Annals of Mathematics **192** (2020), 551–581
https://doi.org/10.4007/annals.2020.192.2.6

A sharp square function estimate for the cone in $\mathbb{R}^3$

By Larry Guth, Hong Wang, and Ruixiang Zhang

Abstract

We prove a sharp square function estimate for the cone in $\mathbb{R}^3$ and consequently the local smoothing conjecture for the wave equation in $2+1$ dimensions.

1. Introduction

1.1. *Main results.* This paper concerns the restriction theory of the cone in $\mathbb{R}^3$. Let Γ be the truncated light cone $\Gamma = \{\xi_1^2 + \xi_2^2 = \xi_3^2, 1/2 \le \xi_3 \le 1\}$, and let $N_{R^{-1}}(\Gamma)$ denote its R^{-1}-neighborhood. Cover $N_{R^{-1}}(\Gamma)$ by finitely overlapping sectors θ of angular width $R^{-1/2}$, where each sector is a rectangular box of dimensions about $R^{-1} \times R^{-1/2} \times 1$. If $\hat{f}$ has support on $N_{R^{-1}}(\Gamma)$, we consider a set of functions $\{f_\theta\}$ such that

(a) $\hat{f}_\theta$ is supported on θ, and
(b) $f = \sum_\theta f_\theta$.

For example,[1] here is a natural way to choose $\{f_\theta\}$: let ψ_θ be a smooth partition of unity subordinate to the covering $\{\theta\}$, and define f_θ by $\hat{f}_\theta = \hat{f}\psi_\theta$. We prove the following sharp square function estimate for this decomposition:

Theorem 1.1 (Square function estimate). *For any $\epsilon > 0$, $R \ge 1$ and any function f whose Fourier transform is supported on $N_{R^{-1}}(\Gamma)$, we have*

$$\|f\|_{L^4(\mathbb{R}^3)} \le C_\epsilon R^\epsilon \left\| \left(\sum_\theta |f_\theta|^2 \right)^{1/2} \right\|_{L^4(\mathbb{R}^3)}.$$

This type of square function estimate was considered by Mockenhaupt [19] who proved that it implies the cone multiplier conjecture in $\mathbb{R}^3$, and by

Keywords: wave equation, local smoothing, square function estimate, incidence estimate
AMS Classification: Primary: 35L05, 42B15; Secondary: 42B25, 42B20.

[1]We remark that what we prove about $\{f_\theta\}$ in this paper is uniform as long as (a) and (b) are satisfied, i.e., it does not depend on the particular choice of $\{f_\theta\}$.

Mockenhaupt–Seeger–Sogge [20] (in a slightly different form) who essentially showed that it implies the local smoothing conjecture for the wave equation in $2+1$ dimensions. Here we recall the local smoothing conjecture, and we refer to [19] and [18] for more information about the cone multiplier conjecture. The local smoothing conjecture was formulated by Sogge in [23]. If u is a solution to the wave equation on $\mathbb{R}^n$, a local smoothing inequality bounds $\|u\|_{L^p(\mathbb{R}^n\times[1,2])}$ in terms of the Sobolev norms of the initial data. In particular, the local smoothing conjecture in $2+1$ dimensions is the following estimate.

THEOREM 1.2 (Local smoothing in 2+1 dimensions). *Suppose that $u(x,t)$ is a solution of the wave equation in $2+1$ dimensions, with initial data $u(x,0)=u_0(x)$ and $\partial_t u(x,0)=u_1(x)$. Then for any $p\geq 4$, and any $\alpha>\frac{1}{2}-\frac{2}{p}$,*

$$\|u\|_{L^p(\mathbb{R}^2\times[1,2])}\leq C_\alpha\left(\|u_0\|_{p,\alpha}+\|u_1\|_{p,-1+\alpha}\right). \tag{1}$$

Theorem 1.2 follows by combining Theorem 1.1 with the arguments in [20].

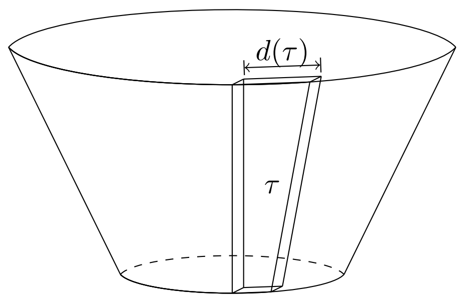
In [23], Sogge formulated the local smoothing conjecture, and he noticed that Bourgain's proof of the boundedness of the circular maximal operator in [1] can be used to establish "local smoothing" estimates with a non-trivial gain of regularity. The critical case of Theorem 1.2 is when $p=4$ and α is close to zero. Mockenhaupt, Seeger, and Sogge [20] proved that (1) holds for $p=4$ with $\alpha>1/8$, and this was improved afterwards by several authors ([24], [27], [17]). In [26], Wolff proved the local smoothing conjecture for $p\geq 74$ in the full range[2] of α. In that paper, Wolff introduced the idea of decoupling. His method was extended to higher dimensions by Łaba–Wolff [16] and refined by Garrigós–Seeger [12], [13] and Garrigós–Schlag–Seeger [11]. Then in [2], Bourgain and Demeter proved a sharp decoupling estimate for the cone in every dimension, in particular, proving the local smoothing conjecture in $2+1$ dimensions for $p\geq 6$ in the full range of α. The sharp decoupling estimate for the cone does not, however, imply the full range of local smoothing estimates; at the end of the introduction we will discuss what the issue is.

In a different direction, Lee and Vargas [18] proved a sharp L^3 square function estimate using multilinear restriction.

1.2. *Proof strategy.* One new feature of our approach is that we prove a stronger estimate that works better for induction on scales. We need a little notation to state this estimate. The precise details and definitions are provided

[2]To be more specific, Sogge originally made the conjecture for α in the range $\alpha>\frac{1}{2}-\frac{2}{p}$ and Wolff confirmed Sogge's conjecture for $p\geq 74$ and α in this range. Later in the work [15] of Heo, Nazarov and Seeger it was conjectured further that when $p>4$, the conjecture should hold for $\alpha\geq\frac{1}{2}-\frac{2}{p}$.

in Section 3. First we recall the locally constant property of f. For each sector θ, we let θ^* denote the dual rectangular box: since θ has dimensions $1 \times R^{-1/2} \times R^{-1}$, θ^* has dimensions $1 \times R^{1/2} \times R$. We call such a θ^* a plank. Recall that $|f_\theta|$ is roughly constant[3] on each translated copy of θ^*. In this paper we tile $\mathbb{R}^3$ with translated copies of θ^*. The restriction of f_θ to one translated copy of θ^* is called a *wave packet*. In addition to the sectors θ, we will consider larger angular sectors τ with any angle between $R^{-1/2}$ and 1. We write $d(\tau)$ to denote this angle, which we call the *aperture* of τ.



For each τ, we define[4] $f_\tau = \sum_{\theta\subset\tau} f_\theta$, and we define τ^* to be the dual rectangle to τ. If $d(\tau) = s$, then τ^* has dimensions $1 \times s^{-1} \times s^{-2}$, and $|f_\tau|$ is roughly constant on each translated copy of τ^*. Next we define $U_{\tau,R}$ to be a scaled copy of τ^* with diameter R. If $d(\tau) = s$, then $U_{\tau,R}$ has dimensions $Rs^2 \times Rs \times R$. Note that if $\theta \subset \tau$ and if T is a translated copy of θ^* that passes through the center of $U_{\tau,R}$, then $T \subset 10U_{\tau,R}$, where $10U_{\tau,R}$ means the dilation of $U_{\tau,R}$ by a factor of 10 with respect to its centroid. For each τ, we tile $\mathbb{R}^3$ by translated copies of $U_{\tau,R}$:

$$\mathbb{R}^3 = \bigsqcup_{U \text{ a translated copy of } U_{\tau,R}} U.$$

This tiling is natural because for each $\theta \subset \tau$, the support of each wave packet of f_θ is essentially contained in ~ 1 tiles U in the tiling. Here the notation for two quantities $A \sim B$ means that $A \le C_1 B \le C_2 A$ for some positive absolute constants C_1 and C_2. We write $\sum_{U /\!/ U_{\tau,R}}$ to denote the sum over all the translated copies U of $U_{\tau,R}$ in the tiling of $\mathbb{R}^3$.

---

[3]Such kind of "locally constant" heuristic will be used a few times in the current paper. To justify this intuition one can use Corollary 4.3 in [3]. See also Lemmas 6.1 and 6.2 in Section 6 of the current paper.

[4]This definition works best if τ is honestly tiled by θ. In general we abuse the notation a bit: Throughout this paper, by writing "summing over $\theta \subset \tau$," we really mean "summing over all $\theta \in A(\tau)$" where the collection $A(\tau)$ is chosen as follows: Each $A(\tau)$ only contains those θ's who intersect τ, and all $A(\tau)$ form a disjoint union $\{\theta\} = \bigsqcup_\tau A(\tau)$.

If U is a translated copy of $U_{\tau,R}$, then we define the square function $S_U f$ associated with U to be

$$S_U f = \left(\sum_{\theta\subset\tau} |f_\theta|^2\right)^{1/2} |_U.$$

We can now state our main estimate.

THEOREM 1.3. *Suppose that f has Fourier support on $N_{R^{-1}}(\Gamma)$. Then*

$$\|f\|^4_{L^4(\mathbb{R}^3)} \leq C_\epsilon R^\epsilon \sum_{R^{-1/2}\leq s\leq 1} \sum_{d(\tau)=s} \sum_{U /\!/ U_{\tau,R}} |U|^{-1} \|S_U f\|^4_{L^2}. \tag{2}$$

Here the sum over s is over dyadic values of s in the range $R^{-1/2} \leq s \leq 1$.

Let us take a moment to digest the right-hand side of this estimate. For this discussion, suppose that f is essentially supported on one B_R. We start with the term where $s = R^{-1/2}$. In this case τ is one of the original sectors θ of aperture $R^{-1/2}$, $U_{\tau,R}$ is equal to θ^*, and $|S_U f| = |f_\theta|\big|_U$. Since $|S_U f| = |f_\theta|$ is roughly constant on U,

$$|U|^{-1}\|S_U f\|^4_{L^2} \sim \|S_U f\|^4_{L^4}.$$

If the functions f_θ are essentially supported on disjoint regions, we would have

$$\|f\|^4_{L^4} \sim \sum_{d(\theta)=R^{-1/2}} \sum_{U /\!/ U_{\tau,R}} \|S_U f\|^4_{L^4},$$

which matches the term $s = R^{-1/2}$ on the right-hand side of (2). Next consider the term where $s = 1$. In this case, there is only one τ that covers all of Γ, and the contribution to the right-hand side is essentially $|B_R|^{-1}\|S_{B_R} f\|^4_{L^2} \sim |B_R|^{-1}\|f\|^4_{L^2(B_R)}$. If $|f|$ is roughly constant on the whole B_R, then we would have

$$\|f\|^4_{L^4(\mathbb{R}^3)} \sim \|f\|^4_{L^4(B_R)} \sim |B_R|^{-1}\|f\|^4_{L^2(B_R)} \sim |B_R|^{-1}\|S_{B_R} f\|^4_{L^2(B_R)},$$

which matches the term $s = 1$ on the right-hand side of (2). Finally we consider the intermediate values of s. It may happen that $f = f_\tau$ for some τ, that f is essentially supported on a particular translated copy U of $U_{\tau,R}$, and that $|f|$ is roughly constant on U. In this case,

$$\|f\|^4_{L^4(\mathbb{R}^3)} \sim \|f_\tau\|^4_{L^4(U)} \sim |U|^{-1}\|f_\tau\|^4_{L^2(U)} \sim |U|^{-1}\|S_U f\|^4_{L^2},$$

which is the term corresponding to U on the right-hand side of (2).

The proof of Theorem 1.3 is based on a new Kakeya-type estimate, which controls the overlapping of the planks in the wave packet decomposition of f.

LEMMA 1.4. *Suppose that $\hat{f}$ has support on $N_{R^{-1}}(\Gamma)$. Let g denote the* (*squared*) *square function $g = \sum_{d(\theta)=R^{-1/2}} |f_\theta|^2$. Then*

$$\int_{\mathbb{R}^3} |g|^2 \lesssim \sum_{R^{-1/2} \le s \le 1} \sum_{d(\tau)=s} \sum_{U /\!/ U_{\tau,R}} |U|^{-1} \|S_U f\|_{L^2}^4,$$

where $A \lesssim B$ means that $A \le CB$ for some absolute positive constant C.

Recall that each function $|f_\theta|$ is morally constant on the translated copies of θ^*, where each θ^* is a $1 \times R^{1/2} \times R$ plank. The estimate in Lemma 1.4 is a Kakeya-type bound on the overlapping of these planks. The new feature of this estimate compared to previous Kakeya-type estimates is the structure of the right-hand side, which is designed to match the right-hand side of Theorem 1.3. The terms on the right-hand side keep track of how planks are packed into the rectangular boxes U. If the planks are spread out in the sense that each box U does not contain too many planks, then it gives a strong bound.

In [26], Wolff connected Kakeya-type estimates for overlapping planks to incidence geometry problems in the spirit of the Szemerédi–Trotter problem. He adapted the cutting method from incidence geometry to this setting and he used it to estimate the overlaps of planks. He applied those geometric estimates at many scales to prove his results on local smoothing. In [2], Bourgain and Demeter apply multilinear Kakeya estimates at many scales to prove decoupling. In this paper, we apply Lemma 1.4 at many scales to prove Theorem 1.3.

Lemma 1.4 is proven using Fourier analysis. By Plancherel, $\int |g|^2 = \int |\hat{g}|^2$. Roughly speaking, we decompose the Fourier space, and the contributions of different regions to $\int |\hat{g}|^2$ correspond to the different terms on the right-hand side of Lemma 1.4. This approach to proving Kakeya-type estimates is based on some work of Orponen in projection theory [21] and is related to Vinh's work [25] about incidence geometry over finite fields. It builds on [14], which applies similar ideas to rectangles and tubes instead of planks.

1.3. *Local estimates.* Our Theorem 1.3 and Lemma 1.4 have "local" counterparts involving polynomially decaying weights that are essentially supported on a given box. For any box B_R of diameter R, define the weight

$$w_{B_R,E}(x) = \left(1 + \frac{\operatorname{dist}(x, B_R)}{R}\right)^{-E}.$$

Here is the local version of Theorem 1.3.

THEOREM 1.5. *If f has Fourier support on $N_{R^{-1}}(\Gamma)$, then for any $E > 0$,*

$$\|f\|_{L^4(B_R)}^4 \le C_{\epsilon,E} R^\epsilon \sum_{R^{-1/2} \le s \le 1} \sum_{d(\tau)=s} \sum_{U /\!/ U_{\tau,R}} |U|^{-1} \|w_{B_R,E} \cdot S_U f\|_{L^2}^4. \tag{3}$$

Here the sum over s is over dyadic values of s in the range $R^{-1/2} \le s \le 1$.

In the above theorem, the sum on the right-hand side is also "morally localized." It is

$$\sum_{R^{-1/2}\le s\le 1}\ \sum_{d(\tau)=s}\ \sum_{U/\!/U_{\tau,R},U\subset 100B_R} |U|^{-1}\|S_U f\|_{L^2}^4$$

plus some decaying error term. To prove Theorem 1.5, we multiply f by a rapidly decaying bump function ϕ_R adapted to B_R such that $|\phi_R| > \frac{1}{C} > 0$ on B_R and $\hat{\phi}_R$ is supported on the ball $B_{R^{-1}}$ centered at the origin, and then we apply Theorem 1.3 to the decomposition $\phi_R f = \sum_\theta \phi_R f_\theta$.

1.4. *Relationship with decoupling.* While working on this project, we were strongly influenced by ideas related to decoupling, but the proof given here does not use the decoupling theorem per se. It does make use of a nice observation that Bourgain and Demeter used to reduce the decoupling theorem for the cone to the decoupling theorem for the paraboloid. (See [2]. Similar ideas can also be traced back to the iteration argument of Pramanik–Seeger [22].) Instead of working with a truncated cone of height 1, Bourgain and Demeter worked with a truncated cone of height $1/K$ for a large constant K, denoted $\Gamma_{\frac{1}{K}}$. This shorter truncated cone can be approximated by a parabola at various scales. We will also work with $\Gamma_{\frac{1}{K}}$, allowing us to bring into play some estimates for the parabola.

As we mentioned above, sharp decoupling theorems do not imply the full range of local smoothing estimates or the square function estimate. Let us explain a little further what the issue is. The decoupling theorem for the cone gives the following bounds, which are sharp for every p between 2 and ∞:

$$\|f\|_{L^p(\mathbb{R}^3)} \le C_\epsilon R^\epsilon \left(\sum_{d(\theta)=R^{-1/2}} \|f_\theta\|_{L^p(\mathbb{R}^3)}^2\right)^{1/2} \quad \text{if } 2\le p\le 6, \tag{4}$$

$$\|f\|_{L^p(\mathbb{R}^3)} \le C_\epsilon R^{\frac14-\frac{3}{2p}+\epsilon} \left(\sum_{d(\theta)=R^{-1/2}} \|f_\theta\|_{L^p(\mathbb{R}^3)}^2\right)^{1/2} \quad \text{if } p\ge 6. \tag{5}$$

For any given p, (5) implies local smoothing for that p. But the inequality (5) cannot hold for any $p<6$ because the power of R would be negative. The power of R in a decoupling inequality cannot be negative because of the following example: suppose that for each θ, $|f_\theta|$ is approximately the characteristic function of B_R, and at each point $|f| \sim \left(\sum_\theta |f_\theta|^2\right)^{1/2}$. In this case, $\|f\|_{L^p} \sim \left(\sum_\theta \|f_\theta\|_{L^p}^2\right)^{1/2}$ for all p. This example is not a counterexample for local smoothing, but to prove local smoothing for some $p<6$ we have to do better than inequality (4) in some scenarios—for instance, if the supports of

f_θ are essentially disjoint at time 0. Roughly speaking, we need to improve the bound (4) when $p < 6$ and when each f_θ is essentially supported on a sparse region of B_R. Theorem 1.3 makes this precise.

There are similar issues in the problem of decoupling into small caps, which was studied in [9]. For instance, consider an exponential sum of the form

$$f(x_1, x_2) = \sum_{j=1}^{N} a_j e\left(\frac{j}{N}x_1 + \frac{j^2}{N^2}x_2\right), \text{ with } |a_j| \le 1 \text{ for all } j. \tag{6}$$

The decoupling theorem for the parabola gives a sharp bound on $\|f\|_{L^p(B_{N^2})}$ for every p. But suppose we want to bound $\|f\|_{L^p(B_R)}$ for some $R < N^2$. If we divide the parabola into arcs θ of length $R^{-1/2}$, then each f_θ is a sum of $\sim NR^{-1/2}$ terms of (6). It is not hard to estimate the largest possible value of $\|f_\theta\|_{L^p(B_R)}$ for each p. Combining this bound for $\|f_\theta\|_{L^p(B_R)}$ with decoupling gives an upper bound for $\|f\|_{L^p(B_R)}$, but it is not sharp. When $\|f_\theta\|_{L^p(B_R)}$ is close to its largest value, then $|f_\theta|$ is concentrated on a sparse region of B_R. The argument in [9] exploits this sparsity to improve the bound from decoupling and give sharp estimates for $\|f\|_{L^p(B_R)}$ for every p. The proof of the main theorem here builds on that proof.

The paper [9] also considers a decoupling problem in which the cone is divided into small squares instead of sectors. This problem was raised by Bourgain and Watt [5] in their work on the Gauss circle problem. The paper [9] shows that the square function estimate Theorem 1.1 implies a sharp estimate for this decoupling problem.

Acknowledgements. We would like to thank Ciprian Demeter for sharing his ideas and for many helpful conversations. He proposed the problem of decoupling into small caps and suggested improving decoupling when each f_θ is concentrated in a sparse region. We would also like to thank Misha Rudnev for sharing thoughtful comments about [14] that helped us in this project. We would like to thank Terence Tao for helpful comments that improved the exposition of the proof of Proposition 3.4. We would like to thank Zhipeng Lu and Xianchang Meng for pointing out several typos in an earlier version. LG was supported by a Simons Investigator Award. HW was supported by the Simons Foundation grant for David Jerison. RZ was supported by the National Science Foundation under Grant Number DMS-1856541. He would like to thank Andreas Seeger for helpful historical remarks about square functions and local smoothing. Part of this work was done when RZ was visiting MIT and he would like to thank MIT for the warm hospitality.

We would like to thank the anonymous referees for their thorough readings and many helpful suggestions.

2. Proof of the square function estimate from Theorem 1.3

In this section, we explain how Theorem 1.3 implies the square function estimate Theorem 1.1, and we discuss how the latter implies the local smoothing Theorem 1.2. First we recall the statement of Theorem 1.1:

THEOREM. *For any function f whose Fourier transform is supported on $N_{R^{-1}}(\Gamma)$, we have*

$$\|f\|_{L^4(\mathbb{R}^3)} \leq C_\epsilon R^\epsilon \| \left(\sum_{d(\theta)=R^{-1/2}} |f_\theta|^2 \right)^{1/2} \|_{L^4(\mathbb{R}^3)}.$$

Proof. Let U be a translated copy of $U_{\tau,R}$. Recall that

$$\|S_U f\|_{L^2}^2 = \int_U \sum_{\theta\subset\tau} |f_\theta|^2.$$

By Cauchy–Schwarz,

$$\|S_U f\|_{L^2}^4 \leq |U| \int_U \left(\sum_{\theta\subset\tau} |f_\theta|^2 \right)^2.$$

Therefore,

$$\sum_{d(\tau)=s} \sum_{U /\!/ U_{\tau,R}} |U|^{-1} \|S_U f\|_{L^2}^4 \leq \sum_{d(\tau)=s} \int_{\mathbb{R}^3} \left(\sum_{\theta\subset\tau} |f_\theta|^2 \right)^2$$
$$\leq \int_{\mathbb{R}^3} \left(\sum_{\theta} |f_\theta|^2 \right)^2. \qquad \square$$

Summing in s (dyadic numbers) contributes an additional $\log R$ factor compared to Theorem 1.3.

Essentially by [20], the square function estimate in Theorem 1.1 implies the local smoothing Theorem 1.2 for the wave equation in $2+1$ dimensions. This implication was sketched in Proposition 6.2 of [24]. One technical difference is that the square function considered in [20] was the one in terms of "small caps" ζ, $R^{-1/2}$-squares on Γ. Instead of the Littlewood–Paley estimate corresponding to equally spaced decompositions in $\mathbb{R}^2$ used in [20] (see (1.9) and the following first two lines on page 214 of [20]), one needs such an estimate for angular decompositions. In the L^4 case, such an angular square function estimate was proved by Córdoba (see ii) on the first page of [8]). Another proof[5] by Carbery–Seeger can be found in [6].

[5]See Proposition 4.6 in [6]. That proposition has two parameters and Córdoba's estimate (up to an R^ε-loss) can be viewed as a simpler one-parameter variant. See also the remark in the end of Section 4 in [6]

3. Outline of the proof of the main theorem

In this section, we give an overview of the proof of Theorem 1.3 and outline the rest of the paper. First we review the statement of Theorem 1.3 and present it in a more detailed way.

Let Γ be the truncated light cone $\Gamma = \{\xi_1^2 + \xi_2^2 = \xi_3^2, 1/2 \leq |\xi_3| \leq 1\}$. We now precisely define the sectors discussed in the introduction. For each point $\xi \in \Gamma$ with $\xi_3 = 1$, we define a basis of $\mathbb{R}^3$ as follows: the core line direction is $\mathbf{c}(\xi) = (\xi_1, \xi_2, 1)$, the normal direction is $\mathbf{n}(\xi) = (\xi_1, \xi_2, -1)$, and the tangent direction is $\mathbf{t}(\tau) = (-\xi_2, \xi_1, 0)$. Now for each such ξ, and each $s < 1$, we define the sector with direction ξ and aperture s as follows:

$$\tau(s,\xi) = \{\omega \in \mathbb{R}^3 : 1 \leq \mathbf{c}(\xi) \cdot \omega \leq 2 \text{ and } |\mathbf{n}(\xi) \cdot \omega| \leq s^2 \text{ and } |\mathbf{t}(\xi) \cdot \omega| \leq s\}.$$

Here $s = d(\tau)$ is the aperture of τ as described in the introduction.

For each s, We choose $10s^{-1}$ evenly spaced ξ in the circle $\Gamma \cap \{\xi_3 = 1\}$, and we let $\mathbf{S}_s$ be the set of $\tau(s,\xi)$ for these ξ. It is straightforward to check that these form a finitely overlapping cover of $N_{s^2}(\Gamma)$.

In the introduction, we considered a finitely-overlapping cover of $N_{R^{-1}}\Gamma$ by sectors θ with dimensions $\sim R^{-1} \times R^{-1/2} \times 1$. The set of these sectors is $\mathbf{S}_{R^{-1/2}}$.

For each $\tau = \tau(s,\xi)$, and each $\rho \geq s^{-2}$, we define a box $U_{\tau,\rho}$ as follows:

$$U_{\tau,\rho} = \{x \in \mathbb{R}^3 : |\mathbf{c}(\xi) \cdot x| \leq \rho s^2 \text{ and } |\mathbf{n}(\xi) \cdot x| \leq \rho \text{ and } |\mathbf{t}(\xi) \cdot x| \leq \rho s\}. \tag{7}$$

The box $U_{\tau,\rho}$ is approximately the convex hull of the union of θ^* over all sectors $\theta \subset \tau$ with $d(\theta) = \rho^{-1/2}$. In other words, $U_{\tau,\rho}$ is approximately the smallest rectangular box such that for any $\rho^{-1/2}$-sector $\theta \subset \tau$, if a translated copy of θ^* intersects $U_{\tau,\rho}$, then it must lie in $10U_{\tau,\rho}$. We tile $\mathbb{R}^3$ by translated copies of $U_{\tau,\rho}$.

If U is a translated copy of $U_{\tau,\rho}$, then we define $S_U f$ by

$$S_U f = \left(\sum_{\theta \in \mathbf{S}_{\rho^{-1/2}} : \theta \subset \tau} |f_\theta|^2 \right)^{1/2} |_U. \tag{8}$$

As written, this definition appears to depend upon U, τ, and ρ. But in fact the parameters ρ and τ can be read off from U. The parameter ρ is the diameter of U. The aperture $d(\tau) = s$ can be read off from the dimensions of U, which are $\rho s^2 \times \rho s \times \rho$. And the direction ξ of τ can be read off from the direction of U. To illustrate this, suppose that U is B_r — a ball of radius r. The diameter of U is r, and so $\rho = r$. The dimensions of U are $r \times r \times r$, and so $d(\tau) = 1$. Since τ has aperture 1, it covers all of Γ. Therefore,

$$S_{B_r} f = \left(\sum_{\theta \in \mathbf{S}_{r^{-1/2}}} |f_\theta|^2 \right)^{1/2} \Big|_{B_r}.$$

In particular, $|S_{B_1} f|$ is just $|f|$ restricted to B_1.

We define $S(r, R)$ as the smallest constant such that for every function f with $\mathrm{supp}\hat{f} \subset N_{R^{-1}}(\Gamma)$,

$$\text{(9)} \quad \sum_{B_r \subset \mathbb{R}^3} |B_r|^{-1} \|S_{B_r} f\|^4_{L^2(B_r)} \leq S(r, R) \sum_{R^{-1/2} \leq s \leq 1} \sum_{\tau \in \mathbf{S}_s} \sum_{U /\!/ U_{\tau,R}} |U|^{-1} \|S_U f\|^4_{L^2}.$$

On the left-hand side of inequality (9), $\sum_{B_r \subset \mathbb{R}^3}$ means the sum over the balls B_r in a finitely overlapping cover of $\mathbb{R}^3$. On the right-hand side of inequality (9), the first sum, $\sum_{R^{-1/2} \leq s \leq 1}$, means the sum over dyadic numbers s between $R^{-1/2}$ and 1. The last sum, $\sum_{U /\!/ U_{\tau,R}}$, means the sum over a set of translates of $U_{\tau,R}$ that tile $\mathbb{R}^3$.

By Hölder's inequality, $S(r, R) < \infty$ for any $0 < r, R < \infty$. We will only consider $S(r, R)$ when $r \leq R$. Theorem 1.3 is equivalent to the bound $S(1, R) \leq C_\epsilon R^\epsilon$ since $|S_{B_1} f| = |f|$ on any B_1 and $|f|$ is morally constant on B_1. We will derive Theorem 1.3 from a series of bounds for $S(r, R)$.

In Section 4, we prove the Kakeya-type estimate Lemma 1.4, and we use it to prove

LEMMA 3.1. *For any* $r \geq 10$, $r_1 \in [r, r^2]$,

$$S(r_1, r^2) \leq C.$$

Next we bring into play a trick from the proof of decoupling for the cone in [2]: instead of working with Γ we work with a subset of Γ that lies close to a short parabolic cylinder. We let P denote an arc of a parabola of length ~ 1 lying in Γ. For any $K \geq 10$, we define $\Gamma_{\frac{1}{K}}$ to be the $1/K$-neighborhood of P in Γ. We will eventually choose K to be a large constant depending on ϵ (which remains fixed as $R \to \infty$). The precise formula for $\Gamma_{\frac{1}{K}}$ is designed to make Lorentz rescaling work in a clean way, and we give the formula in Section 5 when we discuss Lorentz rescaling. We can define a sector $\tau \subset \Gamma_{\frac{1}{K}}$ and its aperture $d(\tau)$ in the same way as before (again see Section 5). Then we define $S_K(r, R)$ as the smallest constant such that (9) holds for every f with supp $\hat{f} \subset N_{R^{-1}}(\Gamma_{\frac{1}{K}})$. Since $\Gamma_{\frac{1}{K}} \subset \Gamma$, $S_K(r, R) \leq S(r, R)$. On the other hand, since K will be a chosen constant, $S_K(r, R)$ is almost equal to $S(r, R)$ and we can use it equally well to prove Theorem 1.3.

If $R = K$, then $N_{R^{-1}}(\Gamma_{\frac{1}{K}})$ is the $1/K$-neighborhood of the parabolic arc P, and the restriction theory for the parabola can be used to study $S_K(1, K)$. In Section 6 we use this idea to prove the following lemma:

Lemma 3.2. *For any* $K \geq 10$, *any* $1 \leq r \leq K$, *and any* $\delta > 0$, *we have* $S_K(r, K) \leq C_\delta K^\delta$.

Theorem 1.3 will follow by combining Lemmas 3.1 and 3.2 with a Lorentz rescaling argument. We review the Lorentz rescaling in Section 5. We use it in Section 7 to prove the following lemma, which relates $S_K(r, R)$ for various values of r, R:

Lemma 3.3. *For any* $r_1 < r_2 \leq r_3$,

$$S_K(r_1, r_3) \leq \log r_2 \cdot S_K(r_1, r_2) \max_{r_2^{-1/2} \leq s \leq 1} S_K(s^2 r_2, s^2 r_3).$$

This lemma is an important motivation for working with $S_K(r, R)$. It allows Lemmas 3.1 and 3.2 to be applied at many different scales. A key point of studying Theorem 1.3 instead of trying to prove Theorem 1.1 directly is that it allows this multiscale analysis to come into play.

Assuming the lemmas, we now prove bounds on $S_K(r, R)$ and use them to deduce Theorem 1.3.

Proposition 3.4. *For any* $\epsilon > 0$, *there exists* $K = K(\epsilon)$ *so that for any* $1 \leq r \leq R$, *we have*

$$S_K(r, R) \leq \widetilde{C}_\epsilon (R/r)^\epsilon.$$

Proof. First we note that if $r > R^{1/2}$, then Lemma 3.1 tells us that $S_K(r, R) \leq S(r, R) \leq C$, and so the conclusion holds.

Let $K = K(\epsilon) > 10$ be a constant depending only on ϵ that we will choose below. (The constant $K(\epsilon)$ will depend on ϵ and on the constants in Lemmas 3.1 and 3.2.)

We apply induction on the ratio R/r.

Our base case is when $R/r \leq \sqrt{K}$. We have already checked the proposition in case $r > R^{1/2}$. If $r \leq R^{1/2}$ and $R/r \leq \sqrt{K}$, then $R \leq K$. In this case, since K is a constant depending only on ϵ, it is straightforward to check that $S_K(r, R)$ is bounded by a constant $\tilde{C}_K = \tilde{C}_\epsilon$. This finishes the base case.

Next we proceed with the induction. Given a pair (r, R), our induction hypothesis is the following: for any pair (r', R') with $R'/r' \leq R/2r$, we have $S_K(r', R') \leq \tilde{C}_\epsilon (R'/r')^\epsilon$.

The proof of the induction has two cases, depending on whether $r \leq K^{1/2}$.

If $r \leq K^{1/2}$, we apply Lemma 3.3 with $r_1 = r$, $r_2 = K^{1/2} r$, and $r_3 = R$, which gives

$$S_K(r, R) \leq \log K \cdot S_K(r, K^{1/2} r) \max_{r_2^{-1/2} \leq s \leq 1} S_K(s^2 K^{1/2} r, s^2 R).$$

We bound the first S_K factor using Lemma 3.2, and we bound the second S_K factor using induction. These bounds give

$$\begin{aligned} S_K(r,R) &\le \log K \cdot S_K(r, K^{1/2}r) \max_{r_2^{-1/2}\le s\le 1} S_K(s^2K^{1/2}r, s^2R) \\ &\le \log K \cdot C_\delta \widetilde{C}_\epsilon K^\delta \left(\frac{R}{K^{1/2}r}\right)^\epsilon. \end{aligned}$$

We choose $\delta = \epsilon/4$, and then we choose $K = K(\epsilon)$ large enough so that $\log K \cdot C_{\epsilon/4}K^{-\epsilon/4} \le 1$, and the induction closes in this case.

Now suppose $r \ge K^{1/2}$. Recall from the start of the proof that we may assume $r \le R^{1/2}$. We apply Lemma 3.3 with $r_1 = r$, $r_2 = r^2$, and $r_3 = R$, which gives

$$S_K(r,R) \le 2\log r \cdot S_K(r, r^2) \max_{r^{-1}\le s\le 1} S_K(s^2r^2, s^2R).$$

We bound the first S_K factor using Lemma 3.1 and we bound the second S_K factor using induction, giving

$$S_K(r,R) \le 2\log r \cdot S_K(r,r^2) \max_{r^{-1}\le s\le 1} S_K(s^2r^2, s^2R) \le 2\log r \cdot C\widetilde{C}_\epsilon \left(\frac{R}{r^2}\right)^\epsilon.$$

We choose $K = K(\epsilon)$ large enough so that for all $r \ge K^{1/2}$, we have $2\log r \cdot Cr^{-\epsilon} \le 1$, and the induction closes in this case. □

Finally we show how Proposition 3.4 implies Theorem 1.3.

Proof. Proposition 3.4 implies that for every $\epsilon > 0$, we can choose $K = K(\epsilon)$ so that $S_K(1,R) \le C_\epsilon R^\epsilon$ for all R. Suppose that the support of $\hat{f}$ is contained in $N_{R^{-1}}(\Gamma_{\frac{1}{K}}) \subset B_3$. Since $|f|$ is morally constant on unit balls, we have[6]

$$\begin{aligned} \int_{\mathbb{R}^3} |f|^4 &\lesssim \sum_{B_1\subset\mathbb{R}^3} \|f\|^4_{L^2(B_1)} = \sum_{B_1\subset\mathbb{R}^3} \|S_{B_1}f\|^4_{L^2(B_1)} \\ &\le C_\epsilon R^\epsilon \sum_{R^{-1/2}\le s\le 1} \sum_{d(\tau)=s} \sum_{U /\!/ U_{\tau,R}} |U|^{-1}\|S_U f\|^4_{L^2}. \end{aligned} \tag{10}$$

This inequality is essentially Theorem 1.3 except that we assumed that $\hat{f}$ is supported on $N_{R^{-1}}(\Gamma_{\frac{1}{K}})$ instead of $N_{R^{-1}}(\Gamma)$. Since $N_{R^{-1}}(\Gamma)$ can be covered by $O(K) = O_\epsilon(1)$ affine copies of $\Gamma_{\frac{1}{K}}$, we can reduce Theorem 1.3 to (10). Here are the details.

Take $\{A_j\}_{1\le j\lesssim K}$ to be a collection of linear transformations such that $\Gamma \subset \bigcup A_j(\Gamma_{\frac{1}{K}})$. Here each A_j is a composition of a scaling by a factor ~ 1 and

[6] Strictly speaking, one need to apply Lemmas 6.1 and 6.2 to justify the first "$\lesssim$" in inequality (10). This is similar to the arguments in Section 6, where we give full details.

a rotation in the (ξ_1, ξ_2)-plane.[7] Similarly, we can arrange that $N_{R^{-1}}(\Gamma) \subset \bigcup A_j\big(N_{R^{-1}}(\Gamma_{\frac{1}{K}})\big)$. Let $\{\psi_j\}$ be a C^∞ partition of unity subordinate to this covering. This partition of unity only depends on K. If f is a function whose Fourier transform is supported on $N_{R^{-1}}(\Gamma)$, then $\hat{f} = \sum_j \psi_j \hat{f}$. Define f_j by $\hat{f}_j = \psi_j \hat{f}$ and $\hat{f}_{j,\theta} = \psi_j \hat{f}_\theta$. The support of $\hat{f}_j$ is contained in $A_j(N_{R^{-1}}(\Gamma_{\frac{1}{K}}))$. Since (10) is invariant under rotations and approximately invariant under rescaling by a factor ~ 1, (10) holds for each function f_j.

Now by the triangle inequality and Hölder's inequality,

$$\begin{aligned}\|f\|_{L^4(\mathbb{R}^3)}^4 &\lesssim K^3 \sum_j \|f_j\|_{L^4(\mathbb{R}^3)}^4 \\ &\lesssim K^3 C_\epsilon R^\epsilon \sum_j \sum_{R^{-1/2}\le s\le 1} \sum_{d(\tau)=s} \sum_{U /\!/ U_{\tau,R}} |U|^{-1} \|S_U f_j\|_{L^2}^4 \\ &\lesssim K^3 C_\epsilon R^\epsilon \sum_{R^{-1/2}\le s\le 1} \sum_{d(\tau)=s} \sum_{U /\!/ U_{\tau,R}} |U|^{-1} \left(\sum_j \|S_U f_j\|_{L^2}^2\right)^2 \\ &\lesssim_K C_\epsilon R^\epsilon \sum_{R^{-1/2}\le s\le 1} \sum_{d(\tau)=s} \sum_{U /\!/ U_{\tau,R}} |U|^{-1} \|S_U f\|_{L^2}^4.\end{aligned}$$

To see the last inequality, note that $f_{j,\theta} = f_\theta * \check{\psi}_j$ and $\check{\psi}_j$ is rapidly decaying outside the ball of radius K centered at the origin. Hence, by Lemma 6.2, each $\|f_{j,\theta}\|_{L^2(B_1)} \lesssim_K \|f_\theta\|_{L^2(w_{B_1,E})}$ for any polynomially decaying weight $w_{B_1,E}$. It suffices to take E large enough.

Since K is a constant only depending on ϵ, this gives Theorem 1.3. □

4. A Kakeya-type estimate

In this section, we prove the Kakeya-type estimate Lemma 1.4, and we use it to prove Lemma 3.1. First we recall the following statement:

LEMMA. *Suppose $\hat{f}$ has support on $N_{r^{-2}}(\Gamma)$. Let g denote the (squared) square function $g = \sum_{\theta\in \mathbf{S}_{r^{-1}}} |f_\theta|^2$. Then*

$$\int_{\mathbb{R}^3} |g|^2 \lesssim \sum_{R^{-1/2}\le s\le 1} \sum_{\tau\in \mathbf{S}_s} \sum_{U /\!/ U_{\tau,R}} |U|^{-1} \|S_U f\|_{L^2}^4.$$

(Comparing with the statement in the introduction, we use r^2 in place of R. This makes the algebra in the proof a little simpler, and it connects with the notation in Lemma 3.1.)

[7] One can choose $\lesssim 1$ rotations R_k such that $\bigcup_k R_k(\Gamma_{\frac{1}{K}})$ covers $\Gamma(h) = \Gamma \cap \{h \le \xi_3 \le h + K/10\}$ for some $h \sim 1$. Then we choose $\lesssim K$ dilations D_l such that $\Gamma \subset \bigcup_l D_l(\Gamma(h))$. We define $A_j = D_l R_k$. for some l and k.

Proof of Lemma 1.4. Suppose that supp $\hat{f} \subset N_{r^{-2}}(\Gamma)$. Recall that

$$g = \sum_{\theta \in \mathbf{S}_{r^{-1}}} |f_\theta|^2.$$

The Fourier transform of $|f_\theta|^2$ is supported on the Minkowski sum $\tilde{\theta} = \theta + (-\theta)$. The set $\tilde{\theta}$ is itself a plank of dimensions $\sim r^{-2} \times r^{-1} \times 1$ centered at the origin. Notice that while the original sectors θ are disjoint, the planks $\tilde{\theta}$ are not disjoint. The way that they overlap plays an important role in the proof.

The Minkowski sum $\tilde{\theta}(\xi) = \theta(\xi) + (-\theta(\xi))$ is approximately equal to the following rectangular box:

$$\tilde{\theta}(\xi) \approx \{\omega \in \mathbb{R}^3 : |\mathbf{c}(\xi) \cdot \omega| \le 1 \text{ and } |\mathbf{n}(\xi) \cdot \omega| \le r^{-2} \text{ and } |\mathbf{t}(\xi) \cdot \omega| \le r^{-1}\},$$

where two convex sets $A \approx B$ means that $A \subset 10B \subset 100A$.

The overlapping of the boxes $\tilde{\theta}$ is best described in terms of similar rectangular boxes at smaller scales. For any dyadic σ in the range $r^{-1} \le \sigma \le 1$, and any ξ as above, we define a box $\Theta = \Theta(\sigma, \xi)$ by

$$\Theta(\sigma, \xi) = \{\omega : |\mathbf{c}(\xi) \cdot \omega| \le \sigma^2 \text{ and } |\mathbf{n}(\xi) \cdot \omega| \le r^{-2} \text{ and } |\mathbf{t}(\xi) \cdot \omega| \le r^{-1}\sigma\}. \tag{11}$$

Notice that $\Theta(1, \xi)$ is equal to $\tilde{\theta}(\xi)$, and for $\sigma < 1$, $\Theta(\sigma, \xi) \subset \tilde{\theta}(\xi)$. At the other extreme, $\Theta(r^{-1}, \xi)$ is essentially the ball of radius r^{-2} centered at the origin, regardless of ξ.

If we intersect $\Theta(\sigma, \xi)$ with the slab $\{(1/2)\sigma^2 \le \omega_3 \le \sigma^2\}$, then it lies in the r^{-2}-neighborhood of the light cone. Let $\Gamma(\sigma^2)$ denote the part of the light cone where $(1/2)\sigma^2 \le \omega_3 \le \sigma^2$. Each $\Theta(\sigma, \xi) \cap \{(1/2)\sigma^2 \le \omega_3 \le \sigma^2\}$ is a sector of $N_{r^{-2}}(\Gamma(\sigma^2))$, just as θ is a sector of $N_{r^{-2}}(\Gamma)$. The number of such sectors needed to cover $N_{r^{-2}}(\Gamma(\sigma^2))$ is $\sim \sigma r$. If $|\xi - \xi'| > \sigma^{-1} r^{-1}$, then $\Theta(\sigma, \xi) \cap \Theta(\sigma, \xi') \cap \{(1/2)\sigma^2 \le \omega_3 \le \sigma^2\}$ is empty. Conversely, if $|\xi - \xi'| < \sigma^{-1} r^{-1}$, then $\Theta(\sigma, \xi) \cap \{(1/2)\sigma^2 \le \omega_3 \le \sigma^2\}$ is comparable to $\Theta(\sigma, \xi') \cap \{(1/2)\sigma^2 \le \omega_3 \le \sigma^2\}$. By symmetry, the same holds when we intersect with $\{-\sigma^2 \le \omega_3 \le -(1/2)\sigma^2\}$ at the other side of the light cone. Now by convexity, we conclude that if $|\xi - \xi'| \le \sigma^{-1} r^{-1}$, then $\Theta(\sigma, \xi) \subset 2\Theta(\sigma, \xi')$.

For each dyadic σ in the range $r^{-1} \le \sigma \le 1$, let $\mathbf{CP}_\sigma$ be a set of $\sim \sigma r$ planks of the form $\Theta(\sigma, \xi)$ with the directions ξ evenly spaced in the circle. (The letters **CP** stand for centered plank.) The size of $\mathbf{CP}_\sigma$ is chosen so that for any $\Theta(\sigma, \xi)$, we can choose $\Theta(\sigma, \xi') \in \mathbf{CP}_\sigma$ so that $\Theta(\sigma, \xi) \subset 2\Theta(\sigma, \xi')$. We define **CP** as a union over dyadic scales: $\mathbf{CP} = \cup_{r^{-1} \le \sigma \le 1} \mathbf{CP}_\sigma$. Since $\Theta(1, \xi)$ is the same as $\tilde{\theta}(\xi)$, $\mathbf{CP}_1 = \mathbf{S}_{r^{-1}}$. On the other hand, $\mathbf{CP}_{r^{-1}}$ is a set with one element, which is essentially the ball of radius r^{-2} around the origin.

For a given $\theta(\xi)$ and a given scale σ, there are ~ 1 $\Theta = \Theta(\sigma, \xi') \in \mathbf{CP}_\sigma$ with $\Theta \subset 2\tilde{\theta}$. To see this, note on the one hand that $\Theta(\sigma, \xi) \subset \tilde{\theta}(\xi)$, and we can choose $\Theta(\sigma, \xi') \in \mathbf{CP}_\sigma$ so that $\Theta(\sigma, \xi') \subset 2\Theta(\sigma, \xi)$. On the other hand,

$\tilde{\theta}(\xi) \cap N_{r^{-2}}(\Gamma(\sigma^2))$ is essentially equal to the sector $\Theta(\sigma, \xi) \cap \{(1/2)\sigma^2 \leq \omega_3 \leq \sigma^2\}$, and so $2\tilde{\theta}(\xi)$ contains $\Theta(\sigma, \xi')$ only if $|\xi - \xi'| \lesssim \sigma^{-1} r^{-1}$.

In our proof, r remains fixed but we have to consider various scales σ. To simplify notation, we abbreviate $\mathbf{S}_{r^{-1}}$ as $\mathbf{S}$. Now for each scale σ, for each $\theta = \theta(\xi) \in \mathbf{S} = \mathbf{S}_{r^{-1}}$, we associate one $\Theta = \Theta(\sigma, \xi') \in \mathbf{CP}_\sigma$ with $|\xi' - \xi| \leq \sigma^{-1} r^{-1}$. For each $\Theta \in \mathbf{CP}_\sigma$, we let $\mathbf{S}_\Theta$ be the set of all $\theta \in \mathbf{S}$ that are associated with Θ. So for each σ, $\mathbf{S} = \bigsqcup_{\Theta \in \mathbf{CP}_\sigma} \mathbf{S}_\Theta$. If $\theta \in \mathbf{S}_\Theta$, then $\Theta \subset 2\tilde{\theta}$.

Let $\Omega = \cup_{\theta \in \mathbf{S}} \tilde{\theta} \sim \cup_{\Theta \in \mathbf{CP}_1} \Theta$. Since $(|f_\theta|^2)^\wedge$ is supported on $\tilde{\theta}$, it follows that $\hat{g}$ is supported on Ω. We break Ω into pieces associated with different scales σ as follows. We define $\Omega_{\leq \sigma} = \cup_{\Theta \in \mathbf{CP}_\sigma} \Theta$. Then we define $\Omega_\sigma = \Omega_{\leq \sigma} \setminus \Omega_{\leq \sigma/2}$ if $\sigma > r^{-1}$, and we define $\Omega_{r^{-1}} = \Omega_{\leq r^{-1}}$, so that

$$\Omega = \bigsqcup_{r^{-1} \leq \sigma \leq 1} \Omega_\sigma.$$

(Here $\bigsqcup$ denotes a disjoint union, and the union is over dyadic σ.)

Now if $\omega \in \Omega_\sigma$, we bound $|\hat{g}(\omega)|$ as follows:

$$\left|\hat{g}(\omega)\right| = \Big|\sum_{\theta \in \mathbf{S}} (|f_\theta|^2)^\wedge(\omega)\Big| \leq \sum_{\Theta \in \mathbf{CP}_\sigma} \Big|\sum_{\theta \in \mathbf{S}_\Theta} (|f_\theta|^2)^\wedge(\omega)\Big|. \tag{12}$$

Lemma 4.1. *If $\Theta \in \mathbf{CP}_\sigma$ makes a non-zero contribution to the right-hand side of (12) for an $\omega \in \Omega_\sigma$, then $\omega \in 4\Theta$.*

Proof. Suppose that $\sum_{\theta \in \mathbf{S}_\Theta} (|f_\theta|^2)^\wedge(\omega)$ is non-zero. Then we must have $\omega \in \tilde{\theta}$ for some $\theta \in \mathbf{S}_\Theta$. Suppose $\theta = \theta(\xi)$ and $\Theta = \Theta(\sigma, \xi')$. Since $\theta \in \mathbf{S}_\Theta$, we know that $|\xi - \xi'| \leq \sigma^{-1} r^{-1}$ and so $\Theta(\sigma, \xi) \subset 2\Theta$.

We claim that $\tilde{\theta} \cap \Omega_{\leq \sigma}$ is contained in $2\Theta(\sigma, \xi)$. This will finish the proof, because $\omega \in \tilde{\theta} \cap \Omega_{\leq \sigma} \subset 2\Theta(\sigma, \xi) \subset 4\Theta(\sigma, \xi')$.

To check the claim, we have to understand the geometry of the set $\Omega_{\leq \sigma}$. To picture the set $\Omega_{\leq \sigma}$, we found it helpful to consider the intersection of $\Theta(\sigma, \xi)$ with the plane $\omega_3 = h$. We assume $|h| \leq \sigma^2$ — otherwise the intersection is empty. The intersection $\Theta(\sigma, \xi) \cap \{\omega_3 = h\}$ is a rectangle with dimensions $r^{-1}\sigma \times \sqrt{2} r^{-2}$, and the long side of the rectangle is tangent to the circle of radius h around the origin at the point $h\xi$. Therefore, $\Theta(\sigma, \xi) \cap \{\omega_3 = h\}$ is contained in the annulus $\{h^2 \leq \omega_1^2 + \omega_2^2 \leq h^2 + r^{-2}\sigma^2\}$. If we rotate ξ, the rectangle $\Theta(\sigma, \xi) \cap \{\omega_3 = h\}$ rotates also, and the union of these rotated rectangles over all ξ is equal to this annulus. Therefore, if $h \leq \sigma^2$, then $\Omega_{\leq \sigma} \cap \{\omega_3 = h\}$ is approximately equal to this annulus:

$$\Omega_{\leq \sigma} \cap \{\omega_3 = h\} \sim \{\omega : \omega_3 = h, h^2 \leq \omega_1^2 + \omega_2^2 \leq h^2 + r^{-2}\sigma^2\}. \tag{13}$$

On the other hand, $\tilde{\theta}(\xi) \cap \{\omega_3 = h\} = \Theta(1, \xi) \cap \{\omega_3 = h\}$ is a rectangle of dimensions $\sim r^{-1} \times r^{-2}$ that is tangent to the circle of radius h at $h\xi$. The intersection of this rectangle with the annulus above is contained in a shorter rectangle with the same center and with dimensions $\sigma r^{-1} \times r^{-2}$, which in turn is contained in $2\Theta(\sigma, \xi) \cap \{\omega_3 = h\}$. Since this holds for every h with $|h| \leq \sigma^2$, we see that $\tilde{\theta}(\xi) \cap \Omega_{\leq \sigma} \subset 2\Theta(\sigma, \xi)$ as claimed. $\square$

Using Lemma 4.1, we can rewrite inequality (12): if $\omega \in \Omega_\sigma$, then

$$|\hat{g}(\omega)| \leq \sum_{\Theta \in \mathbf{CP}_\sigma, \omega \in 4\Theta} \Big| \sum_{\theta \in \mathbf{S}_\Theta} (|f_\theta|^2)^\wedge(\omega) \Big|. \tag{14}$$

LEMMA 4.2. *For any $\omega \in \Omega_\sigma$, the number of $\Theta \in \mathbf{CP}_\sigma$ so that $\omega \in 4\Theta$ is bounded by a constant C.*

Proof. Building on the description of $\Omega_{\leq \sigma}$ in (13) above, we see that if $|h| \leq \sigma^2/4$, then $\Omega_\sigma \cap \{\omega_3 = h\}$ is approximately given by

$$\{h^2 + (1/4) r^{-2} \sigma^2 \leq \omega_1^2 + \omega_2^2 \leq h^2 + r^{-2}\sigma^2\}. \tag{15}$$

If $\sigma^2/4 \leq |h| \leq \sigma^2$, then $\Omega_\sigma \cap \{\omega_3 = h\}$ is approximately given by

$$\{h^2 \leq \omega_1^2 + \omega_2^2 \leq h^2 + r^{-2}\sigma^2\}. \tag{16}$$

Let $C_{h,\rho}$ be the circle defined by $\omega_3 = h$ and $\omega_1^2 + \omega_2^2 = \rho^2$ with $|h| \leq \sigma^2$ and ρ chosen such that $C_{h,\rho}$ lies in (15) or (16). These circles cover Ω_σ. For any ξ, we will compute in the next two paragraphs that the fraction of $C_{h,\rho}$ contained in $4\Theta(\sigma, \xi)$ is $\lesssim \sigma^{-1} r^{-1}$. There are $\sim \sigma r$ different $\Theta(\sigma, \xi) \subset \mathbf{CP}_\sigma$. By circular symmetry, each frequency $\omega \in C_{h,\rho}$ lies in 4Θ for approximately the same number of $\Theta \in \mathbf{CP}_\sigma$, and so each frequency ω lies in 4Θ for $\leq C$ different $\Theta \in \mathbf{CP}_\sigma$.

We first do the case $|h| \leq \sigma^2/4$. Recall that $\Theta(\sigma, \xi) \cap \{\omega_3 = h\}$ is a rectangle with dimensions $r^{-1}\sigma \times r^{-2}$ that is tangent to the circle of radius $|h|$. Suppose for now that $r^{-1}\sigma \leq |h|$. If A, B are the two endpoints of this rectangle and O is the origin, then the angle AOB is approximately $r^{-1}\sigma/|h|$. The angle between the rectangle $\Theta \cap \{\omega_3 = h\}$ and the circle $C_{h,\rho}$ is approximately equal to the angle AOB. Therefore, the arc length of $4\Theta \cap C_{h,\rho}$ is bounded by

$$\text{Length } (4\Theta \cap C_{h,\rho}) \lesssim r^{-1}\sigma^{-1}|h|.$$

Since the length of $C_{h,\rho}$ is $2\pi\rho \sim |h|$, the fraction of $C_{h,\rho}$ contained in 4Θ is $\lesssim r^{-1}\sigma^{-1}$ as desired.

If $|h| < r^{-1}\sigma$, then the angle AOB is ~ 1, and the length of $4\Theta \cap C_{h,\rho}$ is approximately r^{-2}. In this case the length of $C_{h,\rho}$ is $2\pi\rho \sim r^{-1}\sigma$, and so the fraction of $C_{h,\rho}$ covered by 4Θ is still $\lesssim r^{-1}\sigma^{-1}$.

Finally, suppose that $\sigma^2/4 \leq |h| \leq \sigma^2$. In this case $4\Theta \cap C_{h,\rho}$ has arc length $\sim \sigma r^{-1}$ (the long side of the rectangle $\Theta \cap \{\omega_3 = h\}$). Since the length of $C_{h,\rho}$ is $2\pi\rho \sim |h| \sim \sigma^2$, the fraction of $C_{h,\rho}$ covered by 4Θ is again $\lesssim \sigma^{-1} r^{-1}$. $\square$

Remark. If $\omega \in \Omega_\sigma$ and $|\omega_3|$ is much smaller than σ^2, then ω lies in two rather different $\Theta \in \mathbf{CP}_\sigma$, and maybe also on other Θ neighboring these two. This is because a point outside a circle lies on two lines tangent to the circle.

Applying Cauchy–Schwarz to (14) and using Lemma 4.2 we see that if $\omega \in \Omega_\sigma$, then

$$|\hat{g}(\omega)|^2 \lesssim \sum_{\Theta \in \mathbf{CP}_\sigma, \omega \in 4\Theta} \Big| \sum_{\theta \in \mathbf{S}_\Theta} (|f_\theta|^2)^\wedge(\omega) \Big|^2. \tag{17}$$

We let η_Θ be a smooth function that is ≥ 1 on 4Θ and decays rapidly outside 4Θ. Summing over all dyadic σ, we see that for every frequency ω,

$$|\hat{g}(\omega)|^2 \lesssim \sum_{\Theta \in \mathbf{CP}} \Big| \eta_\Theta(\omega) \sum_{\theta \in \mathbf{S}_\Theta} (|f_\theta|^2)^\wedge(\omega) \Big|^2.$$

Now we integrate and use Plancherel, giving

$$\int |g|^2 \lesssim \sum_{\Theta \in \mathbf{CP}} \int \Big| \eta_\Theta^\vee * \sum_{\theta \in \mathbf{S}_\Theta} |f_\theta|^2 \Big|^2.$$

Now we can choose η_Θ so that $|\eta_\Theta^\vee(x)| \lesssim |\Theta^*|^{-1}$ for all x, and $\eta_\Theta^\vee$ is supported on Θ^*. Therefore, it is natural to break up the right integral into translated copies of Θ^*:

$$\int |g|^2 \lesssim \sum_{\Theta \in \mathbf{CP}} \sum_{U /\!/ \Theta^*} \int_U \Big| \eta_\Theta^\vee * \sum_{\theta \in \mathbf{S}_\Theta} |f_\theta|^2 \Big|^2.$$

In the last integral, for each $x \in U$, we have

$$\Big| \eta_\Theta^\vee * \sum_{\theta \in \mathbf{S}_\Theta} |f_\theta|^2(x) \Big| \lesssim |U|^{-1} \int \eta_U \sum_{\theta \in \mathbf{S}_\Theta} |f_\theta|^2,$$

where $\eta_U(z) = |\Theta^*| \cdot \max_{y \in z + \Theta^* - U} |\eta_\Theta^\vee(y)|$ is a bump function with $\|\eta_U\|_\infty \sim 1$ supported on $2U$. We remark that the arguments presented here exploit the locally constant property. We shall discuss another variant of this property in Lemma 6.1.

Therefore,

$$\int |g|^2 \lesssim \sum_{\Theta \in \mathbf{CP}} \sum_{U /\!/ \Theta^*} |U|^{-1} \left(\int \eta_U \sum_{\theta \in \mathbf{S}_\Theta} |f_\theta|^2 \right)^2.$$

We associate $\Theta(\sigma, \xi)$ to $\tau(\sigma^{-1} r^{-1}, \xi)$. This gives a bijection from $\mathbf{CP}_\sigma$ to $\mathbf{S}_s$ with $s = \sigma^{-1} r^{-1}$. If $\Theta(\sigma, \xi) \subset 2\tilde{\theta}(\xi')$, then we saw above that $|\xi - \xi'| \lesssim \sigma^{-1} r^{-1}$, and so $\theta(\xi') \subset 4\tau(\sigma^{-1} r^{-1}, \xi)$. In particular, if $\theta \in \mathbf{S}_\Theta$, then $\theta \subset 4\tau$. Also $\Theta(\sigma, \xi)^*$ is comparable to $U_{\tau(\sigma^{-1} r^{-1}, \xi), r^2}$, which we can see by comparing the definition of U_{τ,r^2} in (7) with the definition of Θ in (11). Rewriting the last inequality in terms of $\tau \in \mathbf{S}_s$ instead of $\Theta \in \mathbf{CP}_\sigma$, we get

$$\int |g|^2 \lesssim \sum_{r^{-1}\leq s\leq 1} \sum_{\tau\in\mathbf{S}_s} \sum_{U /\!/ U_{\tau,r^2}} |U|^{-1} \left(\int \eta_U \sum_{\theta\subset\tau} |f_\theta|^2 \right)^2 .$$

By the definition of $S_U f$,

$$\sum_{U /\!/ U_{\tau,r^2}} \left(\int \eta_U \sum_{\theta\subset\tau} |f_\theta|^2 \right)^2 \lesssim \sum_{U /\!/ U_{\tau,r^2}} \|S_U f\|_{L^2}^4 .$$

Plugging this in, we get

$$\int |g|^2 \lesssim \sum_{r^{-1}\leq s\leq 1} \sum_{d(\tau)=s} \sum_{U /\!/ U_{\tau,r^2}} |U|^{-1} \|S_U f\|_{L^2}^4 .$$

This proves Lemma 1.4 by taking $r = R^{\frac{1}{2}}$. □

We use this Kakeya-type estimate as well as local orthogonality to prove Lemma 3.1. First we recall local orthogonality, and then we recall the statement of Lemma 3.1.

Local orthogonality is written using a weight functions localized a given ball. For a ball B_R of radius R, define the weight

$$w_{B_R,E}(x) = \left(1 + \frac{\mathrm{dist}(x, B_R)}{R}\right)^{-E} .$$

LEMMA 4.3 (Local L^2 orthogonality lemma, essentially Proposition 6.1 in [3]). *Suppose that $f \in L^2(\mathbb{R}^n)$. Suppose that $f = \sum_\theta f_\theta$, where* $\mathrm{supp} \hat{f}_\theta \subset \theta$ *in the Fourier space. In this statement the sets θ are arbitrary. Suppose that $r > 0$ and that each $\xi \in \mathbb{R}^n$ lies in $N_{r^{-1}(\theta)}$ for at most M different sets θ appearing in the sum. Then for any* $E > 0$,

$$\|f\|_{L^2(B_r)}^2 \lesssim_{M,E} \sum_{\theta\in\mathcal{I}} \|f_\theta\|_{L^2(w_{B_r,E})}^2 .$$

To prove Lemma 4.3, it suffices to take a function ψ_{B_r} such that $\psi_{B_r} \gtrsim 1$ on B_r, $|\psi_{B_r}(x)| \leq C_E(1 + r^{-1}\mathrm{dist}(x, B_r))^{-E/2}$, and $\hat{\psi}_{B_r} \subset B(0, r^{-1})$. Then $\|f\|_{L^2(B_r)} \lesssim \|f\psi_{B_r}\|_{L^2}$. We apply Plancherel's theorem and observe that the support of $\widehat{f_\theta} * \widehat{\psi}_{B_r}$ lies in $N_{r^{-1}}(\theta)$.

Now we turn to the proof of Lemma 3.1. Unwinding the definition of $S(r, R)$, Lemma 3.1 says

LEMMA. *If $\hat{f}$ is supported on $N_{r^{-2}}(\Gamma)$ and $r_1 \in [r, r^2]$, then*

$$\sum_{B_{r_1}\subset\mathbb{R}^3} |B_{r_1}|^{-1} \|S_{B_{r_1}} f\|_{L^2(B_{r_1})}^4 \lesssim \sum_{r^{-1}\leq s\leq 1} \sum_{d(\tau)=s} \sum_{U /\!/ U_{\tau,r^2}} |U|^{-1} \|S_U f\|_{L^2}^4 . \tag{18}$$

Proof of Lemma 3.1. As in Lemma 1.4, let $g = \sum_{\theta\in \mathbf{S}_{r^{-1}}} |f_\theta|^2$. The functions f_θ have essentially disjoint Fourier support. Since $r \le r_1$, each point ξ lies in $\lesssim 1$ many $N_{r_1^{-1}}(\theta)$.

We choose E sufficiently large (for instance $E = 10$). Then we apply the local L^2 orthogonality Lemma 4.3, on each B_{r_1}:

$$\|S_{B_{r_1}} f\|^2_{L^2(B_{r_1})} = \int_{B_{r_1}} \sum_{d(\tau)=r_1^{-1/2}} |f_\tau|^2$$
$$\lesssim \int_{\mathbb{R}^3} w_{B_{r_1},E} \cdot \sum_{d(\tau)=r_1^{-1/2}} \sum_{\theta\subset\tau} |f_\theta|^2 \sim \int_{\mathbb{R}^3} w_{B_{r_1},E} \cdot g.$$

By Cauchy–Schwarz, we get

$$|B_{r_1}|^{-1} \|S_{B_{r_1}} f\|^4_{L^2(B_{r_1})} \lesssim \int_{\mathbb{R}^3} w_{B_{r_1},E/2} \cdot |g|^2.$$

Summing over B_{r_1},

$$\sum_{B_{r_1}\subset\mathbb{R}^3} |B_{r_1}|^{-1} \|S_{B_{r_1}} f\|^4_{L^2(B_{r_1})} \lesssim \int_{\mathbb{R}^3} |g|^2.$$

Lemma 1.4 bounds $\int_{\mathbb{R}^3} |g|^2$ by the right-hand side of (18). □

5. The Lorentz rescaling

Lorentz transformations are the symmetries of our problem, and they have been used in many earlier papers on this topic (cf. [26] and [2]). Here we review the Lorentz rescaling and check the properties that we will need in our rescaling argument in the next two sections.

The piece $\Gamma_{\frac{1}{K}}$ is defined to work well with Lorentz transformations, and we now record the formula. This formula and the Lorentz rescaling generally look nicest in a rotated coordinate system where the light cone is given by the equation $2\nu_1\nu_3 = \nu_2^2$. Here $\nu_2 = \xi_1$, $\nu_1 = 2^{-1/2}(\xi_3 - \xi_2)$ and $\nu_3 = 2^{-1/2}(\xi_3 + \xi_2)$. In these coordinates, if we intersect the light cone with the plane $\nu_3 = 1$, then we get the parabola $\nu_1 = (1/2)\nu_2^2$. So the light cone is actually the cone over a parabola.

Now $\Gamma_{\frac{1}{K}}$ is defined as follows:

$$\Gamma_{\frac{1}{K}} = \left\{ 2\nu_1\nu_3 = \nu_2^2, 1 - \frac{1}{K} \le \nu_3 \le 1, \left|\frac{\nu_2}{\nu_3}\right| \le 1 \right\}.$$

For any real number η with $|\eta| < 1$ and $0 < s < 1$ satisfying $-1 \le \eta \pm s \le 1$, we can define a *surface sector* $\Lambda \subseteq \Gamma_{\frac{1}{K}}$ by

$$\Lambda = \Lambda(\eta, s) = \left\{ (\nu_1, \nu_2, \nu_3) \in \Gamma_{\frac{1}{K}} : \left|\frac{\nu_2}{\nu_3} - \eta\right| < s \right\}. \tag{19}$$

Here s is the *aperture* of Λ, also denoted by $d(\Lambda)$. For each Λ, let $\eta(\Lambda)$ denote the η in (19).

Each surface sector Λ is closely associated to a sector $\tau = \tau(\Lambda)$, which is a rectangular box containing Λ with smallest comparable dimensions. The sector $\tau(\Lambda)$ is approximately the convex hull of Λ in the sense that $\frac{1}{10}\tau(\Lambda) \subset \mathrm{ConvexHull}(\Lambda) \subset 10\tau(\Lambda)$. Similarly, starting with any sector τ, there is an associated surface sector $\Lambda_\tau = \tau \cap \Gamma_{\frac{1}{K}}$. The aperture of Λ_τ and the aperture of τ are approximately the same.

For any surface sector $\Lambda \subset \Gamma_{\frac{1}{K}}$, there is a Lorentz transformation $\mathcal{L}$ that maps Λ diffeomorphically onto $\Gamma_{\frac{1}{K}}$. (The precise definition of $\Gamma_{\frac{1}{K}}$ was arranged to make this work.) The formula for $\mathcal{L}$ is as follows.

Let $\mathcal{L} : \Lambda(d(\Lambda), \eta) \to \Gamma_{\frac{1}{K}}$ be defined as follows (away from $\{z = 0\}$):

$$\begin{cases} \nu_3 & \mapsto \nu_3, \\ \frac{\nu_2}{\nu_3} & \mapsto \frac{1}{d(\Lambda)}\left(\frac{\nu_2}{\nu_3} - \eta(\Lambda)\right), \\ \frac{\nu_1}{\nu_3} & \mapsto \frac{1}{d(\Lambda)^2}\left(\frac{\nu_1}{\nu_3} - \eta(\Lambda) \cdot \frac{\nu_2}{\nu_3} + \frac{\eta(\Lambda)^2}{2}\right). \end{cases} \tag{20}$$

We can see that $\mathcal{L}$ is actually a linear transformation:

$$\begin{cases} \nu_3 & \mapsto \nu_3, \\ \nu_2 & \mapsto \frac{1}{d(\Lambda)}(\nu_2 - \eta(\Lambda)\nu_3), \\ \nu_1 & \mapsto \frac{1}{d(\Lambda)^2}\left(\nu_1 - \eta(\Lambda)\nu_2 + \frac{\eta(\Lambda)^2}{2}\nu_3\right). \end{cases} \tag{21}$$

This linear transformation $\mathcal{L}$ is called a Lorentz rescaling.

Suppose that τ is a sector with $d(\tau) = s$, and let $\Lambda = \Lambda_\tau$. We then study the rescaling map $\mathcal{L}$ defined in (21). We will need to keep track of how this change of variables affects the characters in our inequalities, like sectors $\tau' \subset \tau$ and the regions $U_{\tau,R}$.

First, if $\Lambda' \subset \Lambda$ is a smaller surface sector, then $\mathcal{L}(\Lambda')$ is a surface sector of aperture $\sim s^{-1}d(\tau')$.

More precisely, since $\Lambda' \subseteq \Lambda$, we have

$$[\eta(\Lambda') - d(\Lambda'), \eta(\Lambda') + d(\Lambda')] \subseteq [\eta(\Lambda) - d(\Lambda), \eta(\Lambda) + d(\Lambda)]. \tag{22}$$

By the above definition of $\mathcal{L}$, we can see that $\mathcal{L}(\Lambda')$ is defined as

$$\Big\{(\nu_1, \nu_2, \nu_3) \in \Gamma_{\frac{1}{K}} : \frac{\nu_2}{\nu_3} \in \Big[\frac{1}{d(\Lambda)}(\eta(\Lambda') - \eta(\Lambda)) - \frac{d(\Lambda')}{d(\Lambda)}, \frac{1}{d(\Lambda)}(\eta(\Lambda') - \eta(\Lambda)) + \frac{d(\Lambda')}{d(\Lambda)}\Big]\Big\}.$$

We see that (22) implies the above range of ν_2/ν_3 is in $[-1,1]$, and that $\mathcal{L}(\Lambda')$ is a surface sector of aperture $\frac{d(\Lambda')}{d(\Lambda)}$ lying inside the whole $\Gamma_{\frac{1}{K}} = \mathcal{L}(\Lambda)$.

Next we consider how $\mathcal{L}$ affects sectors $\tau' \subset \tau$. Suppose that $\Lambda_{\tau'}$ is a surface sector associated to τ'. Note that τ' is approximately the convex hull of $\Lambda_{\tau'}$. Since taking convex hulls commutes with linear transformations, we see that $\mathcal{L}(\tau')$ is approximately the convex hull of $\mathcal{L}(\Lambda_{\tau'})$, which is a sector of aperture $\sim s^{-1}d(\tau')$.

Next we consider $\mathcal{L}(N_{R^{-1}}(\Lambda))$ for some $R > s^{-2}$. Note that $N_{s^2}(\Lambda)$ is approximately $\tau(\Lambda)$, but if $R > s^{-2}$, then $N_{s^2}(\Lambda)$ is far from being a convex set. The R^{-1}-neighborhood of $\Gamma_{\frac{1}{K}}$ is covered by sectors $\theta \subset \tau$ with $d(\theta) = R^{-1/2}$. Therefore, $\mathcal{L}(N_{R^{-1}}(\Lambda))$ is covered by sectors $\mathcal{L}(\theta)$ with aperture $\sim s^{-1}R^{-1/2}$. The union of these sectors is the $s^{-2}R^{-1}$-neighorhood of $\Gamma_{\frac{1}{K}}$. In summary $\mathcal{L}(N_{R^{-1}}(\Lambda))$ is approximately $N_{s^{-2}R^{-1}}(\Gamma_{\frac{1}{K}})$.

Next we consider how the adjoint transformation, $\mathcal{L}^*$, behaves on physical space. It is standard that the adjoint transformation behaves naturally with respect to taking duals, so, if θ is a sector, then we have $\mathcal{L}(\theta)^* = \mathcal{L}^*(\theta^*)$.

Finally we consider how $\mathcal{L}^*$ affects the sets $U_{\tau,R}$. Recall from (7) that if $\tau = \tau(s,\xi)$, then

$$U_{\tau,R} = \{x \in \mathbb{R}^3 : |\mathbf{c}(\xi)\cdot x| \le Rs^2 \text{ and } |\mathbf{n}(\xi)\cdot x| \le R \text{ and } |\mathbf{t}(\xi)\cdot x| \le Rs\}. \tag{23}$$

There is an equivalent more conceptual description, which is useful for understanding $\mathcal{L}^*(U_{\tau,R})$:

$$U_{\tau,R} \approx \text{Convex Hull } (\cup_{\theta\subset\tau, d(\theta)=R^{-1/2}}\theta^*). \tag{24}$$

Now let τ again denote a fixed sector with $d(\tau) = s$ and let $\mathcal{L}$ be the Lorentz rescaling that takes Λ_τ to $\Gamma_{\frac{1}{K}}$.

LEMMA 5.1. *For any sector $\tau' \subset \tau$ and any $R \ge s^{-2}$,*

$$\mathcal{L}^*(U_{\tau',R}) = U_{\mathcal{L}(\tau'),s^2R}.$$

Proof.

$$\begin{aligned}\mathcal{L}^*(U_{\tau',R}) &\approx \text{ConvexHull}(\cup_{\theta\subset\tau', d(\theta)=R^{-1/2}}\mathcal{L}^*\theta^*)\\ &\approx \text{Convex Hull } (\cup_{\theta\subset\tau', d(\theta)=R^{-1/2}}\mathcal{L}(\theta)^*)\\ &\approx \text{Convex Hull } (\cup_{\theta\subset\mathcal{L}(\tau'), d(\theta)=s^{-1}R^{-1/2}}\theta^*) \approx U_{\mathcal{L}(\tau'),s^2R}.\end{aligned}$$ □

We have now gathered enough background about Lorentz rescaling to carry out our Lorentz rescaling arguments in the next two sections.

6. **The Proof of Lemma 3.2**

In this section, we prove Lemma 3.2. First we prove several lemmas about the "locally constant property" of f_θ.

LEMMA 6.1. *Let $\theta \subset \mathbb{R}^n$ be a compact convex set that is symmetric about a center point $c(\theta)$. If $\mathrm{supp}\hat{f}_\theta \subset \theta$ and $T_\theta = \theta^* = \{x : |x \cdot (y - c(\theta))| \leq 1$ for all $y \in \theta\}$, then there exists a positive function η_{T_θ} satisfying*

(1) *η_{T_θ} is essentially supported on $10T_\theta$ and rapidly decays away from it: for any integer $N \geq 0$, there exists a constant C_N such that $\eta_{T_\theta}(x) \leq C_N(n(x, 10T_\theta))^{-N}$, where $n(x, 10T_\theta)$ is the smallest positive integer n such that $x \in n \cdot 10T_\theta$;*

(2) $\|\eta_{T_\theta}\|_{L^1} \lesssim 1$;

(3) *we have*

$$|f_\theta| \leq \sum_{T /\!/ T_\theta} c_T \chi_T \leq |f_\theta| * \eta_{T_\theta}, \tag{25}$$

where c_T is defined as $\max_{x \in T} |f_\theta|(x)$ and the sum $\sum_{T /\!/ T_\theta}$ is over a finitely overlapping cover $\{T\}$ of $\mathbb{R}^n$ with each $T /\!/ T_\theta$.

Proof. We bound $|f_\theta|$ by

$$|f_\theta| \leq \sum_{T /\!/ T_\theta} c_T \chi_T. \tag{26}$$

Let ϕ_θ be a smooth bump function supported on 2θ and $\phi_\theta = 1$ on θ. Since $\mathrm{supp}\hat{f}_\theta \subset \theta$, we have $\hat{f}_\theta = \hat{f}_\theta \phi_\theta$ and $f_\theta = f_\theta * \phi_\theta^\vee$. Let $\eta_{T_\theta}(x) = \max_{t \in x + 10T_\theta} |\phi_\theta^\vee|(t)$. By non-stationary phase, $\phi_\theta^\vee$ is a function essentially supported on $T_\theta = \theta^*$, $|\phi_\theta^\vee(x)| \leq C_N(n(x, T_\theta))^{-N}$ and $\|\phi_\theta^\vee\|_{L^1} \sim 1$, so η_{T_θ} satisfies (1) and (2).

For any $T /\!/ T_\theta$,

$$\begin{aligned} \max_{x \in T} |f_\theta|(x) &\leq \max_{x \in T} \int |f_\theta|(y) |\phi_\theta^\vee(x - y)| dy \\ &\leq \min_{x \in T} \int |f_\theta|(y) \eta_{T_\theta}(x - y) dy \end{aligned}$$

because for each y, $\max_{x \in T} |\phi_\theta^\vee|(x - y) \leq \min_{x \in T} \max_{t \in x - y + 10T_\theta} |\phi_\theta^\vee|(t)$. □

LEMMA 6.2. *Let η_{T_θ} be defined as in Lemma 6.1 and $T /\!/ T_\theta$. Then for any integer $N > 0$, there exists a positive function $w_T = 1$ on $10T$ and $w_T(x) \leq C_N(1 + \mathrm{dist}(x, T))^{-N}$ such that for any $1 \leq p < \infty$,*

$$\int_T (|f_\theta| * \eta_{T_\theta})^p \lesssim_p \int |f_\theta|^p w_T. \tag{27}$$

Proof. We only need to prove the lemma for N sufficiently large (depending on p).

The function η_{T_θ} satisfies

$$\eta_{T_\theta} \le \sum_{T /\!/ T_\theta} C_T \chi_T, \tag{28}$$

where $C_T \cdot |T| \lesssim_N n(T, T_\theta)^{-N}$ for any large integer $N > 0$ and $n(T, T_\theta)$ is the smallest $n \ge 1$ such that $T \subset nT_\theta$.

By Hölder's inequality,

$$\begin{aligned}\int_T (|f_\theta| * \eta_{T_\theta})^p &\le \int_T \left(\sum_{T' /\!/ T_\theta} |f_\theta| * C_{T'} \chi_{T'} \right)^p \\ &= \int_T \left(\sum_{T' /\!/ T_\theta} n(T', T_\theta \right)^{-\frac{4(p-1)}{p}} \cdot n(T', T_\theta)^{\frac{4(p-1)}{p}} |f_\theta| * C_{T'} \chi_{T'})^p \\ &\lesssim \left(\sum_{T' /\!/ T_\theta} n(T', T_\theta)^{-4} \right)^{p-1} \cdot \sum_{T' /\!/ T_\theta} n(T', T_\theta)^{4(p-1)} \int_T (|f_\theta| * C_{T'} \chi_{T'})^p \\ &\lesssim \sum_{T' /\!/ T_\theta} n(T', T_\theta)^{4(p-1)} \int_T (|f_\theta| * C_{T'} \chi_{T'})^p.\end{aligned}$$

Let $\chi_{T-T'}(x)$ be the characteristic function of the Minkowski sum $T - T' = T + (-T')$. Then by Young's inequality,

$$\begin{aligned}\int_T (|f_\theta| * C_{T'} \chi_{T'})^p &\le \int ((|f_\theta| \chi_{T-T'}) * (C_{T'} \chi_{T'}))^p \\ &\lesssim_N n(T', T_\theta)^{-pN} \cdot \int_{T-T'} |f_\theta|^p\end{aligned}$$

It suffices to choose $w_T(x) \sim_N \sum_{\tilde{T} /\!/ T} n(\tilde{T}, T)^{-N} \chi_{\tilde{T}}(x)$. □

COROLLARY 6.3. *If U is tiled by $T /\!/ T_\theta$, then for any $1 \le p < \infty$,*

$$\int_U (|f_\theta| * \eta_{T_\theta})^p \lesssim_p \int |f_\theta|^p w_U, \tag{29}$$

where $w_U \ge 0$ is essentially supported on $10U$ and rapidly decays away from it.

Remark. It is important that w_U can be taken uniformly independent of the choice of T. To see this, simply notice that if $x \in nU$ and $x \notin (n-1)U$, then x cannot be in $(n-1)T$ for any $T \subset U$. Moreover for any m, a point x lies in mT for $\lesssim m^3$ different T in a given tiling $\{T\}_{T /\!/ T_\theta}$ of $\mathbb{R}^3$.

LEMMA 6.4. *Let* $\theta_1, \theta_2 \subset \tau$ *be two sectors of aperture* $d(\theta_1) = d(\theta_2) = K^{-1/2}$, *and* $\mathrm{dist}(\theta_1, \theta_2) \sim d(\tau) = s > K^{-1/2}$. *Then for any functions* $\mathrm{supp}\hat{f}_{\theta_1} \subset N_{\frac{1}{K}}\Gamma_{\frac{1}{K}} \cap \theta_1$ *and* $\mathrm{supp}\hat{f}_{\theta_2} \subset N_{\frac{1}{K}}\Gamma_{\frac{1}{K}} \cap \theta_2$,

$$\sum_{B_{K^{1/2}} \subset \mathbb{R}^3} \int_{B_{K^{1/2}}} |f_{\theta_1} f_{\theta_2}|^2 \lesssim s^{-1} \sum_{B_K \subset \mathbb{R}^3} |B_K|^{-1} \int |f_{\theta_1}|^2 w_{B_K} \int |f_{\theta_2}|^2 w_{B_K}.$$

Proof. The proof is essentially a bilinear-Kakeya-style[8] estimate in $\mathbb{R}^2$ plus the locally constant property in Lemma 6.1. This proof is a simple case of the ball inflation theorem (Theorem 9.2 in [3]) in the proof of the Bourgain–Demeter decoupling theorem. Since $\mathrm{supp}\hat{f}_{\theta_j} \subset N_{\frac{1}{K}}\Gamma_{\frac{1}{K}} \cap \theta_j$ for $j = 1, 2$, the Fourier support of f_{θ_j} lies inside a box $\tilde{\theta}_j$ of dimensions $K^{-1/2} \times K^{-1} \times K^{-1}$ with a common K^{-1}-side on the ν_3-direction. (Recall the (ν_1, ν_2, ν_3)-coordinate system and the equation of $\Gamma_{\frac{1}{K}}$ from Section 5.) And $T_{\tilde{\theta}_j} = \tilde{\theta}_j^*$ becomes a slab of dimensions $K^{1/2} \times K \times K$. Since $\mathrm{dist}(\tilde{\theta_1}, \tilde{\theta_2}) = \mathrm{dist}(\theta_1, \theta_2) = s$, for each $T_1 /\!/ T_{\tilde{\theta}_1}$, $T_2 /\!/ T_{\tilde{\theta}_2}$ and $T_1, T_2 \subset B_K$, we have $|T_1 \cap T_2| \sim K^{1/2} \cdot (s^{-1} K^{1/2}) \cdot K = s^{-1} K^2$. Hence the key inequality $|T_1 \cap T_2| \sim s^{-1} |B_K|^{-1} |T_1||T_2|$ holds.[9]

Using Lemma 6.1, now we are ready to bound

$$\begin{aligned}
&\sum_{B_{K^{1/2}} \subset B_K} \int_{B_{K^{1/2}}} |f_{\theta_1} f_{\theta_2}|^2 \\
&\quad \le \sum_{\substack{B_{K^{1/2}} \subset B_K \\ T_1 /\!/ T_{\tilde{\theta}_1}, B_{K^{1/2}} \cap T_1 \neq \emptyset \\ T_2 /\!/ T_{\tilde{\theta}_2}, B_{K^{1/2}} \cap T_2 \neq \emptyset}} |B_{K^{1/2}}| c_{T_1}^2 c_{T_2}^2 \\
&\quad \lesssim s^{-1} |B_K|^{-1} \Big(\int_{B_K} \sum_{T_1 /\!/ T_{\tilde{\theta}_1}} c_{T_1}^2 \chi_{T_1} \Big) \Big(\int_{B_K} \sum_{T_2 /\!/ T_{\tilde{\theta}_2}} c_{T_2}^2 \chi_{T_2} \Big) \\
&\quad \le s^{-1} |B_K|^{-1} \int_{B_K} (|f_{\theta_1}| * \eta_{T_{\tilde{\theta}_1}})^2 \int_{B_K} (|f_{\theta_2}| * \eta_{T_{\tilde{\theta}_2}})^2 \\
\text{(Corollary 6.3)} &\quad \lesssim s^{-1} |B_K|^{-1} \int |f_{\theta_1}|^2 w_{B_K} \int |f_{\theta_2}|^2 w_{B_K}.
\end{aligned}$$

□

[8]Bilinear Kakeya is an elementary statement stating the following: Let $|\mathbb{T}_1|$ and $|\mathbb{T}_2|$ be two finite families of infinite strips in $\mathbb{R}^2$ such that each strip has width 1. Assume further that each $T_1 \in \mathbb{T}_1$ and each $T_2 \in \mathbb{T}_2$ have their directions $\sim$ 1-separated. Then $\int_{\mathbb{R}^2} (\sum_{T_1 \in \mathbb{T}_1} \chi_{T_1}) \cdot (\sum_{T_2 \in \mathbb{T}_2} \chi_{T_2}) \lesssim |\mathbb{T}_1| \cdot |\mathbb{T}_2|$.

[9]*Note*: All arguments in this paper work if we dilate a convex body by a constant. If we replace B_K by the slightly bigger B_{10K}, then it is possible for T_1 and T_2 to miss each other, hence we can only obtain "$\lesssim$"instead of the above "$\sim$." However we only use "$\lesssim$" in the inequality below so "$\lesssim$" is sufficient.

LEMMA 6.5. *Let f be a function whose Fourier transform is supported on the $\frac{1}{K}$-neighborhood of $\Gamma_{\frac{1}{K}}$. For any $\delta > 0$,*

$$\|f\|_{L^4(\mathbb{R}^3)}^4 \leq C_\delta K^\delta \sum_{K^{-1/2}\leq s\leq 1} \sum_{d(\tau)=s} \sum_{U/\!/U_{\tau,K}} |U|^{-1} \|S_U f\|_{L^2}^4. \tag{30}$$

Proof. We induct on K. The base case $K \lesssim_\delta 1$ is easy by Hölder's inequality.

Let $1 \ll K_0 \ll K^{\delta/10}$. We tile $N_{\frac{1}{K}}(\Gamma_{\frac{1}{K}})$ with sectors τ of aperture $\frac{1}{K_0}$ and width $\frac{1}{K}$, and we decompose $f = \sum_{d(\tau)=\frac{1}{K_0}} f_\tau$.

Now $N_{\frac{1}{K}}(\Gamma_{\frac{1}{K}})$ is the $\frac{1}{K}$-neighborhood of an arc of a parabola of length 1, and each τ is the $\frac{1}{K}$-neighborhood of an arc of the parabola of length $\frac{1}{K_0}$.

The Bourgain–Guth argument [4] says the following. At each point, $f(x) = \sum_\tau f_\tau(x)$. Let τ^* satisfy $\max_\tau |f_\tau|(x) = |f_{\tau^*}|(x)$. If $|f_{\tau^*}|(x) \geq 1/10|f|(x)$, then $|f|^4(x) \lesssim \sum_\tau |f_\tau|^4(x)$. Otherwise, there exists a τ^{**} such that $\mathrm{dist}(\tau^{**}, \tau^*) \geq 1/K_0$ and $|f_{\tau^*}|(x) \geq |f_{\tau^{**}}|(x) \geq \frac{1}{2K_0}|f|(x)$. Hence,

$$|f|^4 \lesssim \sum_{d(\tau)=1/K_0} |f_\tau|^4 + K_0^4 \sum_{\mathrm{dist}(\tau_1,\tau_2)\geq 1/K_0} |f_{\tau_1} f_{\tau_2}|^2.$$

For the integral of the first term, we rescale τ to be the $K^{-1}K_0^2$-neighborhood of $\Gamma_{1/K}$ (the rescaling argument here is similar to the one in the proof of Lemma 3.3 in Section 7, which we will do with full details), and then we apply the induction hypothesis on the scale $K/K_0^2 < K$.

For the integral of the second term, we decompose

$$f_{\tau_j} = \sum_{\theta_j \subset \tau_j, d(\theta_j)=K^{-1/2}} f_{\theta_j}, \quad j = 1, 2.$$

The functions $f_{\theta_1} f_{\theta_2}$ are essentially orthogonal because they have almost disjoint Fourier support, as in the Fefferman–Córdoba proof of restriction for the parabola [10], [7].

Since $\mathrm{dist}(\tau_1, \tau_2)$ is not less than $\frac{1}{K_0}$, the Minkowski sum $(\theta_1+\theta_2)\cap(\theta_1'+\theta_2')$ is empty for $\theta_j, \theta_j' \subset \tau_j$, $j = 1, 2$, unless $\theta_1' \subset K_0\theta_1$ and $\theta_2' \subset K_0\theta_2$. Hence

$$\sum_{B_{K^{1/2}}\subset\mathbb{R}^3} \int_{B_{K^{1/2}}} |f_{\tau_1} f_{\tau_2}|^2 \leq K_0^2 \sum_{B_{K^{1/2}}\subset\mathbb{R}^3} \sum_{\mathrm{dist}(\theta_1,\theta_2)\geq 1/K_0} \int_{B_{K^{1/2}}} |f_{\theta_1} f_{\theta_2}|^2,$$

$$\text{(Lemma 6.4)} \lesssim K_0^3 \sum_{B_K\subset\mathbb{R}^3} |B_K|^{-1} \sum_{\mathrm{dist}(\theta_1,\theta_2)\geq 1/K_0} \int |f_{\theta_1}|^2 w_{B_K} \int |f_{\theta_2}|^2 w_{B_K}$$

$$\lesssim K_0^3 \sum_{B_K\subset\mathbb{R}^3} |B_K|^{-1} \|S_{B_K} f\|_{L^2}^4. \qquad \square$$

The right-hand side of the final line corresponds to the $s = 1$ term of the right-hand side of (30).

We recall the statement of Lemma 3.2. Unwinding the definition of $S_K(r, K)$ it says the following:

PROPOSITION 6.6. *Let f be a function whose Fourier transform is supported on the $\frac{1}{K}$-neighborhood of $\Gamma_{\frac{1}{K}}$. For any $\delta > 0$ and any $r \le K$,*

$$\sum_{B_r \subset \mathbb{R}^3} |B_r|^{-1} \|S_{B_r} f\|^4_{L^2(B_r)} \le C_\delta K^\delta \sum_{K^{-1/2} \le s \le 1} \sum_{d(\tau)=s} \sum_{U /\!/ U_{\tau,K}} |U|^{-1} \|S_U f\|^4_{L^2}. \tag{31}$$

Proof. We take advantage that $\Gamma_{\frac{1}{K}}$ is well-approximated by a parabola at the scale $1/K$ and use an approach similar to Fefferman–Córdoba's to bound the left-hand side of (31) by (essentially) the left-hand side of (30).[10]

Since the smallest aperture in this proposition is $K^{-1/2}$, we use θ to denote a sector on $\Gamma_{\frac{1}{K}}$ of aperture $K^{-1/2}$ in the current proof.

Let $A_1, \ldots, A_{1000}$ be disjoint sets of θ such that each θ is in one of them and the following property holds:

(∗) Within each A_j, if the Minkowski sum $(\theta_1 + \theta_2) \cap (\theta_1' + \theta_2') \ne \emptyset$, then $(\theta_1, \theta_2) = (\theta_1', \theta_2')$ or (θ_2', θ_1').

Similar to Fefferman–Córdoba's proof, we show that if we take each A_j to be a collection of sectors that are sufficiently separated and on a short enough arc, then (∗) holds. In fact, it suffices to justify (∗) when the constraint $(\theta_1 + \theta_2) \cap (\theta_1' + \theta_2') \ne \emptyset$ is replaced by the weaker one below: $\pi_3((\theta_1 + \theta_2)) \cap \pi_3((\theta_1' + \theta_2')) \ne \emptyset$. Here π_3 is the standard projection to the first two coordinates in the (ν_1, ν_2, ν_3)-coordinate system. But the projection of $\Gamma_{\frac{1}{K}}$ onto the first two coordinates is contained in the $\frac{2}{K}$-neighborhood of the parabola $\nu_2^2 = 2\nu_1$, and the projection of each θ is the corresponding cap inside that neighborhood. We use "Error" to denote a number (the "error term") whose absolute value is $\le 4K^{-1}$. If $x_1 + x_2 = a +$ Error and $x_1^2 + x_2^2 = b +$ Error with $a, b \le 2$, then $(x_1 - x_2)^2 = 2b - a^2 + 7\,\text{Error}$. Hence $|x_1 - x_2| = \sqrt{|2b - a^2|} + 3\sqrt{\text{Error}}$. This would imply that the pair (x_1, x_2) is determined by the pair (a, b), up to a swap in order and up to changing within 100 adjacent caps θ.

We use τ to denote caps with aperture $r^{-1/2} \ge K^{-1/2}$ in the current proof. Consider the decomposition $f_j = \sum_{\theta \in A_j} f_\theta$, and let $f_{j,\tau} = \sum_{\theta \subset \tau, \theta \in A_j} f_\theta$.

[10] Alternatively, one can blackbox the L^4 angular square function estimate by Córdoba [8] and have a slightly shorter proof. We present a self-contained proof here.

By the property $(*)$ and Plancherel, we have for a fixed j,

$$
\begin{aligned}
\int_{\mathbb{R}^3} |f_j|^4 &= \int_{\mathbb{R}^3} |\sum_{\tau} f_{j,\tau}|^4 \\
&= \int_{\mathbb{R}^3} \sum_{\tau_1,\tau_2,\tau_3,\tau_4:(\mathrm{supp} f_{j,\tau_1}+\mathrm{supp} f_{j,\tau_2})\cap(\mathrm{supp} f_{j,\tau_3}+\mathrm{supp} f_{j,\tau_4})\neq\emptyset} f_{j,\tau_1} f_{j,\tau_2} \bar{f}_{j,\tau_3} \bar{f}_{j,\tau_4} \\
&= \int_{\mathbb{R}^3} \sum_{\tau_1,\tau_2} n_{\tau_1,\tau_2} |f_{j,\tau_1} f_{j,\tau_2}|^2 \\
&\sim \int_{\mathbb{R}^3} \left(\sum_{\tau} |f_{j,\tau}|^2\right)^2,
\end{aligned}
\tag{32}
$$

where $n_{\tau_1,\tau_2} = 1$ if $\tau_1 = \tau_2$ and $n_{\tau_1,\tau_2} = 4/2 = 2$ if $\tau_1 \neq \tau_2$.

By (32) we have

$$
\begin{aligned}
\sum_{B_r\subset\mathbb{R}^3} |B_r|^{-1} \|S_{B_r} f\|^4_{L^2(B_r)} &\lesssim \sum_{j=1}^{1000} \sum_{B_r\subset\mathbb{R}^3} |B_r|^{-1} \|S_{B_r} f_j\|^4_{L^2(B_r)} \\
&\leq \sum_{j=1}^{1000} \sum_{B_r\subset\mathbb{R}^3} \|S_{B_r} f_j\|^4_{L^4(B_r)} \\
&= \sum_{j=1}^{1000} \int_{\mathbb{R}^3} \left(\sum_{\tau} |f_{j,\tau}|^2\right)^2 \\
&\sim \sum_{j=1}^{1000} \int_{\mathbb{R}^3} |f_j|^4 \\
(\text{Lemma 6.5}) &\leq C_\delta K^\delta \sum_{j=1}^{1000} \sum_{K^{-1/2}\leq s\leq 1} \sum_{d(\tau)=s} \sum_{U /\!/ U_{\tau,K}} |U|^{-1} \|S_U f_j\|^4_{L^2} \\
&\lesssim C_\delta K^\delta \sum_{K^{-1/2}\leq s\leq 1} \sum_{d(\tau)=s} \sum_{U /\!/ U_{\tau,K}} |U|^{-1} \|S_U f\|^4_{L^2}. \qquad \square
\end{aligned}
$$

7. The proof of Lemma 3.3

Now we prove Lemma 3.3 using the Lorentz rescaling. First we recall the statement following statement:

LEMMA. *For any* $r_1 < r_2 \leq r_3$,

$$S_K(r_1, r_3) \leq \log r_2 \cdot S_K(r_1, r_2) \max_{r_2^{-1/2}\leq s\leq 1} S_K(s^2 r_2, s^2 r_3).$$

Proof. Suppose that $\hat{f}$ is supported on $N_{r_3^{-1}}(\Gamma_{\frac{1}{K}})$. To bound $S_K(r_1, r_3)$, we need to bound

$$\sum_{B_{r_1}\subset\mathbb{R}^3} |B_{r_1}|^{-1}\|S_{B_{r_1}}f\|^4_{L^2(B_{r_1})}.$$

We can apply the definition of $S_K(r_1, r_2)$ and get

$$\sum_{B_{r_1}\subset\mathbb{R}^3} |B_{r_1}|^{-1}\|S_{B_{r_1}}f\|^4_{L^2(B_{r_1})} \leq S_K(r_1, r_2) \sum_{r_2^{-1/2}\leq s\leq 1} \sum_{d(\tau)=s} \sum_{U_1/\!/U_{\tau,r_2}} |U_1|^{-1}\|S_{U_1}f\|^4_{L^2(U_1)}.$$

Recall that if $U/\!/U_{\tau,r}$, then $S_U f = (\sum_{d(\theta')=r^{-1/2},\theta'\subset\tau} |f_{\theta'}|^2)^{\frac{1}{2}}|_U$. In particular, $S_{B_r}f = (\sum_{d(\theta')=r^{-1/2}} |f_{\theta'}|^2)^{\frac{1}{2}}|_{B_r}$.

Using Lorentz rescaling, we will prove the following lemma:

Lemma 7.1. *For any sector τ with $d(\tau) = s$,*

$$\sum_{U_1/\!/U_{\tau,r_2}} |U_1|^{-1}\|S_{U_1}f\|^4_{L^2(U_1)} \leq S_K(s^2r_2, s^2r_3) \sum_{r_3^{-1/2}\leq s'\leq s} \sum_{d(\tau')=s',\tau'\subset\tau} \sum_{U/\!/U_{\tau',r_3}} |U|^{-1}\|S_U f\|^4_{L^2(U)}. \tag{33}$$

We defer the proof of Lemma 7.1 to the end of this section. If we plug in Lemma 7.1 and expand everything, then we get Lemma 3.3:

$$\sum_{B_{r_1}\subset\mathbb{R}^3} |B_{r_1}|^{-1}\|S_{B_{r_1}}f\|^4_{L^2(B_{r_1})} \leq \log r_2 S_K(r_1, r_2) \max_{r_2^{-1/2}\leq s\leq 1} S_K(s^2r_2, s^2r_3) \times \sum_{r_3^{-1/2}\leq s'\leq 1} \sum_{d(\tau')=s'} \sum_{U/\!/U_{\tau',r_3}} |U|^{-1}\|S_U f\|^4_{L^2(U)}.$$

The factor $\log r_2$ appears here for the following reason: after we expand, each sector τ' will appear at most $\log r_2$ times, because τ' lies in τ for at most $\log r_2$ sectors τ with $r_2^{-1/2} \leq d(\tau) \leq 1$. □

Proof of Lemma 7.1. The definition of $S_K(s^2r_2, s^2r_3)$ says that if $\hat{h}$ is supported on $N_{s^{-2}r_3^{-1}}(\Gamma_{\frac{1}{K}})$, then

$$\sum_{B_{s^2r_2}} |B_{s^2r_2}|^{-1}\|S_{B_{s^2r_2}}h\|^4_{L^2(B_{s^2r_2})} \leq S_K(s^2r_2, s^2r_3) \sum_{s^{-1}r_2^{-1/2}\leq d(\tau'')\leq 1} \sum_{U''/\!/U_{\tau'',s^2r_3}} |U''|^{-1}\|S_{U''}h\|^4_{L^2(U'')}. \tag{34}$$

On the other hand, Lemma 7.1 says that if τ is a sector of $\Gamma_{\frac{1}{K}}$ with $d(\tau) = s$, and $\hat{f}_\tau$ is supported on $N_{r_3^{-1}}(\Gamma_{\frac{1}{K}}) \cap \tau$, then

$$\begin{aligned}(35) \qquad &\sum_{U_1 /\!/ U_{\tau,r_2}} |U_1|^{-1} \|S_{U_1} f\|^4_{L^2(U_1)} \\ &\leq S_K(s^2 r_2, s^2 r_3) \sum_{r_3^{-1/2} \leq s' \leq s} \sum_{d(\tau')=s', \tau' \subset \tau} \sum_{U /\!/ U_{\tau',r_3}} |U|^{-1} \|S_U f\|^4_{L^2}.\end{aligned}$$

To connect them, we begin with a Lorentz transformation $\mathcal{L}$ so that $\mathcal{L} : \tau \cap \Gamma_{\frac{1}{K}} \to \Gamma_{\frac{1}{K}}$ is a diffeomorphism. This $\mathcal{L}$ is constructed in Section 5, where it is shown that $\mathcal{L}$ takes $N_{r_3^{-1}}(\Gamma_{\frac{1}{K}}) \cap \tau$ to $N_{s^{-2} r_3^{-1}}(\Gamma_{\frac{1}{K}})$. Now we define h by $\hat{h} = \hat{f}_\tau(\mathcal{L}^{-1}(\cdot))$. Moreover let $\hat{h}_{\tau''} = \hat{f}_{\tau'}(\mathcal{L}^{-1}(\cdot))$ where $\mathcal{L}(\tau') = \tau''$; see item (1) below. We see that $\hat{h}$ is supported on $N_{s^{-2} r_3^{-1}}(\Gamma_{\frac{1}{K}})$ and so h obeys (34). When we unwind the Lorentz transformations, we claim that (34) becomes (35), which proves the lemma. To see that this unwinding works as desired, we check how each piece transforms.

(1) If $\tau' \subset \tau$ is a sector of $\Gamma_{\frac{1}{K}}$ with aperture $d(\tau')$, then $\mathcal{L}(\tau')$ is a sector τ'' of $\Gamma_{\frac{1}{K}}$ with $d(\tau'') = s^{-1} d(\tau')$, as we showed in Section 5. In particular, $\mathcal{L}$ transforms a $\theta' \subset \tau$ with aperture $d(\theta') = r_3^{-1/2}$ into a sector with aperture $s^{-1} r_3^{-1/2}$, which appears in the definition of $S_{U''} h$.
(2) $\mathcal{L}^*(U_{\tau',r_3}) = U_{\tau'',s^2 r_3}$. Since $\tau'' = \mathcal{L}(\tau')$, this follows from Lemma 5.1.
(3) $\mathcal{L}^*(U_{\tau,r_2}) = B_{s^2 r_2}$. Note that $\mathcal{L}(\tau)$ is the sector corresponding to all of $\Gamma_{\frac{1}{K}}$, which is essentially the unit ball. We will denote this sector just by B_1. By Lemma 5.1, $\mathcal{L}^*(U_{\tau,r_2}) = U_{B_1,s^2 r_2}$. By definition, the right-hand side is the convex hull of the union of θ^* over all sectors θ of aperture $\sim s^{-1} r_2^{-\frac{1}{2}}$, and this is approximately the ball of radius $s^2 r_2$.
(4) The Jacobian factors from the change of variables work out the same on the left-hand side and the right-hand side. Since both sides involve a volume to the power -1 times an L^2 norm to the power 4, the Jacobian factors are the same on both sides of the inequality. □

References

[1] J. Bourgain, Averages in the plane over convex curves and maximal operators, *J. Analyse Math.* **47** (1986), 69–85. MR 0874045. Zbl 0626.42012. https://doi.org/10.1007/BF02792533.
[2] J. Bourgain and C. Demeter, The proof of the l^2 decoupling conjecture, *Ann. of Math.* (2) **182** no. 1 (2015), 351–389. MR 3374964. Zbl 1322.42014. https://doi.org/10.4007/annals.2015.182.1.9.

[3] J. Bourgain and C. Demeter, A study guide for the l^2 decoupling theorem, *Chin. Ann. Math. Ser. B* **38** no. 1 (2017), 173–200. MR 3592159. Zbl 1370.42021. https://doi.org/10.1007/s11401-016-1066-1.

[4] J. Bourgain and L. Guth, Bounds on oscillatory integral operators based on multilinear estimates, *Geom. Funct. Anal.* **21** no. 6 (2011), 1239–1295. MR 2860188. Zbl 1237.42010. https://doi.org/10.1007/s00039-011-0140-9.

[5] J. Bourgain and N. Watt, Mean square of zeta function, circle problem and divisor problem revisited, 2017. arXiv 1709.04340.

[6] A. Carbery and A. Seeger, Homogeneous Fourier multipliers of Marcinkiewicz type, *Ark. Mat.* **33** no. 1 (1995), 45–80. MR 1340270. Zbl 0832.42006. https://doi.org/10.1007/BF02559605.

[7] A. Córdoba, A note on Bochner-Riesz operators, *Duke Math. J.* **46** no. 3 (1979), 505–511. MR 0544242. Zbl 0438.42013. https://doi.org/10.1215/S0012-7094-79-04625-8.

[8] A. Córdoba, Geometric Fourier analysis, *Ann. Inst. Fourier* (*Grenoble*) **32** no. 3 (1982), vii, 215–226. MR 0688026. https://doi.org/10.5802/aif.885.

[9] C. Demeter, L. Guth, and H. Wang, Small cap decouplings, 2019. arXiv 1908.09166.

[10] C. Fefferman, A note on spherical summation multipliers, *Israel J. Math.* **15** (1973), 44–52. MR 0320624. Zbl 0262.42007. https://doi.org/10.1007/BF02771772.

[11] G. Garrigós, W. Schlag, and A. Seeger, Improvements in Wolff's inequality for decompositions of cone multipliers, unpublished preprint, available on the website for Gustavo Garrigós.

[12] G. Garrigós and A. Seeger, On plate decompositions of cone multipliers, *Proc. Edinb. Math. Soc.* (2) **52** no. 3 (2009), 631–651. MR 2546636. Zbl 1196.42010. https://doi.org/10.1017/S001309150700048X.

[13] G. Garrigós and A. Seeger, A mixed norm variant of Wolff's inequality for paraboloids, in *Harmonic Analysis and Partial Differential Equations*, *Contemp. Math.* **505**, Amer. Math. Soc., Providence, RI, 2010, pp. 179–197. MR 2664568. Zbl 1202.42021. https://doi.org/10.1090/conm/505/09923.

[14] L. Guth, N. Solomon, and H. Wang, Incidence estimates for well spaced tubes, *Geom. Funct. Anal.* **29** no. 6 (2019), 1844–1863. MR 4034922. Zbl 07136964. https://doi.org/10.1007/s00039-019-00519-y.

[15] Y. Heo, F. Nazarov, and A. Seeger, Radial Fourier multipliers in high dimensions, *Acta Math.* **206** no. 1 (2011), 55–92. MR 2784663. Zbl 1219.42006. https://doi.org/10.1007/s11511-011-0059-x.

[16] I. Łaba and T. Wolff, A local smoothing estimate in higher dimensions, *J. Anal. Math.* **88** (2002), 149–171, dedicated to the memory of Tom Wolff. MR 1956533. Zbl 1042.42006. https://doi.org/10.1007/BF02786576.

[17] J. Lee, A trilinear approach to square function and local smoothing estimates for the wave operator, 2016. arXiv 1607.08426.

[18] S. Lee and A. Vargas, On the cone multiplier in $\mathbb{R}^3$, *J. Funct. Anal.* **263** no. 4 (2012), 925–940. MR 2927399. Zbl 1252.42014. https://doi.org/10.1016/j.jfa.2012.05.010.

[19] G. Mockenhaupt, A note on the cone multiplier, *Proc. Amer. Math. Soc.* **117** no. 1 (1993), 145–152. MR 1098404. Zbl 0771.42006. https://doi.org/10.2307/2159709.

[20] G. Mockenhaupt, A. Seeger, and C. D. Sogge, Wave front sets, local smoothing and Bourgain's circular maximal theorem, *Ann. of Math.* (2) **136** no. 1 (1992), 207–218. MR 1173929. Zbl 0759.42016. https://doi.org/10.2307/2946549.

[21] T. Orponen, On the dimension and smoothness of radial projections, *Anal. PDE* **12** no. 5 (2019), 1273–1294. MR 3892404. Zbl 1405.28011. https://doi.org/10.2140/apde.2019.12.1273.

[22] M. Pramanik and A. Seeger, L^p regularity of averages over curves and bounds for associated maximal operators, *Amer. J. Math.* **129** no. 1 (2007), 61–103. MR 2288738. Zbl 1161.42009. https://doi.org/10.1353/ajm.2007.0003.

[23] C. D. Sogge, Propagation of singularities and maximal functions in the plane, *Invent. Math.* **104** no. 2 (1991), 349–376. MR 1098614. Zbl 0754.35004. https://doi.org/10.1007/BF01245080.

[24] T. Tao and A. Vargas, A bilinear approach to cone multipliers. II. Applications, *Geom. Funct. Anal.* **10** no. 1 (2000), 216–258. MR 1748921. Zbl 0949.42013. https://doi.org/10.1007/s000390050007.

[25] L. A. Vinh, The Szemerédi-Trotter type theorem and the sum-product estimate in finite fields, *European J. Combin.* **32** no. 8 (2011), 1177–1181. MR 2838005. Zbl 1253.11015. https://doi.org/10.1016/j.ejc.2011.06.008.

[26] T. Wolff, Local smoothing type estimates on L^p for large p, *Geom. Funct. Anal.* **10** no. 5 (2000), 1237–1288. MR 1800068. Zbl 0972.42005. https://doi.org/10.1007/PL00001652.

[27] T. Wolff, A sharp bilinear cone restriction estimate, *Ann. of Math.* (2) **153** no. 3 (2001), 661–698. MR 1836285. Zbl 1125.42302. https://doi.org/10.2307/2661365.

(Received: October 9, 2019)
(Revised: June 22, 2020)

Massachusetts Institute of Technology, Cambridge, MA
E-mail: lguth@math.mit.edu

Institute for Advanced Study, Princeton, NJ
E-mail: hongwang@math.ias.edu

University of Wisconsin-Madison, Madison, WI
E-mail: rzhang347@wisc.edu