\tau_k}\|_{\infty} > \lambda$$
Recall that $\psi_T$ has Gaussian decay off of $T$. It follows from (3.7) that
\[
\left| \sum_{\tau_k \subset \tau} \sum_{T \in T_{\tau_k}} \psi_T(x) f_{k+1, \tau_k}(x) \right| \leq R^{-1000}.
\]
Therefore,
\[
\left| \sum_{\tau_k \subset \tau} \sum_{T \in T_{\tau_k}} \psi_T(x) f_{k+1, \tau_k}(x) \right| \leq \lambda^{-1} \sum_{\tau_k \subset \tau} \sum_{T \in T_{\tau_k}} \|\psi_T f_{k+1, \tau_k}\|_{L^\infty(10\log R T_{\tau_k})}^2 + R^{-1000}.
\]

The number of terms in the inner sum is $\lesssim (\log R)^2$. Let $T_{\tau_k}(x)$ be the unique rectangle in the tiling $T_{\tau_k}$ that includes $x$. If $x \in (\log R) T_{\tau_k}$, then $\psi_T$ is essentially supported in $10(\log R) T_{\tau_k}(x)$, and so we have
\[
\sum_{\tau_k \subset \tau} \sum_{T \in T_{\tau_k}} \|\psi_T f_{k+1, \tau_k}\|_{L^\infty(10\log R T_{\tau_k})}^2 \lesssim (\log R)^2 \lambda^{-1} \sum_{\tau_k \subset \tau} \|f_{k+1, \tau_k}\|_{L^\infty(10\log R) T_{\tau_k}(x)}^2 + R^{-1000}.
\]

Recall by Lemma 3.20 that the Fourier transform of $f_{k+1, \tau_k}$ is essentially supported in $(1 + (\log R)^{-10}) \tau_k$, so for $\rho_{\tau_k}$ from Definition 3.10
\[
|f_{k+1, \tau_k}(y)| \leq |f_{k+1, \tau_k} * \hat{\rho}_{\tau_k}(y)| + R^{-1000}.
\]

Now we defined $\varphi_{\tilde{T}_{\tau_k}}(z)$ to be at least $\sup_{z+(\log R)^{1+1/2}} |\hat{\rho}_{\tau_k}|$, so
\[
(3.8) \quad \|f_{k+1, \tau_k}\|_{L^\infty(10\log R) T_{\tau_k}(x)} \leq |f_{k+1, \tau_k} * \varphi_{\tilde{T}_{\tau_k}}(x)|.
\]

Therefore,
\[
\left| \sum_{\tau_k \subset \tau} \sum_{T \in T_{\tau_k}} \psi_T(x) f_{k+1, \tau_k}(x) \right| \lesssim (\log R)^2 \lambda^{-1} \sum_{\tau_k \subset \tau} \left(|f_{k+1, \tau_k} * \varphi_{\tilde{T}_{\tau_k}}(x)|\right)^2 + R^{-1000}.
\]

Applying Cauchy-Schwarz to the integral in the convolution, we get
\[
\leq (\log R)^2 \lambda^{-1} \sum_{\tau_k \subset \tau} |f_{k+1, \tau_k}|^2 * \varphi_{\tilde{T}_{\tau_k}}(x) \varphi_{\tilde{T}_{\tau_k}} \|_{L^1} + R^{-1000}.
\]

Finally, note that $\|\varphi_{\tilde{T}_{\tau_k}}\|_{L^1} \leq (\log R)^c$ by Lemma 3.13 and that
\[
(3.9) \quad \sum_{\tau_k \subset \tau} |f_{k+1, \tau_k}|^2 * \varphi_{\tilde{T}_{\tau_k}}(x) \leq g_k(x)
\]
by Definition 3.21 \qed
3.4. High/low lemmas for $g_k$.

**Definition 3.23** (Definition of $g_k^\ell$ and $g_k^h$). For $k = 1, \ldots, N - 1$ and $\eta_k$ from (3.1), define

$$g_k^\ell := g_k \ast \tilde{\eta}_k \quad \text{and} \quad g_k^h := g_k - g_k^\ell.$$

**Lemma 3.24** (Low lemma).

$$|g_k^\ell| \leq g_{k+1} + R^{-1000}$$

*Proof.* Write

$$|g_k^\ell(x)| = \left| \sum_{\tau_k} |f_{k+1,\tau_k}|^2 \ast \varphi_{\tilde{T}_{\tau_k}} \ast \tilde{\eta}_k(x) \right|$$

where $\eta_k$ is defined in Definition 3.8. Analyze for each $\tau_k$:

$$|f_{k+1,\tau_k}|^2 \ast \tilde{\eta}_k(x) = \int \widehat{f_{k+1,\tau_k}} \ast \widehat{\tilde{T}_{k+1,\tau_k}}(\xi)e^{2\pi i x \cdot \eta_k(\xi)}d\xi = \sum_{\tau_{k+1} \subset \tau_k} \int \tilde{T}_{k+1,\tau_{k+1}} \ast \tilde{T}_{k+1,\tau_{k+1}}(\xi)e^{2\pi i x \cdot \eta_k(\xi)}d\xi.$$

The function $\tilde{T}_{k+1,\tau_{k+1}} \ast \tilde{T}_{k+1,\tau_{k+1}}$ is essentially supported in $(1 + (\log R)^{-8})(\tau_{k+1} - \tau_{k+1})$. By the well-spaced property, we know that this set does not intersect the ball of radius $\frac{1}{2} R^{-1/2}$ (the support of $\eta_k$) unless $\tau_k = \tau_k'$.

By Lemma 3.20, $|f_{k+1,\tau_{k+1}}| \leq |f_{k+2,\tau_{k+1}}|$, and so

$$||f_{k+1,\tau_k}|^2 \ast \tilde{\eta}_k(x)| \leq \sum_{\tau_{k+1} \subset \tau_k} |f_{k+2,\tau_{k+1}}|^2 \ast |\tilde{\eta}_k|(x).$$

Plug this back in to the definition of $g_k^\ell$ to get

$$|g_k^\ell(x)| \leq \sum_{\tau_k} ||f_{k+1,\tau_k}|^2 \ast \tilde{\eta}_k| \ast \varphi_{\tilde{T}_{\tau_k}} \leq \sum_{\tau_{k+1} \subset \tau_k} |f_{k+2,\tau_{k+1}}|^2 \ast |\tilde{\eta}_k| \ast \varphi_{\tilde{T}_{\tau_{k+1}}}.$$

Now $\varphi_{\tilde{T}_{\tau_{k+1}}}$ was defined in Definition 3.12 so that when $\tau_{k+1} \subset \tau_k$,

$$|\tilde{\eta}_k| \ast \varphi_{\tilde{T}_{\tau_k}} \leq \varphi_{\tilde{T}_{\tau_{k+1}}}.$$

Plugging that in again, we get

$$|g_k^\ell| \leq \sum_{\tau_{k+1}} |f_{k+2,\tau_{k+1}}|^2 \ast \varphi_{\tilde{T}_{\tau_{k+1}}} = g_{k+1}.$$

**Lemma 3.25** (High lemma).

$$\int |g_k^h|^2 \lesssim (\log R)^2 \int \sum_{\tau_k} |f_{k+1,\tau_k}|^4 + R^{-1000}.$$
Proof. By Definition 3.21 and 3.23,
\[
\int |g_k^h|^2 = \sum_{\tau_k} \sum_{\tau_k'} \int (|f_{k+1, \tau_k}|^2 \hat{\varphi}_{\tau_k} (1 - \eta_k))(|f_{k+1, \tau_k'}|^2 \hat{\varphi}_{\tau_k'} (1 - \eta_k)).
\]
The Fourier transform of \(|f_{k+1, \tau_k}|^2\) is essentially supported in \(2(\tau_k - \tau_k)\). Recall also from Definition 3.8 that \((1 - \eta_k)\) is supported where \(|\xi| \geq \frac{1}{4 \log R} R_{k+2}^{-1/2}\). The set \(2(\tau_k - \tau_k) \setminus B(\frac{1}{4 \log R} R_{k+2}^{-1/2})\) overlaps at most \(\log RR_{k+2}^{1/2}/R_{k}^{1/2} = (\log R)^{13}\) many of the sets \(2(\tau_k' - \tau_k') \setminus B(\frac{1}{4 \log R} R_{k+2}^{-1/2})\). Thus applying Cauchy-Schwarz to the integral in the convolution, we get
\[
\int |g_k^h|^2 \lesssim (\log R)^{c_1} \int \sum_{\tau_k} \sum_{\tau_k'} \int (|\varphi_{\tau_k}||f_{k+1, \tau_k}|^2 + R^{-1000})
\lesssim (\log R)^{c_1} \int \sum_{\tau_k} \int |\varphi_{\tau_k}||f_{k+1, \tau_k}|^4 + R^{-1000}
\lesssim (\log R)^{c_2} \int \sum_{\tau_k} \int |f_{k+1, \tau_k}|^4 + R^{-1000}
\]
where we used that \(\|\varphi_{\tau_k}\|_{\infty} \lesssim (\log R)^{c_2}\) (by Lemma 3.13). \(\square\)

3.5. The sets \(\Omega_k\). In this subsection, we will decompose the starting set \(Q_R\) into \((Q_R \cap L) \cup (Q_R \cap \Omega_1) \cup \cdots \cup (Q_R \cap \Omega_{N-1})\). On the set \(\Omega_k\), the bilinearized version of \(f\) is basically the same as for \(f_k\) (see Lemma 3.27) and \(g_k\) is high-dominated (see Lemma 3.28).

Definition 3.26 (Definition of \(\Omega_k\)). Recall the parameter \(r > 0\) defined in Definition (3.2). Let \(\Omega_N = \cup Q_N\) be a collection of pairwise disjoint \(R_{N}^{1/2}\)-cubes which have nonempty intersection with \(Q_R\). Let \(\Omega_{N-1}\) be a collection of disjoint \(R_{N-1}^{1/2}\)-cubes \(Q_{N-1}\) in \(\Omega_N\) satisfying
\[
(1 + \delta)r + R^{-500} < \|g_{N-1}\|_{L^\infty((\log R)^9 Q_{N-1})}.
\]
We define \(\Omega_k\) for \(k = N - 2\), then \(k = N - 3\), down to \(k = 1\). To define \(\Omega_k\), partition \(Q_R \setminus (\Omega_{N-1} \sqcup \cdots \sqcup \Omega_{k+1})\) into \(R_{k}^{1/2}\)-cubes \(Q_k\). Define \(\Omega_k\) to be the collection of \(Q_k\) in the partition which satisfy
\[
(1 + \delta)^{N-k}r + (N - k)R^{-500} < \|g_k\|_{L^\infty((\log R)^9 Q_k)}.
\]
Also define
\[
L := Q_R \setminus (\Omega_1 \sqcup \Omega_2 \sqcup \cdots \sqcup \Omega_{N-1}).
\]
Recall that \(\delta = \frac{1}{\log R}\). The absolute constant \(C\) will be determined by the proof of Lemma 3.28.

Lemma 3.27. Suppose \(\tau\) is a \(\rho^{-1/2} \times \rho^{-1}\) rectangle in a \(\rho^{-1}\)-neighborhood of \(\mathbb{P}^1\), at any scale \(1 \leq \rho \leq R\). For \(k = 1, \ldots, N - 1\), if \(R_k \geq \rho\),
\[
|f_\tau(x) - \sum_{\tau_k \subset \tau} f_{k+1, \tau_k}(x)| \leq (\log R)^{-10\alpha} + R^{-500} \quad \forall x \in U_\alpha \cap Q_R \cap \Omega_k.
\]
**Proof.** First note that for each \( l \in \{1, \ldots, N-1\} \), by (3.6),
\[
\sum_{\tau_l \subset \tau} f_{l, \tau_l} = \sum_{\tau_{l-1} \subset \tau} \sum_{\tau_l \subset \tau_{l-1}} f_{l, \tau_l} = \sum_{\tau_{l-1} \subset \tau} f_{l, \tau_{l-1}}.
\]
Then we may decompose the difference as
\[
|f_{\tau} - \sum_{\tau_k \subset \tau} f_{k+1, \tau_k}| = |f_{\tau} - \sum_{\theta \subset \tau} f_{N, \theta}| + |\sum_{\tau_{N-1} \subset \tau} f_{N, \tau_{N-1}} - \sum_{\tau_{N-1} \subset \tau} f_{N-1, \tau_{N-1}}| + \cdots + |\sum_{\tau_{k+1} \subset \tau} f_{k+2, \tau_{k+1}} - \sum_{\tau_{k+1} \subset \tau} f_{k+1, \tau_{k+1}}|.
\]
By Lemma 3.22, this is bounded by
\[
(\log R)^{\delta} \lambda^{-1} (g_N(x) + g_{N-1}(x) + \cdots + g_{k+1}(x)) + (N-k)R^{-1000}.
\]
Finally, use the definition of \( \Omega_k \) for \( k \leq N-1 \) to say
\[
(3.10) \quad |f_{\tau}(x) - \sum_{\tau_k \subset \tau} f_{k+1, \tau_k}| \lesssim (\log R)^{\delta} (N-k)(1+\delta)^{N-k} \lambda^{-1} r + (N-k)^2 R^{-500}
\]
Recall that \( N \leq \log R, (1+\delta)^N \sim 1 \), and recall that \( \lambda \) was defined in Notation 3.17 by
\[
\lambda = (\log R)^m r \alpha^{-1}.
\]
By choosing \( m \) sufficiently large, we can guarantee that the main term on the right-hand side of inequality (3.10) is bounded by \((\log R)^{-10} \alpha \).

**Lemma 3.28** (\( g_k \) is high-dominated on \( \Omega_k \)). Let \( k = 1, \ldots, N-1 \). For each \( R_k^{1/2} \)-cube \( Q_k \subset \Omega_k \),
\[
\|g_k\|_{L^\infty((\log R)^{9}Q_k)} \leq 2(\log R)\|g_k^h\|_{L^\infty((\log R)^{9}Q_k)}.
\]

**Proof.** First consider the \( k = N-1 \) case. By definition of \( \Omega_{N-1} \), for each \( Q_{N-1} \subset \Omega_{N-1} \),
\[
(3.11) \quad (1+\delta)r + R^{-500} < \|g_{N-1}\|_{L^\infty((\log R)^{9}Q_{N-1})}.
\]
Note that
\[
\|g_{N-1}\|_{L^\infty((\log R)^{9}Q_{N-1})} \leq \|g_{N-1}^f\|_{L^\infty((\log R)^{9}Q_{N-1})} + \|g_{N-1}^h\|_{L^\infty((\log R)^{9}Q_{N-1})}
\]
and suppose that
\[
\|g_{N-1}^h\|_{L^\infty((\log R)^{9}Q_{N-1})} \leq \delta \|g_{N-1}^f\|_{L^\infty((\log R)^{9}Q_{N-1})}.
\]
Then by Lemma 3.24 and (3.2),
\[
\|g_{N-1}\|_{L^\infty((\log R)^{9}Q_{N-1})} \leq (1+\delta)\|g_{N-1}^f\|_{L^\infty((\log R)^{9}Q_{N-1})} + \|g_{N-1}^h\|_{L^\infty((\log R)^{9}Q_{N-1})}
\]
\[
\leq (1+\delta)\|g_{N}^f\|_{L^\infty((\log R)^{10}Q_{N-1})} + (1+\delta)R^{-1000}
\]
\[
\leq (1+\delta)r + R^{-500}.
\]
Note that this contradicts (3.11), so we have the desired conclusion.

Now suppose that \( k \in \{1, \ldots, N-2\} \). By definition of \( \Omega_k \),
\[
(3.12) \quad (1+\delta)^{N-k}r + (N-k)R^{-500} < \|g_k\|_{L^\infty((\log R)^{9}Q_k)}.
\]
Note that
\[
\|g_k\|_{L^\infty((\log R)^{9}Q_k)} \leq \|g_k^f\|_{L^\infty((\log R)^{9}Q_k)} + \|g_k^h\|_{L^\infty((\log R)^{9}Q_k)}
\]
and suppose that
\[
(3.13) \quad \|g_k^h\|_{L^\infty((\log R)^{9}Q_k)} \leq \delta \|g_k^f\|_{L^\infty((\log R)^{9}Q_k)}.
\]
Then by Lemma 3.24
\[(3.14)\]
\[\|g_k\|_{L^\infty((\log R)^p Q_k)} \leq (1 + \delta)\|g_k^\prime\|_{L^\infty((\log R)^p Q_k)} \leq (1 + \delta)\|g_{k+1}\|_{L^\infty((\log R)^p Q_k)} + (1 + \delta)R^{-1000}.\]
By the definition of \(\Omega_k\),
\[\|g_{k+1}\|_{L^\infty((\log R)^p Q_k)} \leq (1 + \delta)^{N - k - 1}\tau + (N - k - 1)R^{-500}.\]
This combined with (3.14) gives
\[\|g_k\|_{L^\infty((\log R)^p Q_k)} \leq (1 + \delta)^{N - k - 1}\tau + (N - k)R^{-500},\]
here we used that \((1 + \delta)(N - k - 1)R^{-500} + (1 + \delta)R^{-1000} \leq (N - k)R^{-500}\) since \(N \lesssim \log R/\log \log R\). This contradicts (3.12) and means that (3.13) must be false, so the conclusion follows.
\[\square\]

3.6. Proof of Proposition 3.4. Recall from Note 3.6 that we have replaced \(f\) by a constant multiple \(c\) so that \(\max_\theta \|f\|_{L^\infty(\mathbb{R}^2)} = 1\).

The first step of the proof of Proposition 3.4 involves an application of a local bilinear restriction theorem. We will use the following version.

Theorem 3.29 (Local bilinear restriction). Let \(\tau_k\) and \(\tau'_k\) be nonadjacent \(\sim R_k^{-1/2} \times R_k^{-1}\)-rectangles in an \(R_k^{-1/4}\)-neighborhood of \(\mathbb{P}^1\). Suppose \(j \geq k\) and \(f \in S\) has Fourier support in \(N_{R_j^{-1}}(\mathbb{P}^1)\). Suppose \(T\) is in the range \(R_j \geq T > 10(\log R) R_j^{1/2}/\text{dist}(\tau_k, \tau'_k)\), and that \(Q_T\) is a cube of sidelength \(T\). Then
\[
\int_{Q_T} |f_{\tau_k} f_{\tau'_k}|^2 \lesssim \frac{(\log R)^4 T^{-2}}{\text{dist}(\tau_k, \tau'_k)} \int \sum_{\tau_j \subset \tau_k} |f_{\tau_j}|^2 \omega_{Q_T} \cdot \int \sum_{\tau_j' \subset \tau'_k} |f_{\tau_j'}|^2 \omega_{Q_T} + \frac{R^{-1000}}{}
\]
for a Gaussian weight function \(\omega_{Q_T}\) localized to \(\log R Q_T\) and with Fourier transform essentially supported in a ball of radius \(2\log RT^{-1}\).

Proof. Let \(\phi_{Q_T}\) be a Gaussian bump function adapted to \(Q_T\) as in Definition 3.14. Then the Fourier transform \(\hat{\phi}_{Q_T}\) is essentially supported in a ball of radius \(2\log RT^{-1}\).
\[
\int_{Q_T} |f_{\tau_k} f_{\tau'_k}|^2 \lesssim \int |\sum_{\tau_j \subset \tau_k} f_{\tau_j} f_{\tau_j'}|^2 \phi_{Q_T} \lesssim \int \sum_{\tau_j \subset \tau_k} |f_{\tau_j} f_{\tau_j'}|^2 \phi_{Q_T} + \frac{R^{-1000}}{}
\]
The reason for the last inequality is that for a fixed pair \((\tau_j, \tau'_j)\), the number of pairs \((\tau''_j, \tau'''_j)\) such that
\[(3.15)\]
\[\tau_j + \tau'_j + B_2 \log RT^{-1} \cap \tau''_j + \tau'''_j + B_2 \log RT^{-1} \neq \emptyset\]
is at most \(O(1)\). Here we used the fact that \(T \geq 10(\log R) R_j^{1/2}/\text{dist}(\tau_k, \tau'_k)\). For more details checking (3.15), see the appendix.
It suffices to show that
\[
\int |f_{\tau_j} f_{\tau'_j}|^2 \phi_{Q_T} \lesssim (\log R)^2 T^{-2} \text{dist}(\tau_k, \tau'_k) \int |f_{\tau_j}|^2 \omega_{Q_T} \int |f_{\tau_j'}|^2 \omega_{Q_T}.
\]
For a translate $T_{r_j}$ of $\tau_j^+$ and a translate $T_{r_j}$ of $\tau_j^-$, let $c_{T_{r_j}} = \max_{x \in T_{r_j}} |f_{r_j}|^2(x)$ and
$$c_{T_{r_j}'} = \max_{x \in T_{r_j}'} |f_{r_j}'|^2(x).$$

Since $T_{r_j}$ and $T_{r_j}'$ have angle $\sim \text{dist}(\tau_k, \tau_k')$, we have
$$|T_{r_j} \cap T_{r_j}' \cap \log R Q_T| \leq R_j / \text{dist}(\tau_k, \tau_k'),$$
while $T_{r_j} \cap \log R Q_T \sim R_j^{1/2} \log R T$.}

$$\int |f_{r_j} f_{r_j}'|^2 \omega_{Q_T} \leq \int_{\log R Q_T} |f_{r_j} f_{r_j}'|^2 + R^{-1000}$$
$$\leq \int_{\log R Q_T} \sum_{T_{r_j}} \sum_{T_{r_j}'} c_{T_{r_j}} \chi_{T_{r_j}} \cdot c_{T_{r_j}'} \chi_{T_{r_j}'} + R^{-1000}$$
$$\lesssim (\log R)^2 T^{-2} \cdot \sum_{T_{r_j}} \sum_{T_{r_j}'} c_{T_{r_j}} \chi_{T_{r_j}} \cdot \int_{\log R Q_T} \sum_{T_{r_j}'} c_{T_{r_j}} \chi_{T_{r_j}'} + R^{-1000}$$

The next step is to show that

$$\int_{\log R Q_T} \sum_{T_{r_j}} c_{T_{r_j}} \chi_{T_{r_j}} \lesssim \int |f_{r_j}|^2 w_{Q_T} + R^{-1000}.$$

The Fourier transform of $|f_{r_j}|^2$ is supported on $\bar{\tau}_{j,0} = \tau_j + (-\tau_j)$, which is approximately an $R_j^{-1/2} \times R_j^{-1}$-rectangle. Let $\sum_{\bar{\tau}_{j,i} \subset \log R \bar{\tau}_{j,0}} \hat{\phi}_{\bar{\tau}_{j,i}}$. Then $|\phi_{r_j} - \chi_{\bar{\tau}_{j,0}}(\xi) \leq R^{-1000}$ for $\xi \in \bar{\tau}_{j,0}$ and $|f_{r_j}|^2 = |f_{r_j}|^2 \ast \hat{\phi}_{\bar{\tau}_{j,i}} + O(R^{-1000})$.

Let $\psi_{r_j}(x) = \max_{y \in x + T_{r_j} - T_{r_j}} |\hat{\phi}_{\bar{\tau}_{j,i}}|$, then
$$\sum_{T_{r_j}} c_{T_{r_j}} \chi_{T_{r_j}} \lesssim |f_{r_j}|^2 \ast \psi_{r_j}.$$

We finish the proof since $\psi_{r_j} \ast \chi_{\log R Q_T} \lesssim (\log R)^2 w_{Q_T} + R^{-1000}$ for a Gaussian bump function $w_{Q_T}$ localized at $\log R Q_T$. 

\textbf{Proof of Proposition 3.4.} It suffices to bound $|U_\alpha \cap \Omega_k|$ for $k = 1, \ldots, N - 1$ and $|U_\alpha \cap \Omega_1|$ since there are $\lesssim \log R$ of these sets and $|U_\alpha \cap \Omega_R| \leq \sum_{k=1}^{N-1} |U_\alpha \cap \Omega_k| + |U_\alpha \cap \Omega_L|.$

For $k = 1, \ldots, N - 1$, by Lemma 3.27, if $x \in U_\alpha \cap \Omega_k$, then
$$\max_{\tau, \tau' \text{ non-adj}} |f_{r}(x) f_{r'}(x)| \leq \max_{\tau, \tau' \text{ non-adj}} |f_{k+1, r}(x) f_{r'}(x)| + (\log R)^{-10} \alpha |f_{r'}(x)| + R^{-500},$$
where $f_{k+1, r} = \sum_{\tau_k \subset \tau} f_{k+1, \tau_k}$. By the definition of $U_\alpha$ (Definition 3.1), $|f_{r'}(x)| \leq (\log R)^9 \alpha$, and so
$$\max_{\tau, \tau' \text{ non-adj}} |f_{r}(x) f_{r'}(x)| \leq \max_{\tau, \tau' \text{ non-adj}} |f_{k+1, r}(x) f_{r'}(x)| + (\log R)^{-1} \alpha^2 + R^{-500}$$
$$\leq \max_{\tau, \tau' \text{ non-adj}} |f_{k+1, r}(x) f_{k+1, \tau'}(x)| + (\log R)^{-1} \alpha^2$$
$$+ (\log R)^{-10} \alpha |f_{k+1, \tau}(x)| + 2R^{-500}. $$
Using the definition of $U_\alpha$ as above as well as Lemma 3.27, it follows that $|f_{k+1,\tau}(x)| \leq 2(\log R)^9\alpha$, and so all together

$$\max_{\tau,\tau' \text{ non-adj}} |f_\tau(x)f_{\tau'}(x)| \leq \max_{\tau,\tau' \text{ non-adj}} |f_{k+1,\tau}(x)f_{k+1,\tau'}(x)| + 3(\log R)^{-1}\alpha^2 + 2R^{-500}.$$ 

The $R^{-500}$ error term is negligible given our normalization, as explained in Note 3.6. Since $x \in U_\alpha$, $\max_{\tau,\tau' \text{ non-adj}} |f_\tau(x)f_{\tau'}(x)| \sim \alpha^2$, and so $\max_{\tau,\tau' \text{ non-adj}} |f_{k+1,\tau}(x)f_{k+1,\tau'}(x)| \sim \alpha^2$ as well. Therefore

$$\alpha^4 |U_\alpha \cap \Omega_k| \lesssim \max_{\tau,\tau' \text{ non-adj}} |f_{k+1,\tau}f_{k+1,\tau'}|^{1/2} 1^{\frac{4}{L^4(\tilde{\Omega}_k)}}$$

where $\tau$ and $\tau'$ are nonadjacent $\sim (\log R)^{-6} \times (\log R)^{-12}$ rectangles. We wish to apply the bilinear restriction theorem above, but the functions $f_{k+1,\tau}$ are only essentially supported in $\sim \tau_k$. This just means that we have an error term of $R^{-1000}$ which is negligible given our normalization, as explained in Note 3.6.

For each $Q_k \subset \Omega_k$,

$$\| \max_{\tau,\tau'} |f_{k+1,\tau}f_{k+1,\tau'}|^{1/2} 1^{\frac{4}{L^4(Q_k)}} \| \lesssim \sum_{\tau,\tau' \text{ non-adj}} \int_{Q_k} |f_{k+1,\tau}f_{k+1,\tau'}|^2 \tag{Theorem 3.29}$$

$$\lesssim |Q_k|^{-1} \left( \int \sum_{\tau_k} |f_{k+1,\tau}|^2 \omega_{Q_k} \right)^2 + R^{-1000} \lesssim |Q_k|^{-1} \left( \int g_k \omega_{Q_k} \right)^2 + R^{-1000} |Q_k| + R^{-1000}$$

where $\omega_{Q_k}$ is a weight function localized to $(\log R)^8 Q_k$ and the final inequality follows from (3.8) and (3.9) in the proof of Lemma 3.22.

By Lemma 3.28 and the decay properties of $\omega_{Q_k}$,

$$\int g_k \omega_{Q_k} \lesssim (\log R)^{16} \| g_k \|_{L^\infty((\log R)^9 Q_k)} |Q_k| + R^{-500} \lesssim (\log R)^{17} \| g_k \|_{L^\infty((\log R)^9 Q_k)} |Q_k| + R^{-500} \lesssim (\log R)^{18} \| g_k \|_{L^1(W_{Q_k})} + R^{-500} \lesssim (\log R)^{19} \| g_k \|_{L^2(W_{Q_k})} |Q_k|^{1/2} + R^{-500}$$

where $W_{Q_k}$ is a Gaussian weight function localized to $\sim (\log R)^9 Q_k$ coming from the locally constant property (see (3.16)). Use this in the previous displayed math and add up the contributions from each $Q_k$ to obtain

$$\| \max_{\tau,\tau'} |f_{k,\tau}f_{k,\tau'}|^{1/2} 1^{\frac{4}{L^4(\Omega_k)}} \| \lesssim (\log R)^{16} \int |g_k|^2 \omega_{\Omega_k}$$

where $\omega_{\Omega_k} = \sum_{Q_k} W_{Q_k}$. Note that $\omega_{\Omega_k} \lesssim 1$ and by the high lemma (Lemma 3.25),

$$\int |g_k|^2 \omega_{\Omega_k} \lesssim (\log R)^{2} \int \sum_{\tau_k} |f_{k+1,\tau_k}|^4 + R^{-1000}.$$ 

Then $f_{k+1,\tau} = \sum_{\tau_k \subset \tau_k} f_{k+1,\tau_k + 1}$, so
\[ \int \sum_{\tau_k} |f_{k+1,\tau_k}|^4 \leq \frac{R^{3/2}_{k+1}}{R^{3/2}_k} \int \sum_{\tau_{k+1}} |f_{k+1,\tau_{k+1}}|^4. \]

By Lemma 3.20, we have \( \|f_{k+1,\tau_{k+1}}\|_{L^\infty} \leq C (\log R)^2 \lambda + R^{-1000} \). The \( R^{-1000} \) error term is negligible as explained in Note 3.6, and so we essentially have

\[ \int \sum_{\tau_k} |f_{k+1,\tau_k}|^4 \leq (\log R)^4 \lambda^2 \int \sum_{\tau_{k+1}} |f_{k+1,\tau_{k+1}}|^2. \]

Finally, we have to carefully unwind the definition of \( f_k \) to relate this last quantity to the original \( f_\theta \):

\[ (3.19) \int \sum_{\tau_{k+1}} |f_{k+1,\tau_{k+1}}|^2 \leq \sum_\theta \int |f_\theta|^2 + R^{-500}. \]

First we recall by Lemma 3.20 that \( |f_{k+1,\tau_{k+1}}(x)| \leq |f_{k+2,\tau_{k+1}}(x)| \), and so

\[ (3.20) \sum_{\tau_{k+1}} \int |f_{k+1,\tau_{k+1}}|^2 \leq \sum_{\tau_{k+1}} \int |f_{k+2,\tau_{k+1}}|^2 \]

Next, by Definition 3.19 \( f_{k+2,\tau_{k+1}} = \sum_{\tau_{k+2} \subset \tau_{k+1}} f_{k+2,\tau_{k+2}} \), and so

\[ (3.21) \sum_{\tau_{k+1}} \int |f_{k+1,\tau_{k+1}}|^2 \leq \sum_{\tau_{k+1}} \int | \sum_{\tau_{k+2} \subset \tau_{k+1}} f_{k+2,\tau_{k+2}}|^2. \]

By Lemma 3.20 the Fourier transform of \( f_{k+2,\tau_{k+2}} \) is essentially supported in \((1 + (\log R)^{-8}) \tau_{k+2}\). Since distinct \( \tau_{k+2} \) and \( \tau'_{k+2} \) are \( \geq \frac{1}{2} R^{-1/2} \)-separated, these sets are disjoint. By orthogonality, we get

\[ (3.22) \sum_{\tau_{k+1}} \int \sum_{\tau_{k+2} \subset \tau_{k+1}} f_{k+2,\tau_{k+2}}^2 \leq \sum_{\tau_{k+2}} \int |f_{k+2,\tau_{k+2}}|^2 + R^{-500}. \]

Now we repeat the reasoning in inequalities (3.20), (3.21), and (3.22) at many scales to conclude

\[ \sum_{\tau_{k+1}} \int \sum_{\tau_{k+2} \subset \tau_{k+1}} f_{k+2,\tau_{k+2}}^2 \leq \sum_{\tau_{k+2}} \int |f_{k+2,\tau_{k+2}}|^2 \]
\[ \leq \sum_{\tau_{k+2}} \int |f_{k+3,\tau_{k+2}}|^2 \]
\[ \leq \sum_{\tau_{k+3}} \int |f_{k+3,\tau_{k+3}}|^2 \]
\[ \ldots \]
\[ \leq \sum_\theta \int |f_\theta|^2 \]
\[ \leq \sum_\theta \int |f_\theta|^2. \]
In the above sequence of inequalities, we neglected to include an $R^{-500}$ added error term in each step due to the difference between “essential support” and “actual support.” These error terms are all negligible according to Note 3.6.

The conclusion of this argument is that for $k = 1, \ldots, N - 1$,

$$\alpha^4 |U_\alpha \cap \Omega_k| \lesssim (\log R)^c \frac{r^2}{\alpha^2} \sum_{\theta} \|f_\theta\|^2_{L^2(\mathbb{R}^2)}.$$

Finally we check that this indeed gives the conclusion of Proposition 3.4. Recall that $g_N(x) = \sum_{\theta} |f_\theta|^2 * \varphi_{T_\theta}$. By Lemma 3.13 $\|\varphi_{T_\theta}\|_{L^1} \leq (\log R)^c$. Thus for each $\theta$ and $x \in (\log R)^2 Q_N$ (where $Q_N \cap \Omega_N \neq \emptyset$),

$$|f_\theta|^2 * \varphi_{T_\theta}(x) \lesssim (\log R)^c \|f_\theta\|^2_{L^\infty(\mathbb{R}^2)} + R^{-1000}.$$

It follows that $r \lesssim (\log R)^c \sum_{\theta} \|f_\theta\|^2_{L^\infty(\mathbb{R}^2)} + R^{-1000}$. Plugging this in gives the conclusion of Proposition 3.4.

Finally, it remains to bound $|U_\alpha \cap L|$. The first step is going from $f$ to $f_1$ using Lemma 3.27 (the argument for $\Omega_1$ in (3.17) holds for $L$ as well):

$$\alpha^6 |U_\alpha \cap L| \lesssim \int_{U_\alpha \cap L} \max_{\tau, \tau'} |f_{1, \tau} f_{1, \tau'}|^3 + R^{-1000}$$

$$\lesssim (\log R)^c \int_{U_\alpha \cap L} \left( \sum_{\tau_1} |f_{1, \tau_1}|^2 \right)^3 + R^{-1000}$$

(Lemma 3.20) $$\lesssim (\log R)^c \int_{U_\alpha \cap L} \left( \sum_{\tau_1} |f_{2, \tau_1}|^2 \right)^3 + R^{-1000}$$

$$\leq (\log R)^c \int_{U_\alpha \cap L} g_1^2 \left( \sum_{\tau_1} |f_{2, \tau_1}|^2 \right) + R^{-1000}$$

where the last inequality is due to (3.8). Then by the definition of $L$,

$$\|g_1\|_{L^\infty(U_\alpha \cap L)} \lesssim Cr + NR^{-500}.$$

Finally, by (3.19),

$$\int \sum_{\tau_1} |f_{2, \tau_1}|^2 \leq \int \sum_{\theta} |f_\theta|^2 + R^{-500}.$$

\[\square\]

4. Proof of Theorem 1.2 – the general case

In the last section, we proved Proposition 3.4, which establishes our main theorem in the broad, well-spaced case. In this section, we prove Theorem 1.2 in full generality. We use Proposition 3.4 as a black box, and then we remove the broad hypothesis by using a broad/narrow analysis, and we remove the well-spaced hypothesis by a random sampling argument.

4.1. Removing the broad hypothesis . The following proposition uses a broad/narrow analysis to prove an upper bound for $\|f\|_{L^6(X)}$ using Proposition 3.4.
Proposition 4.1. There exist \( c, C \in (0, \infty) \) such that for all well-spaced collections \( \Theta \) and \( f \in S \) with Fourier support in \( \bigcup_{\theta \in \Theta} \),

\[
\|f\|^6_{L^6(X)} \leq C(\log R)^c \left( \sum_{\theta} \|f_\theta\|_{L^\infty(\mathbb{R}^2)}^2 \right)^2 \sum_{\theta} \|f_\theta\|_{L^2(\mathbb{R}^2)}^2.
\]

First we prove a few technical lemmas.

Lemma 4.2 (Narrow lemma). Suppose that \( \tilde{\tau}_k \) is an arc of length \( R_k^{-1/2} \leq \ell(\tilde{\tau}_k) \leq 3R_k^{-1/2} \).

Let \( \{\tau_{k+1}\} \) be a partition of \( \tilde{\tau}_k \) into \( R_k^{-1/2} \)-arcs. If \( x \) satisfies

\[
|f_{\tilde{\tau}_k}(x)| > \frac{(\log R)^2 R_{k+1}^{1/2}}{R_k^{1/2}} \max_{\tau_{k+1}, \tau_{k+1}' \text{ non-adj}} |f_{\tau_{k+1}}(x) f_{\tau_{k+1}'}(x)|^{1/2},
\]

then there exists an arc \( \tilde{\tau}_{k+1} \) such that \( \ell(\tilde{\tau}_{k+1}) = 3R_k^{-1/2} \) and

\[
|f_{\tilde{\tau}_k}(x)| \leq (1 + \frac{1}{\log R}) |f_{\tilde{\tau}_{k+1}}(x)|.
\]

Proof. Write \( f_{\tilde{\tau}_k} = \sum_{\tau_{k+1}} f_{\tau_{k+1}} \) and let \( \tau_{k+1}^* \) index a summand satisfying

\[
\max_{\tau_{k+1} \subseteq \tilde{\tau}_k} |f_{\tau_{k+1}}(x)| = |f_{\tau^*_k}(x)|.
\]

For each \( \tau_{k+1} \) that is nonadjacent to \( \tau_k^* \),

\[
|f_{\tau_{k+1}}(x)| \leq |f_{\tau_{k+1}}(x) f_{\tau^*_k}(x)|^{1/2} < \frac{R_k^{1/2}}{(\log R)^2 R_{k+1}^{1/2}} |f_{\tilde{\tau}_k}(x)|
\]

using the hypothesis (4.1) about \( x \) in. Then

\[
|f_{\tilde{\tau}_k}(x) - \sum_{\tau_{k+1} \text{ non-adj} \to \tau_k^*} f_{\tau_{k+1}}(x)| > (1 - \#\tau_{k+1} \max_{\tau_{k+1}, \tau_{k+1}' \text{ non-adj}} \frac{R_k^{1/2}}{(\log R)^2 R_{k+1}^{1/2}}) |f(x)|.
\]

The number of \( \tau_{k+1} \) is bounded by \( 3R_k^{1/2}/R_k^{1/2} \). Define \( \tilde{\tau}_{k+1} \) to be \( (\tau_{k+1}^*)_L \cup (\tau_{k+1}^* + 1)_L \cup (\tau_{k+1}^*_R)_R \) where \( (\tau_{k+1}^*)_L \) is the left neighbor of \( \tau_{k+1}^* \) and \( (\tau_{k+1}^*_R)_R \) is the right neighbor of \( \tau_{k+1}^* \).

\[\square\]

Lemma 4.3 (Case 2 in the proof of Proposition 4.1). Suppose that \( \tau_k^* \) is an arc in a neighborhood of \( \mathbb{P}^1 \). Then

\[
\int_{H_{\tau_k^*}} |f_{\tau_k^*}|^6 \leq (\log R)^c \left( \sum_{\theta \subset \tau_k^*} \|f_\theta\|_{L^\infty(\mathbb{R}^2)}^2 \right)^2 \sum_{\theta \subset \tau_k^*} \|f_\theta\|_{L^2(\mathbb{R}^2)}^2
\]

where

\[
H_{\tau_k^*} = \{x \in Q_R : |f_{\tau_k^*}(x)| \leq (\log R)^8 \max_{\tau_{k+1}, \tau_{k+1}' \text{ non-adj}} |f_{\tau_{k+1}}(x) f_{\tau_{k+1}'}(x)|^{1/2}
\]

and

\[
(\log R)^{-9} \left( \sum_{\tau_{k+1} \subset \tau_k^*} |f_{\tau_{k+1}}(x)|^6 \right)^{1/6} \leq \max_{\tau_{k+1}, \tau_{k+1}' \subset \tau_k^* \text{ non-adj}} |f_{\tau_{k+1}}(x) f_{\tau_{k+1}'}(x)|^{1/2}.
\]

Here \( \tau_{k+1} \) are \( R_{k+1}^{-1/2} \times R_{k+1}^{-1} \) rectangles.
\textbf{Proof.} Let \((c, c^2)\) be the center of \(\tau_k^* \cap \mathbb{P}^1\). Define the affine map
\begin{equation}
\ell(\xi_1, \xi_2) = (R_{k}^{1/2}(\xi_1 - c), R_{k}(\xi_2 - 2\xi_1c + c^2)).
\end{equation}
Then \(\ell(\tau_k^*)\) is contained in an \((R/R_k)^{-1}\)-neighborhood of \(\mathbb{P}^1\). The images \(\{\ell(\theta)\}_{\theta \subset \tau_k^*}\) have the spacing property at scales \(R_{k+1}/R_k, \ldots, R/R_k\). Define the function \(h\) as
\[\widehat{h} = \hat{f} \circ \ell^{-1}\]
and note that for each \(R_l^{-1/2} \times R_l^{-1}\)-rectangle \(\tau_l \subset \tau_k^*\),
\begin{equation}
R_k^{-3/2}h_L(\tau_l)(\gamma(x))e^{2\pi i x \cdot (c,c^2)} = f_{\tau_l}(x)
\end{equation}
where \(\ell(\tau_l)\) is approximately \((R_l/R_k)^{-1/2} \times (R_l/R_k)^{-1}\)-rectangle and
\[\gamma(x) = \left(\frac{x_1 + 2cx_2}{R_l^{1/2}}, \frac{x_2}{R_k}\right)\].

In particular,
\begin{equation}
\int_{H_{\tau_k^*}} |f_{\tau_k^*}|^6 = R_k^{-9+3/2} \int_{\gamma(H_{\tau_k^*})} |h_L(\tau_k^*)(x)|^6.
\end{equation}

By dyadic pigeonholing, there exists \(\alpha_{\tau_k^*} > 0\) such that
\[\|h_L(\tau_k^*)\|_{L^6(\gamma(H_{\tau_k^*}))} \lesssim \alpha_{\tau_k^*} \max_{\tau_{k+1}, \tau_{k+1}' \subset \tau_k^*} \|h_L(\tau_{k+1}')(x)\|_{L^2(\mathbb{R}^2)}^{1/2} \sim \alpha_{\tau_k^*}\].

Repeat the proof of Proposition 3.4 to obtain
\begin{equation}
\|h_L(\tau_k^*)\|_{L^6(\gamma(H_{\tau_k^*}))} \lesssim (\log R)^c \left(\sum_{\theta \subset \tau_k^*} \|h_L(\theta)\|_{L^2(\mathbb{R}^2)}^2 \sum_{\theta \subset \tau_k^*} \|h_L(\theta)\|_{L^2(\mathbb{R}^2)}^2\right).
\end{equation}

First observe that
\[\sum_{\theta \subset \tau_k^*} \|h_L(\theta)\|_{L^2(\mathbb{R}^2)}^2 \lesssim R_k^{-3/2} \sum_{\theta \subset \tau_k^*} \|f_{\theta}\|_{L^2(\mathbb{R}^2)}^2\].

Next, note that for each \(\theta \subset \tau_k\)
\[\|h_L(\theta)\|_{L^2(\mathbb{R}^2)}^2 \leq R_k^3 \|f_{\theta}\|_{L^2(\mathbb{R}^2)}^2\].

These observations combined with (4.3) and (4.4) give the desired conclusion. \qed

\textbf{Proof of Proposition 4.1.} Define an iteration using a broad/narrow argument.

\textbf{Initial step:} Define \(S_1\) to be the set
\begin{equation}
S_1 := \{x \in X : |f(x)| \leq (\log R)^8 \max_{\tau, \tau' \non-adj} |f_{\tau}(x) f_{\tau'}(x)|^{1/2} \text{ and }
\left(\sum_{\tau} |f_{\tau}(x)|^6\right)^{1/6} \leq (\log R)^9 \max_{\tau, \tau' \non-adj} |f_{\tau}(x) f_{\tau'}(x)|^{1/2}\}.
\end{equation}

Define \(B_1 = X \setminus S_1\). Split the integral into
\begin{equation}
\int_X |f|^6 = \int_{S_1} |f|^6 + \int_{B_1} |f|^6.
\end{equation}
By the narrow lemma, if \( x \in B_1 \) satisfies \( |f(x)| > (\log R)^8 \max_{\tau, \tau' \text{ non-adj}} |f_\tau(x) f_{\tau'}(x)|^{1/2} \), then for \( \{\tau^{**}\} \) a collection of pairwise disjoint unions of three consecutive \( \tau \),

\[
|f(x)| \leq (1 + \frac{1}{\log R}) \left( \sum_{\tau^{**}} |f_{\tau^{**}}(x)|^6 \right)^{1/6}.
\]

Alternatively, \( x \in B_1 \) satisfies

\[
|f(x)| \leq (\log R)^8 \max_{\tau, \tau' \text{ non-adj}} |f_\tau(x) f_{\tau'}(x)|^{1/2}
\]

but

\[
(\log R)^{-9} \left( \sum_{\tau} |f_\tau(x)|^6 \right)^{1/6} > \max_{\tau, \tau' \text{ non-adj}} |f_\tau(x) f_{\tau'}(x)|^{1/2}.
\]

Putting this together means that

\[
\int_{B_1} |f|^6 \leq (1 + \frac{1}{\log R})^6 \int_{B_1} \sum_{\tau^{**}} |f_{\tau^{**}}|^6 + \frac{1}{(\log R)^6} \int_{B_1} \sum_{\tau} |f_\tau|^6.
\]

Let \( \{\tau^*\} \) denote the collection \( \{\tau\} \) if

\[
\int_{B_1} \sum_{\tau} |f_\tau|^6 \geq \int_{B_1} \sum_{\tau^{**}} |f_{\tau^{**}}|^6
\]

and equal \( \{\tau^{**}\} \) otherwise. Then

\[
\int_{B_1} |f_\tau|^6 \leq (1 + \frac{2}{\log R})^6 \int_{B_1} \sum_{\tau^*} |f_{\tau^*}|^6
\]

(this just means we only have one finer scale to keep track of rather than two almost equivalent scales). Summarizing all of the inequalities, conclude that

\[
\|f\|_{L^6(X)} \leq \int_{S_1} |f|^6 + (1 + \frac{2}{\log R})^6 \sum_{\tau^*} \int_{B(1)} |f_{\tau^*}|^6
\]

For each \( \tau^* \), further decompose \( B_1 \) into

\[
S_{\tau^*} = \{x \in B_1 : |f_{\tau^*}(x)| \leq (\log R)^8 \max_{\tau_2, \tau_2' \subset \tau^* \text{ non-adj}} |f_{\tau_2}(x) f_{\tau_2'}(x)|^{1/2} \text{ and }
\]

\[
\left( \sum_{\tau_2 \subset \tau^*} |f_{\tau_2}(x)|^6 \right)^{1/6} \leq (\log R)^9 \max_{\tau_2, \tau_2' \subset \tau^* \text{ non-adj}} |f_{\tau_2}(x) f_{\tau_2'}(x)|^{1/2}
\]

where \( \ell(\tau_2) = R_2^{-1/2} \). By analogous reasoning as above, conclude this case with the inequality

\[
\|f\|_{L^6(X)}^6 \leq \int_{S_1} |f|^6 + (1 + \frac{2}{\log R})^6 \sum_{\tau^*} \int_{S_{\tau^*}} |f_{\tau^*}|^6 + (1 + \frac{2}{\log R})^{12} \sum_{\tau^*} \int_{B_{\tau^*}} \sum_{\tau_2 \subset \tau^*} |f_{\tau_2}|^6
\]

where \( B_{\tau^*} = B_1 \setminus S_{\tau^*} \).

**Step k**: \((k \geq 2)\) The conclusion of the previous step is

\[(4.7)\]

\[
\|f\|_{L^6(X)}^6 \leq \int_{S_1} |f|^6 + (1 + \frac{2}{\log R})^6 \sum_{\tau^*} \int_{S_{\tau^*}} |f_{\tau^*}|^6 + \cdots + (1 + \frac{2}{\log R})^{6(k-1)} \sum_{\tau_{k-1}^* \subset \tau_{k-1}} \int_{S_{\tau_{k-1}^*}} |f_{\tau_{k-1}^*}|^6
\]

\[
+ (1 + \frac{2}{\log R})^{6k} \sum_{\tau_{k-1} \subset \tau_{k-1}} \sum_{\tau_{k-1} \subset \tau_{k-1}} |f_{\tau_{k-1}}|^6
\]


where for each $\tau_{k-1}^*$, if $\tau_{k-1}^* \subset \tau_{k-2}^* \subset \cdots \subset \tau_2^* \subset \tau$,

$$B_{\tau_{k-1}^*} = B_1 \setminus (S_1 \cup S_{\tau^*} \cup S_{\tau_2^*} \cup \cdots \cup S_{\tau_{k-1}^*}).$$

For each $\tau_k^* \subset \tau_{k-1}^*$, define $S_{\tau_k^*}$ to be the set

$$\{x \in B_{\tau_{k-1}^*} : |f_{\tau_k^*}(x)| \leq (\log R)^8 \max_{\text{non-adj}} \frac{N}{\tau_{k+1} \cap \tau_{k+1}^*} |f_{\tau_{k+1}}(x)f_{\tau_{k+1}^*}(x)|^{1/2} \text{ and }$$

$$\left(\sum_{\tau_{k+1} \subset \tau_k^*} |f_{\tau_{k+1}}(x)|^6\right)^{1/6} \leq (\log R)^9 \max_{\text{non-adj}} \frac{N}{\tau_{k+1} \cap \tau_{k+1}^*} |f_{\tau_{k+1}}(x)f_{\tau_{k+1}^*}(x)|^{1/2} \right\}$$

where $\ell(\tau_{k+1}) = R_{k+1}^{-1/2}$. Define $B_{\tau_k^*} = B_{\tau_{k-1}^*} \setminus S_{\tau_k^*}$. By analogous arguments as above, conclude that (4.7) holds with $k$ replaced by $k + 1$.

Iterate this procedure until Step $N - 1$ where $N \sim \frac{\log R}{\log \log R}$ to obtain inequality (4.7) for $k = n$. Since there are $\sim (\log R)/(\log \log R)$ terms in the right hand side, it suffices to consider cases where $\|f\|_{L^6(Q_R)}^6$ is bounded by $(\log R)$ times one of the terms on the right hand side. Note that the factors $(1 + \frac{2}{\log R})^{6k}$ are $\lesssim 1$ for all $k \leq N$.

Case 1:

$$\|f\|_{L^6(Q_R)}^6 \lesssim (\log R) \int_{S_1} |f|^6.$$

By the definition of $S_1$ in (4.5),

$$\int_{S_1} |f|^6 \leq (\log R)^{48} \int_{S_1} \max_{\tau, \tau'} |f_{\tau} f_{\tau'}|^3.$$

Let $U_s = \{x \in S_1 : \max_{\tau, \tau'} |f_{\tau}(x)f_{\tau'}(x)|^{1/2} \leq R^{-10} \max_{\theta} \|f_0\|_{L^\infty} \}$. Note that

$$\int_{U_s} \max_{\tau, \tau'} |f_{\tau} f_{\tau'}|^3 \leq R^{-55} \max_{\theta} \|f_0\|_{L^\infty(R^2)} \left(\sum_{\theta} \|f_0\|_{L^\infty(R^2)}^2 \right)^2 \int \sum_{\theta} |f_\theta|^2,$$

which is the right hand side of Proposition 4.1.

Then $S_1 \setminus U_s$ can be partitioned into $\lesssim \log R$ sets $U_{\alpha}$ on which max $|f_{\tau} f_{\tau'}|^{1/2} \sim \alpha$ with $R^{-10} \leq \alpha/\max_{\tau, \tau'} \|f_0\|_{L^\infty(R^2)} \leq R$. By pigeonholing,

$$\int_{S_1 \setminus U_s} \max_{\tau, \tau'} |f_{\tau} f_{\tau'}|^3 \lesssim (\log R) \int_{S_1 \setminus U_{\alpha}} \max_{\tau, \tau'} |f_{\tau} f_{\tau'}|^3 \sim (\log R) \alpha^6 |S_1 \cap U_{\alpha}|.$$

Then Proposition 3.4 applies to bound $\alpha^6 |S_1 \cap U_{\alpha}|$.

Case 2:

$$\|f\|_{L^6(X)}^6 \lesssim (\log R) \sum_{\tau_k^*} \int_{S_{\tau_k^*}} |f_{\tau_k^*}|^6.$$
The sets $S_{\tau_k^*}$ are contained in $H_{\tau_k^*}$ from Lemma 4.3. Using Lemma 4.3,

$$\sum_{\tau_k^*} \int_{S_{\tau_k^*}} |f_{\tau_k^*}|^6 \lesssim \sum_{\tau_k^*} (\log R)^c \left( \sum_{\theta \subseteq \tau_k^*} \|f_\theta\|_{L^\infty}^2 \right)^2 \sum_{\theta \subseteq \tau_k^*} \|f_\theta\|_{L^2}^2,$$

$$\leq (\log R)^c \left( \sum_{\theta} \|f_\theta\|_{L^\infty}^2 \right)^2 \sum_{\theta} \|f_\theta\|_{L^2}^2,$$

so we have the desired conclusion.

**Case 3:**

$$\|f\|_{L^6(QR)}^6 \lesssim (\log R) \sum_{\tau_{N-1}} \int_{B_{\tau_{N-1}}} \sum_{\tau_{N} \subseteq \tau_{N-1}} |f_{\tau_N^*}|^6.$$

The size of the $\tau_N^*$ is approximately the size of the $\theta$ (up to a factor of 3), so we may assume $\tau_N^* = \theta$. Then since the $\ell^6$ norm is bounded by the $\ell^2$ norm, the above inequality implies that

$$\alpha^6 |U_\alpha \cap X| \lesssim (\log R) \int_{U_\alpha \cap X} (\sum_\theta |f_\theta|^2)^3 \lesssim (\log R) \left( \sum_\theta \|f_\theta\|_{L^\infty}^2 \right)^2 \sum_\theta \|f_\theta\|_{L^2}^2.$$

4.2. **Removing the well-spaced hypothesis.** This section contains two well-spaced lemmas which we will use to prove Theorem 1.2 in the following section.

**Lemma 4.4.** For each $f \in S$ with Fourier transform supported in $N_{R^{-1}}(\mathbb{P}^1)$, there exists a collection $\Theta_{N-1}$ that is well-spaced at scale $R_{N-1}$ such that

$$(1 - \frac{1}{\log R})^6 \|f\|_{L^6(QR)}^6 \lesssim \frac{\|\hat{f}\|_{L^6(QR)}^6}{\sum_{\theta \in \Theta_{N-1}} \|f_\theta\|_{L^2}^2 \|\hat{f}_\theta\|_{L^2}^2}$$

where $\hat{f} = \sum_{\theta \in \Theta_{N-1}} f_\theta$.

**Proof.** Write $\|\cdot\|_2$ to denote $\|\cdot\|_{L^2(\mathbb{R}^2)}$. Label the $R^{-1/2}$-arcs from left to right by $\theta_1, \ldots, \theta_n$. For each $i = 0, \ldots, (\log R)^6$, define $H^i$ by

$$\sum_{m \in \mathbb{Z}, 1 \leq i + m (\log R)^6 \leq n} f_{\theta_{i + m (\log R)^6}}.$$

Then $f = \sum_i H^i$ where there are $R^{1/2} / R_{N-1}^{1/2} = (\log R)^6$ terms in the sum. Additionally,

$$\sum_\theta \|f_\theta\|_2^2 = \sum_i \sum_\theta \| (H^i)_\theta \|_2^2.$$

Let $i_0$ denote an index satisfying

$$\sum_\theta \|f_\theta\|_2^2 \geq (\log R)^6 \sum_\theta \| (H^{i_0})_\theta \|_2^2.$$
Define $F$ by
\[ f = F + H_{i_0}. \]

Both $F$ and $H_{i_0}$ have the spacing property at scale $R_{N-1}$.

- If $\|H_{i_0}\|_{L^6(QR)} \leq \frac{1}{\log R} \|f\|_{L^6(QR)}$, then define $\hat{f} = f$. Note that
  \[ (1 - \frac{1}{\log R}) \|f\|_{L^6(QR)} \leq \|\hat{f}\|_{L^6(QR)}. \]

- If $\|H_{i_0}\|_{L^6(QR)} > \frac{1}{\log R} \|f\|_{L^6(QR)}$, then define $\hat{f} = H_{i_0}$. From the way we selected $i_0$, we have
  \[ \|f\|_{L^6(QR)}^6 \leq \frac{(\log R)^6 \|\hat{f}\|_{L^6(QR)}^6}{(\log R)^6 \sum_\theta \|\hat{f}_\theta\|_{L^2(\mathbb{R}^2)}^2}. \]

No matter the case above, since for each $\theta$, $f_\theta = F_\theta$ or $H_{i_0}^\theta$, \[ \sum_\theta \|\hat{f}_\theta\|_{L^2(\mathbb{R}^2)}^2 \leq \sum_\theta \|f_\theta\|_{L^2(\mathbb{R}^2)}^2. \]

Lemma 4.5. For each $f \in S$ with Fourier transform supported in $\mathcal{N}_{R-1}(\mathbb{P}^1)$, there exists a well-spaced collection $\tilde{\Theta}$ such that
\[ \frac{\|f\|_{L^6(QR)}^6}{\sum_\theta \|f_\theta\|_{L^2(\mathbb{R}^2)}^2} \lesssim \frac{\|\hat{f}\|_{L^6(QR)}^6}{\sum_\theta \|\hat{f}_\theta\|_{L^2(\mathbb{R}^2)}^2} \]
where $\hat{f} = \sum_{\theta \in \tilde{\Theta}} f_\theta$.

Proof. Define an iterative procedure. Let $\tilde{F}_N = f$. Apply Proposition 4.4 to $\tilde{F}_N$ to obtain $\tilde{F}_{N-1}$.

Obtaining $\tilde{F}_{N-k-1}$ (where $k \geq 2$) from $\tilde{F}_{N-k}$: We have $\tilde{F}_{N-k}$ from the previous step. $\tilde{F}_{N-k}$ has the spacing property at scales $R_{N-1}, \ldots, R_{N-k}$. Furthermore,
\[ (1 - \frac{1}{\log R}) \sum_{\theta} \|f_\theta\|_{L^2(\mathbb{R}^2)}^2 \lesssim \sum_{\theta} \|\hat{f}_\theta\|_{L^2(\mathbb{R}^2)}^2. \]

The $R_{N-k}^{-1/2}$-arcs made up of unions of the $\theta$ are defined by the previous steps. Label them from left to right by $\tau_1, \ldots, \tau_{n_k}$. For $i = 0, \ldots, (\log R)^6$, define $H_{i_{N-k-1}}^\theta$ by
\[ \sum_{1 \leq i + m(\log R)^6 \leq n_k} (\tilde{F}_{N-k})_{\tau_i + m(\log R)^6}. \]

Note that $\tilde{F}_{N-k} = \sum_{i = 1}^k H_{N-k-1}^i$ and for each $\theta$, $(\tilde{F}_{N-k})_\theta$ equals $(H_{i_{N-k-1}}^i)_\theta$ for exactly one $i$.

Let $i_k$ denote the index satisfying
\[ \sum_{\theta} \|((\tilde{F}_{N-k})_\theta)_{L^6(QR)}^2 \geq (\log R)^6 \sum_{\theta} \|(H_{i_{N-k-1}}^i)_\theta\|_{L^6(QR)}^2. \]
Define $F_{N-k-1}$ by
\[ \tilde{F}_{N-k} = F_{N-k-1} + H_{N-k}^{i_k}. \]
Note that since both $F_{N-k-1}$ and $H_{N-k}^{i_k}$ have smaller Fourier support than $\tilde{F}_{N-k}$, they inherit the spacing property at scales $R_{N-1}, \ldots, R_{N-k}$. By construction, $F_{N-k-1}$ and $H_{N-k}^{i_k}$ also have the spacing property at scale $R_{N-k-1}$.

Define $\tilde{F}_{N-k-1}$ as follows:
- If $\|H_{N-k-1}^{i_k}\|_{L^6(Q_R)} \leq \frac{1}{\log R} \|\tilde{F}_{N-k}\|_{L^6(Q_R)}$, then define $\tilde{F}_{N-k-1} = F_{N-k-1}$ and note that (4.8) holds with $k$ replaced by $k + 1$.
- If $\|H_{N-k-1}^{i_k}\|_{L^6(Q_R)} > \frac{1}{\log R} \|\tilde{F}_{N-k}\|_{L^6(Q_R)}$, then define $\tilde{F}_{N-k-1} = H_{N-k}^{i_k}$ and note that because of the way we defined $H_{N-k}^{i_k}$,
\[ \frac{\|\tilde{F}_{N-k}\|_{L^6(Q_R)}}{\sum_\theta \|\tilde{F}_{N-k}\|_{L^6(Q_R)}} \leq \frac{\|\tilde{F}_{N-k-1}\|_{L^6(Q_R)}}{\sum_\theta \|\tilde{F}_{N-k-1}\|_{L^6(Q_R)}}. \]

Iterate this procedure for $N - 1$-steps, until we obtain $\tilde{F}_1$ which is well-spaced along with inequality (4.8) holds for $k = N - 1$. Since $N \sim \frac{\log R}{\log \log R}$, $(1 - \frac{1}{\log R})^{-N} \leq e^{-\frac{C}{\log R}} \lesssim 1$.

\[\Box\]

4.3. Proof of Theorem 1.2. We prove Theorem 1.2 using Proposition 4.1 and Lemma 4.5.

Proof of Theorem 1.2. By Lemma 4.5
\[ \frac{\|f\|_{L^6(Q_R)}}{\sum \|f_\theta\|_{L^2(\mathbb{R}^2)}^2} \lesssim \frac{\|\tilde{f}\|_{L^6(Q_R)}}{\sum \|\tilde{f}_\theta\|_{L^2(\mathbb{R}^2)}^2} \]
where $\tilde{f} = \sum_{\theta \in \Theta} f_\theta$ for a well-spaced $\Theta$. Then by Proposition 4.1
\[ \frac{\|\tilde{f}\|_{L^6(Q_R)}}{\sum \|\tilde{f}_\theta\|_{L^2(\mathbb{R}^2)}^2} \lesssim (\log R)^c (\sum_{\theta \in \Theta} \|f_\theta\|_{L^\infty(\mathbb{R}^2)}^2)^2. \]

Since $\tilde{\Theta} \subset \Theta$, we are done.

\[\Box\]

5. Showing $\text{Dec}_6(R) \lesssim (\log R)^c$ from Theorem 1.2

5.1. Wave packet decomposition and pigeonholing. We will consider the following form of the decoupling inequality:
\[ (5.1) \quad \|f\|_{L^6(Q_R)} \leq \text{Dec}_6(R) \left( \sum_{\theta} \|f_\theta\|_{L^6(\mathbb{R}^2)}^2 \right)^{1/2}. \]

The constant $\text{Dec}_6(R)$ associated to this inequality is comparable to the constant where the $L^6$-norms are both taken over $\mathbb{R}^2$ and to the constant obtained from the $L^6$ norm in the upper bound being some weight function $\omega_{Q_R}$.

Our goal is to begin with $f \in \mathcal{S}$ with Fourier transform supported in $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ and show it suffices to prove the decoupling inequality for a version of $f$ which has relatively constant
amplitudes and number of wave packets in each direction. In the following definitions, $\sim$ means within a factor of 2.

Write

\[ f = \sum_{\theta} \sum_{T \in \mathbb{T}_\theta} \psi_T f_\theta \]

where for each $\theta$, $\{\psi_T\}_{T \in \mathbb{T}_\theta}$ is a Gaussian partition of unity (meaning adds up to 1) adapted to $(\log R)^9 (R \times R^{1/2})$-tubes $T$. Note that this implies that $|\hat{\psi}_T| \lesssim R^{-1000}$ off of $(\log R)^{-3} (\theta - c_\theta)$, where $c_\theta$ is the center of $\theta$.

**Proposition 5.1** (Wave packet decomposition). There exist subsets $\tilde{\Theta} \subset \Theta$ and $\tilde{\mathbb{T}}_\theta \subset \mathbb{T}_\theta$ as well as a constant $C \in [R^{-10^4}, 1]$ with the following properties:

\[ \| \sum_{\theta} \sum_{T \in \mathbb{T}_\theta} \psi_T f_\theta \|_{L^6(Q_R)} \lesssim (\log R)^2 \| \sum_{\theta \in \tilde{\Theta}} \sum_{T \in \mathbb{T}_\theta} \psi_T f_\theta \|_{L^6(Q_R)} + R^{-9.5} \left( \sum_{\theta} \| f_\theta \|_{L^6(\mathbb{R}^2)}^2 \right)^{1/2}, \]

\[ \# \tilde{\mathbb{T}}_\theta \sim \# \tilde{\mathbb{T}}_{\theta'} \text{ for all } \theta, \theta' \in \tilde{\Theta}, \]

\[ \text{and } \| \psi_T f_\theta \|_{L^\infty(\mathbb{R}^2)} \sim C \max_{\theta' \in \tilde{\Theta}} \max_{T' \in \tilde{\mathbb{T}}_{\theta'}} \| \psi_{T'} f_{\theta'} \|_{L^\infty(\mathbb{R}^2)} \text{ for all } \theta \in \tilde{\Theta} \text{ and } T \in \tilde{\mathbb{T}}_{\theta}. \]

**Proof.** Split the sum (5.2) into

\[ f = \sum_{\theta} \sum_{T \in \mathbb{T}_\theta} \psi_T f_\theta + \sum_{\theta} \sum_{T \in \mathbb{T}_\theta^f} \psi_T f_\theta \]

where the close set is

\[ \mathbb{T}_\theta^c := \{ T \in \mathbb{T}_\theta : T \cap R^{10} Q_R \neq \emptyset \} \]

and the far set is

\[ \mathbb{T}_\theta^f := \{ T \in \mathbb{T}_\theta : T \cap R^{10} Q_R = \emptyset \}. \]

Using Cauchy-Schwarz,

\[ \| \sum_{\theta} \sum_{T \in \mathbb{T}_\theta^f} \psi_T f_\theta \|_{L^6(Q_R)} \leq R^{1/2} \left( \sum_{\theta} \left\| \sum_{T \in \mathbb{T}_\theta^f} \psi_T f_\theta \right\|_{L^6(Q_R)}^2 \right)^{1/2} \]

\[ \leq R^{1/2} \left( \sum_{\theta} \left\| \psi_T f_\theta \right\|_{L^6(Q_R)}^2 \right)^{1/2} \]

\[ \leq R^{1/2} \max_{\theta} \left\| \psi_T \right\|_{L^\infty(Q_R)} \left( \sum_{\theta} \left\| f_\theta \right\|_{L^6(\mathbb{R}^2)}^2 \right)^{1/2} \]

\[ \leq \frac{1}{R^{9.5}} \left( \sum_{\theta} \left\| f_\theta \right\|_{L^6(\mathbb{R}^2)}^2 \right)^{1/2} \]

where we could have bounded the $L^\infty$ norm by $C_N R^{-10^N}$ for any $N \in \mathbb{N}$. This takes care of the far portion of $f$ (i.e. the second term on the right hand side of (5.6)).

The close set has cardinality $|\mathbb{T}_\theta^c| \leq R^{32}$. Let $M = \max_{\theta} \max_{T \in \mathbb{T}_\theta^c} \| \psi_T f_\theta \|_{L^\infty(\mathbb{R}^2)}$. By Lemma 3.5

\[ M \leq \max_{\theta} \| f_\theta \|_{L^\infty} \lesssim \left( \sum_{\theta} \| f_\theta \|_{L^6(\mathbb{R}^2)}^2 \right)^{1/2}. \]
Split the remaining term as
\[
(5.8) \quad \sum_{\theta} \sum_{T \in T^c_{\theta\lambda}} \psi_T f_{\theta} = \sum_{\theta} \sum_{R^{-103} \leq \lambda \leq 1} \sum_{T \in T^c_{\theta\lambda}} \psi_T f_{\theta} + \sum_{\theta} \sum_{T \in T^c_{\theta,s}} \psi_T f_{\theta}
\]
where \( \lambda \) is a dyadic number in the range \([R^{-103}, 1] \),
\[
T^c_{\theta,\lambda} := \{ T \in T^c_{\theta} : \|\psi_T f_{\theta}\|_{L^\infty(\mathbb{R}^2)} \sim \lambda M \},
\]
and
\[
T^c_{\theta,s} := \{ T \in T^c_{\theta} : \|\psi_T f_{\theta}\|_{L^\infty(\mathbb{R}^2)} \leq \frac{1}{2} R^{-103} M \}.
\]
Handle the small term from (5.8) by
\[
\| \sum_{\theta} \sum_{T \in T^c_{\theta,s}} \psi_T f_{\theta} \|_{L^6(Q_R)} \leq R^{1/2} \left( \sum_{\theta} \| \sum_{T \in T^c_{\theta,s}} \psi_T f_{\theta} \|_{L^6(Q_R)}^2 \right)^{1/2}
\]
\[
\leq R^{-10} M \leq R^{-10} \left( \sum_{\theta} \| f_{\theta} \|_{L^6(\mathbb{R}^2)}^2 \right)^{1/2}.
\]
Next decompose the remaining term from (5.8) as
\[
(5.9) \quad \sum_{R^{-103} \leq \lambda \leq 1} \sum_{\theta} \sum_{T \in T^c_{\theta,\lambda}} \psi_T f_{\theta} = \sum_{R^{-103} \leq \lambda \leq 1} \sum_{1 \leq j \leq R^{22}} \sum_{\theta \in \Theta_j(\lambda)} \sum_{T \in T^c_{\theta,\lambda}} \psi_T f_{\theta}
\]
where \( j \) is a dyadic number in the range \([1, R^{22}] \) and \( \Theta_j(\lambda) = \{ \theta : |T^c_{\theta,\lambda}| \sim j \} \).
Because \( j \) and \( \lambda \) are dyadic numbers, there is a choice of \((\lambda, j)\) so that
\[
\| \sum_{R^{-103} \leq \lambda \leq 1} \sum_{1 \leq j \leq R^{22}} \sum_{\theta \in \Theta_j(\lambda)} \sum_{T \in T^c_{\theta,\lambda}} \psi_T f_{\theta} \|_{L^6(Q_R)} \lesssim (\log R)^2 \| \sum_{\theta \in \Theta_j(\lambda)} \sum_{T \in T^c_{\theta,\lambda}} \psi_T f_{\theta} \|_{L^6(Q_R)}.
\]
Take \( \tilde{\Theta} = \Theta_j(\lambda) \) and for each \( \theta \in \tilde{\Theta} \), take \( \tilde{T}_{\theta} = T^c_{\theta,\lambda} \).
\]

5.2. Proof of Theorem 1.1

Proof. By Proposition 5.1, we have
\[
\| \sum_{\theta} f_{\theta} \|_{L^6(Q_R)} \lesssim (\log R)^3 \| \sum_{\theta \in \tilde{\Theta}} \sum_{T \in T^c_{\theta}} \psi_T f_{\theta} \|_{L^6(Q_R)} + R^{-3} \left( \sum_{\theta} \| f_{\theta} \|_{L^6(\mathbb{R}^2)}^2 \right)^{1/2}
\]
where \#\(\tilde{T}_{\theta} \sim \#\tilde{T}_{\theta'}\) for all \( \theta, \theta' \in \tilde{\Theta} \) and
\[
(5.10) \quad \| \psi_T f_{\theta} \|_{L^\infty(\mathbb{R}^2)} \sim A := \max_{\theta' \in \tilde{\Theta}} \max_{T \in \tilde{T}_{\theta'}} \| \psi_T f_{\theta'} \|_{L^\infty(\mathbb{R}^2)}.
\]
Since the Fourier transform of \( \sum_{T \in \tilde{T}_{\theta}} \psi_T f_{\theta} \) is essentially supported in \((1 + (\log R)^{-3}) \theta \),
there exists a function \( f'_{\theta} \) with Fourier transform supported in \((1 + (\log R)^{-3}) \theta \) such that
\[
\sum_{T \in \tilde{T}_{\theta}} \psi_T f_{\theta}(x) = f'_{\theta}(x) + O(R^{-998})A.
\]
Thus
\[
(5.11) \quad \| \sum_{\theta \in \tilde{\Theta}} \sum_{T \in T^c_{\theta}} \psi_T f_{\theta} \|_{L^6(Q_R)} \lesssim \| \sum_{\theta \in \tilde{\Theta}} f'_{\theta} \|_{L^6(Q_R)} + R^{-997} A.
\]
The functions $f'_\theta$ have Fourier support in $(1 + (\log R)^{-3}) \theta$. We may split $\tilde{\Theta}$ into $1$ sets $\tilde{\Theta}_i$ where for distinct $\theta, \theta' \in \tilde{\Theta}_i$, 
\[(2\theta) \cap (2\theta') = \emptyset.\]

Then for some $i$,
\[
\| \sum_{\theta \in \tilde{\Theta}} f'_\theta \|_{L^6(Q_R)}^6 \lesssim \| \sum_{\theta \in \tilde{\Theta}_i} f'_\theta \|_{L^6(Q_R)}^6,
\]
and it follows from Theorem 1.2 that
\[
(5.12) \quad \| \sum_{\theta \in \tilde{\Theta}} f'_\theta \|_{L^6(Q_R)}^6 \lesssim (\log R)^c (\sum_{\theta \in \tilde{\Theta}_i} \| f'_\theta \|_{L^\infty(\mathbb{R}^2)}^2)^2 \sum_{\theta \in \tilde{\Theta}_i} \| f'_\theta \|_{L^2(\mathbb{R}^2)}^2.
\]

Note that
\[
(\sum_{\theta \in \tilde{\Theta}_i} \| f'_\theta \|_{L^\infty(\mathbb{R}^2)}^2)^{1/2} \leq \left( \sum_{\theta \in \tilde{\Theta}} \| \sum_{T \in \tilde{T}_\theta} \psi_T f_\theta \|_{L^\infty(\mathbb{R}^2)}^2 \right)^{1/2} + R^{-500} A
\]
and
\[
(\sum_{\theta \in \tilde{\Theta}_i} \| f'_\theta \|_{L^2(\mathbb{R}^2)}^2)^{1/2} \lesssim \left( \sum_{\theta \in \tilde{\Theta}} \| \sum_{T \in \tilde{T}_\theta} \psi_T f_\theta \|_{L^2(\mathbb{R}^2)}^2 \right)^{1/2} + R^{-500} A.
\]

Combining these observations with (5.12) gives
\[
(5.13) \quad \| \sum_{\theta \in \tilde{\Theta}} f'_\theta \|_{L^6(Q_R)}^6 \lesssim (\log R)^c (\sum_{\theta \in \tilde{\Theta}_i} \| \sum_{T \in \tilde{T}_\theta} \psi_T f_\theta \|_{L^\infty(\mathbb{R}^2)}^2)^2 \sum_{\theta \in \tilde{\Theta}_i} \| \sum_{T \in \tilde{T}_\theta} \psi_T f_\theta \|_{L^2(\mathbb{R}^2)}^2 + R^{-2000} A^6.
\]

The second term is bounded by
\[
(5.14) \quad R^{-2000} A^6 \leq R^{-2000} (\sum_{\theta} \| f_\theta \|_{L^6(\mathbb{R}^2)}^2)^3
\]
using Lemma 33 and the fact that $0 \leq \psi_T \leq 1$. It remains to analyze the first term in the upper bound in (5.13). For each $\theta \in \tilde{\Theta}$, we have
\[
\left| \sum_{T \in \tilde{T}_\theta} \psi_T f_\theta(x) \right| \lesssim \left| \sum_{T \in \tilde{T}_\theta} \psi_T f_\theta(x) \right| + R^{-1000} A \leq (\log R)^2 A + R^{-1000} A.
\]

This leads to the following upper bound for the first term on the right hand side of (5.13):
\[
(5.15) \quad \left( \sum_{\theta \in \tilde{\Theta}} \| \sum_{T \in \tilde{T}_\theta} \psi_T f_\theta \|_{L^\infty(\mathbb{R}^2)}^2 \right)^2 \lesssim (\log R)^8 (\# \tilde{\Theta} A^2)^2 + R^{-1000} A^4.
\]

For each $\theta \in \tilde{\Theta}$, we also have
\[
\int \left| \sum_{T \in \tilde{T}_\theta} \psi_T f_\theta \right|^2 dx \lesssim (\log R)^2 \sum_{T \in \tilde{T}_\theta} \int_{\log RT} \left| \psi_T f_\theta \right|^2 dx + R^{-1998} A^2
\]
\[
\lesssim \log R^4 \# \tilde{\Theta} \# \tilde{T}_\theta A^2 |T| + R^{-1998} A^2.
\]

Combining this with (5.15) leads to the upper bound
\[
(5.16) \quad \| \sum_{\theta \in \tilde{\Theta}} f'_\theta \|_{L^6(Q_R)}^6 \lesssim (\log R)^{c+12} \# \tilde{\Theta}^3 \# \tilde{T}_\theta A^6 |T| + R^{-1000} A^6.
\]
Finally, note that for each $\theta \in \tilde{\Theta}$,
\[
\#\tilde{T}_\theta |T| A^6 \lesssim (\log R)^2 \sum_{T \in \tilde{T}_\theta} \int |\psi_T f_\theta|^6 \omega_T + R^{-2000} A^6
\]
where we used the locally constant property. Finally, since the $\ell^6$ norm is bounded by the $\ell^1$ norm and $0 \leq \psi_T \leq 1$,
\[
\sum_{T \in \tilde{T}_\theta} \int |\psi_T f_\theta|^6 \leq \int \left( \sum_{T \in \tilde{T}_\theta} |\psi_T f_\theta|^6 \right) \leq \|f_\theta\|_{L^6(\mathbb{R}^2)}^6.
\]
In summary, we finish the proof by combining (5.16) with (5.10) and
\[
\#\tilde{\Theta}^3 \#\tilde{T}_\theta |T| A^6 \lesssim (\log R)^2 \left( \sum_{\theta \in \tilde{\Theta}} \|f_\theta\|_{L^6(\mathbb{R}^2)}^6 \right)^{\frac{2}{3}}.
\]
\[
\square
\]

**Appendix**

In the appendix, we show that a slight modification of the proof gives Theorem 1.2 and Theorem 1.1 for a set of curves:
\[
\mathcal{K} := \{ (\xi_1, h(\xi_1)) : |\xi_1| \leq 1 \},
\]
where $h$ are $C^2$ functions satisfying $h(0) = h'(0) = 0$, and $1/2 \leq h''(\xi_1) \leq 2$ for $|\xi| \leq 1$. In particular, one can cut the unit circle into $O(1)$ arcs, such that each arc is part of a curve in $\mathcal{K}$ after translation and rotation. The truncated parabola $\mathbb{P}^1$ is also in $\mathcal{K}$.

To prove Theorem 1.2 and Theorem 1.1 for the curves in $\mathcal{K}$, replace the affine map (4.2) in Lemma 4.3 by
\[
(5.17) \quad \ell(\xi_1, \xi_2) = (\nu(\xi_1 - c), \nu^2(\xi_2 - h(c) - h'(c)(\xi_1 - c)))
\]
with $\nu = R_k^{1/2}$. Since the curve \{ $\nu(\xi_1 - c), \nu^2(\xi_2 - h(c) - h'(c)(\xi_1 - c))$ : $|\xi_1 - c| \leq \nu^{-1}$ \} is also in $\mathcal{K}$, the proof of Lemma 4.3 remains unchanged provided that Proposition 3.4 holds for all the curves in $\mathcal{K}$ for a smaller $R$.

Then it suffices to check (3.15) for the curves in $\mathcal{K}$. Let $(\xi, h(\xi))$ be the center of $\tau_j$. Assume that
\[
\xi, \xi'' \in \tau_k, \quad \xi', \xi''' \in \tau'_k,
\]
and
\[
\xi - \xi'' = \xi''' - \xi' + O(R_j^{-1/2}),
\]
then $h(\xi) - h(\xi'') = h'(\xi_1)(\xi - \xi'')$ for some $\xi_1$ between $\xi$ and $\xi''$, and $h(\xi''') - h(\xi') = h'(\xi_2)(\xi''' - \xi')$ for some $\xi_2$ between $\xi'''$ and $\xi'$. The first derivative $h'(\xi_1) + (\xi_2 - \xi_1)/2 \leq h'(\xi_2)$ since $1/2 \leq h''(\xi) \leq 1$. By $(\xi_2 - \xi_1) \geq \text{dist}(\tau_k, \tau'_k)$, we have
\[
(h(\xi) - h(\xi'') - (h(\xi''') - h(\xi'))) \gtrsim \text{dist}(\tau_k, \tau'_k) R_j^{-1/2}
\]
if $\tau_j, \tau'_j, \tau''_j, \tau'''_j$ have pairwise distance $O(R_j^{-1/2})$. So (3.15) is verified.

Decoupling for the circle with explicit decoupling constant $(\log R)^c$ has an application on a problem about sums of two squares. Such problem arises in the study of Laplace eigenfunctions for the standard two dimensional torus.

Let $\Lambda_m$ be the set of Gaussian integers $\lambda = x + \sqrt{-1}y$, $x, y \in \mathbb{Z}$ with norm $\lambda = m$. 

**Problem.** Give a nontrivial upper bound for the number of solutions
\[ \lambda_1 + \lambda_2 + \lambda_3 = \lambda_4 + \lambda_5 + \lambda_6; \lambda_j \in \Lambda_m. \]

**Corollary 5.2.** For any \( \Lambda \subset \Lambda_m \), if \( N = |\Lambda| > (\log m)^{(c+6)/c} \) for some \( \epsilon > 0 \) and the constant \( c \) as in Theorem 7.2, then
\[
N_m(\Lambda) := \# \{ \lambda_1 + \lambda_2 + \lambda_3 = \lambda_4 + \lambda_5 + \lambda_6, \lambda_j \in \Lambda \} \lesssim N^{3+\epsilon}.
\]

**Proof.** Consider the function
\[ g(z) = \sum_{\lambda \in \Lambda} e^{2\pi i \frac{\lambda}{\sqrt m} z} \]
for \( z \in \mathbb{R}^2 \). Let \( Q_0 = [0, m(\log m)^2] \) and \( \{Q\} \) be a tiling of \( \mathbb{R}^2 \) with translates of \( Q_0 \). Let \( \{\varphi_Q\} \) be the Gaussian partition of unity defined as in Definition 3.14. Define the weight function
\[ \psi = \sum_{Q: \text{dist}(Q, Q_0) \leq m(\log m)^2} \varphi_Q. \]
Then \( |\psi - 1| \leq m^{-1000} \) on \( Q_0 \) and \( |\psi| \leq m^{-1000} \) outside of \( (\log m)^2 Q_0 \) and rapidly decays away from it. Moreover, the Fourier transform \( \hat{\psi} \) is essentially supported on \( B(0, m^{-1}) \).

We apply Theorem 1.1 (for the circle) to the function \( f(z) = g(z)\psi \) with \( R = m \) and \( Q_R = [0, m]^2 \). Then \( f_\theta = e^{2\pi i \lambda z/\sqrt m} \) for the (unique) \( \lambda/\sqrt m \in \theta \). Note that \( g(z) \) is periodic
\[ g(z + \sqrt m v) = g(z) \]
for any \( v \in \mathbb{Z}^2 \). Since \( |f - g| \leq m^{-1000} \) on \( Q_R \), we have
\[
N_m(\Lambda) \lesssim (\log m)^{c+6}|\Lambda|^3
\]
where the \( (\log m)^6 \) comes from \( \psi \) a weight function essentially supported in \([0, (\log m)^3 m]^2\). \( \square \)

This problem was studied by Bombieri and Bourgain in [11] using various methods. In particular, they obtained the bound \( O(|\Lambda_m|^{3+\epsilon}) \) assuming the Riemann hypothesis and the Birch and Swinnerton-Dyer conjecture for the \( L \)-functions of elliptic curves over \( \mathbb{Q} \) and for a random \( m \) with \( |\Lambda_m| \sim 2^{\omega(m)}, \omega(m) \sim \frac{\log m}{A \log \log m} \) for some constant \( A \). Based on the Bourgain-Demeter decoupling, Zane Li obtained in [8] the result of Bombieri and Bourgain unconditionally for all \( m \) with \( |\Lambda_m| > \exp((\log m)^{1-o(1)}) \). Corollary 5.2 proves the result for a larger range of \( |\Lambda|: |\Lambda| > (\log m)^{c/\alpha(1)} \).

In [3], it was conjectured that for any \( \Lambda \subset \Lambda_m \), and any \( \epsilon > 0 \), there exists \( C_\epsilon \) independent of \( m \), such that
\[
\# \{ \lambda_1 + \lambda_2 + \lambda_3 = \lambda_4 + \lambda_5 + \lambda_6, \lambda_j \in \Lambda \} \leq C_\epsilon |\Lambda|^{3+\epsilon}.
\]
Corollary 5.2 confirms for \( |\Lambda| > (\log m)^{(c+6)/c} \).

**References**

[1] E. Bombieri and J. Bourgain, *A problem on sums of two squares*, Int. Math. Res. Not. (11): 3343-3407, (2015).
[2] J. Bourgain, *Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations*, Geom. Funct. Anal., 3 (2):107-156, (1993).
[3] J. Bourgain and C. Demeter, *The proof of the \( l^2 \) decoupling conjecture*, Ann. of math. (2) 182 (1): 351-389, (2015).
[4] C. Demeter, L. Guth, and H. Wang, *Small cap decouplings*, in arXiv: 1908.09166 (2019).
[5] L. Guth, H. Wang, and R. Zhang, *A sharp square function estimate for the cone in \( \mathbb{R}^3 \)*, in arXiv: 1909.10693 (2019).
[6] M. Krisnapur, P. Kurlberg, and I. Wigman, *Nodal length fluctuations for arithmetic random waves*, Annals of Mathematics (2) 177, no.2 (2013): 699-737.

[7] Z. Li, *Decoupling for the parabola and connections to efficient congruencing*, Ph.D. thesis, UCLA, 2019.

[8] Z. Li, *Effective $l^2$ decoupling for the parabola*, Mathematika 66(2020), no.3, 681-712.

[9] Z. Li, *An $l^2$ decoupling interpretation of efficient congruencing: the parabola*, in arXiv:1805.10551 (2018).

[10] Z. Rudnick and I. Wigman, *On the volume of nodal sets for eigenfunctions of the Laplacian on the torus*, Ann. Henri Poincaré, 9 (2008), 109-130.

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02142-4307, USA

E-mail address: dmal@mit.edu