SHARP L^p BOUNDS FOR THE HELICAL MAXIMAL FUNCTION

DAVID BELTRAN SHAOMING GUO JONATHAN HICKMAN ANDREAS SEEGER

Abstract. We establish the $L^p(\mathbb{R}^3)$ boundedness of the helical maximal function for the sharp range $p > 3$. Our results improve the previous known bounds for $p > 4$. The key ingredient is a new microlocal smoothing estimate for averages along dilates of the helix, which is established via a square function analysis.

1. Introduction

1.1. **Main results.** For $n \geqslant 2$ let $\gamma\colon I \to \mathbb{R}^n$ be a smooth curve, where $I \subset \mathbb{R}$ is a compact interval, and $\chi \in C^\infty(\mathbb{R})$ be a bump function supported on the interior of I. Given $t > 0$, consider the averaging operator

$$A_t f(x) := \int_{\mathbb{R}} f(x - t\gamma(s))\,\chi(s)\,\mathrm{d}s$$

and define the associated maximal function

$$M_\gamma f(x) := \sup_{t>0} |A_t f(x)|.$$

We are interested in the L^p mapping properties of M_γ. It is well-known that the range of exponents p for which M_γ is bounded on L^p depends on the curvature of the underlying curve. Accordingly, we consider smooth curves $\gamma\colon I \to \mathbb{R}^n$ which are *non-degenerate*, in the sense that there is a constant $c_0 > 0$ such that

$$|\det(\gamma'(s), \cdots, \gamma^{(n)}(s))| \geqslant c_0 \qquad \text{for all } s \in I. \tag{1.1}$$

A celebrated theorem of Bourgain [7, 6] states that if $\gamma\colon I \to \mathbb{R}^2$ is a smooth, non-degenerate plane curve, then M_γ is bounded on $L^p(\mathbb{R}^2)$ if and only if $p > 2$. Here we establish a 3-dimensional variant of this result.

Theorem 1.1. *If $\gamma : I \to \mathbb{R}^3$ is a smooth, non-degenerate space curve, then M_γ is bounded on $L^p(\mathbb{R}^3)$ if and only if $p > 3$.*

In the $n = 3$ case, the condition (1.1) is equivalent to the non-vanishing of the curvature and torsion functions. As a concrete example, Theorem 1.1 implies that the *helical maximal operator*

$$M_{\mathrm{Helix}} f(x) := \sup_{t>0} \Big| \int_0^{2\pi} f(x_1 - t\cos\theta, x_2 - t\sin\theta, x_3 - t\theta)\,\mathrm{d}\theta \Big|$$

is bounded on $L^p(\mathbb{R}^3)$ for all $p > 3$.

A simple Knapp-type example shows L^p boundedness fails for $p \leqslant 3$ (see §12). On the other hand, Pramanik and the fourth author [19] proved that Wolff's decoupling inequality [24] for the light cone implies the boundedness of M_γ for a suitable range of p. The optimal range for Wolff's inequality was obtained by Bourgain and Demeter [8] and the combination of the results in [19] and [8] yields the L^p boundedness of M_γ for the partial range $4 < p \leqslant \infty$. Thus, Theorem 1.1 closes the gap by establishing boundedness for the remaining exponents $3 < p \leqslant 4$.

To prove Theorem 1.1, we follow the basic strategy introduced by Mockenhaupt, Sogge and the fourth author [18] in the context of the classical circular maximal function in the plane. In particular, in [18] the authors gave an alternative proof of Bourgain's maximal theorem, deriving it as a consequence of certain *local smoothing* estimates for the wave propagator. In the case of maximal functions associated to space curves, Theorem 1.1 follows from a local smoothing estimate for a class of Fourier integral operators associated to the averages A_t. To give a simple statement of the key underlying inequality, set $\mathfrak{A}_\gamma f(x,t) := \rho(t)\cdot A_t f(x)$ for some $\rho \in C_c^\infty(\mathbb{R})$ with $\operatorname{supp}\rho \subseteq [1,2]$. Our main theorem then reads as follows.

Theorem 1.2. *Suppose $\gamma : I \to \mathbb{R}^3$ is a smooth, non-degenerate space curve and let $3 \leqslant p \leqslant 4$ and $\sigma < \sigma(p)$ where $\sigma(p) := \frac{1}{5}\big(1+\frac{2}{p}\big)$. Then $\mathfrak{A}_\gamma$ maps $L^p(\mathbb{R}^3)$ boundedly into $L^p_\sigma(\mathbb{R}^4)$.*

Note that $\sigma(p) > 1/p$ for $p > 3$. Thus, by a well-known Sobolev embedding argument, Theorem 1.2 implies Theorem 1.1. For completeness, the details of this implication are presented in §2.

1.2. **Comparison with previous results.** It follows from work of Pramanik and the fourth author [19] (combined with sharp decoupling estimates from [8]) that, for each fixed t, the single average A_t maps $L^p(\mathbb{R}^3)$ boundedly into $L^p_\alpha(\mathbb{R}^3)$ for all $2 \leqslant p \leqslant \infty$ and $\alpha < \alpha(p)$, where

$$\alpha(p) := \begin{cases} \frac{1}{3}\big(\frac{1}{2}+\frac{1}{p}\big) & \text{if } 2 \leqslant p \leqslant 4 \\ \frac{1}{p} & \text{if } p \geqslant 4 \end{cases};$$

indeed in [19] the $\alpha = \alpha(p)$ endpoint estimate is also shown to hold for $p > 4$. Theorem 1.2 represents a gain of $\sigma(p) - \alpha(p) - \varepsilon = \frac{1}{15}\big(\frac{1}{2}+\frac{1}{p}\big) - \varepsilon$ derivatives when integrating locally in time in the range $3 \leqslant p \leqslant 4$. In this sense, Theorem 1.2 is an example of *local smoothing* (see, for instance, [21, 18, 13, 2] for a discussion of the classical local smoothing phenomenon for the wave equation).

Theorem 1.2 complements previous local smoothing estimates from [19], which deal with the supercritical regime $p > 4$ (here we are referring to criticality for the *single average* operator, so that $p = 4$ correspond to the critical point where the behaviour of the $\alpha(p)$ exponent changes). In [19, Theorem 1.4] it is shown that $\mathfrak{A}_\gamma$ maps $L^p(\mathbb{R}^3)$ boundedly into $L^p_\delta(\mathbb{R}^4)$ for all $2 \leqslant p \leqslant \infty$ and $\delta < \delta(p)$, where

$$\delta(p) := \begin{cases} \frac{1}{3}\big(\frac{1}{2}+\frac{1}{p}\big) & \text{if } 2 \leqslant p \leqslant 6 \\ \frac{4}{3p} & \text{if } p \geqslant 6 \end{cases}.$$

Note that this does **not** yield any local smoothing estimates in the subcritical regime $2 \leqslant p \leqslant 4$, where $\alpha(p)$ and $\delta(p)$ agree. Consequently, the local smoothing estimates in [19] only imply $L^p(\mathbb{R}^3)$-boundedness of M_γ for the restricted range $p > 4$.

It is remarked that the (somewhat loosely) related problem of $L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^{n+1})$ bounds for $\mathfrak{A}_\gamma$ (as opposed to Sobolev bounds) was investigated in [15]. This question is significantly easier than establishing local smoothing estimates and, accordingly, in [15] an almost complete characterisation of the $L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^{n+1})$ mapping properties is obtained in all dimensions.

1.3. **Overview of the argument.** For $\gamma\colon I \to \mathbb{R}^n$ a smooth curve let μ denote the pushforward of the measure $\chi(s)\mathrm{d}s$ under γ. Defining the dilates $\langle \mu_t, f\rangle = \langle \mu, f(t\,\cdot)\rangle$, it follows that the underlying averaging operators satisfy $A_t f = f * \mu_t$. Thus, in the frequency domain A_t corresponds to multiplication against the Fourier transform

$$\widehat{\mu}_t(\xi) = \int_{\mathbb{R}} e^{-it\langle\gamma(s),\xi\rangle}\chi(s)\,\mathrm{d}s.$$

Since the main estimate in Theorem 1.2 is an L^p-Sobolev bound, we are led into studying the decay properties of the above oscillatory integral for large ξ.

Suppose $\gamma\colon I \to \mathbb{R}^3$ satisfies the non-degeneracy hypothesis (1.1). This implies $\sum_{j=1}^{3} |\langle \gamma^{(j)}(s), \xi\rangle| \gtrsim |\xi|$ for all $s \in I$ and all $\xi \in \widehat{\mathbb{R}}^3$ and, consequently, a simple van der Corput estimate yields

$$|\widehat{\mu}_t(\xi)| \lesssim_\gamma (1 + t|\xi|)^{-1/3}.$$

However, this slow decay rate only occurs on a small portion of the frequency domain, corresponding to a (neighbourhood of a) codimension 1 cone $\Gamma \subseteq \widehat{\mathbb{R}}^3$ generated by the binormal vector $\mathbf{e}_3(s)$ to the curve γ. In light of this, it is natural to dyadically decompose the frequency domain into conic regions according to the distance to Γ.

The pieces of the decomposition which are supported far away from Γ satisfy improved decay estimates. In one extreme case, the non-degeneracy condition improves to $\sum_{j=1}^{2} |\langle \gamma^{(j)}(s), \xi\rangle| \gtrsim |\xi|$ and the van der Corput estimate therefore becomes

$$|\widehat{\mu}_t(\xi)| \lesssim_\gamma (1 + t|\xi|)^{-1/2}.$$

In this situation, the operator behaves in many ways like the circular average in the plane, and can be estimated using a lifted version of the argument developed to study the 2 dimensional problem in [18] and [24]. In particular, to prove the desired local smoothing estimate in this extreme case, we observe that the Fourier transform of $\mathfrak{A}_\gamma$ in all 4 variables (x, t) is essentially supported in a neighbourhood of a codimension 1 cone $\widetilde{\Gamma}_1 \subseteq \widehat{\mathbb{R}}^4$. This surface is analogous to the light cone in $\widehat{\mathbb{R}}^3$ which is central to the analysis of local smoothing for the circular averages in [18, 24] and, more recently, [13]. Following an argument of Wolff [24], the operator is further decomposed according to plate regions on $\widetilde{\Gamma}_1$ using a decoupling estimate. The individual pieces of this decomposition are then finally amenable to direct estimation.

The method described in the previous paragraph only directly applies very far from the binormal cone (and therefore far from the most singular parts of the operator). However, by using decoupling inequalities and rescaling, it can also be used to study pieces of the decomposition which lie closer to Γ. The key observation is that the pieces of the decomposition which lie close to Γ can be decoupled into smaller pieces which, when rescaled, resemble the part of the decomposition far from Γ. This, roughly speaking, is the approach used in [19] to obtain Theorem 1.1 in the restricted range $4 < p \leqslant \infty$.

In order to prove the full range of L^p-boundedness of Theorem 1.1 a more direct method is required to analyse the pieces of the decomposition which lie close to the binormal cone. For this part of the operator, the microlocal geometry no longer resembles that of the 2-dimensional problem and, consequently, the decoupling and rescaling argument used in [19] is inefficient.

Close to the binormal cone, we observe that the Fourier transform of $\mathfrak{A}_\gamma$ in all 4 variables (x, t) is essentially supported in a neighbourhood of a codimension 2 cone $\widetilde{\Gamma}_2 \subseteq \widehat{\mathbb{R}}^4$. This cone is a lower-dimensional submanifold of the cone $\widetilde{\Gamma}_1$ we encountered earlier. Similarly to the previous case, the operator is further decomposed according to plate regions, now along $\widetilde{\Gamma}_2$. However, in order to efficiently carry out this decomposition, here we use a square function rather than a decoupling inequality, in the spirit of [18]. The required square function estimate is deduced using a 4-linear restriction estimate from [3]. After applying the square function, a series of weighted L^2 inequalities can be brought to bear on the problem to obtain, together with various corresponding Nikodym-type maximal bounds, a favourable estimate for this part of the operator. This final step of the argument is itself somewhat involved and a discussion of the details is beyond the scope of this introduction.

The above discussion focuses on two extreme cases of the problem:

i) Far from the binormal cone Γ, where $\mathfrak{A}_\gamma$ is (x, t)-Fourier localised to a codimension 1 cone $\widetilde{\Gamma}_1$.
ii) Close to the binormal cone Γ, where $\mathfrak{A}_\gamma$ is (x, t)-Fourier localised to a codimension 2 cone $\widetilde{\Gamma}_2$.

For pieces of the decomposition which lie in the intermediate range, both cones $\widetilde{\Gamma}_1$ and $\widetilde{\Gamma}_2$ play a rôle in the analysis. This complicates matters somewhat, since it is necessary to carry out frequency decompositions simultaneously with respect to both geometries.

Outline of the paper. This paper is structured as follows:

- In §2 we show how Theorem 1.2 implies Theorem 1.1.
- In §3 we reduce Theorem 1.2 to its version for band-limited functions, which is Theorem 3.1.
- In §4 we introduce a class of model curves.
- In §5 we state 3 key auxiliary results that feature in the proof of Theorem 3.1: a reverse square function estimate in $\mathbb{R}^{3+1}$, a forward square function estimate in $\mathbb{R}^3$ and a Nikodym maximal operator bound.
- In §§6–8 we present the proof of Theorem 3.1.
- In §9 we present the proof of the reverse square function estimate in $\mathbb{R}^{3+1}$ (Theorem 5.3).
- In §10 we present the proof of the forward square function estimate in $\mathbb{R}^3$ (Proposition 5.4).
- In §11 we present the proof of the Nikodym maximal operator bound (Proposition 5.5).
- In §12 we show the condition $p > 3$ is necessary for the boundedness of the global maximal function.
- Appendix A contains an abstract broad/narrow decomposition lemma which features in the proof of Theorem 5.3.
- There are two further appendices which deal with various auxiliary results and technical lemmas used in the main argument.

Notational conventions. Given a (possibly empty) list of objects L, for real numbers $A_p, B_p \geqslant 0$ depending on some Lebesgue exponent p or dimension parameter n the notation $A_p \lesssim_L B_p$, $A_p = O_L(B_p)$ or $B_p \gtrsim_L A_p$ signifies that $A_p \leqslant CB_p$ for some constant $C = C_{L,p,n} \geqslant 0$ depending on the objects in the list, p and n. In addition, $A_p \sim_L B_p$ is used to signify that both $A_p \lesssim_L B_p$ and $A_p \gtrsim_L B_p$ hold. Given $a, b \in \mathbb{R}$ we write $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$. The length of a multiindex $\alpha \in \mathbb{N}_0^n$ is given by $|\alpha| = \sum_{i=1}^n \alpha_i$.

Acknowledgements. The authors thank the American Institute of Mathematics for funding their collaboration through the SQuaRE program, also supported in part by the National Science Foundation. D.B. was partially supported by NSF grant DMS-1954479. S.G. was partially supported by NSF grants DMS-1800274 and DMS-2044828. A.S. was partially supported by NSF grants DMS-1764295, DMS 2054220 and a Simons fellowship. This material is partly based upon work supported by the National Science Foundation under Grant No. DMS-1440140 while the authors were in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2017 semester. D.B. and J.H. would also like to thank the LMS for funding a research visit through the LMS 'Research in Pairs' Scheme 4 grant (Grant Ref 41802).

Remark. After finishing this paper, we learned that Hyerim Ko, Sanghyuk Lee and Sewook Oh obtained the same result for the maximal operator M_γ albeit with a different proof.

2. Local smoothing vs maximal bounds

For the readers' convenience, here we state and prove a general result relating local smoothing estimates for the operator $\mathfrak{A}_\gamma f(x,t) := \rho(t)\, A_t f(x)$ to L^p estimates for the corresponding maximal function M_γ.

Proposition 2.1. *Let $\gamma\colon I \to \mathbb{R}^n$ be a smooth curve and suppose $\mathfrak{A}_\gamma$ maps $L^p(\mathbb{R}^n)$ boundedly into $L^p_\sigma(\mathbb{R}^{n+1})$ for some $2 \leqslant p < \infty$ and $\sigma > 1/p$. Then M_γ is bounded on $L^p(\mathbb{R}^n)$.*

Observe that the exponent $\sigma(p) := \frac{1}{5}(1+\frac{2}{p})$ satisfies $\sigma(p) > 1/p$ for all $p > 3$. Consequently, Theorem 1.2 combines with Proposition 2.1 to yield Theorem 1.1 in the restricted range $3 < p \leqslant 4$. The remaining estimates follow from interpolation with the trivial L^∞ bound.

Before presenting the proof we introduce a system of Littlewood–Paley functions which will feature throughout the article. Fix $\eta \in C_c^\infty(\mathbb{R})$ non-negative and such that

$$\eta(r) = 1 \quad \text{if } r \in [-1,1] \quad \text{and} \quad \operatorname{supp} \eta \subseteq [-2,2] \tag{2.1}$$

and define β^k, $\tilde{\beta}^k \in C_c^\infty(\mathbb{R})$ by

$$\beta^k(r) := \eta(2^{-k}r) - \eta(2^{-k+1}r) \qquad \text{and} \qquad \tilde{\beta}^k(r) := \eta(2^{-k-1}r) - \eta(2^{-k+2}r) \tag{2.2}$$

for each $k \in \mathbb{Z}$. By a slight abuse of notation we also let η, β^k, $\tilde{\beta}^k \in C_c^\infty(\widehat{\mathbb{R}}^n)$ denote the radial functions obtained by evaluating the corresponding univariate functions at $|\xi|$. Finally, if $k = 0$, then we drop the superscript and simply write $\beta := \beta^0$ and $\tilde{\beta} := \tilde{\beta}^0$. Note that the β^k form a partition of unity of $\widehat{\mathbb{R}}^n$ subordinated to a family of dyadic annuli, and they satisfy the reproducing formula $\beta^k = \tilde{\beta}^k \cdot \beta^k$.

Proof of Proposition 2.1. Decompose the t parameter into dyadic intervals

$$M_\gamma f(x) = \sup_{\ell \in \mathbb{Z}} \sup_{1 \leqslant t \leqslant 2} |A_{2^\ell t} f(x)|.$$

Performing a Littlewood–Paley decomposition on each of the averaging operators one obtains

$$M_\gamma f(x) \leqslant \sum_{k=1}^{\infty} \Big(\sum_{\ell \in \mathbb{Z}} \sup_{1 \leqslant t \leqslant 2} |A_{2^\ell t} \beta_{k-\ell}(D) f(x)|^p \Big)^{1/p} + C M_{\mathrm{HL}} f(x)$$

where M_{HL} is the Hardy–Littlewood maximal function. Indeed, it is not difficult to verify that the pointwise estimate

$$\sup_{\ell \in \mathbb{Z}} \sup_{1 \leqslant t \leqslant 2} |A_{2^\ell t} \eta_{-\ell}(D) f(x)| \leqslant C M_{\mathrm{HL}} f(x);$$

for $1 \leqslant t \leqslant 2$ the function $A_{2^\ell t}\eta_{-\ell}(D)f(x)$ roughly corresponds to an average of f over a ball of radius 2^ℓ centred at x. Thus, by the Hardy–Littlewood maximal theorem and the triangle inequality it suffices to show that

$$\sum_{k=1}^{\infty} \Big(\sum_{\ell \in \mathbb{Z}} \| \sup_{1 \leqslant t \leqslant 2} |A_{2^\ell t} \beta_{k-\ell}(D) f| \|_{L^p(\mathbb{R}^n)}^p \Big)^{1/p} \lesssim_{\gamma,p} \|f\|_{L^p(\mathbb{R}^n)}. \tag{2.3}$$

By a simple scaling argument, one obtains the operator norm identity

$$\| \sup_{1 \leqslant t \leqslant 2} |A_{2^\ell t} \beta_{k-\ell}(D)| \|_{L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)} = \| \sup_{1 \leqslant t \leqslant 2} |A_t \beta_k(D)| \|_{L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)}.$$

Combining this with the hypothesised local smoothing estimate, it follows that

$$\Big(\int_1^2 \| A_{2^\ell t} \beta_{k-\ell}(D) f \|_{L^p(\mathbb{R}^n)}^p \, \mathrm{d}t \Big)^{1/p} \lesssim_{\gamma,p,\sigma} 2^{-\sigma k} \| \tilde{\beta}_{k-\ell}(D) f \|_{L^p(\mathbb{R}^n)}, \tag{2.4}$$

$$\Big(\int_1^2 \| \frac{\partial}{\partial t} A_{2^\ell t} \beta_{k-\ell}(D) f \|_{L^p(\mathbb{R}^n)}^p \, \mathrm{d}t \Big)^{1/p} \lesssim_{\gamma,p,\sigma} 2^{-\sigma k + k} \| \tilde{\beta}_{k-\ell}(D) f \|_{L^p(\mathbb{R}^n)}. \tag{2.5}$$

The second estimate follows by noting that the Fourier multiplier associated to $\partial_t A_{2^\ell t}\beta_{k-\ell}(D)$ is essentially the same as the multiplier associated to $A_{2^\ell t}\beta_{k-\ell}(D)$ but with an extra $|\xi|$ factor. We therefore pick up an additional 2^k owing to the estimate $\||D|\tilde{\beta}_k(D)f\|_{L^p(\mathbb{R}^n)} \lesssim 2^k \|\tilde{\beta}_k(D)f\|_{L^p(\mathbb{R}^n)}$.

Combining (2.4) and (2.5) with the elementary Sobolev embedding

$$\sup_{1 \leqslant t \leqslant 2} |F(t)|^p \leqslant \int_1^2 |F(s)|^p \, \mathrm{d}s + p \Big(\int_1^2 |F'(s)|^p \, \mathrm{d}s \Big)^{1/p} \Big(\int_1^2 |F(s)|^p \, \mathrm{d}s \Big)^{1/p'},$$

it follows that

$$\| \sup_{1 \leqslant t \leqslant 2} |A_{2^\ell t} \beta_{k-\ell}(D) f| \|_{L^p(\mathbb{R}^n)} \lesssim_{\gamma,p,\sigma} 2^{-k(\sigma - 1/p)} \|\tilde{\beta}_{k-\ell}(D) f\|_{L^p(\mathbb{R}^n)}. \tag{2.6}$$

Taking the ℓ^p-norm of both sides of (2.6), we may sum the resulting expression in ℓ using the elementary inequality

$$\Big(\sum_{\ell \in \mathbb{Z}} \|\tilde{\beta}_\ell(D) f\|_{L^p(\mathbb{R}^n)}^p \Big)^{1/p} \lesssim \|f\|_{L^p(\mathbb{R}^n)},$$

valid for $p \geqslant 2$. On the other hand, under the crucial hypothesis $\sigma > 1/p$, we have a geometric decay which allows us to sum in k. Thus, we deduce the desired estimate (2.3). □

3. Reduction to band-limited estimates

We now turn to the proof of Theorem 1.2, which occupies almost the entirety of the article. Since we are interested in $L^p(\mathbb{R}^3) \to L^p_\sigma(\mathbb{R}^{3+1})$ estimates for σ belonging to an *open* range, the problem is immediately reduced to studying $L^p(\mathbb{R}^3) \to L^p(\mathbb{R}^{3+1})$ bounds for band-limited pieces of the operator. In order to describe this reduction in more detail, it is useful to set up some notational conventions.

Given $m \in L^\infty(\hat{\mathbb{R}}^n \times \mathbb{R})$, for each $t \in \mathbb{R}$ let $m(D;t)$ denote the associated multiplier operator

$$m(D;t)f(x) := \frac{1}{(2\pi)^n} \int_{\hat{\mathbb{R}}^n} e^{i\langle x, \xi\rangle} m(\xi;t) \hat{f}(\xi)\, \mathrm{d}\xi,$$

defined initially for functions f belonging to a suitable *a priori* class. With this notation, the averaging operator A_t is given by $A_t = \hat{\mu}_t(D)$ where μ_t is the measure introduced in §1.3.

The multipliers of interest are of the following form. Let $\gamma \colon I \to \mathbb{R}^n$ be a smooth curve and fix χ, $\rho \in C^\infty_c(\mathbb{R})$ supported in the interior of I and $[1/2, 4]$, respectively. Given a symbol $a \in C^\infty(\hat{\mathbb{R}}^n \backslash \{0\} \times \mathbb{R} \times \mathbb{R})$, define

$$m[a](\xi;t) := \int_{\mathbb{R}} e^{-it\langle \gamma(s), \xi\rangle} a(\xi;t;s) \chi(s) \rho(t)\, \mathrm{d}s. \tag{3.1}$$

Taking a in this definition to be identically 1, we recover the (t-localised) multiplier $\rho(t)\hat{\mu}_t(\xi)$. In general, we perform a decomposition on $\hat{\mu}_t$ by choosing a so that $m[a]$ is localised to a particular region of the frequency space.

For $a \in C^\infty(\hat{\mathbb{R}}^n \backslash \{0\} \times \mathbb{R} \times \mathbb{R})$ as above, we form a dyadic decomposition by writing

$$a = \sum_{k=0}^\infty a_k \qquad \text{where} \qquad a_k(\xi;t;s) := \begin{cases} a(\xi;t;s)\, \beta^k(\xi) & \text{for } k \geqslant 1 \\ a(\xi;t;s)\, \eta(\xi) & \text{for } k = 0 \end{cases}. \tag{3.2}$$

Here η and β^k are the functions introduced in (2.1) and (2.2).

With the above definitions, our main result is as follows.

Theorem 3.1. *Let $\gamma : I \to \mathbb{R}^3$ be a smooth curve and suppose $a \in C^\infty(\hat{\mathbb{R}}^3 \backslash \{0\} \times \mathbb{R} \times \mathbb{R})$ satisfies the symbol condition*

$$|\partial_\xi^\alpha \partial_t^i \partial_s^j a(\xi;t;s)| \lesssim_{\alpha,i,j} |\xi|^{-|\alpha|} \qquad \textit{for all } \alpha \in \mathbb{N}_0^3 \textit{ and } i, j \in \mathbb{N}_0$$

and that

$$\sum_{j=1}^3 |\langle \gamma^{(j)}(s), \xi\rangle| \gtrsim |\xi| \qquad \textit{for all } (\xi;s) \in \mathrm{supp}_\xi\, a \times I. \tag{3.3}$$

Let $3 \leqslant p \leqslant 4$, $\varepsilon > 0$ and $k \geqslant 1$. If a_k is defined as in (3.2), then

$$\Big(\int_1^2 \|m[a_k](D;t)f\|_{L^p(\mathbb{R}^3)}^p\,\mathrm{d}t\Big)^{1/p} \lesssim_{\varepsilon,p} 2^{-\frac{k}{5}(1+\frac{2}{p})+k\varepsilon}\|f\|_{L^p(\mathbb{R}^3)}.$$

For $n = 3$, the condition (3.3) is equivalent to the non-degeneracy hypothesis (1.1). Thus, Theorem 3.1 immediately implies Theorem 1.2 via the Littlewood–Paley characterisation of Sobolev spaces.

Under a stronger hypothesis on the phase function, a stronger local smoothing estimate holds, by a combination of the work of Pramanik and the fourth author [19] with the full decoupling theorem for the light cone by Bourgain and Demeter [8]. We remark that the estimates in [19] are stated for $p > 6$, and the version of the result presented here for $2 \leqslant p \leqslant \infty$ follows via interpolation with trivial L^2-estimates.

Theorem 3.2 (*cf.* Theorem 4.1 in [19])**.** *Let $\gamma : I \to \mathbb{R}^3$ be a smooth curve and suppose that $a \in C^\infty(\widehat{\mathbb{R}}^3\backslash\{0\} \times \mathbb{R} \times \mathbb{R})$ satisfies the symbol conditions*

$$|\partial_\xi^\alpha \partial_t^i \partial_s^j a(\xi;t;s)| \lesssim_{\alpha,i,j} |\xi|^{-|\alpha|} \qquad \textit{for all } \alpha \in \mathbb{N}_0^3 \textit{ and } i,\, j \in \mathbb{N}_0$$

and that

$$|\langle\gamma'(s),\xi\rangle| + |\langle\gamma''(s),\xi\rangle| \gtrsim |\xi| \qquad \textit{for all } \ (\xi;s) \in \operatorname{supp}_\xi a \times I. \tag{3.4}$$

Let $2 \leqslant p \leqslant 6$, $\varepsilon > 0$ and $k \geqslant 1$. If a_k is defined as in (3.2), then

$$\Big(\int_1^2 \|m[a_k](D;t)f\|_{L^p(\mathbb{R}^3)}^p\,\mathrm{d}t\Big)^{1/p} \lesssim_{\varepsilon,p} 2^{-\frac{k}{2}(\frac{1}{2}+\frac{1}{p})+k\varepsilon}\|f\|_{L^p(\mathbb{R}^3)}.$$

Owing to the strengthened hypothesis (3.4), Theorem 3.2 alone is insufficient for our purposes. Indeed, Theorem 3.2 only effectively deals with parts of the multiplier which are supported away from the main singularity. However, we still make use of Theorem 3.2 in the proof of Theorem 3.1 to analyse the multiplier in this less singular region, in which it is effective.

4. Symmetries and model curves

A prototypical example of a smooth curve satisfying the non-degeneracy condition (1.1) is the *moment curve* $\gamma_\circ \colon \mathbb{R} \to \mathbb{R}^n$, given by

$$\gamma_\circ(s) := \Big(s, \frac{s^2}{2}, \dots, \frac{s^n}{n!}\Big).$$

Indeed, in this case the determinant appearing in (1.1) is everywhere equal to 1. Moreover, at small scales, any non-degenerate curve can be thought of as a perturbation of an affine image of $\gamma_\circ$. To see why this is so, fix a non-degenerate curve $\gamma \colon I \to \mathbb{R}^n$ and $\sigma \in I$, $\lambda > 0$ such that $[\sigma - \lambda, \sigma + \lambda] \subseteq I$. Denote by $[\gamma]_\sigma$ the $n \times n$ matrix

$$[\gamma]_\sigma := \begin{bmatrix}\gamma^{(1)}(\sigma) & \cdots & \gamma^{(n)}(\sigma)\end{bmatrix},$$

where the vectors $\gamma^{(j)}(\sigma)$ are understood to be *column* vectors. Note that this is precisely the matrix appearing in the definition of the non-degeneracy condition (1.1) and is therefore invertible by our hypothesis. It is also convenient to let $[\gamma]_{\sigma,\lambda}$ denote the $n \times n$ matrix

$$[\gamma]_{\sigma,\lambda} := [\gamma]_\sigma \cdot D_\lambda, \tag{4.1}$$

where $D_\lambda := \operatorname{diag}(\lambda, \dots, \lambda^n)$, the diagonal matrix with eigenvalues λ, $\lambda^2, \dots, \lambda^n$. Consider the portion of the curve γ lying over the subinterval $[\sigma - \lambda, \sigma + \lambda]$. This is parametrised by the map $s \mapsto \gamma(\sigma + \lambda s)$ for $s \in [-1, 1]$. The degree n Taylor polynomial of $s \mapsto \gamma(\sigma + \lambda s)$ around σ is given by

$$s \mapsto \gamma(\sigma) + [\gamma]_{\sigma,\lambda} \cdot \gamma_\circ(s), \tag{4.2}$$

which is indeed an affine image of $\gamma_\circ$. Furthermore, by Taylor's theorem, the original curve γ agrees with the polynomial curve (4.2) to high order at σ.

Inverting the affine transformation $x \mapsto \gamma(\sigma) + [\gamma]_{\sigma,\lambda} \cdot x$ from (4.2), we can map the portion of γ over $[\sigma - \lambda, \sigma + \lambda]$ to a small perturbation of the moment curve.

Definition 4.1. *Let $\gamma \in C^{n+1}(I;\mathbb{R}^n)$ be a non-degenerate curve and $\sigma \in I, \lambda > 0$ be such that $[\sigma - \lambda, \sigma + \lambda] \subseteq I$. The (σ, λ)-rescaling of γ is the curve $\gamma_{\sigma,\lambda} \in C^{n+1}([-1,1];\mathbb{R}^n)$ given by*

$$\gamma_{\sigma,\lambda}(s) := [\gamma]_{\sigma,\lambda}^{-1}\big(\gamma(\sigma + \lambda s) - \gamma(\sigma)\big).$$

It follows from the preceding discussion that

$$\gamma_{\sigma,\lambda}(s) = \gamma_\circ(s) + [\gamma]_{\sigma,\lambda}^{-1}\mathcal{E}_{\gamma,\sigma,\lambda}(s)$$

where $\mathcal{E}_{\gamma,\sigma,\lambda}$ is the remainder term for the Taylor expansion (4.2). In particular, if γ satisfies the non-degeneracy condition (1.1) with constant c_0, then

$$\|\gamma_{\sigma,\lambda} - \gamma_\circ\|_{C^{n+1}([-1,1];\mathbb{R}^n)} \lesssim c_0^{-1}\lambda\, \|\gamma\|_{C^{n+1}(I)}^n.$$

Thus, if $\lambda > 0$ is chosen to be small enough, then the rescaled curve $\gamma_{\sigma,\lambda}$ is a minor perturbation of the moment curve. In particular, given any $0 < \delta < 1$, we can choose λ so as to ensure that $\gamma_{\sigma,\lambda}$ belongs to the following class of *model curves*.

Definition 4.2. *Given $n \geqslant 2$ and $0 < \delta < 1$, let $\mathfrak{G}_n(\delta)$ denote the class of all smooth curves $\gamma\colon [-1,1] \to \mathbb{R}^n$ that satisfy the following conditions:*

i) $\gamma(0) = 0$ *and* $\gamma^{(j)}(0) = \vec{e}_j$ *for* $1 \leqslant j \leqslant n$*;*
ii) $\|\gamma - \gamma_\circ\|_{C^{n+1}([-1,1])} \leqslant \delta$*.*

Here $\vec{e}_j$ denotes the jth standard Euclidean basis vector and

$$\|\gamma\|_{C^{n+1}(I)} := \max_{1 \leqslant j \leqslant n+1} \sup_{s \in I} |\gamma^{(j)}(s)| \qquad \textit{for all } \gamma \in C^{n+1}(I;\mathbb{R}^n).$$

Given any $\gamma \in \mathfrak{G}_n(\delta)$, condition ii) and the multilinearity of the determinant ensures that $\det[\gamma]_s = \det[\gamma_\circ]_s + O(\delta) = 1 + O(\delta)$. Thus, there exists a dimensional constant $c_n > 0$ such that if $0 < \delta < c_n$, then any curve $\gamma \in \mathfrak{G}_n(\delta)$ is non-degenerate and, moreover, satisfies $\det[\gamma]_s \geqslant 1/2$. Henceforth, it is always assumed that any such parameter $\delta > 0$ satisfies this condition, which we express succinctly as $0 < \delta \ll 1$.

5. Key analytic ingredients in the proof

There are three key ingredients in the proof of Theorem 3.1: a square function on $\mathbb{R}^4$, a square function on $\mathbb{R}^3$ and a Nikodym-type maximal operator mapping functions in $\mathbb{R}^4$ to functions in $\mathbb{R}^3$. These operators are formulated in terms of the geometry of the underlying curve $\gamma\colon I \to \mathbb{R}^3$ and, in particular, are defined with respect to the Frenet frame on γ (more precisely, the square function on $\mathbb{R}^4$ is defined with respect to the Frenet frame associated to a lift of γ to $\mathbb{R}^4$). In this section each of the three key operators is introduced and the relevant norm bounds for these objects are stated in Theorem 5.3, Proposition 5.4 and Proposition 5.5 below. In §§7-8, a careful decomposition of the multiplier $m[a_k]$ is carried out which facilitates application of these results in the proof of Theorem 3.1. We return to proofs of Theorem 5.3, Proposition 5.4 and Proposition 5.5 in §9, §10 and §11, respectively.

5.1. **Frenet geometry.** It is convenient to recall some elementary concepts from differential geometry which feature in our proof. Given a smooth non-denegenate curve $\gamma : I \to \mathbb{R}^n$, the Frenet frame is the orthonormal basis resulting from applying the Gram–Schmidt process to the vectors

$$\{\gamma'(s), \dots, \gamma^{(n)}(s)\},$$

which are linearly independent in view of the condition (1.1). Defining the functions

$$\tilde{\kappa}_j(s) := \langle \mathbf{e}_j'(s), \mathbf{e}_{j+1}(s)\rangle \qquad \text{for } j = 1, \dots, n-1,$$

one has the classical Frenet formulæ

$$\begin{aligned} \mathbf{e}_1'(s) &= \tilde{\kappa}_1(s)\mathbf{e}_2(s), \\ \mathbf{e}_i'(s) &= -\tilde{\kappa}_{i-1}(s)\mathbf{e}_{i-1}(s) + \tilde{\kappa}_i(s)\mathbf{e}_{i+1}(s), \ i = 2, \dots, n-1, \\ \mathbf{e}_n'(s) &= -\tilde{\kappa}_{n-1}(s)\mathbf{e}_{n-1}(s). \end{aligned}$$

Note that the $\tilde{\kappa}_j$ depend on the choice of parametrisation and only agree with the (geometric) curvature functions

$$\kappa_j(s) := \frac{\langle \mathbf{e}_j'(s), \mathbf{e}_{j+1}(s)\rangle}{|\gamma'(s)|}$$

if γ is unit speed parametrised (here we do not assume unit speed parametrisation). Repeated application of the above formulæ shows that

$$\mathbf{e}_i^{(k)}(s) \perp \mathbf{e}_j(s) \qquad \text{whenever} \qquad 0 \leqslant k < |i-j|.$$

Consequently, by Taylor's theorem

$$|\langle \mathbf{e}_i(s_1), \mathbf{e}_j(s_2)\rangle| \lesssim_\gamma |s_1 - s_2|^{|i-j|} \qquad \text{for } 1 \leqslant i, j \leqslant n \text{ and } s_1, s_2 \in I.$$

Furthermore, one may deduce from the definition of $\{\mathbf{e}_j(s)\}_{j=1}^n$ that

$$|\langle \gamma^{(i)}(s_1), \mathbf{e}_j(s_2)\rangle| \lesssim_\gamma |s_1 - s_2|^{(j-i)\vee 0} \qquad \text{for } 1 \leqslant i, j \leqslant n \text{ and } s_1, s_2 \in I. \tag{5.1}$$

In this paper, much of the microlocal geometry of the averaging operators A_t is expressed in terms of the Frenet frame. We further introduce the following definitions.

Definition 5.1. *Given $1 \leqslant d \leqslant n-1$ and $0 < r \leqslant 1$, for each $s \in I$ let $\pi_{d-1}(s;\, r)$ denote the set of all $\xi \in \widehat{\mathbb{R}}^n$ satisfying the following conditions:*

$$|\langle \mathbf{e}_j(s), \xi\rangle| \leqslant r^{d+1-j} \qquad \textit{for } 1 \leqslant j \leqslant d, \tag{5.2a}$$

$$1/2 \leqslant |\langle \mathbf{e}_{d+1}(s), \xi\rangle| \leqslant 2 \tag{5.2b}$$

$$|\langle \mathbf{e}_j(s), \xi\rangle| \leqslant 1 \qquad \textit{for } d+2 \leqslant j \leqslant n. \tag{5.2c}$$

Such sets $\pi_{d-1}(s;\, r)$ are referred to as $(d-1, r)$-Frenet boxes.

The relevance of the $d-1$ index is that the $\pi_{d-1}(s; r)$ correspond to plate regions defined with respect to a codimension $d-1$ cone. For $n = 4$ and $d-1 = 2$, this geometric observation is discussed in detail in §9.1.

Definition 5.2. *A collection $\mathcal{P}_{d-1}(r)$ of $(d-1, r)$-Frenet boxes is a Frenet box decomposition along γ if it consists of precisely the $(d-1, r)$-Frenet boxes $\pi_{d-1}(s;\, r)$ for s varying over an r-separated subset of I.*

5.2. **Reverse square function estimates in $\mathbb{R}^{3+1}$.** The most important ingredient in the proof of Theorem 3.1 is the following square function bound.

Theorem 5.3. *Let $0 < r < 1$ and $\mathcal{P}_2(r)$ be a $(2,r)$-Frenet box decomposition along a non-degenerate $\gamma\colon I \to \mathbb{R}^4$. For all $\varepsilon > 0$ the inequality*

$$\Big\| \sum_{\pi \in \mathcal{P}_2(r)} f_\pi \Big\|_{L^4(\mathbb{R}^4)} \lesssim_{\gamma,\varepsilon} r^{-\varepsilon} \Big\| \Big(\sum_{\pi \in \mathcal{P}_2(r)} |f_\pi|^2 \Big)^{1/2} \Big\|_{L^4(\mathbb{R}^4)}$$

holds for any tuple of functions $(f_\pi)_{\pi \in \mathcal{P}_2(r)}$ satisfying $\operatorname{supp} \widehat{f}_\pi \subseteq \pi$.

This bound pertains to curves in $\mathbb{R}^4$ rather than $\mathbb{R}^3$ and therefore does not directly apply to the curve $\gamma\colon I \to \mathbb{R}^3$ featured in the definition of our original helical maximal operator. Rather, in §8.3 we apply Theorem 5.3 to a certain lift of the original curve γ into the spatio-temporal domain $\mathbb{R}^{3+1}$. This is somewhat analogous to the situation in [18] where a square function estimate in $\mathbb{R}^{2+1}$ is used to study the circular maximal function in $\mathbb{R}^2$.

Theorem 5.3 is related to the Lee–Vargas [16] estimate for the Mockenhaupt square function in $\mathbb{R}^3$. In particular, the Mockenhaupt square function corresponds to studying functions frequency localised with repect to a $(1,r)$-Frenet box decomposition in $\mathbb{R}^3$. Moreover, the strategy used to prove Theorem 9.3 mirrors that of [16]. We first obtain a 4-linear variant of Theorem 5.3 via the multilinear Fourier restriction estimates of Bennett–Bez–Flock–Lee [3]. The linear result is then deduced from the 4-linear inequality using a variant of the Bourgain–Guth method [9]. The details of the argument are provided in §9.

5.3. **Forward square function estimates in $\mathbb{R}^3$.** We also make use of a (forward) L^2-weighted square function estimate in $\mathbb{R}^3$. Here the square function estimate is defined in relation to a $(0,r)$-Frenet decomposition. In contrast with Theorem 5.3, we work with an operator-theoretic formulation involving certain projection operators.

As before, let $\eta \in C_c^\infty(\mathbb{R})$ be non-negative and such that $\eta(r) = 1$ if $r \in [-1,1]$ and $\operatorname{supp} \eta \subseteq [-2,2]$ and define $\tilde{\beta} := \eta(2^{-1}\,\cdot\,) - \eta(4\,\cdot\,)$. Give an $(0,r)$-Frenet box $\pi = \pi_{0,\gamma}(s;r)$ let

$$\chi_\pi(\xi) := \eta\big(r^{-1}\langle \mathbf{e}_1(s), \xi\rangle\big)\, \tilde{\beta}\big(\langle \mathbf{e}_2(s), \xi\rangle\big)\, \eta\big(\langle \mathbf{e}_3(s), \xi\rangle\big) \tag{5.3}$$

so that $\chi_\pi(\xi) = 1$ if $\xi \in \pi_{0,\gamma}(s;r)$ and χ_π vanishes outside some fixed dilate of this set.

Proposition 5.4. *Let $0 < r < 1$ and $\mathcal{P}_0(r)$ be a $(0,r)$-Frenet box decomposition for a non-degenerate $\gamma\colon I \to \mathbb{R}^3$. For all $\varepsilon > 0$ the inequality*

$$\int_{\mathbb{R}^3} \sum_{\pi \in \mathcal{P}_0(r)} |\chi_\pi(D) f(x)|^2 w(x)\, \mathrm{d}x \lesssim_\varepsilon r^{-\varepsilon} \int_{\mathbb{R}^3} |f(x)|^2 \widetilde{\mathcal{N}}^{(\varepsilon)}_{\gamma,r} w(x)\, \mathrm{d}x$$

holds for any non-negative $w \in L^1_{\mathrm{loc}}(\mathbb{R}^3)$, where $\widetilde{\mathcal{N}}^{(\varepsilon)}_{\gamma,r}$ is a maximal operator satisfying

$$\|\widetilde{\mathcal{N}}^{(\varepsilon)}_{\gamma,r}\|_{L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3)} \lesssim_{\varepsilon,\varepsilon_\circ} r^{-\varepsilon_\circ} \qquad \textit{for all } \varepsilon_\circ > 0. \tag{5.4}$$

The above proposition is related to a L^2-weighted version of the classical sectorial square function of Córdoba [11], due to Carbery and the fourth author [10, Proposition 4.6]. The proof is presented in § 10 below.

The definition of $\widetilde{\mathcal{N}}^{(\varepsilon)}_{\gamma,r}$ is rather complicated, involving a repeated composition of Nikodym-type maximal operators at different scales. For this reason, we do not provide an explicit description of the operator here. Further details of the definition and basic properties of this operator are provided in § 10.

5.4. **A singular Nikodym-type maximal function.** The bounds on the spatio-temporal frequency localised pieces of our operator $m[a](D;\cdot)$ are reduced to bounding a Nikodym maximal function mapping functions in $\mathbb{R}^4$ to functions in $\mathbb{R}^3$. Given $\mathbf{r} \in (0,1)^3$ and $s \in [-1,1]$, consider the *plates*

$$\mathcal{T}_{\mathbf{r}}(s) := \big\{(y,t) \in \mathbb{R}^3 \times [1,2] : \big|\langle y - t\gamma(s), \mathbf{e}_j(s)\rangle\big| \leqslant r_j \text{ for } j = 1,2,3\big\} \subset \mathbb{R}^4.$$

Using these sets, we define associated averaging and maximal operators

$$\mathcal{A}_{\mathbf{r}}^{\mathrm{sing}} g(x;s) := \fint_{\mathcal{T}_{\mathbf{r}}(s)} g(x-y,t)\,\mathrm{d}y\mathrm{d}t \quad \text{and} \quad \mathcal{N}_{\mathbf{r}}^{\mathrm{sing}} g(x) := \sup_{-1 \leqslant s \leqslant 1} |\mathcal{A}_{\mathbf{r}}^{\mathrm{sing}} g(x;s)|.$$

Note that $\mathcal{N}_{\mathbf{r}}^{\mathrm{sing}}$ takes as its input some $g \in L^1_{\mathrm{loc}}(\mathbb{R}^4)$ and outputs a measurable function on $\mathbb{R}^3$. In particular, there is a discrepancy between the number of input and the number of output variables of the operator.

Proposition 5.5. *If $\mathbf{r} \in (0,1)^3$ satisfies $r_3 \leqslant r_2 \leqslant r_1 \leqslant r_2^{1/2}$ and $r_2 \leqslant r_1^{1/2} r_3^{1/2}$, then*

$$\|\mathcal{N}_{\mathbf{r}}^{\mathrm{sing}} g\|_{L^2(\mathbb{R}^3)} \lesssim |\log r_3|^3 \|g\|_{L^2(\mathbb{R}^4)}.$$

This result can be thought of as a higher dimensional analogue of a Nikodym maximal estimate from [18], which is used to study the circular maximal function in the plane. Note that the parameter triple $\mathbf{r} = (r,r,r)$ for some $0 < r < 1$ satisfies the hypothesis of Proposition 5.5, corresponding to the case of tubes former around the rays $t \mapsto t\gamma(s)$. More relevant to our study, however, is the highly anisotropic situation where $\mathbf{r} = (r, r^2, r^3)$; note that this case is also covered by the proposition. It is remarked that the situation here is somewhat different to that appearing in Proposition 5.4 (which will be defined in §10), owing to the aforementioned disparity between the number of input and output variables. The proof of Proposition 5.5, which is based on an oscillatory integral argument, is presented in §11 below.

6. Proof of Theorem 3.1: the slow decay cone

Throughout the remainder of the paper, we work with some fixed $0 < \delta_0 \ll 1$, chosen to satisfy the forthcoming requirements of the proofs. For the sake of concreteness, the choice of $\delta_0 := 10^{-10}$ is more than enough for our purposes. It suffices to prove Theorem 3.1 in the special case where $\gamma \in \mathfrak{G}_3(\delta_0)$ and $\operatorname{supp} \chi \subseteq I_0 := [-\delta_0, \delta_0]$. Indeed, using the observations of §4, we may decompose and rescale the operator $m[a_k](D;\,\cdot)$ to reduce to this situation.

Suppose $\gamma \in \mathfrak{G}_3(\delta_0)$ and $a \in C^\infty(\widehat{\mathbb{R}}^3 \backslash \{0\} \times \mathbb{R} \times \mathbb{R})$ satisfies the hypotheses Theorem 3.1. In view of Theorem 3.2, we may further assume that

$$\begin{cases} |\langle \gamma^{(3)}(s), \xi\rangle| \geqslant \frac{9}{10}\,|\xi| & \\ |\langle \gamma^{(j)}(s), \xi\rangle| \leqslant 8\delta_0|\xi| & \text{for } j = 1,2 \end{cases} \qquad \text{for all } (\xi;t;s) \in \operatorname{supp} a. \tag{6.1}$$

We note two further consequences of this technical reduction:

- Since $\gamma \in \mathfrak{G}_3(\delta_0)$, we have $\gamma^{(j)}(0) = \vec{e}_j$ for $1 \leqslant j \leqslant 3$ and so (6.1) immediately implies that
$$|\xi_3| \geqslant \tfrac{9}{10}\,|\xi| \quad \text{and} \quad |\xi_j| \leqslant 8\delta_0|\xi| \quad \text{for } j = 1,2, \qquad \text{for all } \xi \in \operatorname{supp}_\xi a.$$
- Since $\gamma \in \mathfrak{G}_3(\delta_0)$, we have $\|\gamma^{(4)}\|_\infty \leqslant \delta_0$. Thus, provided δ_0 is sufficiently small,
$$|\langle \gamma^{(3)}(s), \xi\rangle| \geqslant \tfrac{1}{2}\,|\xi| \qquad \text{for all } (\xi;s) \in \operatorname{supp}_\xi a \times [-1,1]. \tag{6.2}$$

Observe that this inequality holds on the large interval $[-1,1]$, rather than just I_0.

Henceforth, we also assume that $\xi_3 > 0$ for all $\xi \in \operatorname{supp}_\xi a$. In particular,

$$\langle \gamma^{(3)}(s), \xi \rangle > 0 \qquad \text{for all } (\xi; s) \in \operatorname{supp}_\xi a \times [-1, 1] \tag{6.3}$$

and thus, for each $\xi \in \operatorname{supp}_\xi a$, the function $s \mapsto \langle \gamma'(s), \xi \rangle$ is strictly convex on $[-1, 1]$. The analysis for the portion of the symbol supported on the set $\{\xi_3 < 0\}$ follows by symmetry.

The first step is to isolate regions of the frequency space where the multiplier $m[a]$ decays relatively slowly. Owing to stationary phase considerations, this corresponds to a region around the conic variety

$$\Gamma := \{\xi \in \operatorname{supp}_\xi a : \langle \gamma^{(j)}(s), \xi \rangle = 0,\ 1 \leqslant j \leqslant 2, \text{ for some } s \in I_0\}.$$

To analyse this cone, we begin with the following observation.

Lemma 6.1. *If $\xi \in \operatorname{supp}_\xi a$, then the equation $\langle \gamma''(s), \xi \rangle = 0$ has a unique solution in $s \in [-1, 1]$, which corresponds to the unique global minimum of the function $s \mapsto \langle \gamma'(s), \xi \rangle$. Furthermore, the solution has absolute value $O(\delta_0)$.*

Proof. Given $\xi \in \operatorname{supp}_\xi a$, let

$$\phi \colon [-1, 1] \to \mathbb{R}, \quad \phi \colon s \mapsto \langle \gamma'(s), \xi \rangle. \tag{6.4}$$

By (6.3), $\phi''(s) > 0$ for all $s \in [-1, 1]$ and the equation $\phi'(s) = \langle \gamma^{(2)}(s), \xi \rangle = 0$ has at most one solution on that interval.

On the other hand, by the mean value theorem,

$$\phi'(s) = \langle \gamma^{(2)}(s), \xi \rangle = \xi_2 + \omega(\xi; s)\, s,$$

where ω satisfies $|\omega(\xi; s)| \geqslant \frac{1}{2}|\xi| > 0$. As $|\xi_2| \leqslant 8\delta_0|\xi|$, it follows that $|\omega(\xi; s)||s| > |\xi_2|$ if $|s| > 16\delta_0$, and so the equation $\langle \gamma^{(2)}(s), \xi \rangle = 0$ has a unique solution in the interval $[-16\delta_0, 16\delta_0]$. Moreover, it immediately follows from (6.3) that this solution is the unique global minimum of ϕ on $[-1, 1]$. ☐

Using Lemma 6.1, we construct a smooth mapping $\theta_2 \colon \operatorname{supp}_\xi a \to [-1, 1]$ such that

$$\langle \gamma'' \circ \theta_2(\xi), \xi \rangle = 0 \qquad \text{for all } \xi \in \operatorname{supp}_\xi a.$$

It is easy to see that θ_2 is homogeneous of degree 0. This function can be used to construct a natural Whitney decomposition with respect to the cone Γ defined above. In particular, let

$$u(\xi) := \langle \gamma' \circ \theta_2(\xi), \xi \rangle \qquad \text{for all } \xi \in \operatorname{supp}_\xi a. \tag{6.5}$$

This quantity plays a central rôle in our analysis. If $u(\xi) = 0$, then $\xi \in \Gamma$ and so, roughly speaking, $u(\xi)$ measures the distance of ξ from Γ.

Lemma 6.2. *Let $\xi \in \operatorname{supp}_\xi a$ and consider the equation*

$$\langle \gamma'(s), \xi \rangle = 0. \tag{6.6}$$

i) If $u(\xi) > 0$, then the equation (6.6) has no solution on $[-1, 1]$.
ii) If $u(\xi) = 0$, then the equation (6.6) has only the solution $s = \theta_2(\xi)$ on $[-1, 1]$.
iii) If $u(\xi) < 0$, then the equation (6.6) has precisely two solutions on $[-1, 1]$. Both solutions have absolute value $O(\delta_0^{1/2})$.

Proof. Given $\xi \in \operatorname{supp}_\xi a$, define ϕ as in (6.4).

i) In this case, Lemma 6.1 implies that

$$\phi(s) = \langle \gamma'(s), \xi \rangle \geqslant u(\xi) > 0 \quad \text{for all } s \in [-1, 1],$$

and so (6.6) has no solutions.

ii) This case also follows immediately from Lemma 6.1, since $s = \theta_2(\xi)$ is the only global minimum for ϕ on $[-1, 1]$.

iii) Recall, by (6.3), the function ϕ is strictly convex on $[-1, 1]$, and therefore $\phi(s) = 0$ has at most two solutions on that interval.

On the other hand, by (the proof of) Lemma 6.1 we know that $|\theta_2(\xi)| \leqslant 16\delta_0$. Moreover, the mean value theorem implies

$$|u(\xi)| \leqslant |\xi_1| + \sup_{|s| \leqslant 16\delta_0} |\gamma^{(2)}(s)||\xi||\theta_2(\xi)| \leqslant 8\Big(1 + 2 \sup_{|s| \leqslant 16\delta_0} |\gamma^{(2)}(s)|\Big)\delta_0|\xi| \leqslant 40\delta_0|\xi|, \tag{6.7}$$

since $\gamma \in \mathfrak{G}_3(\delta_0)$. By Taylor expansion of ϕ around $\theta_2(\xi)$, one obtains

$$\phi(s) = u(\xi) + \omega(\xi; s)\,(s - \theta_2(\xi))^2, \tag{6.8}$$

where ω arises from the remainder term and satisfies $\omega(\xi; s) \geqslant \frac{1}{4}|\xi|$. Combining (6.7) and (6.8), it follows that if $|s - \theta_2(\xi)| \geqslant 20\delta_0^{1/2}$, then $\phi(s) > 0$. Recall that $\phi \circ \theta_2(\xi) = u(\xi) < 0$. Consequently, the equation $\phi(s) = 0$ has exactly two solutions on the interval

$$[-16\delta_0, 16\delta_0] + [-20\delta_0^{1/2}, 20\delta_0^{1/2}] \subseteq [-36\delta_0^{1/2}, 36\delta_0^{1/2}],$$

as required. □

Using Lemma 6.2, we construct a (unique) pair of smooth mappings

$$\theta_1^{\pm} \colon \{\xi \in \operatorname{supp}_\xi a : u(\xi) < 0\} \to [-1, 1]$$

with $\theta_1^-(\xi) \leqslant \theta_1^+(\xi)$ which satisfies

$$\langle \gamma' \circ \theta_1^{\pm}(\xi), \xi\rangle = 0 \qquad \text{for all } \xi \in \operatorname{supp}_\xi a \text{ with } u(\xi) < 0.$$

Define the functions

$$v^{\pm}(\xi) := \langle \gamma'' \circ \theta_1^{\pm}(\xi), \xi\rangle \qquad \text{for all } \xi \in \operatorname{supp} a \text{ with } u(\xi) < 0.$$

Lemma 6.3. *Let $\xi \in \operatorname{supp} a$ with $u(\xi) < 0$. Then the following hold:*

$$\big|v^{\pm}\big(\tfrac{\xi}{|\xi|}\big)\big| \sim |\theta_1^{\pm}(\xi) - \theta_2(\xi)| \sim |\theta_1^+(\xi) - \theta_1^-(\xi)| \sim \big|u\big(\tfrac{\xi}{|\xi|}\big)\big|^{1/2}.$$

Proof. By Taylor expansion around $\theta_2(\xi)$, we obtain

$$\begin{aligned} v^{\pm}(\xi) &= \omega_1^{\pm}(\xi)\,(\theta_1^{\pm}(\xi) - \theta_2(\xi)), \\ 0 = \langle \gamma' \circ \theta_1^{\pm}(\xi), \xi\rangle &= u(\xi) + \omega_2(\xi)\,(\theta_1^{\pm}(\xi) - \theta_2(\xi))^2 \end{aligned}$$

where $|\omega_1^{\pm}(\xi)| \sim |\omega_2(\xi)| \sim |\xi|$ by (6.2). Similarly, Taylor expansion around $\theta_1^{\pm}(\xi)$ yields

$$0 = \langle \gamma' \circ \theta_1^+(\xi), \xi\rangle = v^-(\xi)\,(\theta_1^+(\xi) - \theta_1^-(\xi)) + \omega_3(\xi)\,(\theta_1^+(\xi) - \theta_1^-(\xi))^2$$

where again the remainder satisfies $|\omega_3(\xi)| \sim |\xi|$. As $\theta_1^+(\xi) \neq \theta_1^-(\xi)$, we can combine the identities above to obtain the desired bounds. □

7. Proof of Theorem 3.1: Local smoothing relative to Γ

For $k \geqslant 1$, consider the frequency localised symbols $a_k := a\,\beta^k$, as introduced in §3. We decompose each a_k with respect to the size of $|u(\xi)|$. In particular, write

$$a_k = \sum_{\ell=0}^{\lfloor k/3 \rfloor} a_{k,\ell} \qquad \text{where} \qquad a_{k,\ell}(\xi; t; s) := \begin{cases} a_k(\xi; t; s)\,\beta\big(2^{-k+2\ell}u(\xi)\big) & \text{if } 0 \leqslant \ell < \lfloor k/3 \rfloor \\ a_k(\xi; t; s)\,\eta\big(2^{-k+2\lfloor k/3 \rfloor}u(\xi)\big) & \text{if } \ell = \lfloor k/3 \rfloor \end{cases} \tag{7.1}$$

and $\lfloor k/3 \rfloor$ denotes the greatest integer less than or equal to $k/3$. We note that here the function β should be defined slightly differently compared with (2.2); in particular, here $\beta(r) := \eta(2^{-2}r) - \eta(r)$ (we ignore this minor change in the notation).

To prove Theorem 3.1, we establish local smoothing estimates for each of the operators $m[a_{k,\ell}](D;\,\cdot\,)$. The main result is as follows.

Proposition 7.1. *Let $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$. For all $2 \leqslant p \leqslant 4$ and $\varepsilon > 0$,*

$$\|m[a_{k,\ell}](D;\,\cdot\,)f\|_{L^p(\mathbb{R}^{3+1})} \lesssim_\varepsilon 2^{-k/p-\ell(1-3/p)} 2^{\varepsilon k} \|f\|_{L^p(\mathbb{R}^3)}.$$

Proposition 7.1 provides an effective bound in the large ℓ regime (in particular, for $\lfloor k/5 \rfloor \leqslant \ell \leqslant \lfloor k/3 \rfloor$). This corresponds to those pieces of the multiplier which are supported close to the binormal cone Γ, and therefore lie in a neighbourhood of the most significant singularity.

In addition to Proposition 7.1, we also use results from [19] to deal with the less singular pieces of the multiplier.

Proposition 7.2 ([19])**.** *Let $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$. For all $2 \leqslant p \leqslant 6$ and $\varepsilon > 0$,*

$$\|m[a_{k,\ell}](D;\,\cdot\,)f\|_{L^p(\mathbb{R}^{3+1})} \lesssim_\varepsilon 2^{-\frac{k-\ell}{2}(\frac{1}{2}+\frac{1}{p})+\varepsilon k} \|f\|_{L^p(\mathbb{R}^3)}.$$

This proposition follows from Theorem 3.2 via the sharp Wolff inequality for the light cone [8] and a rescaling argument (c.f. §1.3). The details of the proof can be found in [19, §5].

Proof of Theorem 3.1, assuming Proposition 7.1. Applying the decomposition (7.1) and the triangle inequality,

$$\|m[a_k](D;\,\cdot\,)f\|_{L^p(\mathbb{R}^{3+1})} \leqslant \sum_{\ell=0}^{\lfloor k/5 \rfloor} \|m[a_{k,\ell}](D;\,\cdot\,)f\|_{L^p(\mathbb{R}^{3+1})} + \sum_{\ell=\lfloor k/5 \rfloor+1}^{\lfloor k/3 \rfloor} \|m[a_{k,\ell}](D;\,\cdot\,)f\|_{L^p(\mathbb{R}^{3+1})}.$$

For $2 \leqslant p \leqslant 4$ we may bound the terms of the first sum using Proposition 7.2 and the terms of the second using Proposition 7.1. If, in addition, we assume $p \geqslant 3$, then the geometric series resulting from the constants can be evaluated to give the desired bound. □

8. Proof of Theorem 3.1: the main argument

By the observations of the previous section, the problem is reduced to establishing Proposition 7.1. In this section we provide the details of the proof, following the scheme sketched in §1.3.

8.1. **Localisation along the curve.** We begin by further decomposing the symbols with respect to the distance of the s-variable to the roots $\theta_1^{\pm}$ and $\theta_2(\xi)$. Here it is convenient to introduce a 'fine tuning' constant $\rho > 0$. This is a small (but absolute) constant which plays a minor technical rôle in the forthcoming arguments: taking $\rho := 10^{-6}$ more than suffices for our purposes.

Recall from Lemma 6.2 that the two distinct roots $\theta_1^{\pm}(\xi)$ only occur when $u(\xi) < 0$. In view of this, let $\beta^{>0}, \beta^{<0} \in C_c^\infty(\mathbb{R})$ be the unique functions with $\operatorname{supp} \beta^{>0} \subset (0,\infty)$ and $\operatorname{supp} \beta^{<0} \subset (-\infty, 0)$ such that $\beta = \beta^{>0} + \beta^{<0}$. This induces a corresponding decomposition $a_{k,\ell} = a_{k,\ell}^{>0} + a_{k,\ell}^{<0}$ for $0 \leqslant \ell < \lfloor k/3 \rfloor$, where $u(\xi)$ is positive (respectively, negative) on the support of $a_{k,\ell}^{>0}$ (respectively, $a_{k,\ell}^{<0}$). Given $\varepsilon > 0$, define

$$a_{k,\ell}^{(\varepsilon),\pm}(\xi;t;s) := a_{k,\ell}^{<0}(\xi;t;s)\,\eta\big(\rho^{-1}2^{(k-\ell)/2}2^{-k\varepsilon}|s - \theta_1^{\pm}(\xi)|\big) \quad \text{if } 0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$$

and

$$a_{k,\ell}^{(\varepsilon)}(\xi;t;s) := \begin{cases} \sum_{\pm} a_{k,\ell}^{(\varepsilon),\pm}(\xi;t;s) & \text{if } 0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon \\ a_{k,\ell}(\xi;t;s)\,\eta\big(\rho 2^{\ell(1-\varepsilon)}|s - \theta_2(\xi)|\big) & \text{if } \lfloor k/3 \rfloor_\varepsilon \leqslant \ell \leqslant \lfloor k/3 \rfloor \end{cases}, \tag{8.1}$$

where $\lfloor k/3 \rfloor_\varepsilon := \lfloor \big(\frac{1-\varepsilon}{3}\big) \cdot k \rfloor$ is a number we think of as being slightly smaller than $\lfloor k/3 \rfloor$. Note that

$$\min_{\pm} |s - \theta_1^{\pm}(\xi)| \lesssim \rho 2^{-(k-\ell)/2 + k\varepsilon} \qquad \text{for all } (\xi; t; s) \in \operatorname{supp} a_{k,\ell}^{(\varepsilon)} \quad \text{if } 0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon.$$

Remark. The symbols $a_{k,\ell}^{(\varepsilon),+}$ and $a_{k,\ell}^{(\varepsilon),-}$ have disjoint supports if $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$. Indeed, the decomposition ensures that $|u(\xi)| \sim 2^{k-2\ell}$ for all $\xi \in \operatorname{supp}_\xi a_{k,\ell}^{(\varepsilon)}$ and so Lemma 6.3 implies

$$|\theta_1^-(\xi) - \theta_1^+(\xi)| \gtrsim 2^{-\ell} \gtrsim 2^{-(k-\ell)/2} 2^{k\varepsilon}.$$

Here we use the hypothesis $\ell < \lfloor k/3 \rfloor_\varepsilon$. Provided ρ is chosen to be sufficiently small, the above separation condition ensures that the disjointness of the supports of $a_{k,\ell}^{(\varepsilon),+}$ and $a_{k,\ell}^{(\varepsilon),-}$. Consequently,

$$\min_{\pm} |s - \theta_1^{\pm}(\xi)| \gtrsim 2^{-(k-\ell)/2 + k\varepsilon} \qquad \text{for all } (\xi; t; s) \in \operatorname{supp} (a_{k,\ell}^{<0} - a_{k,\ell}^{(\varepsilon)})$$

if $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$.

The main contribution to $m[a_{k,\ell}]$ comes from the symbols $a_{k,\ell}^{(\varepsilon)}$.

Lemma 8.1. *Let $2 \leqslant p < \infty$ and $\varepsilon > 0$. For all $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$*

$$\|m[a_{k,\ell} - a_{k,\ell}^{(\varepsilon)}](D; \cdot) f\|_{L^p(\mathbb{R}^{3+1})} \lesssim_{N,\varepsilon,p} 2^{-kN} \|f\|_{L^p(\mathbb{R}^3)} \qquad \textit{for all } N \in \mathbb{N}.$$

Proof. It is clear that the multipliers satisfy a trivial L^∞-estimate with operator norm $O(2^{Ck})$ for some absolute constant $C \geqslant 1$. Thus, by interpolation, it suffices to prove the rapid decay estimate for $p = 2$ only. This amounts to showing that, under the hypotheses of the lemma,

$$\|m[a_{k,\ell} - a_{k,\ell}^{(\varepsilon)}](\,\cdot\,; t)\|_{L^\infty(\hat{\mathbb{R}}^3)} \lesssim_{N,\varepsilon} 2^{-kN} \qquad \text{for all } N \in \mathbb{N} \tag{8.2}$$

uniformly in $1/2 \leqslant t \leqslant 4$.

<u>Case: $\lfloor k/3 \rfloor_\varepsilon \leqslant \ell \leqslant \lfloor k/3 \rfloor$.</u> Here the localisation of the $a_{k,\ell}$ and $a_{k,\ell}^{(\varepsilon)}$ symbols ensures that

$$|u(\xi)| \lesssim 2^{k-2\ell} \quad \text{and} \quad |s - \theta_2(\xi)| \gtrsim \rho^{-1} 2^{-\ell(1-\varepsilon)} \quad \text{for all } (\xi; t; s) \in \operatorname{supp} (a_{k,\ell} - a_{k,\ell}^{(\varepsilon)}), \tag{8.3}$$

where u is the function introduced in (6.5).

Fix $\xi \in \operatorname{supp}_\xi (a_{k,\ell} - a_{k,\ell}^{(\varepsilon)})$ and consider the oscillatory integral $m[a_{k,\ell} - a_{k,\ell}^{(\varepsilon)}](\xi; t)$, which has phase $s \mapsto t\langle \gamma(s), \xi \rangle$. Taylor expansion around $\theta_2(\xi)$ yields

$$\langle \gamma'(s), \xi \rangle = u(\xi) + \omega_1(\xi; s)\,(s - \theta_2(\xi))^2 \tag{8.4}$$

$$\langle \gamma''(s), \xi \rangle = \omega_2(\xi; s)\,(s - \theta_2(\xi)) \tag{8.5}$$

where ω_i arise from the remainder terms and satisfy $|\omega_i(\xi; s)| \sim 2^k$. Provided ρ is sufficiently small, (8.3) implies that the $\omega_1(\xi; s)\,(s - \theta_2(\xi))^2$ term dominates the right-hand side of (8.4) and therefore

$$|\langle \gamma'(s), \xi \rangle| \gtrsim 2^k |s - \theta_2(\xi)|^2 \qquad \text{for all } (\xi; t; s) \in \operatorname{supp} (a_{k,\ell} - a_{k,\ell}^{(\varepsilon)}). \tag{8.6}$$

Furthermore, (8.5), (8.6) and the localisation (8.3) immediately imply

$$|\langle \gamma''(s), \xi \rangle| \lesssim 2^{-k + 3\ell(1-\varepsilon)} |\langle \gamma'(s), \xi \rangle|^2,$$
$$|\langle \gamma^{(j)}(s), \xi \rangle| \lesssim 2^k \lesssim_j 2^{-(k - 3\ell(1-\varepsilon))(j-1)} |\langle \gamma'(s), \xi \rangle|^j \qquad \text{for all } j \geqslant 3$$

and all $(\xi; t; s) \in \operatorname{supp} (a_{k,\ell} - a_{k,\ell}^{(\varepsilon)})$, where in the last inequality we have used $|s - \theta_2(\xi)|^{j-3} \lesssim 1$ for all $j \geqslant 3$.

On the other hand, by the definition of the symbols, (8.6) and the localisation in (8.3),

$$|\partial_s^N (a_{k,\ell} - a_{k,\ell}^{(\varepsilon)})(\xi; s)| \lesssim_N 2^{\ell(1-\varepsilon)N} \lesssim 2^{-(k-3\ell)N - 3\varepsilon\ell N} |\langle \gamma'(s), \xi \rangle|^N \qquad \text{for all } N \in \mathbb{N}$$

and all $(\xi;t;s) \in \operatorname{supp}(a_{k,\ell} - a_{k,\ell}^{(\varepsilon)})$. Thus, by repeated integration-by-parts (via Lemma D.1, with $r = 2^{k-3\ell+3\varepsilon\ell} \geqslant 1$ for $0 \leqslant \ell \leqslant k/3$), one concludes that

$$|m[a_{k,\ell} - a_{k,\ell}^{(\varepsilon)}](\xi;t)| \lesssim_N 2^{-(k-3\ell)N - 3\varepsilon\ell N} \qquad \text{for all } N \in \mathbb{N}$$

uniformly in $1/2 \leqslant t \leqslant 4$. Since $\lfloor k/3 \rfloor_\varepsilon \leqslant \ell \leqslant \lfloor k/3 \rfloor \leqslant k/3$, the desired bound follows.

Case: $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$. If $u(\xi) > 0$, then (6.3) and (8.4) imply

$$|\langle \gamma'(s), \xi \rangle| \gtrsim |u(\xi)| + 2^k |s - \theta_2(\xi)|^2 \qquad \text{for all } (\xi; s) \in \operatorname{supp} a_{k,\ell}^{>0}.$$

Furthermore, the localisation of the symbol $a_{k,\ell}^{>0}$ guarantees that $u(\xi) \sim 2^{k-\ell}$ for all $\xi \in \operatorname{supp} a_{k,\ell}^{>0}$. It is then a straightforward exercise to adapt the argument used in the previous case to show $\|m[a_{k,\ell}^{>0}](\,\cdot\,;t)\|_\infty \lesssim_{N,\varepsilon} 2^{-kN}$, splitting the analysis into the cases $|s - \theta_2(\xi)| \geqslant 2^{-\ell}$ and $|s - \theta_2(\xi)| \leqslant 2^{-\ell}$. Here we use the fact that $2^{-(k-3\ell)} \leqslant 2^{-\varepsilon k}$.

Thus, the problem is reduced to proving

$$\|m[a_{k,\ell}^{<0} - a_{k,\ell}^{(\varepsilon)}](\,\cdot\,;t)\|_{L^\infty(\hat{\mathbb{R}}^3)} \lesssim_{N,\varepsilon} 2^{-kN}.$$

Here the localisation of the $a_{k,\ell}^{<0}$ and $a_{k,\ell}^{(\varepsilon)}$ symbols ensures that

$$|u(\xi)| \sim 2^{k-2\ell} \quad \text{and} \quad \min_{\pm} |s - \theta_1^{\pm}(\xi)| \gtrsim 2^{-(k-\ell)/2 + k\varepsilon} \quad \text{for all } (\xi;t;s) \in \operatorname{supp}(a_{k,\ell}^{<0} - a_{k,\ell}^{(\varepsilon)}), \tag{8.7}$$

where u is the function introduced in (6.5).

Fix $\xi \in \operatorname{supp}_\xi (a_{k,\ell}^{<0} - a_{k,\ell}^{(\varepsilon)})$ and consider the oscillatory integral $m[a_{k,\ell}^{<0} - a_{k,\ell}^{(\varepsilon)}](\xi;t)$, which has phase $s \mapsto t\langle \gamma(s), \xi \rangle$. If we define

$$\phi \colon [-1,1] \to \mathbb{R}, \quad \phi \colon s \mapsto \langle \gamma'(s), \xi \rangle,$$

then, by (6.3), this function is strictly convex. Thus, given $t \in [-1,1]$, the auxiliary function

$$q_t \colon [-1,1] \to \mathbb{R}, \quad q_t \colon s \mapsto \frac{\phi(s) - \phi(t)}{s - t} \quad \text{for } s \neq t \quad \text{and} \quad q_t \colon t \mapsto \phi'(t)$$

is increasing. Setting $t := \theta_1^-(\xi)$ and noting that $\phi \circ \theta_1^-(\xi) = 0$, it follows that

$$\frac{\phi(s)}{s - \theta_1^-(\xi)} \leqslant \frac{\phi \circ \theta_2(\xi)}{\theta_2(\xi) - \theta_1^-(\xi)} = \frac{u(\xi)}{\theta_2(\xi) - \theta_1^-(\xi)} < 0 \qquad \text{for all } -1 \leqslant s \leqslant \theta_2(\xi),$$

where we have used the fact that $u(\xi) < 0$ on the support of $a_{k,\ell}^{<0}$. If $s \in [\theta_2(\xi), 1]$, then we can carry out the same argument with respect to $t = \theta_1^+(\xi)$ to obtain a similar inequality. From this, we deduce the bound

$$|\langle \gamma'(s), \xi \rangle| \geqslant \min_{\pm} \frac{|u(\xi)||s - \theta_1^{\pm}(\xi)|}{|\theta_2(\xi) - \theta_1^{\pm}(\xi)|} \qquad \text{for all } -1 \leqslant s \leqslant 1. \tag{8.8}$$

Recall from (8.7) that $|u(\xi)| \sim 2^{k-2\ell}$ and therefore $|\theta_2(\xi) - \theta_1^{\pm}(\xi)| \sim 2^{-\ell}$ by Lemma 6.3. Substituting these bounds and the second bound in (8.7) into (8.8), we conclude that

$$|\langle \gamma'(s), \xi \rangle| \gtrsim 2^{k-\ell} \min_{\pm} |s - \theta_1^{\pm}(\xi)| \gtrsim 2^{(k-\ell)/2 + \varepsilon k} \qquad \text{for all } (\xi;t;s) \in \operatorname{supp}(a_{k,\ell}^{<0} - a_{k,\ell}^{(\varepsilon)}). \tag{8.9}$$

Furthermore, by the mean value theorem,

$$|\langle \gamma''(s), \xi \rangle| \lesssim \max_{\pm} |v^{\pm}(\xi)| + 2^k \min_{\pm} |s - \theta_1^{\pm}(\xi)| \lesssim 2^{k-\ell} + 2^\ell |\langle \gamma'(s), \xi \rangle| \lesssim 2^{-k\varepsilon} |\langle \gamma'(s), \xi \rangle|^2,$$

where we have used (8.9), the condition $|v^{\pm}(\xi)| \sim 2^{k-\ell}$ for $\xi \in \operatorname{supp} a_{k,\ell}^{<0}$ from Lemma 6.3 and $0 \leqslant \ell \leqslant k/3$ in the last inequality. For higher order derivatives,

$$|\langle \gamma^{(j)}(s), \xi \rangle| \lesssim_j 2^k \lesssim_j 2^{-(j-1)k\varepsilon} |\langle \gamma'(s), \xi \rangle|^j \qquad \text{for all } j \geqslant 3$$

and all $(\xi;t;s) \in \operatorname{supp}(a_{k,\ell}^{<0} - a_{k,\ell}^{(\varepsilon)})$. On the other hand, by the definition of the symbols and (8.9) we have

$$|\partial_s^N (a_{k,\ell} - a_{k,\ell}^{(\varepsilon)})(\xi;s)| \lesssim_N 2^{N(k-\ell)/2} 2^{-Nk\varepsilon} \lesssim 2^{-2Nk\varepsilon} |\langle \gamma'(s), \xi\rangle|^N \qquad \text{for all } N \in \mathbb{N}$$

and all $(\xi;t;s) \in \operatorname{supp}(a_{k,\ell}^{<0} - a_{k,\ell}^{(\varepsilon)})$. Thus, by repeated integration-by-parts (via Lemma D.1, with $r := 2^{k\varepsilon/2} \geqslant 1$), one obtains the desired bound (8.2). □

8.2. **Fourier localisation.** We perform a radial decomposition of the symbols $a_{k,\ell}^{(\varepsilon)}$ with respect to the homogeneous functions θ_2 and $\theta_1^{\pm}$. Fix $\zeta \in C^\infty(\mathbb{R})$ with $\operatorname{supp} \zeta \subseteq [-1,1]$ such that $\sum_{l \in \mathbb{Z}} \zeta(\,\cdot - l) \equiv 1$. For $k \in \mathbb{N}$ and $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$, write

$$a_{k,\ell}^{(\varepsilon)} = \sum_{\pm} \sum_{\nu \in \mathbb{Z}} a_{k,\ell}^{\nu,(\varepsilon),\pm}$$

where

$$a_{k,\ell}^{\nu,(\varepsilon),\pm}(\xi;t;s) := a_{k,\ell}^{(\varepsilon),\pm}(\xi;t;s)\, \zeta\big(\rho^{-1}(2^{(k-\ell)/2}\theta_1^{\pm}(\xi) - \nu)\big) \qquad \text{if } 0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon.$$

Each of the two terms in $\sum_{\pm}$ can be treated analogously. In order to simplify the notation, we drop the symbol $\pm$ from $a_{k,\ell}^{\nu,(\varepsilon)}$ and $\theta_1^{\pm}$ and adopt the convention

$$a_{k,\ell}^{(\varepsilon)} = \sum_{\nu \in \mathbb{Z}} a_{k,\ell}^{\nu,(\varepsilon)}. \tag{8.10}$$

The key properties of this decomposition are that

$$|s - \theta_1(\xi)| \lesssim \rho 2^{-(k-\ell)/2 + k\varepsilon} \quad \text{and} \quad |\theta_1(\xi) - s_\nu| \lesssim \rho 2^{-(k-\ell)/2} \quad \text{for all } (\xi;t;s) \in \operatorname{supp} a_{k,\ell}^{\nu,(\varepsilon)}, \tag{8.11}$$

where $s_\nu := 2^{-(k-\ell)/2}\nu$ and $\theta_1 \in \{\theta_1^+(\xi), \theta_1^-(\xi)\}$. The decomposition (8.10) is extended to the range $\lfloor k/3 \rfloor_\varepsilon \leqslant \ell \leqslant \lfloor k/3 \rfloor$, with

$$a_{k,\ell}^{\nu,(\varepsilon)}(\xi;t;s) := a_{k,\ell}^{(\varepsilon)}(\xi;t;s)\, \zeta(2^\ell \theta_2(\xi) - \nu) \qquad \text{if } \lfloor k/3 \rfloor_\varepsilon \leqslant \ell \leqslant \lfloor k/3 \rfloor. \tag{8.12}$$

In the case $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$ we also consider symbols formed by grouping the $a_{k,\ell}^{\nu,(\varepsilon)}$ into pieces at the larger scale $2^{-\ell}$. Given $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$ we write $\mathbb{Z} = \bigcup_{\mu \in \mathbb{Z}} \mathfrak{N}_\ell(\mu)$, where the sets $\mathfrak{N}_\ell(\mu)$ are disjoint and satisfy

$$\mathfrak{N}_\ell(\mu) \subseteq \{\nu \in \mathbb{Z} : |\nu - 2^{(k-3\ell)/2}\mu| \leqslant 2^{(k-3\ell)/2}\}.$$

For each $\mu \in \mathbb{Z}$, we then define

$$a_{k,\ell}^{*,\mu,(\varepsilon)} := \sum_{\nu \in \mathfrak{N}_\ell(\mu)} a_{k,\ell}^{\nu,(\varepsilon)}$$

and note that $|\theta_1^{\pm}(\xi) - s_\mu| \lesssim 2^{-\ell}$ on $\operatorname{supp}_\xi a_{k,\ell}^{*,\mu,(\varepsilon)}$, where $s_\mu := 2^{-\ell}\mu$. Of course, by the definition of the sets $\mathfrak{N}_\ell(\mu)$,

$$a_{k,\ell}^{(\varepsilon)} = \sum_{\mu \in \mathbb{Z}} a_{k,\ell}^{*,\mu,(\varepsilon)} = \sum_{\mu \in \mathbb{Z}} \sum_{\nu \in \mathfrak{N}_\ell(\mu)} a_{k,\ell}^{\nu,(\varepsilon)}.$$

It is notationally convenient to trivially extend these definitions by setting $\mathfrak{N}_\ell(\mu) := \{\mu\}$ for $\lfloor k/3 \rfloor_\varepsilon \leqslant \ell \leqslant \lfloor k/3 \rfloor$ and, in this case, defining $a_{k,\ell}^{*,\mu,(\varepsilon)} := a_{k,\ell}^{\mu,(\varepsilon)}$ accordingly.

Given $0 < r \leqslant 1$ and $s \in I$, recall the definition of the $(1,r)$-*Frenet boxes* $\pi_1(s;\, r)$ introduced in Definition 5.1:

$$\pi_1(s;\, r) := \big\{\xi \in \widehat{\mathbb{R}}^3 : |\langle \mathbf{e}_j(s), \xi\rangle| \lesssim r^{3-j} \text{ for } j = 1,\, 2, \quad |\langle \mathbf{e}_3(s), \xi\rangle| \sim 1\big\}.$$

It is also convenient to consider 2-parameter variants of the $(0, r)$-Frenet boxes. Given $0 < r_1, r_2$ and $s \in I$, define the set

$$\pi_0(s;\, r_1, r_2) := \big\{\xi \in \widehat{\mathbb{R}}^3 : |\langle \mathbf{e}_1(s), \xi\rangle| \lesssim r_1,\ |\langle \mathbf{e}_2(s), \xi\rangle| \sim 1,\ |\langle \mathbf{e}_3(s), \xi\rangle| \lesssim r_2\big\}.$$

The geometric significance of these sets is made apparent in §8.6 (and, in particular, Lemma 8.9) below.

The multipliers $a_{k,\ell}^{*,\mu,(\varepsilon)}$ and $a_{k,\ell}^{\nu,(\varepsilon)}$ satisfy the following support properties.

Lemma 8.2. *For all $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$, $\varepsilon > 0$ and $\mu, \nu \in \mathbb{Z}$,*

a) If $\nu \in \mathfrak{N}_\ell(\mu)$, then $\mathrm{supp}_\xi\, a_{k,\ell}^{\nu,(\varepsilon)} \subseteq 2^k \cdot \pi_1(s_\mu; 2^{-\ell})$, where $s_\mu := 2^{-\ell}\mu$;

b) If $\ell < \lfloor k/3 \rfloor_\varepsilon$, then $\mathrm{supp}_\xi\, a_{k,\ell}^{\nu,(\varepsilon)} \subseteq 2^{k-\ell} \cdot \pi_0(s_\nu; 2^{-(k-\ell)/2}, 2^\ell)$, where $s_\nu := 2^{-(k-\ell)/2}\nu$.

As an immediate consequence of part a), we see that $\mathrm{supp}_\xi\, a_{k,\ell}^{*,\mu,(\varepsilon)} \subseteq 2^k \cdot \pi_1(s_\mu; 2^{-\ell})$.

Proof of Lemma 8.2. a) For $\xi \in \mathrm{supp}_\xi\, a_{k,\ell}^{\nu,(\varepsilon)}$ observe that the localisation in (7.1) implies

$$|\langle \gamma^{(i)} \circ \theta_2(\xi), \xi\rangle| \lesssim 2^{k-(3-i)\ell} \qquad \text{for } i = 1,\, 2, \qquad |\langle \gamma^{(3)} \circ \theta_2(\xi), \xi\rangle| \sim 2^k.$$

If $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$, then $|s_\nu - \theta_1(\xi)| \lesssim 2^{-(k-\ell)/2}$ and so

$$|s_\mu - \theta_2(\xi)| \leqslant |s_\mu - s_\nu| + |s_\nu - \theta_1(\xi)| + |\theta_1(\xi) - \theta_2(\xi)| \lesssim 2^{-(k-\ell)/2} + 2^{-\ell} \lesssim 2^{-\ell}$$

by Lemma 6.3. Note that the inequality $|s_\mu - \theta_2(\xi)| \lesssim 2^{-\ell}$ also extends to the case $\lfloor k/3 \rfloor_\varepsilon \leqslant \ell \leqslant \lfloor k/3 \rfloor$ in view of the definition of the symbol from (8.12). Taylor expansion around $\theta_2(\xi)$ therefore yields

$$|\langle \gamma^{(i)}(s_\mu), \xi\rangle| \lesssim 2^{k-(3-i)\ell} \qquad \text{for } i = 1,\, 2, \qquad |\langle \gamma^{(3)}(s_\mu), \xi\rangle| \sim 2^k.$$

Since the Frenet vectors $\mathbf{e}_i(s_\mu)$ are obtained from the $\gamma^{(i)}(s_\mu)$ via the Gram–Schmidt process, the matrix corresponding to change of basis from $\big(\mathbf{e}_i(s_\mu)\big)_{i=1}^3$ to $\big(\gamma^{(i)}(s_\mu)\big)_{i=1}^3$ is lower triangular. Furthermore, the initial localisation implies that this matrix is an $O(\delta_0)$ perturbation of the identity. Consequently,

$$|\langle \mathbf{e}_i(s_\mu), \xi\rangle| \lesssim 2^{k-(3-i)\ell} \qquad \text{for } 1 \leqslant i \leqslant 3.$$

Provided the parameter $\delta_0 > 0$ is sufficiently small, the argument can easily be adapted to prove the remaining lower bound $|\langle \mathbf{e}_3(s_\mu), \xi\rangle| \gtrsim 1$.

b) Let $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$. For $\xi \in \mathrm{supp}_\xi\, a_{k,\ell}^{\nu,(\varepsilon)}$ observe that the localisation in (7.1) and Lemma 6.3 imply

$$|\langle \gamma' \circ \theta_1(\xi), \xi\rangle| = 0, \qquad |\langle \gamma'' \circ \theta_1(\xi), \xi\rangle| \sim 2^{k-\ell}, \qquad |\langle \gamma^{(3)} \circ \theta_1(\xi), \xi\rangle| \sim 2^k.$$

It then follows from Taylor expansion around $\theta_1(\xi)$ that

$$|\langle \gamma'(s_\nu), \xi\rangle| \lesssim 2^{(k-\ell)/2}, \quad |\langle \gamma''(s_\nu), \xi\rangle| \sim 2^{k-\ell} \quad \text{and} \quad |\langle \gamma^{(3)}(s_\nu), \xi\rangle| \sim 2^k,$$

provided ρ is chosen sufficiently small. The $\gamma^{(j)}(s_\nu)$ in the above estimates can then be replaced with the Frenet vectors $\mathbf{e}_j(s_\nu)$ by a similar argument to that used in part a). □

8.3. **Spatio-temporal Fourier localisation.** The symbols are further localised with respect to the Fourier transform of the t-variable. In particular, let

$$q(\xi) := \langle \gamma \circ \theta_2(\xi), \xi\rangle \qquad \text{and} \qquad \chi_{k,\ell}^{(\varepsilon)}(\xi, \tau) := \eta\big(2^{-(k-3\ell)-4\varepsilon k}(\tau + q(\xi))\big)$$

and define the multiplier $m_{k,\ell}^{\nu,(\varepsilon)}$ by

$$\mathcal{F}_t\big[m_{k,\ell}^{\nu,(\varepsilon)}(\xi;\,\cdot\,)\big](\tau) := \chi_{k,\ell}^{(\varepsilon)}(\xi, \tau)\, \mathcal{F}_t\big[m[a_{k,\ell}^{\nu,(\varepsilon)}](\xi;\,\cdot\,)\big](\tau).$$

Here $\mathcal{F}_t$ denotes the Fourier transform acting in the t variable. Define $m_{k,\ell}^{*,\mu,(\varepsilon)}$ and $m_{k,\ell}^{(\varepsilon)}$ accordingly by setting

$$m_{k,\ell}^{*,\mu,(\varepsilon)} := \sum_{\nu \in \mathfrak{N}_\ell(\mu)} m_{k,\ell}^{\nu,(\varepsilon)} \qquad \text{and} \qquad m_{k,\ell}^{(\varepsilon)} := \sum_{\mu \in \mathbb{Z}} m_{k,\ell}^{*,\mu,(\varepsilon)}.$$

The main contribution to $m[a_{k,\ell}^{\nu,(\varepsilon)}]$ comes from the multipliers $m_{k,\ell}^{\nu,(\varepsilon)}$.

Lemma 8.3. *Let $1 \leqslant p \leqslant \infty$ and $\varepsilon > 0$. For all $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$,*

$$\big\| \big(m[a_{k,\ell}^{\nu,(\varepsilon)}] - m_{k,\ell}^{\nu,(\varepsilon)}\big)(D;\,\cdot\,) f \big\|_{L^p(\mathbb{R}^{3+1})} \lesssim_{N,\varepsilon} 2^{-kN} \|f\|_{L^p(\mathbb{R}^3)} \qquad \text{for all } N \in \mathbb{N}.$$

Proof. It suffices to show that

$$|\partial_\xi^\alpha \big(m[a_{k,\ell}^{\nu,(\varepsilon)}] - m_{k,\ell}^{\nu,(\varepsilon)}\big)(\xi;t)| \lesssim_{N,\varepsilon} 2^{-kN}(1+|t|)^{-10} \qquad \text{for } \alpha \in \mathbb{N}_0^3,\ |\alpha| \leqslant 10, \text{ and } N \in \mathbb{N}. \tag{8.13}$$

Indeed, if (8.13) holds, then Fourier inversion and repeated integration-by-parts imply

$$|\big(m[a_{k,\ell}^{\nu,(\varepsilon)}] - m_{k,\ell}^{\nu,(\varepsilon)}\big)(D;t) f(x)| \lesssim_{N,\varepsilon} 2^{-kN}(1+|t|)^{-10}(1+|\cdot|)^{-10} * f(x).$$

Taking the $L^p(\mathbb{R}^{3+1})$-norm of both sides of this inequality immediately yields the desired result.

By the Fourier inversion formula

$$\big(m[a_{k,\ell}^{\nu,(\varepsilon)}] - m_{k,\ell}^{\nu,(\varepsilon)}\big)(\xi;t) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{it\tau}\big(1-\chi_{k,\ell}^{(\varepsilon)}(\xi,\tau)\big)\,\mathcal{F}_t\big[m[a_{k,\ell}^{\nu,(\varepsilon)}](\xi;\,\cdot\,)\big](\tau)\,\mathrm{d}\tau.$$

Let $\Xi = (\xi,\tau) \in \widehat{\mathbb{R}}^{3+1}$ denote the spatio-temporal frequency variables. Clearly, there exists a constant $C \geqslant 1$ such that $|\partial_\Xi^\alpha \chi_{k,\ell}^{(\varepsilon)}(\Xi)| \lesssim 2^{Ck}$ for all $\alpha \in \mathbb{N}_0^4$ with $|\alpha| \leqslant 20$. Furthermore, if $(\xi,\tau) \in \operatorname{supp} \partial_\Xi^\alpha\big(1-\chi_{k,\ell}^{(\varepsilon)}\big)$, then $|\tau + q(\xi)| \gtrsim 2^{-k+3\ell+4\varepsilon k}$. Thus, by integration-by-parts in the τ-variable, to prove (8.13) it suffices to show

$$|\partial_\Xi^\alpha \mathcal{F}\big[m[a_{k,\ell}^{\nu,(\varepsilon)}](\xi;\,\cdot\,)\big](\tau)| \lesssim_{N,\varepsilon} 2^{Ck}\big(1+2^{-k+3\ell+3\varepsilon k}|\tau+q(\xi)|\big)^{-N}, \quad \alpha \in \mathbb{N}_0^4,\ |\alpha| \leqslant 20,\ N \in \mathbb{N}, \tag{8.14}$$

for some choice of absolute constant $C \geqslant 1$ (not necessarily the same as above).

By the Leibniz rule,

$$\partial_\Xi^\alpha \mathcal{F}_t\big[m[a_{k,\ell}^{\nu,(\varepsilon)}](\xi;\,\cdot\,)\big](\tau) = \int_{\mathbb{R}} e^{-ir(\tau+q(\xi))} m[b_{k,\ell}^{\nu,(\varepsilon),\alpha}](\xi;r)\,\mathrm{d}r \tag{8.15}$$

where $b_{k,\ell}^{\nu,(\varepsilon),\alpha}(\xi;r;s) := e^{irq(\xi)} a_{k,\ell}^{\nu,(\varepsilon),\alpha}(\xi;r;s)$ for some symbol $a_{k,\ell}^{\nu,(\varepsilon),\alpha}$ satisfying

$$\big|\partial_r^j a_{k,\ell}^{\nu,(\varepsilon),\alpha}(\xi;r;s)\big| \lesssim_j 2^{Ck} \qquad \text{for all } j \in \mathbb{N}_0,\ \alpha \in \mathbb{N}_0^4,\ |\alpha| \leqslant 20,\ |r| \lesssim 1 \tag{8.16}$$

and with $\operatorname{supp} a_{k,\ell}^{\nu,(\varepsilon),\alpha} \subseteq \operatorname{supp} a_{k,\ell}^{\nu,(\varepsilon)}$. Note, in particular, that

$$m[b_{k,\ell}^{\nu,(\varepsilon),\alpha}](\xi;r) = \int_{\mathbb{R}} e^{-ir\langle \gamma(s) - \gamma\circ\theta_2(\xi), \xi\rangle} a_{k,\ell}^{\nu,(\varepsilon),\alpha}(\xi;r;s)\rho(r)\chi(s)\,\mathrm{d}s. \tag{8.17}$$

By Taylor expansion around $\theta_2(\xi)$, the phase in (8.17) can be written as

$$\langle \gamma(s) - \gamma\circ\theta_2(\xi), \xi\rangle = u(\xi)\,(s-\theta_2(\xi)) + \omega(\xi;s)\,(s-\theta_2(\xi))^3 \tag{8.18}$$

where ω arises from the remainder term and satisfies $|\omega(\xi;s)| \sim 2^k$. Recall,

$$|u(\xi)| \lesssim 2^{k-2\ell} \qquad \text{and} \qquad |s-\theta_2(\xi)| \lesssim 2^{-\ell+\varepsilon k} \qquad \text{for all } \ (\xi;r;s) \in \operatorname{supp} a_{k,\ell}^{\nu,(\varepsilon)}, \tag{8.19}$$

which follows from the definition of $a_{k,\ell}^{\nu,(\varepsilon)}$. Here, in the case $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$, we use Lemma 6.3 to deduce that

$$|s-\theta_2(\xi)| \leqslant |s-\theta_1(\xi)| + |\theta_1(\xi)-\theta_2(\xi)| \lesssim 2^{-\ell}.$$

Combining the expansion (8.18) and the localisation (8.19) yields

$$|\langle \gamma(s) - \gamma \circ \theta_2(\xi), \xi\rangle| \lesssim 2^{k-3\ell+3\varepsilon k} \qquad \text{for all } (\xi; r; s) \in \operatorname{supp} a_{k,\ell}^{\nu,(\varepsilon)}. \tag{8.20}$$

By (8.20), (8.16) and integration by parts in (8.15), one obtains

$$|\partial_\Xi^\alpha \mathcal{F}_t\big[m[a_{k,\ell}^{\nu,(\varepsilon)}](\xi;\,\cdot\,)\big](\tau)| \lesssim_M 2^{Ck}|\tau + q(\xi)|^{-M} 2^{(k-3\ell+3\varepsilon k)M} \qquad \text{for all } M \in \mathbb{N}$$

and all $\alpha \in \mathbb{N}_0^4$, $|\alpha| \leqslant 20$. This implies (8.14) and concludes the proof. ☐

To understand the support properties of the multipliers $m_{k,\ell}^{*,\mu,(\varepsilon)}$, we introduce the primitive curve

$$\bar{\gamma}\colon I \to \mathbb{R}^4, \qquad \bar{\gamma}\colon s \mapsto \begin{bmatrix} \int_0^s \gamma \\ s \end{bmatrix}.$$

Here $\int_0^s \gamma$ denotes the vector in $\mathbb{R}^3$ with ith component $\int_0^s \gamma_i$ for $1 \leqslant i \leqslant 3$. Note that $\bar{\gamma}$ is a non-degenerate curve in $\mathbb{R}^4$ and, in particular, $|\det(\bar{\gamma}^{(1)} \cdots \bar{\gamma}^{(4)})| = |\det(\gamma^{(1)} \cdots \gamma^{(3)})|$. Let $(\bar{\mathbf{e}}_j(s))_{j=1}^4$ denote the Frenet frame associated to $\bar{\gamma}$ and consider the $(2,r)$-Frenet boxes for $\bar{\gamma}$

$$\pi_{2,\bar{\gamma}}(s;\, r) := \big\{\Xi = (\xi,\tau) \in \widehat{\mathbb{R}}^3 \times \widehat{\mathbb{R}} : |\langle \bar{\mathbf{e}}_j(s), \Xi\rangle| \lesssim r^{4-j} \text{ for } 1 \leqslant j \leqslant 3,\ |\langle \bar{\mathbf{e}}_4(s), \Xi\rangle| \sim 1\big\},$$

as introduced in Definition 5.1.

Lemma 8.4. *For all $\lceil 4\varepsilon k\rceil \leqslant \ell \leqslant \lfloor k/3\rfloor$ and $\mu \in \mathbb{Z}$,*

$$\operatorname{supp} \mathcal{F}_t\big[m_{k,\ell}^{*,\mu,(\varepsilon)}\big] \subseteq 2^k \cdot \pi_{2,\bar{\gamma}}(s_\mu; 2^{4\varepsilon k}2^{-\ell}),$$

where $s_\mu := 2^{-\ell}\mu$ and $\mathcal{F}_t$ denotes the Fourier transform in the t-variable.

Proof. If $\Xi = (\xi,\tau) \in \operatorname{supp} \mathcal{F}_t\big[m_{k,\ell}^{*,\mu,(\varepsilon)}\big]$, then $\xi \in \operatorname{supp}_\xi a_{k,\ell}^{*,\mu,(\varepsilon)}$ and $|q(\xi) + \tau| \lesssim 2^{4\varepsilon k}2^{k-3\ell}$. The former condition implies $|u(\xi)| \lesssim 2^{k-2\ell}$ and $|s - \theta_2(\xi)| \lesssim 2^{-\ell+\varepsilon k}$ (see (8.19)) and so, by Taylor expansion around $\theta_2(\xi)$,

$$|\langle \gamma(s_\mu), \xi\rangle + \tau| \lesssim |q(\xi) + \tau| + |u(\xi)||s - \theta_2(\xi)| + 2^k|s - \theta_2(\xi)|^3 \lesssim 2^{4\varepsilon k}2^{k-3\ell}. \tag{8.21}$$

Define the lifted curve and frame

$$\gamma_\uparrow\colon I \to \mathbb{R}^4, \quad \gamma_\uparrow\colon s \mapsto \begin{bmatrix} \gamma(s) \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{e}_{j,\uparrow}\colon I \to S^3, \quad \mathbf{e}_{j,\uparrow}\colon s \mapsto \begin{bmatrix} \mathbf{e}_j(s) \\ 0 \end{bmatrix} \quad \text{for } 1 \leqslant j \leqslant 3,$$

respectively. This definition is motivated by our related work on L^p Sobolev estimates for the moment curve in four dimensions [1]. Note that $\bar{\gamma}$ is a primitive for $\gamma_\uparrow$ in the sense that $\bar{\gamma}' = \gamma_\uparrow$. By the definition of the Frenet frame, it follows that

$$\bar{\mathbf{e}}_j(s) \in \langle \gamma_\uparrow(s), \gamma_\uparrow'(s), \dots, \gamma_\uparrow^{(j-1)}(s)\rangle \qquad \text{and} \qquad \gamma_\uparrow^{(i)}(s) \in \langle \mathbf{e}_{1,\uparrow}(s), \dots, \mathbf{e}_{i,\uparrow}(s)\rangle$$

for $1 \leqslant i < j \leqslant 4$. Thus, one readily deduces that

$$|\langle \bar{\mathbf{e}}_j(s), \Xi\rangle| \lesssim |\langle \gamma_\uparrow(s), \Xi\rangle| + \sum_{i=1}^{j-1} |\langle \mathbf{e}_i(s), \xi\rangle| \qquad \text{for } \Xi = (\xi,\tau) \in \widehat{\mathbb{R}}^{3+1} \text{ and } 1 \leqslant j \leqslant 4.$$

If $\Xi = (\xi,\tau) \in \operatorname{supp} \mathcal{F}_t\big[m_{k,\ell}^{*,\mu,(\varepsilon)}\big]$, then it follows from (8.21) that

$$|\langle \gamma_\uparrow(s_\mu), \Xi\rangle| = |\langle \gamma(s_\mu), \xi\rangle + \tau| \lesssim 2^{4\varepsilon k}2^{k-3\ell}.$$

On the other hand, since $\xi \in \operatorname{supp}_\xi a_{k,\ell}^{*,\mu,(\varepsilon)}$, Lemma 8.2 yields

$$|\langle \mathbf{e}_i(s_\mu), \xi\rangle| \lesssim 2^{k-(3-i)\ell} \qquad \text{for } i = 1, 2, \qquad |\langle \mathbf{e}_3(s_\mu), \xi\rangle| \sim 2^k.$$

Combining these observations, $|\langle \bar{\mathbf{e}}_j(s), \Xi\rangle| \lesssim 2^{4\varepsilon k}2^{k-(4-j)\ell}$ for $1 \leqslant j \leqslant 3$ and therefore it suffices to prove $|\langle \bar{\mathbf{e}}_4(s_\mu), \Xi\rangle| \sim 2^k$. Since our hypothesis $\ell \geqslant \lceil 4\varepsilon k\rceil$ implies that $2^{4\varepsilon k}2^{k-(3-i)\ell} \leqslant 2^k$ for

$0 \leqslant i \leqslant 2$, the above argument directly yields the upper bound. On the other hand, since $\gamma \in \mathfrak{G}_3(\delta_0)$ and we are localised to $|s_\mu| \lesssim \delta_0$, the change of basis mapping $(\bar{\mathbf{e}}_j(s_\mu))_{j=1}^4$ to $(\gamma_\uparrow^{(j-1)}(s_\mu))_{j=1}^4$ is an $O(\delta_0)$ perturbation of the identity. In view of this, the above argument can also be adapted to give the required lower bound. □

8.4. **Reverse square function estimates in $\mathbb{R}^{3+1}$.** In view of the Fourier localisation described in the previous subsection, Theorem 5.3 implies the following reverse square function estimate.

Proposition 8.5. *Let $k \in \mathbb{N}$, $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$. For all $2 \leqslant p \leqslant 4$ and $\varepsilon > 0$, one has*

$$\|m[a_{k,\ell}^{(\varepsilon)}](D;\,\cdot\,)f\|_{L^p(\mathbb{R}^{3+1})} \lesssim_{\varepsilon,N} 2^{(k-3\ell)/4}2^{O(\varepsilon k)}\Big\|\Big(\sum_{\nu\in\mathbb{Z}}|m[a_{k,\ell}^{\nu,(\varepsilon)}](D;\,\cdot\,)f|^2\Big)^{1/2}\Big\|_{L^p(\mathbb{R}^{3+1})} + 2^{-kN}\|f\|_{L^p(\mathbb{R}^3)}.$$

Proof. First suppose $\lceil 4\varepsilon k\rceil \leqslant \ell$ so that Lemma 8.4 applies. Thus,

$$m_{k,\ell}^{(\varepsilon)}(D;\,\cdot\,)f = \sum_{\mu\in\mathbb{Z}} m_{k,\ell}^{*,\mu,(\varepsilon)}(D;\,\cdot\,)f$$

where each $m_{k,\ell}^{*,\mu,(\varepsilon)}(D;\,\cdot\,)f$ has spatio-temporal Fourier support in $2^k\cdot\pi_{2,\bar{\gamma}}(s_\mu; 2^{4\varepsilon k}2^{-\ell})$. The family of sets $\pi_{2,\bar{\gamma}}(s_\mu; 2^{4\varepsilon k}2^{-\ell})$ for $|\mu| \leqslant 2^\ell$ may be partitioned into $O(2^{4\varepsilon k})$ subfamilies, each forming a $(2, 2^{4\varepsilon k}2^{-\ell})$-Frenet box decomposition for the non-degenerate curve $\bar{\gamma}$ in $\mathbb{R}^4$. Consequently, by Theorem 5.3 and pigeonholing,

$$\|m_{k,\ell}^{(\varepsilon)}(D;\,\cdot\,)f\|_{L^p(\mathbb{R}^4)} \lesssim_\varepsilon 2^{O(\varepsilon k)}\Big\|\Big(\sum_{\mu\in\mathbb{Z}}|m_{k,\ell}^{*,\mu,(\varepsilon)}(D;\,\cdot\,)f|^2\Big)^{1/2}\Big\|_{L^p(\mathbb{R}^4)}.$$

By a pointwise application of the Cauchy–Schwarz inequality, using the fact that $\#\mathfrak{N}_\ell(\mu) \lesssim 2^{(k-3\ell)/2}$ for all $\mu \in \mathbb{Z}$, we conclude that

$$\|m_{k,\ell}^{(\varepsilon)}(D;\,\cdot\,)f\|_{L^p(\mathbb{R}^{3+1})} \lesssim_\varepsilon 2^{(k-3\ell)/4}2^{O(\varepsilon k)}\Big\|\Big(\sum_{\nu\in\mathbb{Z}}|m_{k,\ell}^{\nu,(\varepsilon)}(D;\,\cdot\,)f|^2\Big)^{1/2}\Big\|_{L^p(\mathbb{R}^{3+1})}. \tag{8.22}$$

The desired estimate, involving the $m[a_{k,\ell}^{\nu,(\varepsilon)}]$ multipliers rather than the $m_{k,\ell}^{\nu,(\varepsilon)}$, now follows by combining (8.22) and Lemma 8.3.

On the other hand, if $0 \leqslant \ell \leqslant \lceil 4\varepsilon k\rceil$, then the result follows directly from the Cauchy–Schwarz inequality. □

Remark. The above square function estimate is not very effective *away* from the binormal cone ($\ell = 0$ or small values of ℓ), as in that case it essentially amounts to a trivial application of the Cauchy–Schwarz inequality. However, as noted in §7, Proposition 7.1 is only used *close* to the binormal cone ($\ell = \lfloor k/3\rfloor$ or large values of ℓ), for which Proposition 8.5 is most effective. The small values of ℓ in proving Theorem 3.1 are handled via Proposition 7.2.

For $p = 2$ a stronger square function estimate is available simply due to Plancherel's theorem. In particular, this avoids the loss induced by the Cauchy–Schwarz inequality in the proof above.

Lemma 8.6. *Let $k \in \mathbb{N}$, $0 \leqslant \ell \leqslant \lfloor k/3\rfloor$. For all $\varepsilon > 0$,*

$$\|m[a_{k,\ell}^{(\varepsilon)}](D;\,\cdot\,)f\|_{L^2(\mathbb{R}^{3+1})} \lesssim_\varepsilon \Big\|\Big(\sum_{\nu\in\mathbb{Z}}|m[a_{k,\ell}^{\nu,(\varepsilon)}](D;\,\cdot\,)f|^2\Big)^{1/2}\Big\|_{L^2(\mathbb{R}^{3+1})}.$$

Proof. This is simply a consequence of Plancherel's theorem and the fact that the symbols $a_{k,\ell}^{\nu,(\varepsilon)}$ are supported on the essentially disjoint sets $2^{k-\ell}\cdot\pi_0(s_\nu; 2^{-(k-\ell)/2}, 2^\ell)$ by Lemma 8.2. □

8.5. **Kernel estimates.** Given a symbol $a \in C^\infty(\widehat{\mathbb{R}}^n\backslash\{0\} \times \mathbb{R} \times \mathbb{R})$, define the associated convolution kernel

$$K[a](x,t) := \frac{1}{(2\pi)^n} \int_{\widehat{\mathbb{R}}^n} e^{i\langle x, \xi\rangle} m[a](\xi;t)\,\mathrm{d}\xi.$$

Each of the localised symbols $a_{k,\ell}^{\nu,(\varepsilon)}$ satisfies the following kernel estimate, which yields a gain due to the localisation of the symbols in the s-variable introduced in (8.1).

Lemma 8.7. *For $k \in \mathbb{N}$ and $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$,*

$$|K[a_{k,\ell}^{\nu,(\varepsilon)}](x,t)| \lesssim 2^{-(k-\ell)/2} 2^{O(\varepsilon k)}\, \psi_{\mathcal{T}_{k,\ell}(s_\nu)}(x,t)\,\rho(t)$$

where

$$\psi_{\mathcal{T}_{k,\ell}(s_\nu)}(x,t) := 2^{(5k-3\ell)/2}\Big(1 + \sum_{j=1}^{3} 2^{j(k-\ell)/2 \wedge k} |\langle x - t\gamma(s_\nu), \mathbf{e}_j(s_\nu)\rangle|\Big)^{-100}.$$

Proof. Let $\nabla_{\boldsymbol{v}_j}$ denote the directional derivative with respect to the ξ variable in the direction of the vector $\boldsymbol{v}_j := \mathbf{e}_j(s_\nu)$, so that

$$\Big(\frac{1}{i\langle x - t\gamma(s_\nu), \mathbf{e}_j(s_\nu)\rangle} \nabla_{\boldsymbol{v}_j} - 1\Big) e^{i\langle x - t\gamma(s_\nu), \xi\rangle} = 0.$$

Thus, by repeated integration-by-parts, it follows that

$$\begin{aligned}|K[a_{k,\ell}^{\nu,(\varepsilon)}](x,t)| &\leqslant |\langle x - t\gamma(s_\nu), \mathbf{e}_j(s_\nu)\rangle|^{-N} \int_{\widehat{\mathbb{R}}^3} \big|\nabla_{\boldsymbol{v}_j}^N \big[e^{it\langle\gamma(s_\nu),\xi\rangle} m[a_{k,\ell}^{\nu,(\varepsilon)}](\xi;t)\big]\big|\,\mathrm{d}\xi \\ &\lesssim 2^{(5k-3\ell)/2} 2^{O(\varepsilon k)} |\langle x - t\gamma(s_\nu), \mathbf{e}_j(s_\nu)\rangle|^{-N} \sup_{\xi \in \widehat{\mathbb{R}}^3} \big|\nabla_{\boldsymbol{v}_j}^N \big[e^{it\langle\gamma(s_\nu),\xi\rangle} m[a_{k,\ell}^{\nu,(\varepsilon)}](\xi;t)\big]\big|;\end{aligned}$$

here the second inequality follows from the ξ-support properties of the symbols $a_{k,\ell}^{\nu,(\varepsilon)}$ from Lemma 8.2 b) if $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$ (in which case there is no $2^{O(\varepsilon k)}$ loss) or Lemma 8.2 a) if $\lfloor k/3 \rfloor_\varepsilon \leqslant \ell \leqslant \lfloor k/3 \rfloor$. Observe that

$$e^{it\langle\gamma(s_\nu),\xi\rangle} m[a_{k,\ell}^{\nu,(\varepsilon)}](\xi;t) = \int_{\mathbb{R}} e^{-it\langle\gamma(s)-\gamma(s_\nu),\xi\rangle} a_{k,\ell}^{\nu,(\varepsilon)}(\xi;s)\chi(s)\rho(t)\,\mathrm{d}s.$$

Passing the differential operator $\nabla_{\boldsymbol{v}_j}$ into the s-integral, we therefore have

$$\big|\nabla_{\boldsymbol{v}_j}^N \big[e^{it\langle\gamma(s_\nu),\xi\rangle} m[a_{k,\ell}^{\nu,(\varepsilon)}](\xi;t)\big]\big| \lesssim 2^{-(k-\ell)/2} 2^{O(\varepsilon k)} \sup_{s \in \mathbb{R}} \big|\nabla_{\boldsymbol{v}_j}^N \big[e^{-it\langle\gamma(s)-\gamma(s_\nu),\xi\rangle} a_{k,\ell}^{\nu,(\varepsilon)}(\xi;t;s)\big]\big|\,\rho(t). \tag{8.23}$$

Here we have used the s-support properties of $a_{k,\ell}^{\nu,(\varepsilon)}$; in particular, the definition (8.1).

Consider the oscillatory factor $e^{-it\langle\gamma(s)-\gamma(s_\nu),\xi\rangle}$ on the right-hand side of (8.23). The ξ derivatives of this function can be controlled on $\operatorname{supp} a_{k,\ell}^{\nu,(\varepsilon)}$ by noting that

$$|\langle\gamma(s) - \gamma(s_\nu), \boldsymbol{v}_j\rangle| \leqslant \int_{s_\nu}^{s} |\langle\gamma'(\sigma), \boldsymbol{v}_j\rangle|\,\mathrm{d}\sigma \lesssim |s - s_\nu|^j \lesssim 2^{-j(k-\ell)/2} 2^{O(\varepsilon k)} \qquad \text{for } 1 \leqslant j \leqslant 3,$$

where we have used (5.1) and triangle inequality and (8.11) in the last inequality. Thus, by the Leibniz rule, the problem is reduced to showing

$$|\nabla_{\boldsymbol{v}_j}^N a_{k,\ell}^{\nu,(\varepsilon)}(\xi;t;s)| \lesssim_N 2^{-(j(k-\ell)/2 \wedge k)N} 2^{\varepsilon\ell N} \qquad \text{for all } 1 \leqslant j \leqslant 3 \text{ and all } N \in \mathbb{N}. \tag{8.24}$$

For all $N \in \mathbb{N}$, we claim the following:

- For $\xi \in \operatorname{supp}_\xi a_{k,\ell}^{\nu,(\varepsilon)}$ with $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$,

$$2^{\ell}|\nabla_{\boldsymbol{v}_j}^N \theta_2(\xi)|, \quad 2^{-k+2\ell}|\nabla_{\boldsymbol{v}_j}^N u(\xi)| \quad \lesssim_N \quad 2^{-(j(k-\ell)/2 \wedge k)N} 2^{\varepsilon\ell N}; \tag{8.25}$$

- For $\xi \in \mathrm{supp}_\xi\, a_{k,\ell}^{\nu,(\varepsilon)}$ with $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$,

$$2^{(k-\ell)/2} |\nabla_{\boldsymbol{v}_j}^N \theta_1(\xi)| \lesssim_N 2^{-(j(k-\ell)/2 \wedge k)N}. \tag{8.26}$$

Assuming that this is so, the derivative bounds (8.24) follow directly from the chain and Leibniz rule, applying (8.25) and (8.26).

The claimed bound (8.25) follows from repeated application of the chain rule, provided

$$|\langle \gamma^{(3)} \circ \theta_2(\xi), \xi \rangle| \gtrsim 2^k, \tag{8.27a}$$
$$|\langle \gamma^{(K)} \circ \theta_2(\xi), \xi \rangle| \lesssim_K 2^{k+\ell(K-3)}, \tag{8.27b}$$
$$|\langle \gamma^{(K)} \circ \theta_2(\xi), \boldsymbol{v}_j \rangle| \lesssim_K 2^{-(j(k-\ell)/2 \wedge k)+k+\ell(K-3)} 2^{\varepsilon \ell} \tag{8.27c}$$

hold for all $K \geqslant 2$ and all $\xi \in \mathrm{supp}_\xi\, a_{k,\ell}^{\nu,(\varepsilon)}$. In particular, assuming (8.27a), (8.27b) and (8.27c), the bounds in (8.25) are then a consequence of Lemma C.1 in the appendix: (8.25) corresponds to (C.13) and (C.15) whilst the hypotheses in the above display correspond to (C.16) and (C.17) (see Example C.2). Here the parameters featured in the appendix are chosen as follows:

g	h	A	B	M_1	M_2	$\mathbf{e}$
γ''	γ'	$2^{k-\ell}$	$2^{k-2\ell}$	$2^{-(j(k-\ell)/2 \wedge k)} 2^{\varepsilon \ell}$	2^ℓ	$\boldsymbol{v}_j$

The conditions (8.27a), (8.27b) and (8.27c) are direct consequences of the support properties of the $a_{k,\ell}^{\nu,(\varepsilon)}$. Indeed, (8.27a) and the $K \geqslant 3$ case of (8.27b) are trivial consequences of the localisation of the symbol a_k. The remaining $K = 2$ case of (8.27b) follows immediately since $\langle \gamma'' \circ \theta_2(\xi), \xi \rangle = 0$. Finally, the right-hand side of (8.27c) is always greater than 1 unless $j = 3$ and $K = 2$, and so we can immediately reduce to this case. If $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$, then (5.1) together with Lemma 6.3 and the θ_1 localisation in (8.11) implies

$$|\langle \gamma^{(2)} \circ \theta_2(\xi), \boldsymbol{v}_3 \rangle| \lesssim |\theta_2(\xi) - s_\nu| \leqslant |\theta_2(\xi) - \theta_1(\xi)| + |\theta_1(\xi) - s_\nu| \lesssim 2^{-\ell}.$$

On the other hand, if $\lfloor k/3 \rfloor_\varepsilon \leqslant \ell \leqslant \lfloor k/3 \rfloor$, then, by a similar argument, $|\langle \gamma^{(2)} \circ \theta_2(\xi), \boldsymbol{v}_3 \rangle| \lesssim 2^{-\ell(1-\varepsilon)}$. This concludes the proof of (8.27c).

Similarly, the claimed bound (8.26) follows from repeated application of the chain rule, provided

$$|\langle \gamma^{(2)} \circ \theta_1(\xi), \xi \rangle| \gtrsim 2^{k-\ell}, \tag{8.28a}$$
$$|\langle \gamma^{(K)} \circ \theta_1(\xi), \xi \rangle| \lesssim_K 2^{K(k-\ell)/2}, \tag{8.28b}$$
$$|\langle \gamma^{(K)} \circ \theta_1(\xi), \boldsymbol{v}_j \rangle| \lesssim_K 2^{-(j(k-\ell)/2 \wedge k)+K(k-\ell)/2} \tag{8.28c}$$

hold for all $K \geqslant 2$ and all $\xi \in \mathrm{supp}\, a_{k,\ell}^{\nu,(\varepsilon)}$ when $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$. This again follows by Lemma C.1 in the appendix. Here the parameters are chosen as follows:

g	A	M_1	M_2	$\mathbf{e}$
γ'	$2^{(k-\ell)/2}$	$2^{-(j(k-\ell)/2 \wedge k)}$	$2^{(k-\ell)/2}$	$\boldsymbol{v}_j$

The conditions (8.28a), (8.28b) and (8.28c) are direct consequences of the support properties of the $a_{k,\ell}^{\nu,(\varepsilon)}$ for $0 \leqslant \ell < \lfloor k/3 \rfloor_\varepsilon$. Indeed, (8.28a) and the $K = 2$ case of (8.28b) is just a restatement of the condition $|v^\pm(\xi)| \sim 2^{k-\ell}$, which holds due to Lemma 6.3. The $K \geqslant 3$ case of (8.28b) follows immediately from the localisation of the symbols a_k. Finally, the right-hand side of (8.28c) is always greater than 1 unless $j = 3$ and $K = 2$, and so we can immediately reduce to this case. However, (5.1) together with the θ_1 localisation in (8.11) implies

$$|\langle \gamma^{(2)} \circ \theta_1(\xi), \boldsymbol{v}_3\rangle| \lesssim |\theta_1(\xi) - s_\nu| \lesssim 2^{-(k-\ell)/2} 2^{\varepsilon k} \lesssim 2^{-\ell},$$

which concludes the proof of (8.28c). □

8.6. **Localising the input function.** At this juncture it is useful to note some further geometric properties of the support of the multipliers $m[a_{k,\ell}^{\nu,(\varepsilon)}]$ featured in the decomposition.

Recall from Lemma 8.2 a) that

$$\operatorname{supp}_\xi a_{k,\ell}^{\nu,(\varepsilon)} \subseteq 2^k \cdot \pi_1(s_\mu; 2^{-\ell}) \qquad \text{for all } \nu \in \mathfrak{N}_\ell(\mu), \tag{8.29}$$

where $s_\mu := 2^{-\ell}\mu$. The right-hand set is contained in a certain sector in the frequency space. In particular, given $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$ and $m \in \mathbb{Z}$ define

$$\Delta_{k,\ell}(m) := \big\{\xi \in \widehat{\mathbb{R}}^3 : |\xi_2 - \xi_3 2^{-\ell} m| \leqslant C 2^{-\ell}\xi_3 \text{ and } C^{-1}2^k \leqslant \xi_3 \leqslant C 2^k\big\}, \tag{8.30}$$

where $C \geqslant 1$ is an absolute constant, chosen sufficiently large so as to satisfy the requirements of the forthcoming argument.

Lemma 8.8. *If μ, $\nu \in \mathbb{Z}$ satisfy $\nu \in \mathfrak{N}_\ell(\mu)$, then there exists some $m(\mu) \in \mathbb{Z}$ such that*

$$2^k \cdot \pi_1(s_\mu; 2^{-\ell}) \subseteq \Delta_{k,\ell}\big(m(\mu)\big). \tag{8.31}$$

Furthermore, for each fixed k and ℓ, given $m \in \mathbb{Z}$ there are only $O(1)$ values of $\mu \in \mathbb{Z}$ such that $m = m(\mu)$.

Proof. Define $G\colon I_0 \to \mathbb{R}^3$ by $G(s) := \mathbf{e}_{33}(s)^{-1}\mathbf{e}_3(s)$. As a consequence of the Frenet equations, the vectors $G'(s)$, $G''(s)$ span $\mathbb{R}^2 \times \{0\}$. Given $\xi \in \widehat{\mathbb{R}}^3$, it follows that there exist η_1, $\eta_2 \in \mathbb{R}$ such that

$$\xi - \xi_3 G(s) = \sum_{j=1}^{2} 2^{-\ell j}\eta_j G^{(j)}(s). \tag{8.32}$$

Taking the inner product of both sides of this identity with respect to the $\mathbf{e}_j(s)$ for $j = 1$, 2 and applying the Frenet equations

$$\begin{bmatrix} \langle \xi, \mathbf{e}_1(s)\rangle \\ \langle \xi, \mathbf{e}_2(s)\rangle \end{bmatrix} = \begin{bmatrix} 0 & \langle G^{(2)}(s), \mathbf{e}_1(s)\rangle \\ \langle G^{(1)}(s), \mathbf{e}_2(s)\rangle & \langle G^{(2)}(s), \mathbf{e}_2(s)\rangle \end{bmatrix} \begin{bmatrix} 2^{-\ell}\eta_1 \\ 2^{-2\ell}\eta_2 \end{bmatrix} \tag{8.33}$$

where the anti-diagonal entries of the right-hand 2×2 matrix have size ~ 1 (a similar computation is carried out in more detail in §9.1).

Let $\xi \in 2^k \cdot \pi_1(s; 2^{-\ell})$ so that $|\langle \xi, \mathbf{e}_1(s_\mu)\rangle| \lesssim 2^{k-2\ell}$ and $|\langle \xi, \mathbf{e}_2(s_\mu)\rangle| \lesssim 2^{k-\ell}$. Combining these bounds with (8.32) and (8.33), it follows that

$$|\xi_2 - \xi_3 G_2(s_\mu)| \leqslant |\xi - \xi_3 G(s_\mu)| \lesssim 2^{k-\ell} \sim 2^{-\ell}\xi_3.$$

If we take $m(\mu)$ to be the integer which minimises $|2^{-\ell}m - G_2(s_\mu)|$, then we obtain (8.31). On the other hand, the Frenet equations ensure that $G_2(s) = \mathbf{e}_{33}(s)^{-1}\mathbf{e}_{32}(s)$ satisfies $|G_2'(s)| \sim 1$ for all $s \in I_0$. Consequently, the assignment $\mu \mapsto m(\mu)$ is $O(1)$-to-1, as claimed. □

For each $\mu \in \mathbb{Z}$ define the smooth cutoff function

$$\chi_{k,\ell}^{*,\mu}(\xi) := \eta\big(C^{-1}|2^\ell \xi_2/\xi_3 - m(\mu)|\big)\,\big(\eta(C^{-1}2^{-k}\xi_3) - \eta(2C2^k\xi_3)\big).$$

If $\xi \in \operatorname{supp}_\xi a_{k,\ell}^{\nu,(\varepsilon)}$ for $\nu \in \mathfrak{N}(\mu)$, then (8.29) and Lemma 8.8 imply $\chi_{k,\ell}^{*,\mu}(\xi) = 1$. Thus, if we define the corresponding frequency projection

$$f_{k,\ell}^{*,\mu} := \chi_{k,\ell}^{*,\mu}(D) f,$$

it follows that

$$m[a_{k,\ell}^{\nu,(\varepsilon)}](D;\,\cdot\,)f = m[a_{k,\ell}^{\nu,(\varepsilon)}](D;\,\cdot\,)f_{k,\ell}^{*,\mu} \qquad \text{for all } \nu \in \mathfrak{N}_\ell(\mu).$$

Recall from Lemma 8.2 b) that we also have

$$\operatorname{supp}_\xi a_{k,\ell}^{\nu,(\varepsilon)} \subseteq 2^{k-\ell}\cdot \pi_0(s_\nu; 2^{-(k-\ell)/2}, 2^\ell), \qquad \text{where } s_\nu := 2^{-(k-\ell)/2}\nu. \tag{8.34}$$

Fix some $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$ and $\mu \in \mathbb{Z}$ with $s_\mu := 2^{-\ell}\mu \in [-1,1]$. To simplify notation, let $\sigma := s_\mu$, $\lambda := 2^{-\ell}$ and let $\tilde{\gamma} := \gamma_{\sigma,\lambda}$ denote the rescaled curve, as defined in Definition 4.1, so that

$$\tilde{\gamma}(s) := \big([\gamma]_{\sigma,\lambda}\big)^{-1}\big(\gamma(\sigma+\lambda s) - \gamma(\sigma)\big). \tag{8.35}$$

Let $(\tilde{\mathbf{e}}_j)_{j=1}^4$ denote the Frenet frame defined with respect to $\tilde{\gamma}$. Given $0 < r \leqslant 1$ and $s \in I_0$, recall the definition of the $(0,r)$-*Frenet boxes* (with respect to $(\tilde{\mathbf{e}}_j)_{j=1}^3$) introduced in Definition 5.1:

$$\pi_{0,\tilde{\gamma}}(s;r) := \big\{\xi \in \hat{\mathbb{R}}^3 : |\langle \tilde{\mathbf{e}}_1(s),\xi\rangle| \lesssim r, \quad |\langle \tilde{\mathbf{e}}_2(s),\xi\rangle| \sim 1, \quad |\langle \tilde{\mathbf{e}}_3(s),\xi\rangle| \lesssim 1\big\}.$$

Note that all these definitions depend of the choice of μ and ℓ, but we suppress this dependence in the notation.

Lemma 8.9. *With the above setup, and $\nu \in \mathfrak{N}(\mu)$,*

$$[\gamma]_{\sigma,\lambda}^\top \cdot 2^{k-\ell}\cdot \pi_{0,\gamma}(s_\nu; 2^{-(k-\ell)/2}, 2^\ell) \subseteq 2^{k-3\ell}\cdot \pi_{0,\tilde{\gamma}}(\tilde{s}_\nu; 2^{-(k-3\ell)/2}),$$

where $\tilde{s}_\nu := 2^\ell(s_\nu - s_\mu)$ for $s_\nu := 2^{-(k-\ell)/2}\nu$.

Proof. Let $\xi \in 2^{k-\ell}\cdot\pi_{0,\gamma}(s_\nu; 2^{-(k-\ell)/2}, 2^\ell)$ so that

$$|\langle \mathbf{e}_1(s_\nu),\xi\rangle| \lesssim 2^{(k-\ell)/2} \quad , \quad |\langle \mathbf{e}_2(s_\nu),\xi\rangle| \sim 2^{k-\ell}, \quad |\langle \mathbf{e}_4(s_\nu),\xi\rangle| \sim 2^{k}.$$

Since the matrix corresponding to the change of basis from $\big(\mathbf{e}_j(s_\nu)\big)_{j=1}^3$ to $\big(\gamma^{(j)}(s_\nu)\big)_{j=1}^3$ is lower triangular and an $O(\delta_0)$ perturbation of the identity, provided δ_0 is sufficiently small,

$$|\langle \gamma^{(1)}(s_\nu),\xi\rangle| \lesssim 2^{(k-\ell)/2}, \quad |\langle \gamma^{(2)}(s_\nu),\xi\rangle| \sim 2^{k-\ell}, \quad |\langle \gamma^{(3)}(s_\nu),\xi\rangle| \sim 2^{k}.$$

Now define $\tilde{\xi} := \big([\gamma]_{\sigma,\lambda}\big)^\top\cdot\xi$. Since $\lambda := 2^{-\ell}$, it follows from the definition of $\tilde{\gamma}$ from (8.35) that

$$\langle \tilde{\gamma}^{(j)}(\tilde{s}_\nu),\tilde{\xi}\,\rangle = 2^{-j\ell}\langle \gamma^{(j)}(s_\nu),\xi\rangle \qquad \text{for } j \geqslant 1.$$

Combining the above observations,

$$|\langle \tilde{\gamma}^{(1)}(\tilde{s}_\nu),\tilde{\xi}\,\rangle| \lesssim 2^{(k-3\ell)/2}, \quad |\langle \tilde{\gamma}^{(2)}(\tilde{s}_\nu),\tilde{\xi}\,\rangle| \sim 2^{k-3\ell}, \quad |\langle \tilde{\gamma}^{(3)}(\tilde{s}_\nu),\tilde{\xi}\,\rangle| \sim 2^{k-3\ell}.$$

Provided δ_0 is sufficiently small, the desired result now follows since the matrix corresponding to the change of basis from $\big(\tilde{\mathbf{e}}_i(\tilde{s}_\nu)\big)_{i=1}^3$ to $\big(\tilde{\gamma}^{(i)}(\tilde{s}_\nu)\big)_{i=1}^3$ is also lower triangular and an $O(\delta_0)$ perturbation of the identity. ☐

For $\nu \in \mathfrak{N}_\ell(\mu)$ define the smooth cutoff

$$\chi_{k,\ell}^\nu(\xi) := \chi_{\tilde{\pi}}\big(C^{-1}2^{-(k-3\ell)}[\gamma]_{\sigma,\lambda}^\top\cdot\xi\big) \tag{8.36}$$

where $\chi_{\tilde{\pi}}$ is as defined in (5.3) for $\tilde{\pi} := \pi_{0,\tilde{\gamma}}(\tilde{s}_\nu; 2^{-(k-3\ell)/2})$ as above. If $\xi \in \operatorname{supp}_\xi a_{k,\ell}^{\nu,(\varepsilon)}$, then (8.34) and Lemma 8.9 imply $\chi_{k,\ell}^\nu(\xi) = 1$. Thus if we define the corresponding frequency projection

$$f_{k,\ell}^\nu := \chi_{k,\ell}^\nu(D)f_{k,\ell}^{*,\mu},$$

it follows that

$$m[a_{k,\ell}^{\nu,(\varepsilon)}](D;\,\cdot\,)f = m[a_{k,\ell}^{\nu,(\varepsilon)}](D;\,\cdot\,)f_{k,\ell}^{*,\mu} = m[a_{k,\ell}^{\nu,(\varepsilon)}](D;\,\cdot\,)f_{k,\ell}^\nu \qquad \text{for all } \nu \in \mathfrak{N}_\ell(\mu).$$

8.7. **L^2-weighted bounds in $\mathbb{R}^{3+1}$.** We apply a standard duality argument to analyse the square function appearing in Proposition 8.5. In particular, we use an approach based on weighted L^2-norm inequalities; the key ingredient is the Nikodym-type maximal inequality from §5.4.

By duality, there exists a non-negative $g \in L^2(\mathbb{R}^{3+1})$ with $\|g\|_{L^2(\mathbb{R}^{3+1})} = 1$ such that

$$\Big\|\Big(\sum_{\nu\in\mathbb{Z}} |m[a_{k,\ell}^{\nu,(\varepsilon)}](D;\,\cdot\,)f|^2\Big)^{1/2}\Big\|_{L^4(\mathbb{R}^{3+1})}^2 = \sum_{\nu\in\mathbb{Z}} \int_{\mathbb{R}^{3+1}} |m[a_{k,\ell}^{\nu,(\varepsilon)}](D;t)f(x)|^2 g(x;t)\,\mathrm{d}x\mathrm{d}t.$$

By the observations of the previous subsection,

$$m[a_{k,\ell}^{\nu,(\varepsilon)}](D;t)f = m[a_{k,\ell}^{\nu,(\varepsilon)}](D;t)f_{k,\ell}^{\nu}.$$

Let $\psi_{\mathcal{T}_{k,\ell}(s_\nu)}$ be the weight introduced in Lemma 8.7. Since the $\psi_{\mathcal{T}_{k,\ell}(s_\nu)}(\,\cdot\,;t)$ are L^1-normalised uniformly in t, it follows from Lemma 8.7 and the Cauchy–Schwarz inequality that

$$|m[a_{k,\ell}^{\nu,(\varepsilon)}](D;t)f(x)|^2 \lesssim 2^{-(k-\ell)}2^{O(\varepsilon k)}\psi_{\mathcal{T}_{k,\ell}(s_\nu)}(\,\cdot\,;t) * |f_{k,\ell}^{\nu}|^2(x)\,\rho(t). \tag{8.37}$$

Define the Nikodym-type maximal operator

$$\widetilde{\mathcal{N}}_{k,\ell}^{\mathrm{sing}}\, g(x) := \max_{\nu\in\mathbb{Z}\,:\,|s_\nu|\leqslant\delta_0} \int_{\mathbb{R}^4} |g(x-y,t)|\psi_{\mathcal{T}_{k,\ell}(s_\nu)}(y,t)\,\rho(t)\,\mathrm{d}y\mathrm{d}t.$$

By (8.37) and Fubini's theorem, it follows that

$$\sum_{\nu\in\mathbb{Z}} \int_{\mathbb{R}^{3+1}} |m[a_{k,\ell}^{\nu,(\varepsilon)}](D;t)f(x)|^2 g(x;t)\,\mathrm{d}x\mathrm{d}t \lesssim 2^{-(k-\ell)}2^{O(\varepsilon k)} \int_{\mathbb{R}^3} \sum_{\nu\in\mathbb{Z}} |f_{k,\ell}^{\nu}(x)|^2 \widetilde{\mathcal{N}}_{k,\ell}^{\mathrm{sing}}\, g(x)\,\mathrm{d}x.$$

Note that $\widetilde{\mathcal{N}}_{k,\ell}^{\mathrm{sing}}$ is essentially a smooth version of the maximal operator $\mathcal{N}_{\mathbf{r}}^{\mathrm{sing}}$ from §5.4 with parameters $r_1 := 2^{-(k-\ell)/2}$, $r_2 := 2^{-(k-\ell)}$ and $r_3 := 2^{-k}$. By the restriction $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$, it follows that this choice of $\mathbf{r}$ satisfies the hypotheses

$$r_3 \leqslant r_2 \leqslant r_1 \leqslant r_2^{1/2} \qquad \text{and} \qquad r_2 \leqslant r_1^{1/2} r_3^{1/2}$$

from the statement of Proposition 5.5. Thus, by pointwise dominating $\psi_{\mathcal{T}_{k,\ell}(s_\nu)}$ by a weighted series of indicator functions and applying Proposition 5.5, one readily deduces the norm bound

$$\|\widetilde{\mathcal{N}}_{k,\ell}^{\mathrm{sing}}\|_{L^2(\mathbb{R}^{3+1})\to L^2(\mathbb{R}^3)} \lesssim_\varepsilon 2^{\varepsilon k}.$$

Combine the above observations with an application of the Cauchy–Schwarz inequality to obtain

$$\Big\|\Big(\sum_{\nu\in\mathbb{Z}} |m[a_{k,\ell}^{\nu,(\varepsilon)}](D;\,\cdot\,)f|^2\Big)^{1/2}\Big\|_{L^4(\mathbb{R}^{3+1})} \lesssim 2^{-(k-\ell)/2+O(\varepsilon k)}\Big\|\Big(\sum_{\nu\in\mathbb{Z}} |f_{k,\ell}^{\nu}|^2\Big)^{1/2}\Big\|_{L^4(\mathbb{R}^3)}. \tag{8.38}$$

It remains to bound the right-hand square function, which involves only functions of 3 variables.

8.8. **L^2-weighted bounds in $\mathbb{R}^3$.** A similar L^2-weighted approach is now applied one dimension lower to estimate the square function appearing in the right-hand side of (8.38).

Proposition 8.10. *Let $k \in \mathbb{N}$, $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$ and $\varepsilon > 0$. Then*

$$\Big\|\Big(\sum_{\nu\in\mathbb{Z}} |f_{k,\ell}^{\nu}|^2\Big)^{1/2}\Big\|_{L^4(\mathbb{R}^3)} \lesssim_\varepsilon 2^{O(\varepsilon k)}\|f\|_{L^4(\mathbb{R}^3)}. \tag{8.39}$$

Proof. By duality, there exists a non-negative $w \in L^2(\mathbb{R}^3)$ with $\|w\|_{L^2(\mathbb{R}^3)} = 1$ such that

$$\Big\|\Big(\sum_{\nu\in\mathbb{Z}} |f_{k,\ell}^{\nu}|^2\Big)^{1/2}\Big\|_{L^4(\mathbb{R}^3)}^2 = \sum_{\mu\in\mathbb{Z}} \int_{\mathbb{R}^3} \sum_{\nu\in\mathfrak{N}_\ell(\mu)} |f_{k,\ell}^{\nu}(x)|^2 w(x)\,\mathrm{d}x. \tag{8.40}$$

Recall that the $f_{k,\ell}^{\nu}$ are defined by

$$f_{k,\ell}^{\nu} := \chi_{k,\ell}^{\nu}(D) f_{k,\ell}^{*,\mu} \qquad \text{for } \nu \in \mathfrak{N}_\ell(\mu)$$

where the smooth cutoff function $\chi_{k,\ell}^{\nu}$ is as defined in (8.36). Fix μ and, as in §8.6, let $\sigma := s_\mu$ and $\lambda := 2^{-\ell}$. Define $\tilde{f}_{k,\ell}^{\nu} := f_{k,\ell}^{\nu} \circ [\gamma]_{\sigma,\lambda}$, $\tilde{f}_{k,\ell}^{*,\mu} := f_{k,\ell}^{*,\mu} \circ [\gamma]_{\sigma,\lambda}$ and $\tilde{w} := w \circ [\gamma]_{\sigma,\lambda}$ so, by a change of variables,

$$\int_{\mathbb{R}^3} \sum_{\nu \in \mathfrak{N}_\ell(\mu)} |f_{k,\ell}^{\nu}(x)|^2\, w(x)\,\mathrm{d}x = |\det[\gamma]_{\sigma,\lambda}| \int_{\mathbb{R}^3} \sum_{\nu \in \mathfrak{N}_\ell(\mu)} |\tilde{f}_{k,\ell}^{\nu}(x)|^2\, \tilde{w}(x)\,\mathrm{d}x. \tag{8.41}$$

By the definition of $\chi_{k,\ell}^{\nu}$ and Lemma 8.9, each of the $\tilde{f}_{k,\ell}^{\nu}$ is Fourier supported in a $2^{k-3\ell}$ dilate of a $(0, 2^{-(k-3\ell)/2})$-Frenet box. In view of this, we may apply Proposition 5.4 to deduce that

$$\int_{\mathbb{R}^3} \sum_{\nu \in \mathfrak{N}_\ell(\mu)} |f_{k,\ell}^{\nu}(x)|^2\, w(x)\,\mathrm{d}x \lesssim_\varepsilon 2^{\varepsilon k} |\det[\gamma]_{\sigma,\lambda}| \int_{\mathbb{R}^3} |\tilde{f}_{k,\ell}^{*,\mu}(x)|^2\, \widetilde{\mathcal{N}}_{k,\ell}^{\mu,(\varepsilon)} \tilde{w}(x)\,\mathrm{d}x \tag{8.42}$$

where the operator $\widetilde{\mathcal{N}}_{k,\ell}^{\mu,(\varepsilon)}$ is defined by

$$\widetilde{\mathcal{N}}_{k,\ell}^{\mu,(\varepsilon)} := \mathrm{Dil}_{2^{k-3\ell}} \circ \widetilde{\mathcal{N}}_{\tilde{\gamma},\tilde{r}}^{(\varepsilon)} \circ \mathrm{Dil}_{2^{-(k-3\ell)}}$$

for $\tilde{r} := 2^{-(k-3\ell)/2}$ and $\mathrm{Dil}_\rho \colon L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3)$ the dilation operator $\mathrm{Dil}_\rho f := f(\rho\,\cdot\,)$ for $\rho > 0$. Here $\widetilde{\mathcal{N}}_{\tilde{\gamma},\tilde{r}}^{(\varepsilon)}$ is the maximal operator featured in the statement of Proposition 5.4 (the precise definition is given in §10). Note that $\widetilde{\mathcal{N}}_{k,\ell}^{\mu,(\varepsilon)}$ depends on the choice of μ. By reversing the change of variables in (8.41), we can show the following

Claim. *There exists a maximal function $\mathcal{N}_{k,\ell}^{(\varepsilon)}$, independent of μ, such that*

$$\big([\gamma]_{\sigma,\lambda}\big)^{-1} \circ \widetilde{\mathcal{N}}_{k,\ell}^{\mu,(\varepsilon)} \circ [\gamma]_{\sigma,\lambda} \cdot w(x) \lesssim_\gamma \mathcal{N}_{k,\ell}^{(\varepsilon)} w(x) \qquad \text{for all } x \in \mathbb{R}^3,$$

where $[\gamma]_{\sigma,\lambda} \cdot f := f \circ [\gamma]_{\sigma,\lambda}$, and

$$\|\mathcal{N}_{k,\ell}^{(\varepsilon)}\|_{L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3)} \lesssim_\varepsilon 2^{\varepsilon k}. \tag{8.43}$$

The proof of the above claim requires additional information on the form of the maximal operators arising from Proposition 5.4. Since the definitions involved are somewhat unwieldy, the details are postponed until §10.5.

Assuming the claim, changing variables in (8.42) yields

$$\int_{\mathbb{R}^3} \sum_{\nu \in \mathfrak{N}_\ell(\mu)} |f_{k,\ell}^{\nu}(x)|^2\, w(x)\,\mathrm{d}x \lesssim 2^{\varepsilon k} \int_{\mathbb{R}^3} |f_{k,\ell}^{*,\mu}(x)|^2\, \mathcal{N}_{k,\ell}^{(\varepsilon)} w(x)\,\mathrm{d}x.$$

Recalling (8.40), one can sum the above inequality in $\mu \in \mathbb{Z}$, and use (8.43) and the Cauchy–Schwarz inequality to obtain

$$\Big\|\Big(\sum_{\nu \in \mathbb{Z}} |f_{k,\ell}^{\nu}|^2\Big)^{1/2}\Big\|_{L^4(\mathbb{R}^3)}^2 \lesssim_\varepsilon 2^{2\varepsilon k} \Big\|\Big(\sum_{\mu \in \mathbb{Z}} |f_{k,\ell}^{*,\mu}|^2\Big)^{1/2}\Big\|_{L^4(\mathbb{R}^3)}^2. \tag{8.44}$$

Recall that each $f_{k,\ell}^{*,\mu}$ corresponds to a (smooth) frequency projection of f onto the set $\Delta_\ell(m(\mu))$, as defined in (8.30). Furthermore, by Lemma 8.8 the assignment $m \mapsto m(\mu)$ is $O(1)$-to-1. Thus, the right-hand square function in (8.44) falls under the scope of the classical sectorial square function of Córdoba [11]. In particular, by [11, Theorem 1] (see also [10]) and a Fubini argument, we have

$$\Big\|\Big(\sum_{\mu \in \mathbb{Z}} |f_{k,\ell}^{*,\mu}|^2\Big)^{1/2}\Big\|_{L^4(\mathbb{R}^3)}^2 \lesssim_\varepsilon 2^{O(\varepsilon k)} \|f\|_{L^4(\mathbb{R}^3)}. \tag{8.45}$$

The inequalities (8.44) and (8.45) imply the desired estimate (8.39). □

8.9. **Putting everything together.** We combine our observations to establish favourable L^4 and L^2 estimates for the localised multipliers $m[a_{k,\ell}]$.

L^4 estimates. By Lemma 8.1,

$$\|m[a_{k,\ell}](D;\,\cdot\,)f\|_{L^4(\mathbb{R}^{3+1})} \lesssim_{\varepsilon,N} \|m[a_{k,\ell}^{(\varepsilon)}](D;\,\cdot\,)f\|_{L^4(\mathbb{R}^{3+1})} + 2^{-kN}\|f\|_{L^4(\mathbb{R}^3)}.$$

Decompose each $m[a_{k,\ell}^{(\varepsilon)}]$ as a sum of multipliers $m[a_{k,\ell}^{\nu,(\varepsilon)}]$ as defined in §8.2. By Proposition 8.5, it follows that

$$\|m[a_{k,\ell}](D;\,\cdot\,)f\|_{L^4(\mathbb{R}^{3+1})} \lesssim_{\varepsilon,N} 2^{(k-3\ell)/4}2^{O(\varepsilon k)}\Big\|\Big(\sum_{\nu\in\mathbb{Z}}|m[a_{k,\ell}^{\nu,(\varepsilon)}](D;\,\cdot\,)f|^2\Big)^{1/2}\Big\|_{L^4(\mathbb{R}^{3+1})} + 2^{-kN}\|f\|_{L^4(\mathbb{R}^3)}.$$

Thus, (8.38) and (8.39) combine with the previous display to yield the L^4 estimate

$$\|m[a_{k,\ell}](D;\,\cdot\,)f\|_{L^4(\mathbb{R}^{3+1})} \lesssim_{\varepsilon} 2^{(k-3\ell)/4}2^{-(k-\ell)/2}2^{O(\varepsilon k)}\|f\|_{L^4(\mathbb{R}^3)}.$$

Since $\varepsilon > 0$ may be chosen arbitrarily, this corresponds to the $p = 4$ case of Proposition 7.1.

L^2 estimates. Arguing as in the proof of the L^4 estimate, but now using Lemma 8.6 rather than Proposition 8.5, it follows that

$$\|m[a_{k,\ell}](D;\,\cdot\,)f\|_{L^2(\mathbb{R}^{3+1})} \lesssim_{\varepsilon,N} \Big\|\Big(\sum_{\nu\in\mathbb{Z}}|m[a_{k,\ell}^{\nu,(\varepsilon)}](D;\,\cdot\,)f|^2\Big)^{1/2}\Big\|_{L^2(\mathbb{R}^{3+1})} + 2^{-kN}\|f\|_{L^2(\mathbb{R}^3)}.$$

Recall from (8.37) that

$$|m[a_{k,\ell}^{\nu,(\varepsilon)}](D;t)f(x)|^2 \lesssim 2^{-(k-\ell)/2}2^{O(\varepsilon k)}\psi_{\mathcal{T}_{k,\ell}(s_\nu)}(\,\cdot\,;t) * |f_{k,\ell}^{\nu}|^2.$$

Thus, by Young's convolution inequality and the fact that the $\psi_{\mathcal{T}_{k,\ell}(s_\nu)}(\,\cdot\,;t)$ are L^1-normalised,

$$\|m[a_{k,\ell}](D;\,\cdot\,)f\|_{L^2(\mathbb{R}^{3+1})} \lesssim_{\varepsilon,N} 2^{-(k-\ell)/2}2^{O(\varepsilon k)}\Big\|\Big(\sum_{\nu\in\mathbb{Z}}|f_{k,\ell}^{\nu}|^2\Big)^{1/2}\Big\|_{L^2(\mathbb{R}^3)} + 2^{-kN}\|f\|_{L^2(\mathbb{R}^3)}.$$

Finally, as the $f_{k,\ell}^{\nu}$ have essentially disjoint Fourier support, by Plancherel's theorem,

$$\|m[a_{k,\ell}](D;\,\cdot\,)f\|_{L^2(\mathbb{R}^{3+1})} \lesssim_{\varepsilon} 2^{-(k-\ell)/2}2^{O(\varepsilon k)}\|f\|_{L^2(\mathbb{R}^3)}.$$

Since $\varepsilon > 0$ may be chosen arbitrarily, this corresponds to the $p = 2$ case of Proposition 7.1.

Interpolating the above estimates, given $2 \leqslant p \leqslant 4$ and $\varepsilon > 0$, it follows that

$$\|m[a_{k,\ell}](D;\,\cdot\,)f\|_{L^p(\mathbb{R}^{3+1})} \lesssim_{\varepsilon,N} 2^{-(k-\ell)/2}2^{(k-3\ell)(1/2-1/p)}2^{\varepsilon k}\|f\|_{L^p(\mathbb{R}^3)},$$

which is precisely the desired inequality from Proposition 7.1.

9. Proof of the reverse square function inequality in $\mathbb{R}^{3+1}$

9.1. **Geometric observations.** The first step is to relate the Frenet boxes $\pi_{2,\gamma}(s;r)$ to a codimension 2 cone $\widetilde{\Gamma}_2$ in the (ξ,τ)-space.

The underlying cone. Let $\gamma \in \mathfrak{G}_4(\delta_0)$ for $0 < \delta_0 \ll 1$ and $\mathbf{e}_j\colon [-1,1] \to S^3$ for $1 \leqslant j \leqslant 4$ be the associated Frenet frame. Without loss of generality, in proving Theorem 5.3 we may always localise so that we only consider the portion of the curve lying over the interval $I_0 = [-\delta_0, \delta_0]$. In this case

$$\mathbf{e}_j(s) = \vec{e}_j + O(\delta_0) \qquad \text{for } 1 \leqslant j \leqslant 4 \tag{9.1}$$

where, as usual, the $\vec{e}_j$ denote the standard basis vectors.

We consider the conic surface $\widetilde{\Gamma}_2$ 'generated' over the curve $s \mapsto \mathbf{e}_4(s)$. This is similar to the analysis of [19], where a cone in $\mathbb{R}^3$ generated by the binormal $\mathbf{e}_3$ features prominently in the arguments. Define $G\colon I_0 \to \mathbb{R}^4$ by $G(s) := \mathbf{e}_{44}(s)^{-1}\mathbf{e}_4(s)$ for all $s \in I_0$ (note that $\mathbf{e}_{44}(s)$ is bounded away from 0 by (9.1)), so that G is of the form

$$G(s) = \begin{bmatrix} g(s) \\ 1 \end{bmatrix} \qquad \text{for} \qquad g(s) := \Big(\frac{\mathbf{e}_{41}(s)}{\mathbf{e}_{44}(s)}, \frac{\mathbf{e}_{42}(s)}{\mathbf{e}_{44}(s)}, \frac{\mathbf{e}_{43}(s)}{\mathbf{e}_{44}(s)}\Big)^\top.$$

For $U := [1/4, 4] \times I_0$, the 2-dimensional cone $\widetilde{\Gamma}_2$ is parametrised by the function

$$\widetilde{\Gamma}_2 \colon U \to \mathbb{R}^4, \qquad (\rho, s) \mapsto \rho\, G(s).$$

Non-degeneracy conditions. We claim that the curve $g\colon I_0 \to \mathbb{R}^3$ is non-degenerate. To see this, first note that

$$G^{(i)}(s) \in \langle \mathbf{e}_4(s), \mathbf{e}_4^{(1)}(s), \dots, \mathbf{e}_4^{(i)}(s)\rangle$$

where the right-hand expression denotes the linear span of the vectors $\mathbf{e}_4(s), \mathbf{e}_4^{(1)}(s), \dots, \mathbf{e}_4^{(i)}(s)$. Thus, one concludes from the Frenet formulæ that

$$G^{(i)}(s) \in \langle \mathbf{e}_{4-i}(s), \dots, \mathbf{e}_4(s)\rangle \qquad \text{for } 0 \leqslant i \leqslant 3. \tag{9.2}$$

On the other hand, the Frenet formulæ together with the Leibniz rule show that

$$\langle G^{(i)}(s), \mathbf{e}_{4-i}(s)\rangle = (-1)^i \Big(\prod_{\ell=4-i}^{3} \tilde{\kappa}_\ell(s)\Big)\, \mathbf{e}_{44}(s)^{-1}$$

and, consequently,

$$|\langle G^{(i)}(s), \mathbf{e}_{4-i}(s)\rangle| \sim 1 \qquad \text{for all } 1 \leqslant i \leqslant 3. \tag{9.3}$$

Thus, combining (9.2) and (9.3), it follows that the vectors $G^{(i)}(s)$, $1 \leqslant i \leqslant 3$, are linearly independent. From this, we conclude that

$$|\det[g]_s| \gtrsim 1$$

for all $s \in I_0$, which is the claimed non-degeneracy condition.

Frenet boxes revisited. By the preceding observations, the vectors $G^{(i)}(s)$ for $1 \leqslant i \leqslant 3$ form a basis of $\mathbb{R}^3 \times \{0\}$. Fixing $\xi \in \widehat{\mathbb{R}}^3$ and $r > 0$, one may write

$$\xi - \xi_4 G(s) = \sum_{i=1}^{3} r^i \eta_i G^{(i)}(s) \tag{9.4}$$

for some vector of coefficients $(\eta_1, \eta_2, \eta_3) \in \mathbb{R}^3$. The powers of r appearing in the above expression play a normalising rôle below. For each $1 \leqslant k \leqslant 3$ form the inner product of both sides of the above identity with the Frenet vector $\mathbf{e}_k(s)$. Combining the resulting expressions with the linear independence relations inherent in (9.2), the coefficients η_k can be related to the numbers $\langle \xi, \mathbf{e}_k(s)\rangle$ via a lower anti-triangular transformation, viz.

$$\begin{bmatrix} \langle \xi, \mathbf{e}_1(s)\rangle \\ \langle \xi, \mathbf{e}_2(s)\rangle \\ \langle \xi, \mathbf{e}_3(s)\rangle \end{bmatrix} = \begin{bmatrix} 0 & 0 & \langle G_{\mathbf{a}}^{(3)}(s), \mathbf{e}_1(s)\rangle \\ 0 & \langle G_{\mathbf{a}}^{(2)}(s), \mathbf{e}_2(s)\rangle & \langle G_{\mathbf{a}}^{(3)}(s), \mathbf{e}_2(s)\rangle \\ \langle G_{\mathbf{a}}^{(1)}(s), \mathbf{e}_3(s)\rangle & \langle G_{\mathbf{a}}^{(2)}(s), \mathbf{e}_3(s)\rangle & \langle G_{\mathbf{a}}^{(3)}(s), \mathbf{e}_3(s)\rangle \end{bmatrix} \begin{bmatrix} r\eta_1 \\ r^2\eta_2 \\ r^3\eta_3 \end{bmatrix}. \tag{9.5}$$

Recall that

$$\pi_{2,\bar{\gamma}}(s;\, r) := \big\{\xi \in \widehat{\mathbb{R}}^4 : |\langle \mathbf{e}_j(s), \xi\rangle| \lesssim r^{4-j} \text{ for } 1 \leqslant j \leqslant 3,\ |\langle \mathbf{e}_4(s), \xi\rangle| \sim 1\big\}.$$

Thus, if $\xi \in \pi_{2,\gamma}(s; r)$, then it follows from combining the above definition and (9.3) with (9.5) that $|\eta_i| \lesssim_\gamma 1$ for $1 \leqslant i \leqslant 3$. Similarly, the localisation (9.1) implies that

$$\pi_{2,\gamma}(s; r) \subseteq \mathcal{R} := [-2, 2]^3 \times [1/4, 4].$$

The identity (9.4) can be succinctly expressed using matrices. In particular, for $s \in I_0$ and $r > 0$, define the 4×4 matrix

$$[g]_{\mathcal{C},s,r} := \begin{pmatrix} [g]_{s,r} & g(s) \\ 0 & 1 \end{pmatrix}. \tag{9.6}$$

Here the block $[g]_{s,r}$ is the 3×3 matrix as defined in (4.1). With this notation, the identity (9.4) may be written as

$$\xi = [g]_{\mathcal{C},s,r} \cdot \eta \qquad \text{where } \eta = (\eta_1, \eta_2, \eta_3, \xi_4).$$

Moreover, if $\xi \in \pi_{2,\gamma}(s;r)$, then the preceding observations show that η in the above equation may be taken to lie in a bounded region and so

$$\pi_{2,\gamma}(s;r) \subseteq [g]_{\mathcal{C},s,Cr}\big([-2,2]^4\big) \cap \mathcal{R}, \tag{9.7}$$

where $C \geqslant 1$ is a suitably large dimensional constant.

9.2. **A square function estimate for cones generated by non-degenerate curves.** Here the geometric setup described in §9.1 is abstracted.

Definition 9.1. *For $g\colon [-1,1] \to \mathbb{R}^3$ a smooth curve, let Γ_g denote the codimension 2 cone in $\mathbb{R}^4$ parametrised by*

$$(\rho, s) \mapsto \rho \begin{pmatrix} g(s) \\ 1 \end{pmatrix} \qquad \text{for } (\rho, s) \in U := [1/4, 4] \times [-1,1].$$

In this case, Γ_g is referred to as the cone generated by g.

In view of (9.7), one wishes to establish a reverse square function estimate with respect to the *r-plates*

$$\theta(s;r) := [g]_{\mathcal{C},s,r}\big([-2,2]^4\big) \cap \mathcal{R}.$$

In some cases it will be useful to highlight the choice of function g by writing $\theta(g;s;r)$ for $\theta(s;r)$. Note that each of these plates lies in a neighbourhood of the cone Γ_g. We think of the union of all plates $\theta(s;r)$ as s varies over the domain $[-1,1]$ as forming an anisotropic neighbourhood of Γ_g.

Definition 9.2. *A collection $\Theta(r)$ of r-plates is a plate family for Γ_g if it consists of $\theta(g;s;r)$ for s varying over an r-separated subset of $[-1,1]$.*

In view of the preceding observations, Theorem 5.3 is a consequence of the following result.

Theorem 9.3. *Suppose $g\colon [-1,1] \to \mathbb{R}^3$ is a smooth, non-degenerate curve and $\Theta(r)$ is an r-plate family for Γ_g for some dyadic $0 < r \leqslant 1$. For all $\varepsilon > 0$ the inequality*

$$\Big\| \sum_{\theta \in \Theta(r)} f_\theta \Big\|_{L^4(\mathbb{R}^4)} \lesssim_\varepsilon r^{-\varepsilon} \Big\| \Big(\sum_{\theta \in \Theta(r)} |f_\theta|^2 \Big)^{1/2} \Big\|_{L^4(\mathbb{R}^4)}$$

holds whenever $(f_\theta)_{\theta \in \Theta(r)}$ is a sequence of functions satisfying $\operatorname{supp} \widehat{f}_\theta \subseteq \theta$ for all $\theta \in \Theta(r)$.

9.3. **Multilinear estimates.** The proof of Theorem 9.3 follows an argument of Lee–Vargas [16] which relies on first establishing a multilinear variant of the desired square function inequality.

Let $\mathfrak{I}$ denote the collection of all dyadic subintervals of $[-1,1]$ and for any dyadic number $0 < r \leqslant 1$ let $\mathfrak{I}(r)$ denote the subset of $\mathfrak{I}$ consisting of all intervals of length r. Given any pair of dyadic scales $0 < \lambda_1 \leqslant \lambda_2 \leqslant 1$ and $J \in \mathfrak{I}(\lambda_2)$, let $\mathfrak{I}(J;\, \lambda_1)$ denote the collection of all $I \in \mathfrak{I}(\lambda_1)$ which satisfy $I \subseteq J$.

Fix $0 < r \leqslant 1$ and for each $0 \leqslant \lambda \leqslant 1$ decompose $\Theta(r)$ as a disjoint union of subsets $\Theta(I;\, r)$ for $I \in \mathfrak{I}(\lambda)$ such that:

i) If $\theta(s;r) \in \Theta(I;\, r)$, then $s \in I$;
ii) If $r \leqslant \lambda_1 \leqslant \lambda_2$ and $J \in \mathfrak{I}(\lambda_2)$, then $\Theta(J;\, r) = \bigcup_{I \in \mathfrak{I}(J;\lambda_1)} \Theta(I;\, r)$.

Thus, if for all $r \leqslant \lambda \leqslant 1$ we define

$$f_I := \sum_{\theta \in \Theta(I;\,r)} f_\theta \qquad \text{for all } I \in \mathfrak{I}(\lambda), \tag{9.8}$$

then for all $r \leqslant \lambda_1 \leqslant \lambda_2$ it follows that

$$f_J = \sum_{I \in \mathfrak{I}(J;\,\lambda_1)} f_I \qquad \text{for all } J \in \mathfrak{I}(\lambda_2).$$

For each dyadic number $0 < \lambda \leqslant 1$ let $\mathfrak{I}^4_{\mathrm{sep}}(\lambda)$ denote the collection of 4-tuples of intervals $\vec{I} = (I_1, \dots, I_4) \in \mathfrak{I}(\lambda)^4$ which satisfy the separation condition

$$\mathrm{dist}(I_1, \dots, I_4) := \min_{1 \leqslant \ell_1 < \ell_2 \leqslant 4} \mathrm{dist}(I_{\ell_1}, I_{\ell_2}) \geqslant \lambda.$$

Proposition 9.4. *Let $0 < r \leqslant \lambda < 1$ be dyadic. If $(I_1, \dots, I_4) \in \mathfrak{I}^4_{\mathrm{sep}}(\lambda)$ and $\varepsilon > 0$, then*

$$\Big\| \prod_{\ell=1}^4 \Big| \sum_{\theta \in \Theta(I_\ell;r)} f_\theta \Big|^{1/4} \Big\|_{L^4(\mathbb{R}^4)} \lesssim_\varepsilon M(\lambda) r^{-\varepsilon} \prod_{\ell=1}^4 \Big\| \Big(\sum_{\theta \in \Theta(I_\ell;r)} |f_\theta|^2 \Big)^{1/2} \Big\|_{L^4(\mathbb{R}^4)}^{1/4}$$

holds whenever $(f_\theta)_{\theta \in \Theta(r)}$ is a sequence of functions satisfying $\mathrm{supp}\, \widehat{f}_\theta \subseteq \theta$ *for all $\theta \in \Theta(r)$, where* $\sup_{\lambda \in [\lambda_0, 1]} M(\lambda) < \infty$ *for all $\lambda_0 > 0$.*

Using a standard argument, Proposition 9.4 will follow from a 4-linear Fourier restriction estimate. To state the latter inequality, given an interval $J \subseteq [-1,1]$ let Γ_J denote the image of $\Gamma_g \colon (\rho, s) \mapsto \rho\,(g(s), 1)^\top$ restricted to the set $U_J := [1/4, 4] \times J$ and, for $r > 0$, let $N_r\Gamma_J$ denote the r-neighbourhood of Γ_J.

Proposition 9.5. *If $(I_1, \dots, I_4) \in \mathfrak{I}^4_{\mathrm{sep}}(\lambda)$, then for all $0 < r \leqslant \lambda$ and all $\varepsilon > 0$ the inequality*

$$\Big\| \prod_{\ell=1}^4 |F_\ell|^{1/4} \Big\|_{L^4(\mathbb{R}^4)} \lesssim_\varepsilon M(\lambda) r^{1-\varepsilon} \prod_{\ell=1}^4 \|F_\ell\|_{L^2(\mathbb{R}^4)}^{1/4}$$

holds for all $F_\ell \in L^2(\mathbb{R}^4)$ with $\mathrm{supp}\, \widehat{F}_\ell \subseteq N_r\Gamma_{I_\ell}$ *for $1 \leqslant \ell \leqslant 4$.*

Given an interval $J \subset [-1,1]$, define the extension operator

$$E_J f(x) := \int_{U_J} e^{i\langle \Gamma(u), x\rangle} f(u)\,\mathrm{d}u \qquad \text{for all } f \in L^1(U_J),$$

where $U_J := [1/4, 4] \times J$ as above. By standard uncertainty principle techniques and Plancherel's theorem (see, for instance, [5] or [23, Appendix]), Proposition 9.5 is a consequence of the following multilinear extension estimate.

Proposition 9.6. *If $(I_1, \dots, I_4) \in \mathfrak{I}^4_{\mathrm{sep}}(\lambda)$, then for all $R \geqslant 1$ and all $\varepsilon > 0$ the inequality*

$$\Big\| \prod_{\ell=1}^4 |E_{I_\ell} f_\ell|^{1/4} \Big\|_{L^4(B_R)} \lesssim_\varepsilon M(\lambda) R^\varepsilon \prod_{\ell=1}^4 \|f_\ell\|_{L^2(U_{I_\ell})}^{1/4}$$

holds for all $f_\ell \in L^2(U)$ for $1 \leqslant \ell \leqslant 4$, where B_R denotes a ball of radius R.

We refer to the above references for the argument used to pass from Proposition 9.6 to Proposition 9.5 and turn to the proof of the extension estimate.

Proof of Proposition 9.6. This inequality is a special case (via a compactness argument) of the recent generalisation of the Bennett–Carbery–Tao restriction theorem [5] due to Bennett–Bez–Flock–Lee [3, Theorem 1.3]; an improved version with R^ε replaced by $(\log R)^{O(d)}$ has also been obtained by Zhang [25, (1.8)], although the R^ε loss suffices for our purposes. In order to see this, we must verify a certain linear-algebraic condition on the tangent planes to Γ. The setup is recalled presently.

Fix $u_\ell = (\rho_\ell, s_\ell) \in U_{I_\ell}$ for $1 \leqslant \ell \leqslant 4$. We construct a *Brascamp–Lieb datum* $(\mathbf{L}, \mathbf{p})$ by taking

$$\mathbf{L} := (\pi_1, \dots, \pi_4) \qquad \text{and} \qquad \mathbf{p} := (p_1, \dots, p_4) := (1/2, \dots, 1/2)$$

where each $\pi_\ell \colon \mathbb{R}^4 \to V_\ell$ is the orthogonal projection map from $\mathbb{R}^4$ to the 2-dimensional tangent space V_ℓ to Γ at $\Gamma(u_\ell)$. With this definition, the problem is to show that $\mathrm{BL}(\mathbf{L}, \mathbf{p}) < \infty$, where the *Brascamp–Lieb constant* $\mathrm{BL}(\mathbf{L}, \mathbf{p})$ is as defined in, for instance, [3]. By the characterisation of finiteness of the Brascamp–Lieb constant from [4] and our choice of datum, it suffices to verify the following two conditions:

i) $\displaystyle\sum_{\ell=1}^{4} (\dim \operatorname{Im} \pi_\ell)\, p_\ell = 4.$

ii) $\displaystyle\dim V \leqslant \frac{1}{2} \sum_{\ell=1}^{4} \dim \big(\pi_\ell V\big)$ holds for all linear subspaces $V \subseteq \mathbb{R}^4$.

The scaling condition i) is immediate from the choice of datum and it remains to prove the dimension condition ii).

Clearly one may replace π_ℓ with the linear map associated to the 2×4 Jacobian matrix $\mathrm{d}\Gamma|_{(\rho_\ell, s_\ell)}$. By subtracting the first column from the third column and applying the fundamental theorem of calculus,

$$\det \begin{bmatrix} g(s_{\ell_1}) & g'(s_{\ell_1}) & g(s_{\ell_2}) & g'(s_{\ell_2}) \\ 1 & 0 & 1 & 0 \end{bmatrix} = - \int_{s_{\ell_1}}^{s_{\ell_2}} \det \begin{bmatrix} g'(s_{\ell_1}) & g'(s) & g'(s_{\ell_2}) \end{bmatrix} \,\mathrm{d}s.$$

Furthermore, by repeated application of column reduction and the fundamental theorem of calculus, it follows from the non-degeneracy hypothesis and the initial localisation that

$$\big| \det \begin{bmatrix} g'(s_{\ell_1}) & g'(s) & g'(s_{\ell_2}) \end{bmatrix} \big| \gtrsim |s_{\ell_2} - s_{\ell_1}||s - s_{\ell_1}||s_{\ell_2} - s|;$$

see, for instance, [12, Proposition 4.1]. Consequently, the determinant has constant sign and

$$\big| \det \begin{bmatrix} \mathrm{d}\Gamma|_{(\rho_{\ell_1}, s_{\ell_1})} & \mathrm{d}\Gamma|_{(\rho_{\ell_2}, s_{\ell_2})} \end{bmatrix} \big| \gtrsim |\rho_{\ell_1}||\rho_{\ell_2}||s_{\ell_2} - s_{\ell_1}|^4 \gtrsim \lambda^4, \tag{9.9}$$

where the final bound is due to the separation between the I_ℓ. Note that (9.9) is equivalent to the geometric condition that $V_{\ell_1} + V_{\ell_2} = \mathbb{R}^4$ and therefore

$$V_{\ell_1}^\perp \cap V_{\ell_2}^\perp = \big(V_{\ell_1} + V_{\ell_2}\big)^\perp = \{0\}. \tag{9.10}$$

With this observation, it is now a simple matter to verify the dimension condition ii) above.

- If $\dim V = 4$ or $\dim V = 0$, then ii) is trivial.
- If $\dim V = 1$, then it suffices to show that $\dim \pi_\ell V = 1$ for at least two values of ℓ. Suppose $\dim \pi_{\ell_1} V = \dim \pi_{\ell_2} V = 0$ for some $1 \leqslant \ell_1 < \ell_2 \leqslant 4$, so that

$$V \subseteq \ker \pi_{\ell_1} \cap \ker \pi_{\ell_2} = V_{\ell_1}^\perp \cap V_{\ell_2}^\perp.$$

  However, in this case it follows from (9.10) that $V = \{0\}$, which contradicts our dimension hypothesis. Thus, $\dim \pi_\ell V = 0$ for at most a single value of ℓ, which more than suffices for our purpose.

- If $\dim V = 2$, then we may assume that $\dim \pi_{\ell_0} V = 0$ for some $1 \leqslant \ell_0 \leqslant 4$, since otherwise ii) is immediate. By dimensional considerations, it follows that $V = V_{\ell_0}^{\perp}$. Now let $1 \leqslant \ell \leqslant 4$ with $\ell \neq \ell_0$. By (9.10), it follows that $V \cap V_{\ell}^{\perp} = \{0\}$. Thus, by the rank-nullity theorem applied to the mapping $\pi_\ell|_V : V \to V_\ell$, we deduce that $\dim \pi_\ell V = 2$. Since this is true for three distinct values of ℓ, property ii) holds.
- If $\dim V = 3$, then it is clear that $\dim \pi_\ell V \geqslant 1$ for all $1 \leqslant \ell \leqslant 4$. Suppose there exist $1 \leqslant \ell_1 < \ell_2 \leqslant 4$ such that $\dim(\pi_{\ell_1} V) = \dim(\pi_{\ell_2} V) = 1$. In this case, by the rank-nullity theorem applied to $\pi_{\ell_i}|_V : V \to V_{\ell_i}$ and dimensional considerations,
$$V_{\ell_1}^{\perp} + V_{\ell_2}^{\perp} = \ker \pi_{\ell_1} + \ker \pi_{\ell_2} \subseteq V.$$
However, in this case it follows from (9.10) that $V = \mathbb{R}^4$, which contradicts our dimension hypothesis. Thus, $\dim \pi_\ell V = 1$ for at most a single value of ℓ, and for the remaining values of ℓ the dimension is at least 2. This again more than suffices for our purpose.

This establishes the finiteness of the Brascamp–Lieb constant and concludes the proof. □

Having established the multilinear restriction estimate, it is a simple matter to deduce the desired multilinear square function bound.

Proof of Proposition 9.4. Let B be a ball of radius r^{-1} in $\mathbb{R}^4$ with centre x_0. Fix $\eta \in \mathcal{S}(\mathbb{R}^4)$ with $\operatorname{supp} \hat{\eta} \subset B(0,1)$ and $|\eta(x)| \gtrsim 1$ on $B(0,1)$ and define $\eta_B(x) := \eta\big(r(x - x_0)\big)$. By the rapid decay of η, it suffices to show that
$$\Big\| \prod_{\ell=1}^{4} \Big| \sum_{\theta \in \Theta(I_\ell; r)} f_\theta \Big|^{1/4} \Big\|_{L^4(B)} \lesssim_\varepsilon M(\lambda) r^{-\varepsilon} \prod_{\ell=1}^{4} \Big\| \Big(\sum_{\theta \in \Theta(I_\ell; r)} |f_\theta|^2 \Big)^{1/2} |\eta_B|^{1/2} \Big\|_{L^4(\mathbb{R}^4)}^{1/4}.$$
Indeed, once established, this inequality can be summed over a collection of finitely-overlapping balls B which cover $\mathbb{R}^4$ to obtained the desired global estimate.

For $1 \leqslant \ell \leqslant 4$ define
$$F_\ell := \sum_{\theta \in \Theta(I_\ell; r)} f_\theta \, \eta_B$$
so that each F_ℓ is Fourier supported in an $O(r)$-neighbourhood of Γ_{I_ℓ}. Applying Proposition 9.5 to these functions, it follows that
$$\Big\| \prod_{\ell=1}^{4} \Big| \sum_{\theta \in \Theta(I_\ell; r)} f_\theta \Big|^{1/4} \Big\|_{L^4(B)} \lesssim \Big\| \prod_{\ell=1}^{4} |F_\ell|^{1/4} \Big\|_{L^4(\mathbb{R}^4)} \lesssim_\varepsilon M(\lambda) r^{1-\varepsilon} \prod_{j=1}^{4} \Big\| \sum_{\theta \in \Theta(I_\ell; r)} f_\theta \, \eta_B \Big\|_{L^2(\mathbb{R}^4)}^{1/4}.$$
Note that the functions $f_\theta \, \eta_B$ appearing in the right-hand sum have essentially disjoint Fourier support. Consequently, by Plancherel's theorem and Hölder's inequality,
$$\begin{aligned} \Big\| \sum_{\theta \in \Theta(I_\ell; r)} f_\theta \, \eta_B \Big\|_{L^2(\mathbb{R}^4)} &\lesssim \Big\| \Big(\sum_{\theta \in \Theta(I_\ell; r)} |f_\theta|^2 \Big)^{1/2} |\eta_B| \Big\|_{L^2(\mathbb{R}^4)} \\ &\lesssim r^{-1} \Big\| \Big(\sum_{\theta \in \Theta(I_\ell; r)} |f_\theta|^2 \Big)^{1/2} |\eta_B|^{1/2} \Big\|_{L^4(\mathbb{R}^4)}. \end{aligned}$$
Combining the previous two displays completes the proof. □

9.4. **Rescaling.** By combining Proposition 9.6 with an affine rescaling argument, one may deduce a useful refined version of the multilinear inequality. This improves the dependence on separation parameter λ under an additional localisation hypothesis on the intervals $J_1, \dots, J_4$.

Given dyadic scales $0 < \lambda_1 \leqslant \lambda_2 \leqslant 1$ and $J \in \mathfrak{I}(\lambda_2)$, let $\mathfrak{I}^4_{\mathrm{sep}}(J; \lambda_1)$ denote the collection of all 4-tuples of intervals $\vec{I} = (I_1, \dots, I_4) \in \mathfrak{I}^4_{\mathrm{sep}}(\lambda_1)$ such that $I_\ell \subseteq J$ for all $1 \leqslant \ell \leqslant 4$.

With this definition, the refined version of Proposition 9.4 reads as follows.

Corollary 9.7. *Fix dyadic scales $0 < r \leqslant \lambda_1 \leqslant \lambda_2 \leqslant 1$. If $J \in \mathfrak{I}(\lambda_2)$, $(I_1, \dots, I_4) \in \mathfrak{I}^4_{\mathrm{sep}}(J;\, \lambda_1)$ and $\varepsilon > 0$, then*

$$\Big\| \prod_{\ell=1}^{4} \Big| \sum_{\theta \in \Theta(I_\ell; r)} f_\theta \Big|^{1/4} \Big\|_{L^4(\mathbb{R}^4)} \lesssim_\varepsilon M(\lambda_1/\lambda_2) r^{-\varepsilon} \prod_{j=1}^{4} \Big\| \Big(\sum_{\theta \in \Theta(I_\ell; r)} |f_\theta|^2 \Big)^{1/2} \Big\|_{L^4(\mathbb{R}^4)}^{1/4}$$

holds whenever $(f_\theta)_{\theta \in \Theta(r)}$ is a sequence of functions satisfying $\operatorname{supp} \widehat{f}_\theta \subseteq \theta$ *for all $\theta \in \Theta(r)$.*

Proof. The result is a consequence of Proposition 9.4 and a rescaling argument. Let $J = [\sigma - \lambda_2, \sigma + \lambda_2] \subseteq [-1, 1]$ and recall the definition of the rescaled curve

$$g_{\sigma,\lambda_2}(\tilde{s}) := \big([g]_{\sigma,\lambda_2}\big)^{-1}\big(g(\sigma + \lambda_2 \tilde{s}) - g(\sigma)\big).$$

Differentiating this expression, it follows that $g^{(j)}_{\sigma,\lambda_2}(\tilde{s}) = \lambda_2^j \big([g]_{\sigma,\lambda_2}\big)^{-1} g^{(j)}(\sigma + \lambda_2 \tilde{s})$ for $j \geqslant 1$ and so

$$[g_{\sigma,\lambda_2}]_{\tilde{s},\tilde{r}} = \big([g]_{\sigma,\lambda_2}\big)^{-1} \circ [g]_{s,r} \qquad \text{where } s = \sigma + \lambda_2 \tilde{s} \text{ and } r = \lambda_2 \tilde{r}.$$

From this and the definition (9.6), it is not difficult to deduce that

$$[g_{\sigma,\lambda_2}]_{\mathcal{C},\tilde{s},\tilde{r}} = \big([g]_{\mathcal{C},\sigma,\lambda_2}\big)^{-1} \circ [g]_{\mathcal{C},s,r}.$$

Suppose $\theta \in \Theta(J;\, r)$ and $\operatorname{supp} \widehat{F}_\theta \subseteq \theta$. If $\theta = \theta(s, r)$, then

$$\operatorname{supp} \widehat{F}_\theta \circ [g]_{\mathcal{C},\sigma,\lambda_2} \subseteq \tilde{\theta}(\tilde{s}, \tilde{r})$$

where $\tilde{\theta}(\tilde{s}, \tilde{r})$ is the $\tilde{r}$-plate centred at $\tilde{s}$ defined with respect to $\tilde{g} := g_{\sigma,\lambda_2}$. Finally, note that the above rescaling maps the intervals $(I_1, \dots, I_4) \in \mathfrak{I}^4_{\mathrm{sep}}(J; \lambda_1)$ to intervals $(\tilde{I}_1, \dots \tilde{I}_4) \in \mathfrak{I}^4_{\mathrm{sep}}(\lambda_1/\lambda_2)$. □

9.5. **Broad/narrow analysis.** Here arguments from [14] are adapted to pass from the multilinear estimates of Proposition 9.4 (or, more precisely, Corollary 9.7) to the linear estimates in Theorem 9.3.

The key ingredient is the following decomposition lemma, which follows by iteratively applying the decomposition scheme discussed in [14].

Lemma 9.8. *Let $\varepsilon > 0$ and $r > 0$. There exist dyadic numbers $C_\varepsilon \geqslant 1$, r_{n} and r_{b} satisfying*

$$r < r_{\mathrm{n}} \lesssim_{\varepsilon,1} r, \qquad r < r_{\mathrm{b}} \leqslant 1 \tag{9.11}$$

such that

$$\Big\| \sum_{\theta \in \Theta(r)} f_\theta \Big\|_{L^4(\mathbb{R}^4)} \lesssim_\varepsilon r^{-\varepsilon} \Big(\sum_{I \in \mathfrak{I}(r_{\mathrm{n}})} \|f_I\|^4_{L^4(\mathbb{R}^4)} \Big)^{1/4} + r^{-\varepsilon} \Big(\sum_{\substack{J \in \mathfrak{I}(C_\varepsilon r_{\mathrm{b}}) \\ \vec{I} \in \mathfrak{I}^4_{\mathrm{sep}}(J;\, r_{\mathrm{b}})}} \Big\| \prod_{\ell=1}^{4} |f_{I_\ell}|^{1/4} \Big\|^4_{L^4(\mathbb{R}^4)} \Big)^{1/4} \tag{9.12}$$

holds whenever $(f_\theta)_{\theta \in \Theta(r)}$ is a sequence of functions satisfying $\operatorname{supp} \widehat{f}_\theta \subseteq \theta$ *for all $\theta \in \Theta(r)$.*

We provide a proof of (an abstract version of) the above lemma in Appendix A (more precisely, Lemma 9.8 follows from applying Lemma A.2 to the decomposition $f := \sum_{\theta \in \Theta(r)} f_\theta$ for a fixed dyadic scale $0 < r \leqslant 1$ and $\varepsilon > 0$).

We are now in position to prove the desired reverse square function estimate.

Proof of Theorem 9.3. Fix $0 < r \leqslant 1$ a choice of dyadic scale and $\varepsilon > 0$, and apply Lemma 9.8. The analysis splits into two cases depending on which of the right-hand terms in (9.12) dominates. We refer to the first term as the *narrow* term and to the second term as the *broad* term.

The narrow case. Suppose the narrow term dominates the right-hand side of (9.12) in the sense that

$$\Big\|\sum_{\theta\in\Theta(r)} f_\theta\Big\|_{L^4(\mathbb{R}^4)} \lesssim_\varepsilon r^{-\varepsilon}\Big(\sum_{I\in\mathfrak{I}(r_{\mathrm{n}})}\Big\|\sum_{\theta\in\Theta(I;r)} f_\theta\Big\|_{L^4(\mathbb{R}^4)}^4\Big)^{1/4}.$$

This case is dealt with using a trivial argument. If $I \in \mathfrak{I}(r_{\mathrm{n}})$, then

$$\Big|\sum_{\theta\in\Theta(I;r)} f_\theta\Big| \lesssim_\varepsilon \Big(\sum_{\theta\in\Theta(I;r)} |f_\theta|^2\Big)^{1/2} \tag{9.13}$$

by Cauchy–Schwarz, since the condition $r_{\mathrm{n}} \sim r$ from (9.11) implies that there are only $O_\varepsilon(1)$ intervals belonging to $\Theta(I;\, r)$. Thus,

$$\Big\|\sum_{\theta\in\Theta(r)} f_\theta\Big\|_{L^4(\mathbb{R}^4)} \lesssim_\varepsilon r^{-\varepsilon}\Big\|\Big(\sum_{I\in\mathfrak{I}(r_{\mathrm{n}})}\sum_{\theta\in\Theta(I;\,r)}|f_\theta|^2\Big)^{1/2}\Big\|_{L^4(\mathbb{R}^4)} = r^{-\varepsilon}\Big\|\Big(\sum_{\theta\in\Theta(r)}|f_\theta|^2\Big)^{1/2}\Big\|_{L^4(\mathbb{R}^4)},$$

where the first step follows from (9.13) and the embedding $\ell^2 \hookrightarrow \ell^4$ and the last step from the definition of $\mathfrak{I}(r_{\mathrm{n}})$ and $\Theta(I;r)$.

The broad case. Suppose the broad term dominates the right-hand side of (9.12) in the sense that

$$\Big\|\sum_{\theta\in\Theta(r)} f_\theta\Big\|_{L^4(\mathbb{R}^4)} \lesssim_\varepsilon r^{-\varepsilon}\Big(\sum_{\substack{J\in\mathfrak{I}(C_\varepsilon r_{\mathrm{b}})\\ \vec{I}\in\mathfrak{I}^4_{\mathrm{sep}}(J;\,r_{\mathrm{b}})}}\Big\|\prod_{\ell=1}^4\Big|\sum_{\theta\in\Theta(I_\ell;r)} f_\theta\Big|^{1/4}\Big\|_{L^4(\mathbb{R}^4)}^4\Big)^{1/4}.$$

This case is treated using the rescaled multilinear inequality from Corollary 9.7. Since $\#\mathfrak{I}^4(J;\, r_{\mathrm{b}}) \lesssim_\varepsilon 1$ for each $J \in \mathfrak{I}(C_\varepsilon r_{\mathrm{b}})$, by Hölder's inequality

$$\Big\|\sum_{\theta\in\Theta(r)} f_\theta\Big\|_{L^4(\mathbb{R}^4)} \lesssim_\varepsilon r^{-\varepsilon}\Big(\sum_{J\in\mathfrak{I}(C_\varepsilon r_{\mathrm{b}})}\Big(\sum_{\vec{I}\in\mathfrak{I}^4_{\mathrm{sep}}(J;\,r_{\mathrm{b}})}\Big\|\prod_{\ell=1}^4\Big|\sum_{\theta\in\Theta(I_\ell;r)} f_\theta\Big|^{1/4}\Big\|_{L^4(\mathbb{R}^4)}^{16}\Big)^{1/4}\Big)^{1/4}.$$

Applying Corollary 9.7 with $\lambda_1 := r_{\mathrm{b}}$ and $\lambda_2 := C_\varepsilon r_{\mathrm{b}}$, one deduces that

$$\Big\|\sum_{\theta\in\Theta(r)} f_\theta\Big\|_{L^4(\mathbb{R}^4)} \lesssim_\varepsilon r^{-\varepsilon}\Big(\sum_{J\in\mathfrak{I}(C_\varepsilon r_{\mathrm{b}})}\Big(\sum_{\vec{I}\in\mathfrak{I}^4_{\mathrm{sep}}(J;\,r_{\mathrm{b}})}\prod_{\ell=1}^4\Big\|\Big(\sum_{\theta\in\Theta(I_\ell;\,r)}|f_\theta|^2\Big)^{1/2}\Big\|_{L^4(\mathbb{R}^4)}^4\Big)^{1/4}\Big)^{1/4}.$$

Relaxing the inner range of summation to all $\vec{I} \in \mathfrak{I}(J;\, r_{\mathrm{b}})^4$ (that is, dropping the separation condition),

$$\Big\|\sum_{\theta\in\Theta(r)} f_\theta\Big\|_{L^4(\mathbb{R}^4)} \lesssim_\varepsilon r^{-\varepsilon}\Big(\sum_{I\in\mathfrak{I}(r_{\mathrm{b}})}\Big\|\Big(\sum_{\theta\in\Theta(I;\,r)}|f_\theta|^2\Big)^{1/2}\Big\|_{L^4(\mathbb{R}^4)}^4\Big)^{1/4}.$$

Arguing as in the last steps of the narrow case, using the embedding $\ell^2 \hookrightarrow \ell^4$, now concludes the argument. □

10. Proof of the forward square function inequality in $\mathbb{R}^3$

In this section we establish the L^2 weighted forward square function estimate from Proposition 5.4. Before we commence, it is useful to recall the basic setup. Let $\gamma \in \mathfrak{G}_3(\delta_0)$ for $0 < \delta_0 \ll 1$ and $\mathbf{e}_j : [-1, 1] \to S^3$ for $1 \leqslant j \leqslant 3$ be the associated Frenet frame. Recall that this satisfies

$$\mathbf{e}_j(s) = \vec{e}_j + O(\delta_0) \qquad \text{for } 1 \leqslant j \leqslant 3 \text{ and } s \in I_0 = [-\delta_0, \delta_0], \tag{10.1}$$

where the $\vec{e}_j$ denote the standard basis vectors. For $0 < r \leqslant 1$, recall that a $(0, r)$-*Frenet box* is a set of the form

$$\pi_{0,\gamma}(s;\, r) := \big\{\xi \in \widehat{\mathbb{R}}^3 : |\langle \mathbf{e}_1(s), \xi\rangle| \leqslant r,\ 1/2 \leqslant |\langle \mathbf{e}_2(s), \xi\rangle| \leqslant 1,\ |\langle \mathbf{e}_3(s), \xi\rangle| \leqslant 1\big\}$$

for some $s \in [-1,1]$. Proposition 5.4 concerns smooth frequency projections $\chi_\pi(D)$ where χ_π is a bump function adapted to a $(0,r)$-Frenet box π.

10.1. **Geometric observations.** We begin by reparametrising the sets $\pi_{0,\gamma}(s;r)$ using an argument similar to that of §9.1. Define the functions $g_j : I_0 \to \mathbb{R}^3$ by $g_j(s) := -\mathbf{e}_{1j}(s)\,\mathbf{e}_{11}(s)^{-1}$ for $j = 2,3$ (note that $\mathbf{e}_{1,1}(s)$ is bounded away from 0 by (10.1)) so that

$$\langle \mathbf{e}_1(s), \xi\rangle = \mathbf{e}_{11}(s)\big(\xi_1 - \xi_2 g_2(s) - \xi_3 g_3(s)\big). \tag{10.2}$$

Thus, we have the containment property

$$\pi_{0,\gamma}(s;\, r) \subseteq \theta(s;\, Cr) \tag{10.3}$$

where $\theta(s;\, r)$ is the region

$$\theta(s;\, r) := \Big\{\xi \in \widehat{\mathbb{R}}^3 : \big|\xi_1 - \sum_{j=2}^{3} \xi_j g_j(s)\big| < r \text{ and } 1/4 \leqslant |\xi_2| \leqslant 4,\ |\xi_3| \leqslant 4\Big\}.$$

We refer to the sets $\theta(s;\, r)$ as 'plates'.

It is useful to note that the curves $g_j\colon I_0 \to \mathbb{R}^3$ satisfy a certain regularity condition. In particular, for each $\mathbf{a} = (a_2, a_3) \in \mathbb{R}^2$ define the function $g_{\mathbf{a}}(s) := a_2 g_2(s) + a_3 g_3(s)$. By differentiating (10.2) with respect to s and evaluating the result at $\xi = (0, a_2, a_3)$, provided the parameter $\delta_0 > 0$ featured in (10.1) is chosen sufficiently small, it follows that

$$|g_{\mathbf{a}}'(s)| \sim 1 \qquad \text{for all } \mathbf{a} \in [1/4, 4] \times [-1,1]. \tag{10.4}$$

Indeed, this is a simple consequence of the Frenet equations.

We also observe a dual version of the containment condition (10.3). In particular, if we define the dual Frenet box and dual plate

$$\begin{aligned}\pi^*_{0,\gamma}(s;r) &:= \big\{x \in \mathbb{R}^3 : |\langle \mathbf{e}_1(s), x\rangle| \leqslant r^{-1} \text{ and } |\langle \mathbf{e}_j(s), x\rangle| \leqslant 1 \text{ for } j = 2,3\big\},\\ \theta^*(s;r) &:= \big\{x \in \mathbb{R}^3 : |x_1| \leqslant r^{-1} \text{ and } |x_j + g_j(s)x_1| \leqslant 4 \text{ for } j = 2,3\big\},\end{aligned}$$

then it follows that $\pi^*_{0,\gamma}(s;r) \subseteq \theta^*(s; C^{-1}r)$. To this, we first observe the identity

$$\begin{bmatrix}\langle x, \mathbf{e}_2(s)\rangle\\ \langle x, \mathbf{e}_3(s)\rangle\end{bmatrix} = \begin{bmatrix}\mathbf{e}_{22}(s) & \mathbf{e}_{23}(s)\\ \mathbf{e}_{32}(s) & \mathbf{e}_{33}(s)\end{bmatrix}\begin{bmatrix}x_2 + g_2(s)x_1\\ x_3 + g_3(s)x_1\end{bmatrix}, \tag{10.5}$$

which follows from the orthogonality between the Frenet vectors $\big(\mathbf{e}_j(s)\big)_{j=1}^3$. Since the right-hand 2×2 matrix is a small perturbation of the identity, the claimed containment property follows.

10.2. **The iteration scheme.** Our proof of Proposition 5.4 uses an iteration argument. This is based on the approach of Carbery and the fourth author in [10, Proposition 4.6], where a related inequality for the Córdoba sectorial square function was obtained. Driving the iteration scheme is an elementary pointwise square function bound due to Rubio de Francia [20]. Here it is convenient to state a slight generalisation of this result.

Lemma 10.1. *Let $\psi \in \mathscr{S}(\widehat{\mathbb{R}}^n)$, $A \in \mathrm{GL}(\mathbb{R}, n)$ and $G\colon \mathbb{Z}^m \to \mathbb{R}^n$. For all $N \in \mathbb{N}$ the pointwise inequality*

$$\sum_{\nu \in \mathbb{Z}^m} \big|\psi\big(AD - G(\nu)\big)f(x)\big|^2 \lesssim_{\psi,N} \sup_{\nu_2 \in \mathbb{Z}^m} \sum_{\nu_1 \in \mathbb{Z}^m} e^{-|G(\nu_1) - G(\nu_2)|/2} \int_{\mathbb{R}^n} |f(x - A^\top y)|^2 (1 + |y|)^{-N}\,\mathrm{d}y$$

holds for all $f \in \mathscr{S}(\mathbb{R}^n)$.

Proof. The case where $G\colon \mathbb{Z}^n \to \mathbb{Z}^n$ is the identity map is proven in [20]. The argument can be generalised to prove the above lemma, by replacing an application of Plancherel's theorem with a T^*T argument involving the Schur test. For convenience, the details of the argument are presented in Appendix B. □

To describe the iteration step, we first define smooth cutoff functions adapted to the plates θ defined above. As usual, let $\eta \in C_c^\infty(\mathbb{R})$ satisfy $\eta(u) = 1$ for $|u| \leqslant 1/2$ and $\operatorname{supp} \eta \subseteq [-1, 1]$ and define the multipliers

$$m_r^\nu(\xi) := \eta\Big(r^{-1}\big(\xi_1 - \sum_{j=2}^{3} \xi_j g_j(s_\nu)\big)\Big) \qquad \text{for } \nu \in \mathbb{Z} \text{ and } s_\nu := r\nu. \tag{10.6}$$

Let $b(\xi) = \tilde{\beta}(4^{-1}\xi_2)\,\eta(4^{-1}\xi_3)$ where here $\tilde{\beta}$ is as defined in (5.3) so that $(m_r^\nu \cdot b)(\xi) = 1$ if $\xi \in \theta(s_\nu; r)$. For the iteration scheme, we in fact work with truncated versions of the plates. Given $K \geqslant 1$, $-1 \leqslant s \leqslant 1$, $0 < r \leqslant 1$ and $\mathbf{a} = (a_2, a_3) \in \mathbb{R}^2$, consider the truncated plate

$$\theta^{\mathbf{a},K}(s;\, r) := \Big\{\xi \in \widehat{\mathbb{R}}^3 : \big|\xi_1 - \sum_{j=2}^{3} \xi_j g_j(s)\big| \leqslant r \text{ and } |\xi_j - a_j| \leqslant K^{-1} \text{ for } j = 2, 3\Big\}.$$

Correspondingly, we let $\zeta \in C_c^\infty(\mathbb{R})$ satisfy $\operatorname{supp}\zeta \subseteq [-1,1]$ and $\sum_{k\in\mathbb{Z}} \zeta(\,\cdot - k) \equiv 1$ and decompose

$$b = \sum_{\mathbf{a} \in K^{-1}\mathbb{Z}^2} b_{\mathbf{a}} \qquad \text{where} \qquad b_{\mathbf{a}}(\xi) := \prod_{j=2}^{3} \zeta\big(K(\xi_j - a_j)\big)\, b(\xi). \tag{10.7}$$

For $\mathbf{r} := (r_1, r_2, r_3) \in (0,1]^3$ and $s \in [-1,1]$ let $T_{\mathbf{e},\mathbf{r}}(s)$ denote the parallelepiped consisting of all vectors $x \in \mathbb{R}^3$ satisfying $|\langle x, \mathbf{e}_j(s)\rangle| \leqslant r_j^{-1}$ for $1 \leqslant j \leqslant 3$. These sets should be thought of scaled versions of the dual Frenet box $\pi_{0,\gamma}^*(s; r)$ introduced in §10.1. Consider the weighted averaging and Nikodym-type maximal operators associated to these sets, given by

$$\widetilde{\mathcal{A}}_{\mathbf{e},\mathbf{r}} g(x; s) := \int_{\mathbb{R}^3} g(x-y)\psi_{T_{\mathbf{e},\mathbf{r}}(s)}(y)\,\mathrm{d}y \qquad \text{and} \qquad \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}} g(x) := \sup_{s\in[-1,1]} |\widetilde{\mathcal{A}}_{\mathbf{e},\mathbf{r}} g(x;s)| \tag{10.8}$$

where

$$\psi_{T_{\mathbf{e},\mathbf{r}}(s)}(x) := \big(\prod_{j=1}^{3} r_j\big)\,\big(1 + \sum_{j=1}^{3} r_j |\langle \mathbf{e}_j(s), y\rangle|\big)^{-300}. \tag{10.9}$$

Here the subscript $\mathbf{e}$ refers to the Frenet frame $\mathbf{e} := (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$.

With the above definitions, the key iteration step is as follows.

Proposition 10.2. *Let $0 < r < 1$, $K \geqslant 1$, $\tilde{r} = Kr$, $\mathbf{r} := (r, K^{-1}, K^{-1})$ and $\mathbf{a} = (a_2, a_3) \in [1/4, 4] \times [-1, 1]$. With the above definitions,*

$$\int_{\mathbb{R}^3} \sum_{\nu\in\mathbb{Z}} \big|(m_r^\nu \cdot b_{\mathbf{a}})(D) f(x)\big|^2 w(x)\,\mathrm{d}x \lesssim \int_{\mathbb{R}^3} \sum_{\tilde{\nu}\in\mathbb{Z}} \big|(m_{\tilde{r}}^{\tilde{\nu}} \cdot b_{\mathbf{a}})(D) f(x)\big|^2\, \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}} \circ \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}}\, w(x)\,\mathrm{d}x$$

for any non-negative $w \in L^1_{\mathrm{loc}}(\mathbb{R}^3)$.

Proof. The proof is based on the following simple geometric observation, which motivates the use of the truncation. If $|s - \tilde{s}| \leqslant Kr$, then the plates $\theta^{\mathbf{a},K}(s;\, r)$, $\theta^{\mathbf{a},K}(\tilde{s};\, r)$ are essentially parallel translates of one another. More precisely, if $\xi \in \theta^{\mathbf{a},K}(s;\, r)$, then

$$\Big|\xi_1 - \sum_{j=2}^{3} a_j\big(g_j(s) - g_j(\tilde{s})\big) - \sum_{j=2}^{3} \xi_j g_j(\tilde{s})\Big| \leqslant \Big|\xi_1 - \sum_{j=2}^{3} \xi_j g_j(s)\Big| + \sum_{j=2}^{3} |a_j - \xi_j||g_j(s) - g_j(\tilde{s})| \lesssim_{\mathbf{g}} r$$

and, consequently, there exists some constant $C_{\mathbf{g}}$ such that

$$\theta^{\mathbf{a},K}(s;\, r) - \sum_{j=2}^{3} a_j\big(g_j(s) - g_j(\tilde{s})\big)\vec{e}_1 \subseteq \theta^{\mathbf{a},K}\big(\tilde{s};\, C_{\mathbf{g}} r\big). \tag{10.10}$$

In light of this observation, define the multipliers

$$m^{\tilde{\nu},\nu}_{\tilde{r},r}(\xi) := \frac{\eta\Big((2C_{\mathbf{g}}r)^{-1}\Big(\xi_1 - \sum_{j=2}^{3} a_j\big(g_j(s_\nu) - g_j(\tilde{s}_{\tilde{\nu}})\big) - \sum_{j=2}^{3}\xi_j g_j(\tilde{s}_{\tilde{\nu}})\Big)\Big)\,\tilde{b}_{\mathbf{a}}(\xi)}{\sum_{i=-1}^{1} m^{\tilde{\nu}+i}_{\tilde{r}}(\xi)} \tag{10.11}$$

for $\tilde{\nu}$, $\nu \in \mathbb{Z}$ and $\tilde{s}_{\tilde{\nu}} := \tilde{r}\tilde{\nu}$, $s_\nu := r\nu$ and $\tilde{r} = Kr$, where $\tilde{b}_{\mathbf{a}}(\xi) := \prod_{j=2}^{3}\eta(K(\xi_j - a_j))$ so that $b_{\mathbf{a}} = \tilde{b}_{\mathbf{a}} \cdot b_{\mathbf{a}}$. Thus, in view of (10.10), we have

$$m^{\nu}_{r} \cdot b_{\mathbf{a}} = m^{\tilde{\nu},\nu}_{\tilde{r},r} \cdot m^{\nu}_{r} \cdot b_{\mathbf{a}} \sum_{i=-1}^{1} m^{\tilde{\nu}+i}_{\tilde{r}}(\xi) \qquad \text{whenever } |s_\nu - \tilde{s}_{\tilde{\nu}}| \leqslant Kr =: \tilde{r}. \tag{10.12}$$

Furthermore, since for fixed $\tilde{\nu}$ the multipliers $m^{\tilde{\nu},\nu}_{\tilde{r},r}$ correspond to essentially parallel frequency regions for $|s_\nu - \tilde{s}_{\tilde{\nu}}| \leqslant 5\tilde{r}$, Lemma 10.1 implies they satisfy a weighted L^2 inequality. Indeed, recall from (10.4) that the functions $g_{\mathbf{a}}(s) := a_2 g_2(s) + a_3 g_3(s)$ satisfy the uniform regularity condition $|g_{\mathbf{a}}'(s)| \sim 1$; recall that $\mathbf{a} = (a_2, a_3) \in [1/4, 4] \times [-1, 1]$. From this we deduce that

$$\sup_{\nu_2 \in \mathbb{Z}} \sum_{\nu_1 \in \mathbb{Z}} e^{-r^{-1}|g_{\mathbf{a}}(r\nu_1) - g_{\mathbf{a}}(\tilde{r}\nu_2)|/2} \lesssim 1,$$

where the above inequality holds with a constant uniform in both r and $\mathbf{a}$. Thus, recalling the definition of the multipliers $m^{\tilde{\nu},\nu}_{\tilde{r},r}$ from (10.11), Lemma 10.1 implies that for fixed $\tilde{\nu} \in \mathbb{Z}$,

$$\int_{\mathbb{R}^3} \sum_{\substack{\nu \in \mathbb{Z} \\ |s_\nu - \tilde{s}_{\tilde{\nu}}| \leqslant 5\tilde{r}}} |m^{\tilde{\nu},\nu}_{\tilde{r},r}(D) f(x)|^2 w(x)\,\mathrm{d}x \lesssim \int_{\mathbb{R}^3} |f(x)|^2 \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}} w(x)\,\mathrm{d}x; \tag{10.13}$$

indeed the inequality holds with $\widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}}w(x)$ replaced by the single average $\widetilde{\mathcal{A}}_{\mathbf{e},\mathbf{r}}w(x; \tilde{s}_{\tilde{\nu}})$, but there is no loss in taking supremum over $s \in [-1, 1]$ in view of other appearances of $\widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}}$ (see (10.14) below). From (10.12) we get

$$\sum_{\nu \in \mathbb{Z}} \big|(m^{\nu}_{r} \cdot b_{\mathbf{a}})(D) f(x)\big|^2 \lesssim \sum_{\substack{\tilde{\nu},\nu \in \mathbb{Z} \\ |s_\nu - \tilde{s}_{\tilde{\nu}}| \leqslant 5\tilde{r}}} \big|(m^{\nu}_{r} \cdot \tilde{b}_{\mathbf{a}})(D) \circ m^{\tilde{\nu},\nu}_{\tilde{r},r}(D) \circ (m^{\tilde{\nu}}_{\tilde{r}} \cdot b_{\mathbf{a}})(D) f(x)\big|^2.$$

By the Schwartz decay property of $\check{\eta}$, the convolution kernel associated to the multiplier operator $(m^{\nu}_{r} \cdot \tilde{b}_{\mathbf{a}})(D)$ satisfies

$$|(m^{\nu}_{r} \cdot \tilde{b}_{\mathbf{a}})\check{\ }(x)| \lesssim_N rK^{-2}\left(1 + r|x_1| + K^{-1}\sum_{j=2}^{3}|x_j + x_1 g_j(s)|\right)^{-100} \lesssim \psi_{T_{\mathbf{e},\mathbf{r}}(s)}(x)$$

where the function $\psi_{T_{\mathbf{e},\mathbf{r}}(s)}(x)$ is the L^1-normalised smooth cutoff defined in (10.9). To justify the second inequality in the above display we use (10.5), which allows us to deduce that $\sum_{j=2}^{3}|x_j + x_1 g_j(s)| \gtrsim \sum_{j=2}^{3}|\langle \mathbf{e}_j(s), x\rangle|$. Combining the preceding observations with a simple Cauchy–Schwarz and Fubini argument,

$$\int_{\mathbb{R}^3} \sum_{\nu \in \mathbb{Z}} \big|(m^{\nu}_{r} \cdot b_{\mathbf{a}})(D) f(x)\big|^2 w(x)\,\mathrm{d}x \lesssim \sum_{\tilde{\nu} \in \mathbb{Z}} \int_{\mathbb{R}^3} \sum_{\substack{\nu \in \mathbb{Z} \\ |s_\nu - \tilde{s}_{\tilde{\nu}}| \leqslant 5\tilde{r}}} \big|m^{\tilde{\nu},\nu}_{\tilde{r},r}(D) \circ (m^{\tilde{\nu}}_{\tilde{r}} \cdot b_{\mathbf{a}})(D) f(x)\big|^2 \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}} w(x)\,\mathrm{d}x. \tag{10.14}$$

On the other hand, (10.13) implies

$$\int_{\mathbb{R}^3} \sum_{\substack{\nu \in \mathbb{Z} \\ |s_\nu - \tilde{s}_{\tilde{\nu}}| \leqslant 5\tilde{r}}} \big|m_{\tilde{r},r}^{\tilde{\nu},\nu}(D)\circ(m_{\tilde{r}}^{\tilde{\nu}}\cdot b_{\mathbf{a}})(D)f(x)\big|^2 \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}}\, w(x)\,\mathrm{d}x \lesssim \int_{\mathbb{R}^3} \big|(m_{\tilde{r}}^{\tilde{\nu}}\cdot b_{\mathbf{a}})(D)f(x)\big|^2 \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}}\circ\widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}}\, w(x)\,\mathrm{d}x.$$

The two previous displays combine to give the desired estimate. □

10.3. **Proof of the L^2-weighted estimate.** Lemma 10.2 is now repeatedly applied to prove Proposition 5.4.

Proof of Proposition 5.4. First observe that by the definition of π in (5.3), the containment property (10.3) and the definition of m_r^ν in (10.6), for each $\pi \in \mathcal{P}_0(r)$ there is an associated $\nu \in \mathbb{Z}$ such that $m_r^\nu(\xi) = 1$ for $\xi \in \operatorname{supp}\chi_\pi$. Thus, a simple Cauchy–Schwarz and Fubini argument yields

$$\int_{\mathbb{R}^3} \sum_{\pi \in \mathcal{P}_0(r)} |\chi_\pi(D)f(x)|^2 w(x)\,\mathrm{d}x \lesssim \int_{\mathbb{R}^3} \sum_{\nu \in \mathbb{Z}} |(m_r^\nu \cdot b)(D)f(x)|^2 \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}_*} w(x)\,\mathrm{d}x,$$

where $\mathbf{r}_* := (r,1,1)$. Take $K := r^{-\varepsilon/8}$ and decompose $b = \sum_{\mathbf{a} \in K^{-1}\mathbb{Z}^2} b_{\mathbf{a}}$ as in (10.7). By a pigeonholing, it follows that there exists a choice of $\mathbf{a} \in [1/4, 4] \times [-1, 1]$ satisfying

$$\int_{\mathbb{R}^3} \sum_{\pi \in \mathcal{P}_0(r)} |\chi_\pi(D)f(x)|^2 w(x)\,\mathrm{d}x \lesssim r^{-\varepsilon/2} \int_{\mathbb{R}^3} \sum_{\nu \in \mathbb{Z}} |(m_r^\nu \cdot b_{\mathbf{a}})(D)f(x)|^2 \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}_*} w(x)\,\mathrm{d}x.$$

Define the sequence

$$\mathbf{r}_M := (r_M, K^{-1}, K^{-1}) \quad \text{where} \quad r_M := K^M r \quad \text{for} \quad M \geqslant 0$$

and recursively define a sequence of maximal operators by

$$\widetilde{\mathcal{N}}^0_{\mathbf{e},\mathbf{r}} := \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}_0} \circ \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}_0} \circ \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}_*} \qquad \text{and} \qquad \widetilde{\mathcal{N}}^M_{\mathbf{e},\mathbf{r}} := \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}_M} \circ \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}_M} \circ \widetilde{\mathcal{N}}^{M-1}_{\mathbf{e},\mathbf{r}} \quad \text{for } M \geqslant 1.$$

We now repeatedly apply Proposition 10.2 to deduce that

$$\int_{\mathbb{R}^3} \sum_{\nu \in \mathbb{Z}} |(m_r^\nu \cdot b_{\mathbf{a}})(D)f(x)|^2 \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}_*} w(x)\,\mathrm{d}x \leqslant C^M \int_{\mathbb{R}^3} \sum_{\nu \in \mathbb{Z}} |(m_{r_M}^\nu \cdot b_{\mathbf{a}})(D)f(x)|^2 \widetilde{\mathcal{N}}^{M-1}_{\mathbf{e},\mathbf{r}} w(x)\,\mathrm{d}x, \tag{10.15}$$

provided $r_M \leqslant 1$. In particular, if $M := \lfloor 8/\varepsilon \rfloor - 1$, then $r^{\varepsilon/8} \leqslant r_M \leqslant 1$ and, consequently, there are only $O(r^{-\varepsilon/8})$ values of ν which contribute to the right-hand sum in (10.15). Thus, one readily deduces that

$$\int_{\mathbb{R}^3} \sum_{\nu \in \mathbb{Z}} |(m_{r_M}^\nu \cdot b_{\mathbf{a}})(D)f(x)|^2 \widetilde{\mathcal{N}}^{M-1}_{\mathbf{e},\mathbf{r}} w(x)\,\mathrm{d}x \lesssim r^{-\varepsilon/8} \int_{\mathbb{R}^3} |f(x)|^2 \widetilde{\mathcal{N}}^{(\varepsilon)}_{\gamma,r}\, w(x)\,\mathrm{d}x$$

where $\widetilde{\mathcal{N}}^{(\varepsilon)}_{\gamma,r} := \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}_M} \circ \widetilde{\mathcal{N}}^{M-1}_{\mathbf{e},\mathbf{r}}$. Combining the preceding observations concludes the proof of the L^2 weighted inequality, with the above choice of maximal operator.

It remains to show that the iterated maximal operator $\widetilde{\mathcal{N}}^{(\varepsilon)}_{\gamma,r}$ satisfies the L^2 bound from (5.4). However, this is an immediate consequence of Proposition 10.3 of the following subsection. □

10.4. **Boundedness of the maximal functions.** From the proof of Proposition 5.4, we see that the maximal function $\widetilde{\mathcal{N}}^{(\varepsilon)}_{\gamma,r}$ is obtained by repeatedly composing operators of the form $\widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}}$, as defined in (10.8), where:

- The family of curves $\mathbf{e}$ corresponds to the Frenet frame $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ associated to γ;
- The scales $\mathbf{r} = (r_1, r_2, r_3)$ depend on r and ε and vary over the different factors of the composition. Each featured tuple $\mathbf{r} = (r_1, r_2, r_3)$ satisfies

$$\operatorname{ecc}(\mathbf{r}) \leqslant r^{-1}$$

where the *eccentricity* $\operatorname{ecc}(\mathbf{r})$ is the ratio of $\max_j r_j$ and $\min_j r_j$.

In particular, to prove the L^2 bound (5.4) it suffices to show that, for all $\varepsilon_\circ > 0$,

$$\|\widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}}\|_{L^2(\mathbb{R}^3)\to L^2(\mathbb{R}^3)} \lesssim_{\varepsilon_\circ} \mathrm{ecc}(\mathbf{r})^{\varepsilon_\circ}. \tag{10.16}$$

To prove (10.16), we will in fact work with a more general setup, replacing $\mathbf{e}$ with a general family of smooth curves in $\mathbb{R}^n$ satisfying a non-degeneracy hypothesis. Let $\mathbf{e} := (\mathbf{e}_1, \dots, \mathbf{e}_n)$ where $\mathbf{e}_j \colon [-1,1] \to S^{n-1}$ is a smooth curve in the unit sphere in $\mathbb{R}^n$ for $1 \leqslant j \leqslant n$. Suppose these curves satisfy

$$\Big|\bigwedge_{j=1}^n \mathbf{e}_j(s)\Big| \gtrsim 1 \qquad \text{for all } s \in [-1,1].$$

Note that the $\mathbf{e}_j$ notation was previously reserved for the Frenet frame. In applications, we always take the $\mathbf{e}_j$ to be the Frenet vectors, and therefore there should be no conflict in the above choice of notation.

Given a tuple $\mathbf{r} := (r_1, \dots, r_n) \in (0,\infty)^n$ and $s \in [-1,1]$ define the parallelepiped

$$T_{\mathbf{e},\mathbf{r}}(s) := \Big\{x \in \mathbb{R}^n : x = \sum_{j=1}^n \lambda_j \mathbf{e}_j(s) \text{ where } \lambda_j \in [-r_j^{-1}, r_j^{-1}] \text{ for } 1 \leqslant j \leqslant n\Big\}.$$

Associated to these sets are the averaging operators and the maximal operator

$$\mathcal{A}_{\mathbf{e},\mathbf{r}} f(x;s) := \fint_{T_{\mathbf{e},\mathbf{r}}(s)} f(x-y)\,\mathrm{d}y \quad \text{and} \quad \mathcal{N}_{\mathbf{e},\mathbf{r}} f(x) := \sup_{s\in[-1,1]} |\mathcal{A}_{\mathbf{e},\mathbf{r}} f(x;s)| \tag{10.17}$$

defined for $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. The $\mathcal{N}_{\mathbf{e},\mathbf{r}}$ satisfy favourable L^2 estimates.

Proposition 10.3. *With the above definitions, for all $\varepsilon > 0$ we have the norm bound*

$$\|\mathcal{N}_{\mathbf{e},\mathbf{r}} f\|_{L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)} \lesssim_{\mathbf{e},\varepsilon} \mathrm{ecc}(\mathbf{r})^{\varepsilon},$$

where the eccentricity $\mathrm{ecc}(\mathbf{r}) \geqslant 1$ is defined to be the ratio of $\max_j r_j$ and $\min_j r_j$.

This proposition is based on a classical maximal bound due to Córdoba [11]. The details of the proof are provided below.

We generalise the weighted operators introduced in (10.8) by setting

$$\widetilde{\mathcal{A}}_{\mathbf{e},\mathbf{r}} f(x;s) := \int_{\mathbb{R}^n} f(x-y)\psi_{T_{\mathbf{e},\mathbf{r}}(s)}(y)\,\mathrm{d}y \quad \text{and} \quad \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}} f(x) := \sup_{s\in[-1,1]} |\widetilde{\mathcal{A}}_{\mathbf{e},\mathbf{r}} f(x;s)| \tag{10.18}$$

where $\psi_{T_{\mathbf{e},\mathbf{r}}(s)}$ is a smooth weight function adapted to the parallelepiped $T_{\mathbf{e},\mathbf{r}}(s)$, given by

$$\psi_{T_{\mathbf{e},\mathbf{r}}(s)}(y) := \Big(\prod_{j=1}^n r_j\Big)\Big(1 + \sum_{j=1}^n r_j |(\boldsymbol{E}(s)^{-1}y)_j|\Big)^{-100n} \tag{10.19}$$

where $\boldsymbol{E}(s)$ denotes the $n \times n$ matrix whose jth column is $\mathbf{e}_j(s)$ for $1 \leqslant j \leqslant n$. If $(\mathbf{e}_j(s))_{j=1}^n$ forms an orthonormal frame, then $(\boldsymbol{E}(s)^{-1}y)_j = (\boldsymbol{E}(s)^\top y)_j = \langle \mathbf{e}_j(s), y\rangle$ and so (10.19) generalises the definition (10.9). Note that the operators in (10.18) correspond to weighted version of the averaging operator and Nikodym maximal function in (10.17). Moreover, by dominating $\psi_{T_{\mathbf{e},\mathbf{r}}(s)}$ by a weighted sum of characteristic functions, it is clear that Proposition 10.3 implies analogous L^2 bounds for the $\widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}}$ operators.

In view of the preceding discussion, the estimate (5.4) for the maximal function $\widetilde{\mathcal{N}}^{(\varepsilon)}_{\gamma,r}$ appearing in Proposition 5.4 follows as a consequence of Proposition 10.3.

Proof of Proposition 10.3. Write $R := \mathrm{ecc}(\mathbf{r})$ and let $\varepsilon > 0$ be given. We begin with some basic reductions. By pigeonholing, it suffices to show

$$\|\mathcal{N}_{\mathbf{e},\mathbf{r}}\|_{L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)} \lesssim_\varepsilon R^{\varepsilon/2}$$

where now the maximal operator $\mathcal{N}_{\mathbf{e},\mathbf{r}}$ is redefined so that the supremum is taken over some subinterval $I_\varepsilon \subseteq [-1,1]$ of length $R^{-\varepsilon/2}$ rather than the whole of $[-1,1]$. Furthermore, if $|s_1 - s_2| \leqslant R^{-1}$, then $T_{\mathbf{e},\mathbf{r}}(s_1)$ and $T_{\mathbf{e},\mathbf{r}}(s_2)$ define essentially the same parallelepiped, and therefore we may further restrict the supremum to some dyadic R^{-1}-net $\mathfrak{S}_\varepsilon$ in I_ε.

Let $a \in [-1,1]$ denote the centre of the interval I_ε and $N := \lceil 1/\varepsilon \rceil$. For $1 \leqslant j \leqslant n$ let p_j denote the degree $N-1$ Taylor polynomial of $\mathbf{e}_j$ centred at a and define $\boldsymbol{p} := (p_1, \dots, p_n)$. By Taylor's theorem,

$$|p_j(s) - \mathbf{e}_j(s)| \lesssim_\gamma R^{-N\varepsilon} \leqslant R^{-1} \qquad \text{for all } s \in I_\varepsilon$$

and therefore there exists a constant $C \geqslant 1$, independent of $\mathbf{r}$, such that

$$T_{\boldsymbol{p},C^{-1}\mathbf{r}}(s) \subseteq T_{\mathbf{e},\mathbf{r}}(s) \subseteq T_{\boldsymbol{p},C\mathbf{r}}(s) \qquad \text{for all } s \in I_\varepsilon.$$

In light of this observation, henceforth we may assume without loss of generality that the $\mathbf{e}_j$ are all polynomial mappings. Under this hypothesis, the $\mathbf{e}_j$ no longer map into the sphere; however, we may assume that over the domain I_ε they map into, say, a $1/10$-neighbourhood of S^{n-1}.

Since the operators are all positive, it suffices to show

$$\|\sup_{s\in\mathfrak{S}_\varepsilon} |\mathcal{A}_{\mathbf{e},\mathbf{r}} f(\,\cdot\,;s)|\|_{L^2(\mathbb{R}^n)} \lesssim_\varepsilon R^\varepsilon \|f\|_{L^2(\mathbb{R}^n)}$$

for all $f \in L^2(\mathbb{R}^n)$ continuous and non-negative. Fixing such an f, define the averages

$$\mathcal{A}_{\omega,r} f(x) := \int_{\mathbb{R}} f(x - t\omega)\chi_r(t)\,\mathrm{d}t \qquad \text{for } \omega \in \mathbb{R}^n \text{ with } \big||\omega| - 1\big| < 1/10 \text{ and } r > 0,$$

where $\chi_r(t) := r^{-1}\chi_1(r^{-1}t)$ for some $\chi_1 \in C_c^\infty(\mathbb{R})$ non-negative which satisfies $\chi_1(s) = 1$ for $|s| \leqslant 1$. Thus, by the Fubini–Tonelli theorem,

$$\mathcal{A}_{\mathbf{e},\mathbf{r}} f(x;s) \lesssim \mathcal{A}_{\mathbf{e}_n(s),r_n} \circ \cdots \circ \mathcal{A}_{\mathbf{e}_1(s),r_1} f(x). \tag{10.20}$$

Writing $\mathcal{A}_{\mathbf{e}_j} f(x;s) := \mathcal{A}_{\mathbf{e}_j(s),1} f(x)$, we may combine (10.20) with a simple scaling argument the reduce to problem to showing

$$\|\sup_{s\in\mathfrak{S}_\varepsilon} |\mathcal{A}_{\mathbf{e}_j} f(\,\cdot\,;s)|\|_{L^2(\mathbb{R}^n)} \lesssim (\log R)\, \|f\|_{L^2(\mathbb{R}^n)} \qquad \text{for } 1 \leqslant j \leqslant n. \tag{10.21}$$

The previous display is essentially a consequence of a maximal estimate proved in [11, p.223]. There similar maximal operators are considered for smooth curves $\gamma\colon [-1,1] \to S^{n-1}$ under the key hypothesis that γ crosses any affine hyperplane a bounded number of times. Since we are considering polynomial curves $\mathbf{e}_j$, the fundamental theorem of algebra ensures that either:

a) The curve $\mathbf{e}_j$ crosses any affine hyperplane a bounded number of times, where the bound depends on the degrees of the component polynomials, or
b) There exists an affine hyperplane which contains the image of $\mathbf{e}_j$.

In the former case, we may deduce (10.21) directly through an appeal to the result from [11, p.223] (we remark that the argument in [11] carries through for a curve which maps into a 1/10-neighbourhood of the sphere (rather than the sphere itself), provided the curve satisfies the finite crossing property). In the latter case, we may apply the maximal bound from [11] over a lower dimensional affine subspace and combine this with a Fubini argument to again deduce the desired result. □

10.5. **Scaling properties.** We conclude this section with a discussion of the scaling properties of the maximal function $\widetilde{\mathcal{N}}_{\gamma,r}^{(\varepsilon)}$ and, in particular, fill in the gap in proof of Proposition 8.10 by proving the Claim therein.

We begin by introducing a general setup for rescaling the operators $\widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}}$ when defined with respect to a Frenet frame; as in the previous subsection, here we work in general dimensions. Fix $\gamma\colon[-1,1]\to\mathbb{R}^n$ a non-degenerate curve with $\gamma\in\mathfrak{G}(\delta)$ and $\sigma\in[-1,1]$, $0<\lambda<1$ be such that $[\sigma-\lambda,\sigma+\lambda]\subseteq[-1,1]$. Consider the rescaled curve

$$\gamma_{\sigma,\lambda}(\tilde{s}) := \big([\gamma]_{\sigma,\lambda}\big)^{-1}\big(\gamma(\sigma+\lambda\tilde{s})-\gamma(\sigma)\big)$$

as defined in Definition 4.1. Let $\mathbf{e}=(\mathbf{e}_1,\dots,\mathbf{e}_n)$ denote the Frenet frame defined with respect to γ and $\tilde{\mathbf{e}}=(\tilde{\mathbf{e}}_1,\dots,\tilde{\mathbf{e}}_n)$ denote the Frenet frame defined with respect to $\tilde{\gamma}:=\gamma_{\sigma,\lambda}$. We suppose $\mathbf{r}=(r_1,\dots,r_n)\in(0,1]^n$ satisfies

$$r_i \leqslant \lambda r_{i+1} \qquad \text{for } 1\leqslant i\leqslant n-1 \tag{10.22}$$

and define $\tilde{\mathbf{r}}:=D_\lambda\cdot\mathbf{r}$ where $D_\lambda:=\mathrm{diag}(\lambda,\dots,\lambda^n)$ is as in (4.1).

Lemma 10.4. *If $f\in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is non-negative, then, with the above definitions,*

$$\big([\gamma]_{\sigma,\lambda}\big)^{-1}\circ\widetilde{\mathcal{N}}_{\tilde{\mathbf{e}},\tilde{\mathbf{r}}}\circ[\gamma]_{\sigma,\lambda}\cdot f(x) \lesssim_\gamma \widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}}f(x) \qquad \textit{for all } x\in\mathbb{R}^n. \tag{10.23}$$

Here we think of a matrix $M\in\mathrm{GL}(\mathbb{R},n)$ as acting on $L^2(\mathbb{R}^n)$ by $M\cdot f:=f\circ M$ for all $f\in L^2(\mathbb{R}^n)$. Thus, the left-hand side corresponds to the operator $\widetilde{\mathcal{N}}_{\tilde{\mathbf{e}},\tilde{\mathbf{r}}}$ conjugated by the invertible operator $[\gamma]_{\sigma,\lambda}\colon L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)$.

Before presenting the proof of Lemma 10.4, we use the result to verify the rescaling step in the proof of Proposition 8.10. In view of the discussion in §10.3 and by a simple rescaling argument, we know that the maximal function

$$\widetilde{\mathcal{N}}_{k,\ell}^{\mu,(\varepsilon)} := \mathrm{Dil}_{2^{k-3\ell}}\circ\widetilde{\mathcal{N}}_{\tilde{\gamma},\tilde{r}}^{(\varepsilon)}\circ\mathrm{Dil}_{2^{-(k-3\ell)}}$$

corresponds to a repeated composition of operators of the form $\widetilde{\mathcal{N}}_{\tilde{\mathbf{e}},\tilde{\mathbf{r}}}$ where the $\tilde{\mathbf{r}}=(\tilde{r}_1,\tilde{r}_2,\tilde{r}_3)$ satisfy

$$\tilde{r}_1\leqslant\tilde{r}_2\leqslant\tilde{r}_3 \qquad\text{and}\qquad \mathrm{ecc}(\tilde{\mathbf{r}})\lesssim 2^{(k-3\ell)/2}$$

(recall that in the setup in Proposition 8.10 we have $\tilde{\gamma}:=\gamma_{\sigma,\lambda}$, where $\sigma:=2^{-\ell}\mu$ and $\lambda:=2^{-\ell}$, and $\tilde{r}:=2^{-(k-3\ell)/2}$). Consequently, by Lemma 10.4, the conjugate

$$\big([\gamma]_{\sigma,\lambda}\big)^{-1}\circ\widetilde{\mathcal{N}}_{k,\ell}^{\mu,(\varepsilon)}\circ[\gamma]_{\sigma,\lambda}$$

is dominated by a maximal function $\widetilde{\mathcal{N}}_{k,\ell}^{(\varepsilon)}$ given by a repeated composition of operators of the form $\widetilde{\mathcal{N}}_{\mathbf{e},\mathbf{r}}$ where each $\mathbf{r}=(r_1,r_2,r_3)$ satisfies

$$r_1\leqslant\lambda r_2\leqslant\lambda^2 r_3 \qquad\text{and}\qquad \mathrm{ecc}(\mathbf{r})\lesssim 2^{(k+\ell)/2}.$$

Furthermore, there are only $O_\varepsilon(1)$ factors in this composition. The just given definition for $\widetilde{\mathcal{N}}_{k,\ell}^{(\varepsilon)}$ is independent of μ and, by Proposition 10.3, for all $\varepsilon_\circ>0$ the operator $\widetilde{\mathcal{N}}_{k,\ell}^{(\varepsilon)}$ is bounded on $L^2(\mathbb{R}^3)$ with operator norm $O_\varepsilon(2^{\varepsilon k})$. Thus, we have verified all the outstanding claims in the proof of Proposition 8.10.

Proof of Lemma 10.4. Consider the conjugated operator on the left-hand side of (10.23). By applying a change of variables to the integral defining the underlying averages, the problem is quickly reduced to the pointwise estimate

$$|\det[\gamma]_{\sigma,\lambda}|^{-1}\cdot\psi_{T_{\tilde{\mathbf{e}},\tilde{\mathbf{r}}}(\tilde{s})}\circ\big([\gamma]_{\sigma,\lambda}\big)^{-1}(y)\lesssim\psi_{T_{\mathbf{e},\mathbf{r}}(s)}(y)$$

for the weight functions as defined in (10.19), where $s = \sigma + \lambda\tilde{s}$. Suppose $y \in \mathbb{R}^n$ satisfies

$$R \leqslant \sum_{j=1}^{n} r_j |\langle \mathbf{e}_j(s), y\rangle| \leqslant 2R$$

for some $R \geqslant 1$. From the definition of the weight function from (10.19), and the orthonormality of the Frenet frame, the problem is further reduced to showing

$$\sum_{j=1}^{n} \tilde{r}_j |\langle \tilde{\mathbf{e}}_j(\tilde{s}), \tilde{y}\rangle| \gtrsim R \qquad \text{where } \tilde{y} := \big([\gamma]_{\sigma,\lambda}\big)^{-1}(y). \tag{10.24}$$

Let $\alpha = \big([\gamma]_{s,\lambda}\big)^{-1}(y)$ so that, by the definition of the matrix $[\gamma]_{s,\lambda}$, we have

$$y = \sum_{j=1}^{n} \lambda^j \alpha_j \gamma^{(j)}(s).$$

Taking the inner product of both sides of this identity with respect to the vectors $\mathbf{e}_j(s)$, it follows that the vectors $\big(\langle \mathbf{e}_j(s), y\rangle\big)_{j=1}^n$ and $\big(\lambda^j \alpha_j\big)_{j=1}^n$ are related by an *upper-triangular* matrix transformation, which is also an $O(\delta)$ perturbation of the identity. For this observation, we use the fact that $\langle \mathbf{e}_1(s), \dots, \mathbf{e}_j(s)\rangle = \langle \gamma^{(1)}(s), \dots, \gamma^{(j)}(s)\rangle$ for $1 \leqslant j \leqslant n$, owing to the definition of the Frenet frame.

In view of the hypothesis (10.22) which, in particular, implies $r_i \leqslant r_{i+1}$ for $1 \leqslant i \leqslant n-1$, the above observation yields that

$$r_j \lambda^j |\alpha_j| \lesssim R \qquad \text{for } 1 \leqslant j \leqslant n. \tag{10.25}$$

Furthermore, by pigeonholing, there exists some $1 \leqslant J \leqslant n$ such that

$$r_J |\langle \mathbf{e}_J(s), y\rangle| \geqslant R/n \quad \text{and} \quad r_j |\langle \mathbf{e}_j(s), y\rangle| < R/n \qquad \text{for } J+1 \leqslant j \leqslant n.$$

Thus, by the same argument used to show (10.25), provided δ is chosen sufficiently small,

$$r_J \lambda^J |\alpha_J| \sim R. \tag{10.26}$$

Since $\tilde{\gamma}^{(j)}(\tilde{s}) = \lambda^j \big([\gamma]_{\sigma,\lambda}\big)^{-1} \gamma^{(j)}(s)$ for $j \geqslant 1$, it follows that $[\tilde{\gamma}]_{\tilde{s}} = \big([\gamma]_{\sigma,\lambda}\big)^{-1} \circ [\gamma]_{s,\lambda}$ and, consequently,

$$\tilde{y} = \big([\gamma]_{\sigma,\lambda}\big)^{-1}(y) = \big([\gamma]_{\sigma,\lambda}\big)^{-1} \circ [\gamma]_{\lambda,s}(\alpha) = [\tilde{\gamma}]_{\tilde{s}}(\alpha).$$

Thus, we have $\alpha = \big([\tilde{\gamma}]_{\tilde{s}}\big)^{-1}(\tilde{y})$ and, arguing as before, this implies the vectors $\big(\langle \tilde{\mathbf{e}}_j(\tilde{s}), \tilde{y}\rangle\big)_{j=1}^n$ and α are also related by an upper-triangle matrix transformation, which is again an $O(\delta)$ perturbation of the identity. From this observation, provided δ is chosen sufficiently small, we see that

$$\tilde{r}_J |\langle \tilde{\mathbf{e}}_J(\tilde{s}), \tilde{y}\rangle| \gtrsim r_J \lambda^J |\alpha_J| - \delta \sum_{j=J+1}^{n} \big(r_J \lambda^{J-j} r_j^{-1}\big) r_j \lambda^j |\alpha_j| \gtrsim R,$$

where the final inequality uses the hypothesis (10.22) together with (10.25) and (10.26). This implies the desired bound (10.24).

□

11. Proof of the $\mathbb{R}^{3+1} \to \mathbb{R}^3$ Nikodym maximal estimate

In this section we establish Proposition 5.5. We begin by recalling the basic setup. Let $\gamma\colon [-1,1] \to \mathbb{R}^3$ be a smooth, non-degenerate curve with Frenet frame $(\mathbf{e}_j)_{j=1}^3$. Given $\mathbf{r} \in (0,1)^3$ and $s \in [-1,1]$, consider the *plates*

$$\mathcal{T}_{\mathbf{r}}(s) := \big\{(y,t) \in \mathbb{R}^3 \times [1,2] : \big|\langle y - t\gamma(s), \mathbf{e}_j(s)\rangle\big| \leqslant r_j \ \text{ for } j = 1,2,3\big\}$$

and the associated averaging and maximal operators

$$\mathcal{A}^{\mathrm{sing}}_{\mathbf{r}} g(x;s) := \fint_{\mathcal{T}_{\mathbf{r}}(s)} g(x-y,t)\,\mathrm{d}y\,\mathrm{d}t \quad \text{and} \quad \mathcal{N}^{\mathrm{sing}}_{\mathbf{r}} g(x) := \sup_{-1\leqslant s\leqslant 1} |\mathcal{A}^{\mathrm{sing}}_{\mathbf{r}} g(x;s)|.$$

We assume the exponents satisfy the conditions

$$r_3 \leqslant r_2 \leqslant r_1 \leqslant r_2^{1/2} \qquad \text{and} \qquad r_2 \leqslant r_1^{1/2} r_3^{1/2}$$

and the goal is to establish the L^2 bound

$$\|\mathcal{N}^{\mathrm{sing}}_{\mathbf{r}} g\|_{L^2(\mathbb{R}^3)} \lesssim |\log r_3|^3 \|g\|_{L^2(\mathbb{R}^4)}. \tag{11.1}$$

To prove this norm inequality we will rely on the Fourier transform and reduce the problem to certain oscillatory integral estimates. The argument is a (significant) elaboration of that used to establish a lower dimensional variant of (11.1) in [18]. We shall make heavy use of the frequency decomposition used to analyse the helical averaging operator in §8.

Proof of Proposition 5.5. The argument is somewhat involved and is therefore broken into steps.

Initial reductions. Let $0 < \delta_0 \ll 1$ be a small parameter, as introduced at the beginning of §6. By familiar localisation and rescaling arguments, we may assume γ satisfies $\gamma(\,\cdot\,) - \gamma(0) \in \mathfrak{G}_3(\delta_0)$. Further, we may replace $\mathcal{A}^{\mathrm{sing}}_{\mathbf{r}} g(x;s)$ with the localised version $\mathcal{A}^{\mathrm{sing}}_{\mathbf{r}} g(x;s)\chi(s)$, where $\chi \in C^\infty_c(\mathbb{R})$ is supported in $I_0 := [-\delta_0, \delta_0]$. Note that this model situation is already enough for our application in §8.7.

Fourier representation. The first step is to derive an alternative representation of the averages $\mathcal{A}^{\mathrm{sing}}_{\mathbf{r}} g$ in terms of an oscillatory integral operator. Given $a \in C^\infty_c(\hat{\mathbb{R}}^3 \times \mathbb{R} \times \mathbb{R})$, define

$$\begin{aligned}\mathcal{A}[a]g(x;s) &:= \frac{1}{(2\pi)^3}\int_1^2 \int_{\mathbb{R}^3}\int_{\hat{\mathbb{R}}^3} e^{i\langle x-y-t\gamma(s),\xi\rangle} a(\xi;s;t)\,\mathrm{d}\xi\, g(y,t)\mathrm{d}y\,\mathrm{d}t \\ &= \frac{1}{(2\pi)^3}\int_{\hat{\mathbb{R}}^3} e^{i\langle x,\xi\rangle}\int_1^2 e^{-it\langle\gamma(s),\xi\rangle} a(\xi;s;t)\tilde{g}(\xi,t)\,\mathrm{d}t\,\mathrm{d}\xi,\end{aligned}$$

where $\tilde{g}$ denotes the Fourier transform of g with respect to the y-variable only. The associated maximal operator is then defined by

$$\mathcal{N}[a]g(x) := \sup_{-1\leqslant s\leqslant 1} |\mathcal{A}[a]g(x;s)|.$$

Without loss of generality, to prove Proposition 5.5 it suffices to consider the estimate for g Schwartz and taking values in $[0,\infty)$. Fix $\psi \in C^\infty_c(\hat{\mathbb{R}})$ with $\operatorname{supp}\psi \subseteq [-1,1]$ such that $\check{\psi}$ takes values in the positive real line and $\check{\psi}(y) \gtrsim 1$ for $|y| \leqslant 1$. Define

$$a_{\mathbf{r}}(\xi;s) := \prod_{j=1}^3 \psi\big(r_j\langle \xi, \mathbf{e}_j(s)\rangle\big)\,\chi(s)$$

so that, by integral formula for the inverse Fourier transform and a change of variable,

$$\frac{1}{|\mathcal{T}_{\mathbf{r}}(s)|}\mathbb{1}_{\mathcal{T}_{\mathbf{r}}(s)}(y,t)\chi(s) \lesssim \prod_{j=1}^3 r_j^{-1}\check{\psi}\big(r_j^{-1}\langle y - t\gamma(s), \mathbf{e}_j(s)\rangle\big)\chi(s) = \frac{1}{(2\pi)^3}\int_{\hat{\mathbb{R}}^3} e^{i\langle y-t\gamma(s),\xi\rangle} a_{\mathbf{r}}(\xi;s;t)\,\mathrm{d}\xi.$$

Thus, the pointwise inequality

$$|\mathcal{A}^{\mathrm{sing}}_{\mathbf{r}} g(x;s)| \lesssim |\mathcal{A}[a_{\mathbf{r}}]g(x;s)|$$

holds and therefore it suffices to bound the operator $\mathcal{N}[a_{\mathbf{r}}]$.

Sobolev embedding. Given $a \in C_c^\infty(\widehat{\mathbb{R}}^3 \times \mathbb{R} \times \mathbb{R})$, by elementary Sobolev embedding,

$$\|\mathcal{N}[a]g\|_{L^2(\mathbb{R}^3)}^2 \leqslant \|\mathcal{A}[a]g\|_{L^2(\mathbb{R}^{3+1})}^2 + 2 \prod_{\iota \in \{0,1\}} \|\partial_s^\iota \mathcal{A}[a]g\|_{L^2(\mathbb{R}^{3+1})}; \tag{11.2}$$

indeed, this bound is a simple and standard consequence of the fundamental theorem of calculus and the Cauchy–Schwarz inequality (see for instance [22, Chapter XI, §3.2]). Observe that $\partial_s \mathcal{A}[a]$ is an operator of the same form as $\mathcal{A}[a]$ and, in particular,

$$\partial_s \mathcal{A}[a] = \mathcal{A}[\mathfrak{d}_s a] \qquad \text{where} \qquad \mathfrak{d}_s a(\xi; s; t) := -it\langle \gamma'(s), \xi\rangle\, a(\xi; s; t) + \partial_s a(\xi; s; t). \tag{11.3}$$

These observations reduce the problem to proving estimates of the form

$$\|\mathcal{A}[\mathfrak{d}_s^\iota\, a]g\|_{L^2(\mathbb{R}^{3+1}) \to L^2(\mathbb{R}^{3+1})} \leqslant B^{(\iota - 1/2)} \qquad \text{for } \iota \in \{0, 1\} \tag{11.4}$$

for suitable symbols a and constants $B \geqslant 1$. In particular, it suffices to decompose the original symbol $a_{\mathbf{r}}$ into $O(|\log r_3|^3)$ many pieces and show that (11.4) holds for some choice of $B \geqslant 1$ on each piece.

Reduction to oscillatory integral estimates. Continuting to work with a general $a \in C_c^\infty(\widehat{\mathbb{R}}^3 \times \mathbb{R} \times \mathbb{R})$, it follows from Plancherel's theorem in the x-variable and the Cauchy–Schwarz inequality that

$$\begin{aligned} \|\mathcal{A}[a]g\|_{L^2(\mathbb{R}^{3+1})}^2 &\leqslant \int_{\widehat{\mathbb{R}}^3} \int_{\mathbb{R}} |T_\xi[a]\tilde{g}(\xi;\,\cdot\,)(t)\tilde{g}(\xi; t)|\, \mathrm{d}t\, \mathrm{d}\xi \\ &\leqslant \int_{\widehat{\mathbb{R}}^3} \|T_\xi[a]\tilde{g}(\xi;\,\cdot\,)\|_{L^2(\mathbb{R})} \|\tilde{g}(\xi;\,\cdot\,)\|_{L^2(\mathbb{R})}\, \mathrm{d}\xi \end{aligned} \tag{11.5}$$

where, for each $\xi \in \widehat{\mathbb{R}}^3$, the operator $T_\xi[a]$ acts on univariate functions by integrating (in the t'-variable) against the kernel

$$\mathcal{K}[a](t, t'; \xi) := \int_{\mathbb{R}} e^{i(t-t')\langle \gamma(s), \xi\rangle} \overline{a(\xi; s; t)} a(\xi; s, t') \mathbb{1}_{[1,2]^2}(t, t')\, \mathrm{d}s. \tag{11.6}$$

It suffices to show that

$$\|T_\xi[\mathfrak{d}_s^\iota a]\tilde{g}(\xi;\cdot)\|_{L^2(\mathbb{R})} \leqslant B^{2\iota - 1} \|\tilde{g}(\xi;\cdot)\|_{L^2(\mathbb{R})} \qquad \text{for } \iota \in \{0, 1\} \tag{11.7}$$

holds uniformly in $\xi \in \widehat{\mathbb{R}}^3$. Indeed, in this case the norm bound (11.4) would follow via (11.5) and a further application of Plancherel's theorem in the ξ-variable. By the Schur test, the inequality (11.7) is reduced to verifying the oscillatory integral estimates

$$\sup_{t' \in [1,2]} \int_1^2 |\mathcal{K}[\mathfrak{d}_s^\iota a](t, t'; \xi)|\, \mathrm{d}t, \quad \sup_{t \in [1,2]} \int_1^2 |\mathcal{K}[\mathfrak{d}_s^\iota a](t, t'; \xi)|\, \mathrm{d}t' \leqslant B^{2\iota - 1}, \qquad \iota \in \{0, 1\} \tag{11.8}$$

hold uniformly over all $\xi \in \hat{\mathbb{R}}^3$.

Initial decomposition. In order to obtain favourable estimates, it is necessary to first decompose the original symbol $a_{\mathbf{r}}$ into a number of localised pieces. This decomposition is similar to that used in §8 and is described in detail presently. Later in the proof, the kernel estimates (11.8) are verified for each piece of the decomposition and the resulting norm bounds are combined to estimate the entire operator.

Define $\delta_1 := \delta_0^3$, $\delta_2 := \delta_0$ and $\delta_3 := 9/10$ and for $1 \leqslant J \leqslant 3$ let Ω_J denote the set of $\xi \in \hat{\mathbb{R}}^3$ satisfying

$$\begin{aligned} &\inf_{s \in I_0} |\langle \gamma^{(J)}(s), \xi\rangle| \geqslant \delta_J |\xi|, \\ &\inf_{s \in I_0} |\langle \gamma^{(j)}(s), \xi\rangle| \leqslant \delta_j |\xi| \qquad \text{for } 1 \leqslant j \leqslant J - 1. \end{aligned}$$

Provided $\delta_0 > 0$ is chosen sufficiently small, the condition $\gamma(\,\cdot\,) - \gamma(0) \in \mathfrak{G}_3(\delta_0)$ ensures that these sets partition $\hat{\mathbb{R}}^3$. By pigeonholing, it suffices to work with the symbols $a^J_{\mathbf{r}}(\xi; s) := a_{\mathbf{r}}(\xi; s)\mathbb{1}_{\Omega_J}(\xi)$ for $1 \leqslant J \leqslant 3$ (as we are interested in L^2 estimates here, we are free to decompose the symbol using the rough partition of unity $1 \equiv \mathbb{1}_{\Omega_1} + \mathbb{1}_{\Omega_2} + \mathbb{1}_{\Omega_3}$).

Decompose the symbol into dyadic frequency bands by writing

$$a_{\mathbf{r}} = \sum_{k=0}^{\infty} a_{\mathbf{r},k} \qquad \text{where} \qquad a_{\mathbf{r},k}(\xi; s) := \begin{cases} a^J_{\mathbf{r}}(\xi; s) \cdot \beta^k(\xi) & \text{for } k \geqslant 1 \\ a^J_{\mathbf{r}}(\xi; s) \cdot \eta(\xi) & \text{for } k = 0 \end{cases}.$$

Here, for notational convenience, we suppress the choice of J in the notation. Since $r_3 \leqslant r_1, r_2$, only the first $O(|\log r_3|)$ terms of the above sum are non-zero, so it suffices to show

$$\|\mathcal{N}[a_{\mathbf{r},k}]\|_{L^2(\mathbb{R}^4) \to L^2(\mathbb{R}^3)} \lesssim k^2 \qquad \text{for all } k \in \mathbb{N}_0. \tag{11.9}$$

In particular, note that $2^k \lesssim r_3^{-1}$.

<u>$J = 1$ case.</u> Suppose $\mathrm{supp}_\xi\, a_{\mathbf{r},k} \subseteq \Omega_1$. Here a simple integration-by-parts argument yields

$$\sup_{t' \in [1,2]} \int_1^2 |\mathcal{K}[\mathfrak{d}_s^\iota a_{\mathbf{r},k}](t, t'; \xi)|\, \mathrm{d}t, \quad \sup_{t \in [1,2]} \int_1^2 |\mathcal{K}[\mathfrak{d}_s^\iota a_{\mathbf{r},k}](t, t'; \xi)|\, \mathrm{d}t' \lesssim 2^{k(2\iota - 1)} \qquad \text{for } \iota \in \{0, 1\}.$$

In view of our earlier observations, the bound (11.9) therefore holds in this case with a uniform bound in k.

<u>$J = 2$ case.</u> Suppose $\mathrm{supp}_\xi\, a_{\mathbf{r},k} \subseteq \Omega_2$. If $\xi \in \Omega_2$, then the equation $\langle \gamma'(s), \xi \rangle = 0$ has a unique solution in $\frac{5}{4} \cdot I_0$ which we denote by $\theta(\xi)$. Indeed, this follows from a simple calculus exercise, similar to the proof of Lemma 6.1.

Further decomposition. Here the symbol $a_{\mathbf{r},k}$ is further decomposed by writing

$$a_{\mathbf{r},k} = \sum_{\ell=0}^{\lfloor k/2 \rfloor} a_{\mathbf{r},k,\ell} \qquad \text{where} \qquad a_{\mathbf{r},k,\ell}(\xi; s) := \begin{cases} a_{\mathbf{r},k}(\xi; s)\beta\big(2^\ell |s - \theta(\xi)|\big) & \text{if } 0 \leqslant \ell < \lfloor k/2 \rfloor \\ a_{\mathbf{r},k}(\xi; s)\eta\big(2^{\lfloor k/2 \rfloor} |s - \theta(\xi)|\big) & \text{if } \ell = \lfloor k/2 \rfloor \end{cases}.$$

Since $|\langle \gamma''(s), \xi \rangle| \sim 2^k$ for all $(\xi; s) \in \mathrm{supp}\, a_{\mathbf{r},k,\ell}$, one has the relation $2^k \leqslant r_2^{-1}$.

Kernel estimates. The kernels are analysed using stationary phase techniques.

Lemma 11.1. *If $k \in \mathbb{N}$, $0 \leqslant \ell \leqslant \lfloor k/2 \rfloor$ and $\iota \in \{0, 1\}$, then*

$$\sup_{t' \in [1,2]} \int_1^2 |\mathcal{K}[\mathfrak{d}_s^\iota a_{\mathbf{r},k,\ell}](t, t'; \xi)|\, \mathrm{d}t, \quad \sup_{t \in [1,2]} \int_1^2 |\mathcal{K}[\mathfrak{d}_s^\iota a_{\mathbf{r},k,\ell}](t, t'; \xi)|\, \mathrm{d}t' \lesssim 2^{(k-\ell)(2\iota - 1)}. \tag{11.10}$$

Proof. If $\ell = \lfloor k/2 \rfloor$, then the localisation of the symbol ensures that $|s - \theta(\xi)| \lesssim 2^{-\ell}$ for all $(\xi; s) \in \mathrm{supp}\, a_{\mathbf{r},k,\ell}$. The bound for $\iota = 0$ then follows immediately from the size of the s-support of $a_{\mathbf{r},k,\ell}$. For $\iota = 1$, note that by the mean value theorem, we may write

$$\langle \gamma'(s), \xi \rangle = \omega(\xi; s)\,(s - \theta(\xi)) \tag{11.11}$$

where $|\omega(\xi; s)| \sim 2^k$ on $\mathrm{supp}\, a_{k,\ell}$. Consequently,

$$|\langle \gamma'(s), \xi \rangle| \lesssim 2^{k/2} \qquad \text{for all } (\xi; s) \in \mathrm{supp}\, a_{\mathbf{r},k,\ell}. \tag{11.12}$$

Furthermore, by the definition of $a_{\mathbf{r}}$ and of the Frenet frame $\{\mathbf{e}_j(r)\}_{j=1}^3$, the relation $r_3 \leqslant r_2 \leqslant r_1 \lesssim r_2^{1/2} \leqslant 2^{-k/2}$ implies

$$|\partial_s a_{\mathbf{r},k,\ell}(\xi; s)| \lesssim 2^{k/2}. \tag{11.13}$$

In view of the definition of $\mathfrak{d}_s$ in (11.3), the bounds (11.12) and (11.13) immediately imply that $|\mathfrak{d}_s a_{\mathbf{r},k,\ell}(\xi; s)| \lesssim 2^{k/2}$ and the bound for $\iota = 1$ now follows immediately from the size of the s-support of $a_{\mathbf{r},k,\ell}$ and the definition of $\mathcal{K}$ in (11.6).

If $0 \leqslant \ell < \lfloor k/2 \rfloor$, then the localisation of the symbols ensures that

$$|s - \theta(\xi)| \sim 2^{-\ell} \qquad \text{for all } (\xi; s) \in \operatorname{supp} a_{\mathbf{r},k,\ell}. \tag{11.14}$$

Consequently, by directly applying (11.14) in (11.11), we have the bounds

$$|\langle \gamma'(s), \xi \rangle| \sim 2^{k-\ell}, \quad |\langle \gamma^{(N)}(s), \xi \rangle| \lesssim 2^{k} \qquad \text{for } N \geqslant 2, \ (\xi; s) \in \operatorname{supp} a_{\mathbf{r},k,\ell}. \tag{11.15}$$

Moreover, by the definition of $a_{\mathbf{r}}$, the first relation above immediately implies $2^{k-\ell} \leqslant r_1^{-1}$; recall that $r_2, r_3 \leqslant 2^{-k}$. Thus, by the definition of the Frenet frame $\{\mathbf{e}_j(s)\}_{j=1}^3$, the symbol satisfies

$$|\partial_s^N a_{\mathbf{r},k,\ell}(\xi; s)| \lesssim 2^{\ell N} = 2^{-(k-2\ell)N} 2^{(k-\ell)N} \qquad \text{for all } N \in \mathbb{N}_0. \tag{11.16}$$

Thus, we may bound the kernel via repeated integration-by-parts. In particular, applying Lemma D.1 with $\phi(s) := (t - t')\langle \gamma(s), \xi \rangle$ and $R := 2^{k-2\ell}|t - t'|$, we deduce that

$$|\mathcal{K}[\mathfrak{d}_s^{\iota} a_{\mathbf{r},k,\ell}](\xi; t, t')| \lesssim_N 2^{2(k-\ell)\iota} 2^{-\ell} \big(1 + 2^{k-2\ell}|t - t'|\big)^{-N}, \qquad \text{for } \iota \in \{0, 1\}.$$

The additional $2^{2(k-\ell)}$ factor arises in the bound for the derived operator owing to the formula (11.3) for the corresponding symbol (and in particular, due to the first bound in (11.15), the bounds in (11.16) and the relation $0 \leqslant \ell < \lfloor k/2 \rfloor$) and the form of the kernel as described in (11.6). Integrating both sides of the above display in either t or t', the desired estimate (11.10) follows. □

Putting everything together. In view of the kernel estimates from Lemma 11.1 and the discussion at the beginning of the proof, it follows that

$$\|\mathcal{A}[\mathfrak{d}_s^{\iota} a_{\mathbf{r},k,\ell}] g\|_{L^2(\mathbb{R}^4) \to L^2(\mathbb{R}^{3+1})} \lesssim 2^{(k-\ell)(\iota - 1/2)} \qquad \text{for all } 0 \leqslant \ell \leqslant \lfloor k/2 \rfloor \text{ and } \iota \in \{0, 1\}.$$

Combining these bounds with (11.2), it follows that

$$\|\mathcal{N}[a_{\mathbf{r},k,\ell}] g\|_{L^2(\mathbb{R}^4) \to L^2(\mathbb{R}^3)} \lesssim 1 \quad \text{for all } 0 \leqslant \ell \leqslant \lfloor k/2 \rfloor.$$

The frequency localised maximal bound (11.9) immediately follows (with linear dependence on k) from the triangle inequality.

<u>$J = 3$ case.</u> Suppose $\operatorname{supp}_\xi a_{\mathbf{r},k} \subseteq \Omega_3$. As in Lemma 6.1, if $\xi \in \Omega_3$, then the equation $\langle \gamma''(s), \xi \rangle = 0$ has a unique solution in $[-1, 1]$, which we denote by $\theta_2(\xi)$. As in Lemma 6.2, if $u(\xi) < 0$, where

$$u(\xi) := \langle \gamma' \circ \theta_2(\xi), \xi \rangle,$$

then the equation $\langle \gamma'(s), \xi \rangle = 0$ has a precisely two solutions in $[-1, 1]$, which we denote by $\theta_1^{\pm}(\xi)$. We will further assume without loss of generality that $\langle \gamma^{(3)}(s), \xi \rangle > 0$ for all $\xi \in \operatorname{supp}_\xi a_{\mathbf{r},k}$.

Further decomposition. Here the symbol $a_{\mathbf{r},k}$ is decomposed in a manner similar (but not quite identical) to that used in §8. First perform a dyadic decomposition of $u(\xi)$ by writing

$$a_{\mathbf{r},k} = \sum_{\ell=0}^{\lfloor k/3 \rfloor} a_{\mathbf{r},k,\ell} + \sum_{\ell=0}^{\lfloor k/3 \rfloor - 1} a_{\mathbf{r},k,\ell}^{+}$$

where

$$a_{\mathbf{r},k,\ell}(\xi; s) := \begin{cases} a_{\mathbf{r},k}(\xi; s) \beta^{-}\big(2^{-k+2\ell} u(\xi)\big) & \text{if } 0 \leqslant \ell < \lfloor k/3 \rfloor \\ a_{\mathbf{r},k}(\xi; s) \eta\big(2^{-k+2\lfloor k/3 \rfloor} u(\xi)\big) & \text{if } \ell = \lfloor k/3 \rfloor \end{cases}$$

and the $a_{\mathbf{r},k,\ell}^{+}$ are defined similarly but with β^{+} in place of β^{-}. Here $\beta = \beta^{-} + \beta^{+}$ is the decomposition of the bump function described in §8.1. The symbols $a_{\mathbf{r},k,\ell}^{+}$ are relatively easy to analyse, and are dealt with using an argument similar to that of the $J = 2$ case. Henceforth, we focus exclusively on the $a_{\mathbf{r},k,\ell}$.

We further decompose each $a_{\mathbf{r},k,\ell}$ with respect to the distance of the s-variable to the root $\theta_2(\xi)$. Once again it is convenient to introduce a fine tuning constant $\rho > 0$. Similar to (8.1), define

$$a_{\mathbf{r},k,\ell,0}(\xi;s) := a_{\mathbf{r},k,\ell}(\xi;s)\eta\big(\rho 2^{\ell}|s-\theta_2(\xi)|\big) \qquad \text{for } 0 \leqslant \ell \leqslant \lfloor k/3 \rfloor. \tag{11.17}$$

Note, in contrast with (8.1), we have not decomposed with respect to $|s-\theta_1^{\pm}(\xi)|$ for $\ell < \lfloor k/3 \rfloor$. Such a decomposition does appear later: here it is necessary to localise simultaneously with respect to *both* roots $\theta_2(\xi)$ and $\theta_1^{\pm}(\xi)$. Also in contrast with the analysis of §8, here it is not possible to reduce the problem to studying the s-localised pieces in (11.17). Consequently, we also consider the s-localisation of the symbol to the remaining dyadic shells, viz.

$$a_{\mathbf{r},k,\ell,m}(\xi;s) := a_{\mathbf{r},k,\ell}(\xi)\beta\big(\rho 2^{\ell-m}|s-\theta_2(\xi)|\big) \qquad \text{for } 0 \leqslant \ell \leqslant \lfloor k/3 \rfloor,\ 0 < m \leqslant \ell.$$

The most difficult terms to estimate correspond to $0 \leqslant \ell < \lfloor k/3 \rfloor$ and $m = 0$. These symbols require a further decomposition. In particular, for $0 \leqslant \ell < \lfloor k/3 \rfloor$ let

$$b_{\mathbf{r},k,\ell,m}(\xi;s) := \begin{cases} a_{\mathbf{r},k,\ell,0}(\xi;s)\eta\big(\rho^{-1}2^{(k-\ell)/2}\min_{\pm}|s-\theta_1^{\pm}(\xi)|\big) & \text{if } m = 0 \\ a_{\mathbf{r},k,\ell,0}(\xi;s)\beta\big(\rho^{-1}2^{(k-\ell)/2-m}\min_{\pm}|s-\theta_1^{\pm}(\xi)|\big) & \text{if } 1 \leqslant m < \lfloor \frac{k-3\ell}{2} \rfloor \\ a_{\mathbf{r},k,\ell,0}(\xi;s)\big(1-\eta\big(\rho^{-1}2^{(k-\ell)/2-m}\min_{\pm}|s-\theta_1^{\pm}(\xi)|\big)\big) & \text{if } m = \lfloor \frac{k-3\ell}{2} \rfloor \end{cases}.$$

Observe that Lemma 6.3 already implies that $|s-\theta_1^{\pm}(\xi)| \lesssim \rho^{-1}2^{-\ell}$ for $(\xi;s) \in \operatorname{supp} a_{\mathbf{r},k,\ell,0}$. Thus, $\rho 2^{-\ell} \lesssim |s-\theta_1^{\pm}(\xi)| \lesssim \rho^{-1}2^{-\ell}$ for $(\xi;s) \in \operatorname{supp} b_{\mathbf{r},k,\ell,m}$ for $m = \lfloor \frac{k-3\ell}{2} \rfloor$.

Combining the above definitions and observations, the symbol may be written as

$$a_{\mathbf{r},k} = \sum_{\ell=0}^{\lfloor k/3 \rfloor}\sum_{m=0}^{\ell} a_{\mathbf{r},k,\ell,m} = \sum_{(\ell,m)\in\Lambda_a(k)} a_{\mathbf{r},k,\ell,m} + \sum_{(\ell,m)\in\Lambda_b(k)} b_{\mathbf{r},k,\ell,m}$$

where

$$\Lambda_a(k) := \big\{(\ell,m) \in \mathbb{N}_0^2 : 0 \leqslant \ell \leqslant \lfloor \tfrac{k}{3} \rfloor \text{ and } 1 \leqslant m \leqslant \ell\big\} \cup \big\{\big(\lfloor \tfrac{k}{3} \rfloor, 0\big)\big\},$$
$$\Lambda_b(k) := \big\{(\ell,m) \in \mathbb{N}_0^2 : 0 \leqslant \ell < \lfloor \tfrac{k}{3} \rfloor \text{ and } 0 \leqslant m \leqslant \lfloor \tfrac{k-3\ell}{2} \rfloor\big\}.$$

Note that the range of m in the definition of $\Lambda_a(k)$ is restricted since $a_{\mathbf{r},k,\ell,m}$ is identically zero whenever $m > \ell$.

Kernel estimates. The kernels are analysed using stationary phase techniques.

Lemma 11.2. *Let $k \in \mathbb{N}$ and $\iota \in \{0,1\}$.*

a) If $(\ell,m) \in \Lambda_a(k)$, then

$$\sup_{t'\in[1,2]}\int_1^2 |\mathcal{K}[\mathfrak{d}_s^{\iota} a_{\mathbf{r},k,\ell,m}](t,t';\xi)|\,\mathrm{d}t, \quad \sup_{t\in[1,2]}\int_1^2 |\mathcal{K}[\mathfrak{d}_s^{\iota} a_{\mathbf{r},k,\ell,m}](t,t';\xi)|\,\mathrm{d}t' \lesssim 2^{(k-2\ell+2m)(2\iota-1)}. \tag{11.18}$$

b) If $(\ell,m) \in \Lambda_b(k)$, then

$$\sup_{t'\in[1,2]}\int_1^2 |\mathcal{K}[\mathfrak{d}_s^{\iota} b_{\mathbf{r},k,\ell,m}](t,t';\xi)|\,\mathrm{d}t, \quad \sup_{t\in[1,2]}\int_1^2 |\mathcal{K}[\mathfrak{d}_s^{\iota} b_{\mathbf{r},k,\ell,m}](t,t';\xi)|\,\mathrm{d}t' \lesssim 2^{((k-\ell)/2+m)(2\iota-1)}. \tag{11.19}$$

Proof. The argument is similar to that used to prove Lemma 8.1.

a) Let $(\ell,m) \in \Lambda_a(k)$. If $(\ell,m) = (\lfloor k/3 \rfloor, 0)$, then the localisation of the $a_{\mathbf{r},k,\ell,m}$ symbols ensures that

$$|u(\xi)| \lesssim 2^{k/3} \quad \text{and} \quad |s-\theta_2(\xi)| \lesssim \rho^{-1}2^{-k/3} \quad \text{for all } (\xi;s) \in \operatorname{supp} a_{\mathbf{r},k,\ell,m}. \tag{11.20}$$

The bound (11.18) for $\iota = 0$ follows immediately from the size of the s-support of $a_{\mathbf{r},k,\ell,m}$. For $\iota = 1$, apply the familiar Taylor expansion to write

$$\begin{aligned}\langle \gamma'(s), \xi\rangle &= u(\xi) + \omega_1(\xi; s)\,(s - \theta_2(\xi))^2,\\ \langle \gamma''(s), \xi\rangle &= \omega_2(\xi; s)\,(s - \theta_2(\xi))\end{aligned} \tag{11.21}$$

where $|\omega_j(\xi; s)| \sim 2^k$ on $\operatorname{supp} a_{\mathbf{r},k,\ell,m}$ for $j = 1, 2$. Consequently, by directly applying (11.20), we have the upper bounds

$$|\langle \gamma'(s), \xi\rangle| \lesssim \rho^{-2}2^{k/3}, \quad |\langle \gamma''(s), \xi\rangle| \lesssim \rho^{-1}2^{2k/3}, \qquad \text{for all } (\xi; s) \in \operatorname{supp} a_{\mathbf{r},k,\ell,m}. \tag{11.22}$$

Note that the relations $r_2 \leqslant r_1 \leqslant r_2^{1/2}$ and $r_3 \leqslant r_2 \leqslant r_1^{1/2} r_3^{1/2}$ imply, in particular, $r_1 \leqslant r_3^{1/3} \lesssim 2^{-k/3}$ and $r_2 \leqslant r_3^{2/3} \lesssim 2^{-2k/3}$. It then follows from the definitions of $a_{\mathbf{r}}$ and of the Frenet frame $\{\mathbf{e}_j(r)\}_{j=1}^3$ that

$$|\partial_s a_{\mathbf{r},k,\ell,m}(\xi; s)| \lesssim 2^{k/3}. \tag{11.23}$$

In view of the definition of $\mathfrak{d}_s$ in (11.3), the first bound in (11.22) and (11.23) immediately imply that $|\mathfrak{d}_s a_{\mathbf{r},k,\ell,m}(\xi; s)| \lesssim 2^{k/3}$, and the bound for $\iota = 1$ now follows immediately from the size of the s-support of $a_{\mathbf{r},k,\ell,m}$ and the definition of $\mathcal{K}$ in (11.7).

Now suppose $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$ and $1 \leqslant m \leqslant \ell$. Then the localisation of the $a_{\mathbf{r},k,\ell,m}$ symbols ensures that

$$|u(\xi)| \lesssim 2^{k-2\ell} \quad \text{and} \quad |s - \theta_2(\xi)| \sim \rho^{-1}2^{-\ell+m} \quad \text{for all } (\xi; s) \in \operatorname{supp} a_{\mathbf{r},k,\ell,m}. \tag{11.24}$$

Provided ρ is chosen sufficiently small, by directly applying (11.24) in (11.21), we have the bounds

$$|\langle \gamma'(s), \xi\rangle| \sim \rho^{-2}2^{k-2\ell+2m}, \quad |\langle \gamma''(s), \xi\rangle| \sim \rho^{-1}2^{k-\ell+m}, \quad |\langle \gamma^{(N)}(s), \xi\rangle| \lesssim_N 2^k \qquad \text{for } N \geqslant 3. \tag{11.25}$$

By the definition of $a_{\mathbf{r}}$, the first and second bounds above immediately imply $2^{k-2\ell+2m} \leqslant r_1^{-1}$ and $2^{k-\ell+m} \leqslant r_2^{-1}$, whilst $2^k \leqslant r_3^{-1}$. Thus, by the definition of the Frenet frame $\{\mathbf{e}_j(s)\}_{j=1}^3$ and the bounds (11.25), the symbol satisfies

$$|\partial_s^N a_{\mathbf{r},k,\ell,m}(\xi; s)| \lesssim 2^{(\ell-m)N} = 2^{-(k-3\ell+3m)N}2^{(k-2\ell+2m)N} \qquad \text{for all } N \in \mathbb{N}_0. \tag{11.26}$$

Thus, we may bound the kernel via repeated integration-by-parts. In particular, applying Lemma D.1 with $\phi(s) := (t - t')\langle \gamma(s), \xi\rangle$ and $R := 2^{k-3\ell+3m}|t - t'|$, we deduce that

$$|\mathcal{K}[\mathfrak{d}_s^\iota a_{\mathbf{r},k,\ell,m}](\xi; t, t')| \lesssim_N 2^{2(k-2\ell+2m)\iota}2^{-\ell+m}\big(1 + 2^{k-3\ell+3m}|t - t'|\big)^{-N}.$$

The additional $2^{2(k-2\ell+2m)\iota}$ arises in the bound for the derived operator $\mathfrak{d}_s$ owing to the formula (11.3) for the corresponding symbol (and in particular, due to the bounds in (11.25) and in (11.26) and the relation $0 \leqslant \ell - m \leqslant \ell \leqslant \lfloor k/3 \rfloor$) and the form of the kernel $\mathcal{K}$ as described in (11.6). Finally, by integrating both sides of the above display in either t or t', the desired estimate (11.18) follows.

b) Let $(\ell, m) \in \Lambda_b(k)$. If $m = 0$, then the localisation of the $b_{\mathbf{r},k,\ell,m}$ symbols ensures that

$$|u(\xi)| \sim 2^{k-2\ell} \quad \text{and} \quad \min_{\pm} |s - \theta_1^{\pm}(\xi)| \lesssim \rho 2^{-(k-\ell)/2} \quad \text{for all } (\xi; s) \in \operatorname{supp} b_{\mathbf{r},k,\ell,m}. \tag{11.27}$$

The bound (11.19) for $\iota = 0$ follows immediately from the size of the s-support of $b_{\mathbf{r},k,\ell,m}$. For $\iota = 1$, apply the familiar Taylor expansion to write

$$\begin{aligned}\langle \gamma'(s), \xi\rangle &= v^{\pm}(\xi)\,(s - \theta_1^{\pm}(\xi)) + \omega_1^{\pm}(\xi; s)\,(s - \theta_1^{\pm}(\xi))^2,\\ \langle \gamma''(s), \xi\rangle &= v^{\pm}(\xi) + \omega_2^{\pm}(\xi; s)\,(s - \theta_1^{\pm}(\xi))\end{aligned} \tag{11.28}$$

where $|\omega_j^{\pm}(\xi; s)| \sim 2^k$ on $\operatorname{supp} b_{\mathbf{r},k,\ell,m}$ for $j = 1, 2$. Consequently, in view of Lemma 6.3 and (11.27), and provided $\rho > 0$ is chosen sufficiently small, we have the bounds,

$$|\langle \gamma'(s), \xi\rangle| \lesssim \rho 2^{(k-\ell)/2}, \quad |\langle \gamma''(s), \xi\rangle| \sim 2^{k-\ell} \qquad \text{for all } (\xi; s) \in \operatorname{supp} b_{\mathbf{r},k,\ell,m}, \tag{11.29}$$

using the relation $0 \leqslant \ell \leqslant \lfloor k/3 \rfloor$.

By the definition of $a_{\mathbf{r}}$, the second bound above implies $r_2 \lesssim 2^{-(k-\ell)}$ and therefore $r_1 \leqslant r_2^{1/2} \lesssim 2^{-(k-\ell)/2}$, whilst $r_3 \lesssim 2^{-k}$. Thus, by the definition of the Frenet frame $\{\mathbf{e}_j(s)\}_{j=1}^3$ and the bounds (11.29), the symbol satisfies

$$|\partial_s b_{\mathbf{r},k,\ell,m}(\xi; s)| \lesssim_N 2^{(k-\ell)/2}, \tag{11.30}$$

using the relation $0 \leqslant \ell < \lfloor k/3 \rfloor$. In view of the definition of $\mathfrak{d}_s$ in (11.3), the first bound in (11.29) and (11.30) immediately implies that $|\mathfrak{d}_s b_{\mathbf{r},k,\ell,m}(\xi; s)| \lesssim 2^{(k-\ell)/2}$, and the bound for $\iota = 1$ now follows immediately from the size of the s-support of $b_{\mathbf{r},k,\ell,m}$ and the definition of $\mathcal{K}$ in (11.6).

Now suppose $0 < m < \lfloor \frac{k-3\ell}{2} \rfloor$. Then the localisation of the $b_{\mathbf{r},k,\ell,m}$ symbols ensures that

$$|u(\xi)| \sim 2^{k-2\ell} \quad \text{and} \quad \min_{\pm} |s - \theta_1^{\pm}(\xi)| \sim \rho 2^{-(k-\ell)/2+m} \quad \text{for all } (\xi; s) \in \operatorname{supp} b_{\mathbf{r},k,\ell,m}. \tag{11.31}$$

Using the convexity argument from the proof of Lemma 8.1, we may bound

$$|\langle \gamma'(s), \xi \rangle| \geqslant \min_{\pm} \frac{|u(\xi)||s - \theta_1^{\pm}(\xi)|}{|\theta_2(\xi) - \theta_1^{\pm}(\xi)|} \qquad \text{for all } (\xi; s) \in \operatorname{supp} b_{\mathbf{r},k,\ell,m}. \tag{11.32}$$

Consequently, using Lemma 6.3 and (11.31) in (11.28) and (11.32), and provided $\rho > 0$ is chosen sufficiently small,

$$|\langle \gamma'(s), \xi \rangle| \sim \rho 2^{(k-\ell)/2+m}, \quad |\langle \gamma''(s), \xi \rangle| \sim 2^{k-\ell} \quad \text{and} \quad |\langle \gamma^{(N)}(s), \xi \rangle| \lesssim_N 2^k \qquad \text{for all } N \geqslant 3. \tag{11.33}$$

For the upper bound in the first derivative in the above display, we use the restriction $m \leqslant \lfloor \frac{k-3\ell}{2} \rfloor$. It is for this reason that we simultaneously localise with respect to *both* $\theta_2(\xi)$ and $\theta_1^{\pm}(\xi)$. In particular,

$$|\langle \gamma^{(N)}(s), \xi \rangle| \lesssim 2^k \sim 2^{k-((k-\ell)/2+m)N} |\langle \gamma'(s), \xi \rangle|^N \lesssim 2^{-2m(N-1)} |\langle \gamma'(s), \xi \rangle|^N \quad \text{for all } N \geqslant 3,$$

where in the last inequality one uses the restriction $m \leqslant \lfloor \frac{k-3\ell}{2} \rfloor$ and the fact $N \geqslant 3$.

By the definition of $a_{\mathbf{r}}$, the first and second bounds in (11.33) imply $r_1 \leqslant 2^{-(k-\ell)/2-m}$ and $r_2 \leqslant 2^{-(k-\ell)}$, whilst $r_3 \leqslant 2^{-k}$. Thus, by the definition of the Frenet frame $\{\mathbf{e}_j(s)\}_{j=1}^3$ and the bounds (11.33), the symbol satisfies

$$|\partial_s^N b_{\mathbf{r},k,\ell,m}(\xi; s)| \lesssim 2^{((k-\ell)/2-m)N} = 2^{-2mN} 2^{((k-\ell)/2+m)N} \qquad \text{for all } N \in \mathbb{N}_0, \tag{11.34}$$

using the restriction $m \leqslant \lfloor \frac{k-3\ell}{2} \rfloor$. Thus, we may bound the kernel via repeated integration-by-parts. In particular, applying Lemma D.1 with $\phi(s) := (t - t')\langle \gamma(s), \xi \rangle$ and $R := 2^{2m}|t - t'|$, we deduce that

$$|\mathcal{K}[\mathfrak{d}_s^{\iota} b_{\mathbf{r},k,\ell,m}](\xi; t, t')| \lesssim_N 2^{(k-\ell+2m)\iota} 2^{-(k-\ell)/2+m} \big(1 + 2^{2m}|t - t'|\big)^{-N}.$$

The additional $2^{(k-\ell+2m)\iota}$ arises in the bound for the derived operator $\mathfrak{d}_s$ owing to the formula (11.3) for the corresponding symbol (and in particular, due to the bounds in (11.33) and in (11.34)) and the form of the kernel $\mathcal{K}$ as described in (11.6). Finally, by integrating both sides of the above display in either t or t', the desired estimate (11.18) follows.

Finally, consider the case $m = \lfloor \frac{k-3\ell}{2} \rfloor$. Then the localisation of the $b_{\mathbf{r},k,\ell,m}$ symbols ensures that

$$|u(\xi)| \sim_\rho 2^{k-2\ell} \quad \text{and} \quad \min_{\pm} |s - \theta_1^{\pm}(\xi)| \sim_\rho 2^{-\ell} \quad \text{for all } (\xi; s) \in \operatorname{supp} b_{\mathbf{r},k,\ell,m}. \tag{11.35}$$

Using Lemma 6.3 and (11.35) in (11.28) and (11.32), we have the bounds

$$|\langle \gamma'(s), \xi \rangle| \sim 2^{k-2\ell}, \quad |\langle \gamma''(s), \xi \rangle| \lesssim 2^{k-\ell}, \quad \text{and} \quad |\langle \gamma^{(N)}(s), \xi \rangle| \lesssim_N 2^k \qquad \text{for all } N \geqslant 3.$$

By the definition of $a_{\mathbf{r}}$, the first bound above implies $r_1 \lesssim_\rho 2^{-(k-2\ell)}$ and, as $r_3 \lesssim 2^{-k}$, one has $r_2 \leqslant r_1^{1/2} r_3^{1/2} \lesssim_\rho 2^{k-\ell}$. Thus, by the definition of the Frenet frame $\{\mathbf{e}_j(s)\}_{j=1}^3$ and the bounds (11.33), the symbol satisfies

$$|\partial_s^N b_{\mathbf{r},k,\ell,m}(\xi;s)| \lesssim 2^{\ell N} = 2^{-(k-3\ell)N} 2^{(k-2\ell)N} \qquad \text{for all } N \in \mathbb{N}_0.$$

We may then bound the kernel via repeated integration-by-parts. In particular, applying Lemma D.1 with $\phi(s) := (t-t')\langle \gamma(s), \xi\rangle$ and $R := 2^{k-3\ell}|t-t'|$, we deduce that

$$|\mathcal{K}[\mathfrak{d}_s^\iota b_{\mathbf{r},k,\ell,m}](\xi;t,t')| \lesssim_N 2^{2(k-2\ell)\iota} 2^{-\ell}\big(1 + 2^{k-3\ell}|t-t'|\big)^{-N}.$$

The additional $2^{2(k-2\ell)\iota}$ arises in the bound for the derived operator $\mathfrak{d}_s$ owing to the formula (11.3) for the corresponding symbol (and in particular, due to the bounds in (11.33) and in (11.34) and the restriction $\ell \leqslant \lfloor k/3 \rfloor$) and the form of the kernel $\mathcal{K}$ as described in (11.6). Finally, by integrating both sides of the above display in either t or t', the desired estimate (11.18) follows. □

Putting everything together. In view of the kernel estimates from Lemma 11.2 and the discussion at the beginning of the proof, it follows that

$$\|\mathcal{A}[\mathfrak{d}_s^\iota a_{\mathbf{r},k,\ell,m}]g\|_{L^2(\mathbb{R}^4)\to L^2(\mathbb{R}^{3+1})} \lesssim 2^{(k-2\ell+2m)(\iota-1/2)} \quad \text{for all } (\ell,m) \in \Lambda_a(k),$$
$$\|\mathcal{A}[\mathfrak{d}_s^\iota b_{\mathbf{r},k,\ell,m}]g\|_{L^2(\mathbb{R}^4)\to L^2(\mathbb{R}^{3+1})} \lesssim 2^{((k-\ell)/2+m)(\iota-1/2)} \quad \text{for all } (\ell,m) \in \Lambda_b(k),$$

for $\iota \in \{0,1\}$. Combining these bounds with (11.2), it follows that

$$\|\mathcal{N}[a_{\mathbf{r},k,\ell,m}]g\|_{L^2(\mathbb{R}^4)\to L^2(\mathbb{R}^3)} \lesssim 1 \quad \text{for all } (\ell,m) \in \Lambda_a(k),$$
$$\|\mathcal{N}[b_{\mathbf{r},k,\ell,m}]g\|_{L^2(\mathbb{R}^4)\to L^2(\mathbb{R}^3)} \lesssim 1 \quad \text{for all } (\ell,m) \in \Lambda_b(k).$$

Since the cardinalities of $\Lambda_a(k)$ and $\Lambda_b(k)$ are $O(k^2)$, the frequency localised maximal bound (11.9) immediately follows from the triangle inequality. Summing over k then concludes the proof of the proposition. □

12. Necessary conditions

In this final section we show that the condition $p > 3$ in Theorem 1.1 is necessary. Moreover, we prove the following result, which is valid in arbitrary dimensions $n \geqslant 2$.

Proposition 12.1. *If $n \geqslant 2$ and $\gamma\colon I \to \mathbb{R}^n$ is a smooth non-degenerate curve, then*

$$\|M_\gamma\|_{L^p(\mathbb{R}^n)\to L^p(\mathbb{R}^n)} = \infty \qquad \text{for } 1 \leqslant p \leqslant n.$$

Proof. By localisation of the operator and applying the rescaling from §4, it suffices to consider the case where

$$\gamma(\,\cdot\,) - \gamma(0) \in \mathfrak{G}_n(\delta_0) \qquad \text{and} \qquad \langle \gamma(0), \vec{e}_n\rangle \neq 0$$

for $\delta_0 := 10^{-n}$, say. By reparametrising the curve, we may also assume that the first component of $\gamma\colon [-1,1] \to \mathbb{R}^n$ is of the form $\gamma_1(s) = s + a_1$ for some $a_1 \in \mathbb{R}$.

By a simple projection argument, it suffices to study the boundedness of a maximal operator defined over the Euclidean plane. In particular, fix $a = (a_1,a_2) \in \mathbb{R}^2$ with $a_2 \neq 0$ and a smooth function $h\colon [-1,1] \to \mathbb{R}$ satisfying

$$h^{(j)}(0) = 0 \quad \text{for } 0 \leqslant j \leqslant n-1 \quad \text{and} \quad h^{(n)}(0) \neq 0. \tag{12.1}$$

Define the maximal operator

$$\mathcal{M}_h f(x) := \sup_{t>0} \Big| \int_{\mathbb{R}} f\big(x_1 - t(s+a_1), x_2 - t(h(s)+a_2)\big)\chi(s)\,\mathrm{d}s \Big|$$

where $\chi \in C_c^\infty(\mathbb{R})$ is non-negative, satisfies $\chi(s) = 1$ for $|s| \leqslant 1/2$ and has support contained in the interior of $[-1,1]$. To prove the proposition, it suffices to show

$$\|\mathcal{M}_h\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2)} = \infty \qquad \text{for } 1 \leqslant p \leqslant n. \tag{12.2}$$

Furthermore, since the maximal operator is trivially bounded on L^∞ it suffices to consider the case $p = n$ only.

By Taylor expansion and (12.1), we have

$$|h(s)| \leqslant D_h \cdot |s|^n \qquad \text{for } |s| \leqslant 1 \quad \text{where} \quad D_h := \frac{1}{n!}\sup_{|s|\leqslant 1}|h^{(n)}(s)|.$$

For $0 < r < 1$ let $f_r := \mathbb{1}_{K(r)}$ denote the indicator function of the set

$$K(r) := \big\{y = (y_1, y_2) \in \mathbb{R}^2 : |y_1 - a_1| \leqslant r \text{ and } |y_2 - a_2| \leqslant D_h \cdot r^n\big\}$$

and observe that

$$\|f_r\|_{L^n(\mathbb{R}^2)} \sim_h r^{(n+1)/n}. \tag{12.3}$$

Now let $\delta^n \leqslant \lambda \leqslant 1$ be a dyadic number and suppose $x \in E_\lambda(r)$ where

$$E_\lambda(r) := \Big\{x = (x_1, x_2) \in \mathbb{R}^2 : \Big|x_1 - \frac{a_1}{a_2}x_2\Big| \leqslant \frac{r}{2} \text{ and } \lambda \leqslant \frac{x_2}{a_2} - 1 < 2\lambda\Big\}.$$

If we define $t_x := a_2^{-1}x_2 - 1 \in [\lambda, 2\lambda]$, then for any $s \in \mathbb{R}$ satisfying $|s| \leqslant \frac{1}{2}\cdot\lambda^{(n-1)/n}r$ we have

$$|x_1 - t_x(t_x^{-1}s + a_1) - a_1| \leqslant \Big|x_1 - \frac{a_1}{a_2}x_2\Big| + |s| \leqslant r,$$
$$|x_2 - t_x(h(t_x^{-1}s) + a_2) - a_2| = |t_x||h(t_x^{-1}s)| \leqslant D_h \cdot \lambda^{-(n-1)}|s|^n \leqslant D_h \cdot r^n.$$

From these observations, we conclude that

$$\text{if } x \in E_\lambda(r) \text{ and } |s| \leqslant \frac{1}{2}\lambda^{(n-1)/n}r, \text{ then } \big(x_1 - t_x(t_x^{-1}s + a_1), x_2 - t_x(h(t_x^{-1}s) + a_2)\big) \in K(r).$$

Performing a change of variables in the underlying averaging operator, we deduce that

$$\mathcal{M}_h f_r(x) \gtrsim \lambda^{-1/n}r \qquad \text{for all } x \in E_\lambda(r),$$

where here we pick up an extra factor of λ^{-1} owing to the Jacobian. Consequently,

$$\|\mathcal{M}_h f_r\|_{L^n(\mathbb{R}^2)} \gtrsim \Big(\sum_{\substack{\lambda:\,\text{dyadic}\\ r^n\leqslant\lambda\leqslant 1}} \lambda^{-1}r^n|E_\lambda(r)|\Big)^{1/n} \sim_a |\log r|^{1/n}r^{(n+1)/n}. \tag{12.4}$$

Comparing (12.3) and (12.4), we see that the ratio of $\|\mathcal{M}_h f_r\|_{L^n(\mathbb{R}^2)}$ and $\|f_r\|_{L^n(\mathbb{R}^2)}$ is unbounded in r and therefore (12.2) holds for $p = n$, as desired. □

Appendix A. An abstract broad/narrow decomposition

Here we provide an abstract version of the broad/narrow decomposition in Lemma 9.8. For the sake of self-containedness of this appendix, we recall some of the definitions introduced in §9.3.

Let $\mathfrak{I}$ denote the collection of all dyadic subintervals of $[-1,1]$ and for any dyadic number $0 < r \leqslant 1$ let $\mathfrak{I}(r)$ denote the subset of $\mathfrak{I}$ consisting of all intervals of length r. Let $\mathfrak{I}_{\geqslant r}$ denote the union of the $\mathfrak{I}(\lambda)$ over all dyadic λ satisfying $r \leqslant \lambda \leqslant 1$. Given any pair of dyadic scales $0 < \lambda_1 \leqslant \lambda_2 \leqslant 1$ and $J \in \mathfrak{I}(\lambda_2)$, let $\mathfrak{I}(J;\lambda_1)$ denote the collection of all $I \in \mathfrak{I}(\lambda_1)$ which satisfy $I \subseteq J$.

For $d \in \mathbb{N}$ and each dyadic number $0 < \lambda \leqslant 1$ let $\mathfrak{I}^d_{\text{sep}}(\lambda)$ denote the collection of d-tuples of intervals $\vec{I} = (I_1, \dots, I_d) \in \mathfrak{I}(\lambda)^d$ which satisfy the separation condition

$$\text{dist}(I_1, \dots, I_d) := \min_{1\leqslant \ell_1 < \ell_2 \leqslant d} \text{dist}(I_{\ell_1}, I_{\ell_2}) \geqslant \lambda.$$

Given dyadic scales $0 < \lambda_1 \leqslant \lambda_2 \leqslant 1$ and $J \in \mathfrak{I}(\lambda_2)$, let $\mathfrak{I}^d_{\mathrm{sep}}(J;\lambda_1)$ denote the collection of all d-tuples of intervals $\vec{I} = (I_1, \dots, I_d) \in \mathfrak{I}^d_{\mathrm{sep}}(\lambda_1)$ which satisfy $I_\ell \subseteq J$ for all $1 \leqslant \ell \leqslant d$.

The dyadic decomposition from (9.8) is one instance of an 'abstract' notion of dyadic decomposition, introduced in the following definition.

Definition A.1. *Let (X, μ) be a measure space and $F \colon X \to \mathbb{C}$ a measurable function and $0 < r \leqslant 1$. A sequence $(F_I)_{I \in \mathfrak{I}_{\geqslant r}}$ of measurable functions $F_I \colon X \to \mathbb{C}$ is said to be a dyadic decomposition of F up to scale r if it satisfies*

$$F_{[0,1]} = F \qquad \textit{and} \qquad F_J = \sum_{I \in \mathfrak{I}(J;\lambda_1)} F_I \qquad \textit{for all } J \in \mathfrak{I}(\lambda_2)$$

whenever $0 < r \leqslant \lambda_1 \leqslant \lambda_2 \leqslant 1$ are dyadic. Here the identities are understood to hold μ almost everywhere.

The broad/narrow decomposition result from which Lemma 9.8 follows is the following.

Lemma A.2. *Let (X, μ) be a measure space, $k \in \mathbb{N}$ with $k \geqslant 2$ and $\varepsilon > 0$. For all $r > 0$ there exist dyadic numbers r_{n} and r_{b} satisfying*

$$r < r_{\mathrm{n}} \lesssim_{\varepsilon,k} r, \qquad r < r_{\mathrm{b}} \leqslant 1 \tag{A.1}$$

such that the following holds. If $F \in L^p(X)$ for some $1 \leqslant p < \infty$ and $(F_I)_{I \in \mathfrak{I}}$ is a dyadic decomposition of F up to scale r, then

$$\|F\|_{L^p(X)} \lesssim_{\varepsilon,k} r^{-\varepsilon}\Big(\sum_{I \in \mathfrak{I}(r_{\mathrm{n}})} \|F_I\|_{L^p(X)}^p\Big)^{1/p} + r^{-\varepsilon}\Big(\sum_{\substack{J \in \mathfrak{I}(Cr_{\mathrm{b}}) \\ \vec{I} \in \mathfrak{I}^k_{\mathrm{sep}}(J;r_{\mathrm{b}})}} \Big\|\prod_{j=1}^k |F_{I_j}|^{1/k}\Big\|_{L^p(X)}^p\Big)^{1/p}, \tag{A.2}$$

where $C = C_{\varepsilon,k} \geqslant 1$ is a dyadic number depending only on ε and k.

The intervals $I \in \mathfrak{I}(r_{\mathrm{n}})$ are referred to as *narrow intervals* whilst the k-tuples of intervals $\vec{I} \in \mathfrak{I}^k_{\mathrm{sep}}(I; r_{\mathrm{b}})$ are referred to as *broad interval tuples.*

The key ingredient in the proof of Lemma A.2 is a 1-parameter variant of the Bourgain–Guth decomposition from [9] due to Ham–Lee [14].

Lemma A.3 (Ham–Lee [14])**.** *Let $1 \leqslant p < \infty$ and $k \in \mathbb{N}$ with $k \geqslant 2$. Suppose $0 < \ell_1, \dots, \ell_{k-1} \leqslant 1$ are dyadic numbers such that*

$$1 =: \ell_0 \geqslant \ell_1 \geqslant \dots \geqslant \ell_{k-1} > 0.$$

If (X, μ) is a measure space and $(F_I)_{I \in \mathfrak{I}}$ is a dyadic decomposition of $F \in L^p(X)$, then for any $\ell > 0$,

$$\begin{aligned}\Big(\sum_{J \in \mathfrak{I}(\ell)} \|F_J\|_{L^p(X)}^p\Big)^{1/p} \leqslant 4 \sum_{i=1}^{k-1} \ell_{i-1}^{-2(i-1)} \Big(\sum_{I \in \mathfrak{I}(\ell_i \ell)} \|F_I\|_{L^p(X)}^p\Big)^{1/p} \\ + \ell_{k-1}^{-2(k-1)} \Big(\sum_{\substack{J \in \mathfrak{I}(\ell) \\ \vec{I} \in \mathfrak{I}^k_{\mathrm{sep}}(J;\ell_{k-1}\ell)}} \Big\|\prod_{j=1}^k |F_{I_j}|^{1/k}\Big\|_{L^p(X)}^p\Big)^{1/p}.\end{aligned}$$

Rather than working in the relatively abstract setting of dyadic decompositions of measurable functions, Ham–Lee [14, Lemma 2.8] apply the decomposition only in the concrete setting of Fourier extension operators associated to space curves. However, the proof is elementary, relying little on the exact form of the extension operator, and can easily be adapted to yield Lemma A.3. For completeness, the details are presented at the end of the section.

Lemma A.2 is deduced by applying Lemma A.3 iteratively, for appropriately chosen dyadic scales $\ell_1, \dots, \ell_{k-1}$.

Proof of Lemma A.2. Fix $\varepsilon > 0$ and $k \in \mathbb{N}$ with $N \geqslant 2$. Define the dyadic scales $1 \geqslant \ell_1 \geqslant \cdots \geqslant \ell_{k-1} > 0$ recursively so as to satisfy

$$\frac{\log\big(4 \vee (k-1)\big)}{\log \ell_1^{-1}} \leqslant \frac{\varepsilon}{6}, \qquad \frac{\log \ell_{j-1}^{-2(j-1)}}{\log \ell_j^{-1}} \leqslant \frac{\varepsilon}{6} \quad \text{for } 2 \leqslant j \leqslant k-1.$$

Now fix $r > 0$, $F \in L^p(X)$ and $(F_I)_{I \in \mathfrak{I}}$ a dyadic decomposition of F. If $r \gtrsim_{\varepsilon,k} 1$, then the desired result immediately follows from the triangle inequality and so r may be assumed to be smaller than some small constant $c_{\varepsilon,k}$, depending only on ε and k and chosen for the purposes of the forthcoming argument; in particular we can assume $\ell_{k-1} > r$.

Let $\mathcal{W}$ denote the set of all finite words formed from the alphabet $\{1, \dots, k-1\}$. Given any $w \in \mathcal{W}$ and $1 \leqslant j \leqslant k-1$ write $[w]_j$ for the number of occurrences of j in w and $|w| := [w]_1 + \cdots + [w]_{k-1}$ for the length of the word.

Let $\ell^w := \prod_{j=1}^{k-1} \ell_j^{[w]_j}$ for any $w \in \mathcal{W}$ and define

$$\mathcal{A}(r) := \big\{\alpha \in \mathcal{W} : r < \ell^\alpha \leqslant r/\ell_{k-1}\big\}, \quad \mathcal{B}(r) := \{\beta \in \mathcal{W} : \ell^\beta > r/\ell_{k-1}\}.$$

Finally, for each $N \in \mathbb{N}_0$ define

$$\begin{aligned} \mathcal{A}_{\leqslant N}(r) &:= \big\{\alpha \in \mathcal{A}(r) : |\alpha| \leqslant N\big\}, \quad \mathcal{B}_{\leqslant N}(r) := \{\beta \in \mathcal{B}(r) : |\beta| \leqslant N\}, \\ \mathcal{A}_N(r) &:= \{\alpha \in \mathcal{A}(r) : |\alpha| = N\}, \qquad \mathcal{B}_N(r) := \{\beta \in \mathcal{B}(r) : |\beta| = N\}. \end{aligned}$$

An iterative application of Lemma A.3 yields the following key claim.

Claim. *For all $N \in \mathbb{N}_0$,*

$$\begin{aligned} \|F\|_{L^p(X)} \leqslant &\sum_{\alpha \in \mathcal{A}_{\leqslant N}(r) \cup \mathcal{B}_N(r)} M^\alpha_{\varepsilon,k} \Big(\sum_{I \in \mathfrak{I}(\ell^\alpha)} \|F_I\|^p_{L^p(X)} \Big)^{1/p} \\ &+ \ell_{k-1}^{-2(k-1)} \sum_{\beta \in \mathcal{B}_{\leqslant N-1}(r)} M^\beta_{\varepsilon,k} \Big(\sum_{\substack{J \in \mathfrak{I}(\ell^\beta) \\ \vec{I} \in \mathfrak{I}^k_{\mathrm{sep}}(J; \ell_{k-1}\ell^\beta)}} \Big\| \prod_{j=1}^k |F_{I_j}|^{1/k} \Big\|^p_{L^p(X)} \Big)^{1/p}. \end{aligned} \tag{A.3}$$

where $M^\alpha_{\varepsilon,k} := 4^{|\alpha|} \prod_{j=1}^{k-1} \ell_{j-1}^{-2(j-1)[\alpha]_j}$.

Proof (of Claim). The proof is by induction on N. The case $N = 0$ is vacuous and thus one may assume, by way of induction hypothesis, that (A.3) holds for some $N \geqslant 0$. It remains to establish the inductive step.

Consider the terms on the right-hand side of (A.3) of the form

$$\Big(\sum_{I \in \mathfrak{I}(\ell^\beta)} \|F_I\|^p_{L^p(X)} \Big)^{1/p} \qquad \text{for } \beta \in \mathcal{B}_N(r).$$

Applying Lemma A.3 to each of these terms,

$$\begin{aligned}\|F\|_{L^p(X)} \leqslant \sum_{\alpha \in \mathcal{A}_{\leqslant N}(r)} M^{\alpha}_{\varepsilon,k}\Big(\sum_{I \in \mathfrak{I}(\ell^{\alpha})} \|F_I\|^p_{L^p(X)}\Big)^{1/p} \\ + \sum_{\beta \in \mathcal{B}_N(r)} M^{\beta}_{\varepsilon,k} 4 \sum_{i=1}^{k-1} \ell_{i-1}^{-2(i-1)}\Big(\sum_{I \in \mathfrak{I}(\ell_i \ell^{\beta})} \|F_I\|^p_{L^p(X)}\Big)^{1/p} \\ + \ell_{k-1}^{-2(k-1)} \sum_{\beta \in \mathcal{B}_{\leqslant N}(r)} M^{\beta}_{\varepsilon,k}\Big(\sum_{\substack{J \in \mathfrak{I}(\ell^{\beta}) \\ \vec{I} \in \mathfrak{I}^k_{\mathrm{sep}}(J;\ell_{k-1}\ell^{\beta})}} \|\prod_{j=1}^{k} |F_{I_j}|^{1/k}\|^p_{L^p(X)}\Big)^{1/p}.\end{aligned}$$

From the definitions,

$$\mathcal{A}_{\leqslant N+1}(r) \cup \mathcal{B}_{N+1}(r) = \mathcal{A}_{\leqslant N}(r) \cup \mathcal{A}_{N+1}(r) \cup \mathcal{B}_{N+1}(r),$$

where the union is disjoint. Furthermore, the set $\mathcal{A}_{N+1}(r) \cup \mathcal{B}_{N+1}(r)$ precisely corresponds to the set of words obtained by adding a single letter to one of the words in $\mathcal{B}_N(r)$. Combining these observations, the induction readily closes. □

Using the claim, the proof of Lemma A.2 quickly follows from the choice of scales ℓ_j. Indeed, first observe that for $N := \max_{\alpha \in \mathcal{A}(r)} |\alpha|$ it follows that $\mathcal{B}_N(r) = \varnothing$ and thus

$$\mathcal{A}_{\leqslant N}(r) \cup \mathcal{B}_N(r) = \mathcal{A}(r) \qquad \text{and} \qquad \mathcal{B}_{\leqslant N-1}(r) = \mathcal{B}(r).$$

Note that each $w \in \mathcal{A}(r) \cup \mathcal{B}(r)$ satisfies $(\ell^w)^{-1} < r^{-1}$ and therefore

$$|w| \log \ell_1^{-1} \leqslant \sum_{j=1}^{k-1} [w]_j \log \ell_j^{-1} \leqslant \log r^{-1}. \tag{A.4}$$

By the choice of ℓ_1, it follows that

$$4^{|w|} \leqslant 4^{\log r^{-1}/\log \ell_1^{-1}} = r^{-\log 4/\log \ell_1^{-1}} \leqslant r^{-\varepsilon/6}, \tag{A.5}$$

whilst, similarly,

$$\#\mathcal{A}(r) \cup \mathcal{B}(r) \leqslant \#\{w \in \mathcal{W} : |w| \leqslant \log r^{-1}/\log \ell_1^{-1}\} \leqslant r^{-\log(k-1)/\log \ell_1^{-1}} \leqslant r^{-\varepsilon/6}.$$

On the other hand, as a further consequence of (A.4) and the choice of scales ℓ_j, if $w \in \mathcal{A}(r) \cup \mathcal{B}(r)$, then

$$\log \prod_{j=1}^{k-1} \ell_{j-1}^{-2(j-1)[w]_j} = \sum_{j=1}^{k-1} [w]_j \log \ell_j^{-1} \frac{\log \ell_{j-1}^{-2(j-1)}}{\log \ell_j^{-1}} \leqslant \log r^{-\varepsilon/6}. \tag{A.6}$$

The estimates (A.5) and (A.6) imply that

$$M^{\alpha}_{\varepsilon,k} = 4^{|\alpha|} \prod_{j=1}^{k-1} \ell_{j-1}^{-2(j-1)[w]_j} \leqslant r^{-\varepsilon/3},$$

where $M^{\alpha}_{\varepsilon,k}$ are the constants appearing in the above claim. Combining these observations with (A.3) for the choice of N as above,

$$\begin{aligned}\|F\|_{L^p(X)} \leqslant r^{-\varepsilon/3} \sum_{\alpha \in \mathcal{A}(r)} \Big(\sum_{I \in \mathfrak{I}(\ell^{\alpha})} \|F_I\|^p_{L^p(X)}\Big)^{1/p} \\ + r^{-\varepsilon/3} \ell_{k-1}^{-2(k-1)} \sum_{\beta \in \mathcal{B}(r)} \Big(\sum_{\substack{J \in \mathfrak{I}(\ell^{\beta}) \\ \vec{I} \in \mathfrak{I}^k_{\mathrm{sep}}(J;\ell_{k-1}\ell^{\beta})}} \|\prod_{j=1}^{k} |F_{I_j}|^{1/k}\|^p_{L^p(X)}\Big)^{1/p}.\end{aligned}$$

Finally, since $\ell_{k-1}^{-1} \lesssim_{\varepsilon,k} 1$ and $\#\mathcal{A}(r), \#\mathcal{B}(r), \leqslant r^{-\varepsilon/6}$, by pigeonholing there exists some $\alpha_{\mathrm{n}} \in \mathcal{A}(r)$ and $\beta_{\mathrm{b}} \in \mathcal{B}(r)$ such that, if $r_{\mathrm{n}} := \ell^{\alpha_{\mathrm{n}}}$ and $r_{\mathrm{b}} := \ell_{k-1}\ell^{\beta_{\mathrm{b}}}$ and $C_{\varepsilon,k} := \ell_{k-1}^{-1}$, then the desired inequality (A.2) holds. It is easy to see that these parameters also satisfy (A.1) directly from the relevant definitions. □

To close this section, the proof of Lemma A.3 is presented, following the argument in [14].

Proof of Lemma A.3. For notational convenience, given $m \in \mathbb{N}$ and $J \in \mathfrak{I}(\ell)$ define

$$\pi_J^m(F)(x) := \max_{\vec{I}^{k-1} \in \mathfrak{I}^m_{\mathrm{sep}}(J;\ell_{k-1}\ell)} \prod_{j=1}^m |F_{I_j^{m-1}}(x)|^{1/m}.$$

When $m = 1$ this reduces to $\pi_J^m(F)(x) = |F_J(x)|$. The main step in the proof of Lemma A.3 is the following pointwise estimate.

Claim. *For all $m \in \mathbb{N}$ and $J \in \mathfrak{I}(\ell)$, the pointwise estimate*

$$\pi_J^m(F)(x) \leqslant 4 \max_{I^m \in \mathfrak{I}(J;\ell_m\ell)} |F_{I^m}(x)| + \ell_m^{-2}\, \pi_J^{m+1}(F)(x)$$

holds for μ-almost all $x \in X$.

Proof. Fix $x \in X$ and $\vec{I}^{m-1} = (I_1^{m-1}, \dots, I_m^{m-1}) \in \mathfrak{I}^m(\ell_{m-1}\ell)$ with $I_j^{m-1} \subset J$ for $1 \leqslant j \leqslant m$. For each j there exists an interval $I_j^{m,*} \in \mathfrak{I}(\ell_m\ell)$ satisfying

$$I_j^{m,*} \subset I_j^{m-1} \quad \text{and} \quad |F_{I_j^{m,*}}(x)| = \max_{I_j^m \in \mathfrak{I}(I_j^{m-1};\ell_m\ell)} |F_{I_j^m}(x)|.$$

There are two cases to consider:

Narrow case: Either one of the following two conditions hold:

i) For all $1 \leqslant j \leqslant m$, if $I_j^m \in \mathfrak{I}(\ell_m\ell)$ satisfies $I_j^m \subset I_j^{m-1}$ and $\mathrm{dist}(I_j^m, I_j^{m,*}) \geqslant \ell_m\ell$, then

$$|F_{I_j^m}(x)| \leqslant \Big(\frac{\ell_m}{\ell_{m-1}}\Big)|F_{I_j^{m,*}}(x)|.$$

ii) The selected interval $I_j^{m,*} \in \mathfrak{I}(\ell_m\ell)$ above satisfies

$$\min_{1\leqslant j\leqslant m} |F_{I_j^{m,*}}(x)| \leqslant \Big(\frac{\ell_m}{\ell_{m-1}}\Big)^m \max_{1\leqslant j\leqslant m} |F_{I_j^{m,*}}(x)|.$$

Broad case: The conditions of the narrow case fail.

The narrow case. If condition i) of the narrow case holds, then

$$|F_{I_j^{m-1}}(x)| \leqslant 3|F_{I_j^{m,*}}(x)| + \sum_{\substack{I_j^m \in \mathfrak{I}(\ell_m\ell),\, I_j^m \subset I_j^{m-1} \\ \mathrm{dist}(I_j^m, I_j^{m,*}) \geqslant \ell_m\ell}} |F_{I_j^m}(x)| \leqslant 4|F_{I_j^{m,*}}(x)|,$$

since there are at most ℓ_{m-1}/ℓ_m intervals $I_j^m \in \mathfrak{I}(\ell_m\ell)$ contained in I_j^{m-1}. Thus, in this case,

$$\prod_{j=1}^m |F_{I_j^{m-1}}(x)|^{1/m} \leqslant 4 \max_{I^m \in \mathfrak{I}(J;\ell_m\ell)} |F_{I^m}(x)|. \tag{A.7}$$

Now suppose that condition ii) of the narrow case holds. Thus,

$$\prod_{j=1}^m |F_{I_j^{m-1}}(x)|^{1/m} \leqslant \Big(\frac{\ell_{m-1}}{\ell_m}\Big) \prod_{j=1}^m |F_{I_j^{m,*}}(x)|^{1/m} \leqslant \max_{1\leqslant j\leqslant m} |F_{I_j^{m,*}}(x)|$$

where the first inequality follows since there are at most ℓ_{m-1}/ℓ_m intervals $I_j^m \in \mathfrak{I}(\ell_m\ell)$ contained in I_j^{m-1}. Once again, (A.7) holds (in fact, it holds with a constant 1 rather 4). Hence, a favourable estimate holds in the narrow case.

The broad case. Suppose the broad case holds. By definition, condition i) from the narrow fails. Consequently, there exists some $1 \leqslant j_0 \leqslant m$ and an interval $I_{j_0}^{m,**} \in \mathfrak{I}(\ell_m\ell)$ satisfying

$$I_{j_0}^{m,**} \subseteq I_{j_0}^{m-1}, \qquad \operatorname{dist}(I_j^{m,**}, I_j^{m,*}) \geqslant \ell_m\ell \qquad \text{and} \qquad |F_{I_{j_0}^{m,*}}(x)| \leqslant \Big(\frac{\ell_{m-1}}{\ell_m}\Big)|F_{I_{j_0}^{m,**}}(x)|.$$

On the other hand, condition ii) from the narrow case also fails and, consequently,

$$\max_{1\leqslant j\leqslant m} |F_{I_j^{m,*}}(x)| \leqslant \Big(\frac{\ell_{m-1}}{\ell_m}\Big)^m |F_{I_{j_0}^{m,*}}(x)| \leqslant \Big(\frac{\ell_{m-1}}{\ell_m}\Big)^{m+1}|F_{I_{j_0}^{m,**}}(x)|.$$

Thus, for each $1 \leqslant j \leqslant m$, it follows that

$$|F_{I_j^{m,*}}(x)|^{1/m} \leqslant \Big(\frac{\ell_{m-1}}{\ell_m}\Big)^{1/m} |F_{I_j^{m,*}}(x)|^{1/(m+1)}|F_{I_{j_0}^{m,**}}(x)|^{1/m(m+1)}.$$

Finally, taking the product of the above estimate over all j, one deduces that

$$\begin{aligned}\prod_{j=1}^m |F_{I_j^{m-1}}(x)|^{1/m} &\leqslant \Big(\frac{\ell_{m-1}}{\ell_m}\Big)\prod_{j=1}^m |F_{I_j^{m,*}}(x)|^{1/m}\\ &\leqslant \Big(\frac{\ell_{m-1}}{\ell_m}\Big)^2\Big(\prod_{j=1}^m |F_{I_j^{m,*}}(x)|^{1/(m+1)}\Big)|F_{I_{j_0}^{m,**}}(x)|^{1/(m+1)}\\ &\leqslant \ell_m^{-2}\,\pi_J^{m+1}(F)(x),\end{aligned}$$

where in the last inequality we use the separation condition. Hence, in the broad case a favourable estimate also holds. □

By repeated application of the claim and the relation $\ell_1 \geqslant \cdots \geqslant \ell_{k-1}$,

$$|F_J(x)| \leqslant 4\sum_{m=1}^{k-1} \ell_{m-1}^{-2(m-1)} \max_{I^m\in\mathfrak{I}(J;\ell_m\ell)} |F_{I^m}(x)| + \ell_{k-1}^{-2(k-1)}\pi_J^k(F)(x)$$

for μ-almost every $x \in X$. Bounding all the maxima in the above display by the corresponding ℓ^p expressions and integrating over $x \in X$, one deduces that

$$\begin{aligned}\|F_J\|_{L^p(X)} \leqslant 4\sum_{m=1}^{k-1}\ell_{m-1}^{-2(m-1)}&\Big(\sum_{I^m\in\mathfrak{I}(J;\ell_m\ell)}\|F_{I^m}\|_{L^p(X)}^p\Big)^{1/p}\\ &+\ell_{k-1}^{-2(k-1)}\Big(\sum_{\vec{I}\in\mathfrak{I}^k_{\mathrm{sep}}(J;\ell_{k-1}\ell)}\Big\|\prod_{j=1}^k|F_{I_j}|^{1/k}\Big\|_{L^p(X)}^p\Big)^{1/p}\end{aligned}$$

Finally, taking a ℓ^p sum over J of both sides of the above inequality and applying the triangle inequality concludes the proof. □

Appendix B. A pointwise square function inequality

Here we provide the simple proof of Lemma 10.1, which is a slight extension of an argument due to Rubio de Francia [20]. Given $G\colon \mathbb{Z}^m \to \mathbb{R}^n$ define

$$|||G||| := \sup_{\nu_2\in\mathbb{Z}^m}\sum_{\nu_1\in\mathbb{Z}^m} e^{-|G(\nu_1)-G(\nu_2)|/2}.$$

By rescaling and a simple limiting argument, Lemma 10.1 is a consequence of the following pointwise bound.

Lemma B.1. *Let $\psi \in \mathscr{S}(\widehat{\mathbb{R}}^n)$ and $G\colon \mathbb{Z}^m \to \mathbb{R}^n$. For all M, $N \in \mathbb{N}$ the pointwise inequality*

$$\sum_{\nu \in \mathbb{Z}^m \cap [-M,M]^m} \big|\psi\big(D - G(\nu)\big) f(x)\big|^2 \lesssim_{\psi,N} |||G||| \int_{\mathbb{R}^n} |f(x-y)|^2 (1+|y|)^{-N}\, \mathrm{d}y$$

holds for all $f \in \mathscr{S}(\mathbb{R}^n)$, with an implied constant independent of M.

Proof. Let $a = (a_\nu)_{\nu \in \mathbb{Z}^m}$ be a sequence supported in $\mathbb{Z}^m \cap [-M,M]^m$ satisfying $\|a\|_{\ell^2} = 1$. Consider the function

$$\sum_{\nu \in \mathbb{Z}^m} a_\nu \psi\big(D - G(\nu)\big) f(x) = \mathcal{K} * f(x)$$

where the kernel $\mathcal{K}$ is given by

$$\mathcal{K}(x) := \frac{1}{(2\pi)^n} \int_{\widehat{\mathbb{R}}^n} e^{i\langle x, \xi\rangle} \sum_{\nu \in \mathbb{Z}^m} a_\nu \psi\big(\xi - G(\nu)\big)\, \mathrm{d}\xi = \Big[\sum_{\nu \in \mathbb{Z}^m} a_\nu e^{i\langle x, G(\nu)\rangle}\Big] \check{\psi}(x).$$

By duality, it suffices to show

$$|\mathcal{K} * f(x)|^2 \leqslant |||G||| \int_{\mathbb{R}^n} |f(x-y)|^2 (1+|y|)^{-N}\, \mathrm{d}y.$$

Applying the Cauchy–Schwarz inequality,

$$|\mathcal{K} * f(x)|^2 \leqslant \int_{\mathbb{R}^n} \Big| \sum_{\nu \in \mathbb{Z}^m} a_\nu e^{i\langle y, G(\nu)\rangle} \Big|^2 |\check{\psi}(y)|\, \mathrm{d}y \int_{\mathbb{R}^n} |f(x-y)|^2 |\check{\psi}(y)|\, \mathrm{d}y$$

and so, in view of the rapid decay of $\check{\psi}$, the problem is further reduced to showing

$$\int_{\mathbb{R}^n} \Big| \sum_{\nu \in \mathbb{Z}^m} a_\nu e^{i\langle y, G(\nu)\rangle} \Big|^2 |\check{\psi}(y)|\, \mathrm{d}y \lesssim |||G|||.$$

Since $\psi \in \mathscr{S}(\widehat{\mathbb{R}}^n)$ we have $|\check{\psi}(y)| \lesssim \phi(y)$ where $\phi(z) := (1+z^2)^{-n-1}$. Consider $\phi(z)$ for $|\operatorname{Im}(z)| \leqslant 1/2$ and observe that, by contour integration, $|\widehat{\phi}(\xi)| \lesssim e^{-|\xi|/2}$ for $\xi \in \widehat{\mathbb{R}}^n$. Hence

$$\begin{aligned}\int_{\mathbb{R}^n} \Big| \sum_{\nu \in \mathbb{Z}^m} a_\nu e^{i\langle y, G(\nu)\rangle} \Big|^2 |\check{\psi}(y)|\, \mathrm{d}y &\lesssim \sum_{\nu_1, \nu_2 \in \mathbb{Z}^m} \overline{a_{\nu_1}} a_{\nu_2} \widehat{\phi}\big(G(\nu_1) - G(\nu_2)\big) \\ &\lesssim_\psi \sum_{\nu_1, \nu_2 \in \mathbb{Z}^m} |a_{\nu_1}||a_{\nu_2}| e^{-|G(\nu_1) - G(\nu_2)|/2}.\end{aligned}$$

The right-hand side of the above inequality is then bounded by $|||G|||$ via the Cauchy–Schwarz inequality and the Schur test, as $\|a\|_{\ell^2} = 1$. ☐

Appendix C. Derivative bounds for implicitly defined functions

Let Ω, $I \subseteq \mathbb{R}$ be open intervals and $G\colon \Omega \times I \to \mathbb{C}$ a C^∞ mapping. Suppose $\partial_y G(x,y)$ is non-vanishing on $\Omega \times I$ and $y\colon \Omega \to I$ is a C^∞ mapping such that

$$G(x, y(x)) = 0 \qquad \text{for all } x \in \Omega.$$

Lemma C.1. *Let $G\colon \Omega \times I \to \mathbb{C}$ and $y\colon \Omega \to I$ be as above and suppose $A, M_1, M_2 > 0$ are constants such that*

$$\begin{cases} \big|(\partial_y G)(x, y(x))\big| \geqslant A M_2, & \\ \big|(\partial_x^{\alpha_1} \partial_y^{\alpha_2} G)(x, y(x))\big| \lesssim_\alpha A M_1^{\alpha_1} M_2^{\alpha_2} & \text{for all } \alpha \in \mathbb{N}_0^2 \backslash \{0\}. \end{cases} \tag{C.1}$$

Then the function y satisfies

$$|y^{(j)}(x)| \lesssim_j M_1^j M_2^{-1} \qquad \textit{for all } j \in \mathbb{N}. \tag{C.2}$$

Consequently, for all C^∞ functions $H\colon \Omega \times I \to \mathbb{C}$ for which there exists some constant $B > 0$ such that

$$|(\partial_x^{\alpha_1}\partial_y^{\alpha_2} H)(x, y(x))| \lesssim_N BM_1^{\alpha_1}M_2^{\alpha_2} \qquad \textit{for all } \alpha \in \mathbb{N}_0^2\backslash\{0\}, \tag{C.3}$$

one has

$$\Big|\frac{\mathrm{d}^N}{\mathrm{d}x^N} H(x, y(x))\Big| \lesssim_N BM_1^N \qquad \textit{for all } N \in \mathbb{N}. \tag{C.4}$$

Before giving the proof of Lemma C.1, we make some preliminary observations. A simple induction argument shows that there exists a sequence of coefficients $(C_{\alpha,d})_{d\in\mathbb{N}_0^j}$, depending only on j and α, such that for all C^∞ functions $H\colon \Omega \times I \to \mathbb{C}$ the identity

$$\frac{\mathrm{d}^j}{\mathrm{d}x^j} H(x, y(x)) = \sum_{\substack{\alpha\in\mathbb{N}_0^2\backslash\{0\}\\ \alpha_1,\alpha_2\leqslant j}} (\partial_x^{\alpha_1}\partial_y^{\alpha_2} H)(x, y(x)) \sum_{\substack{d_1+\dots+jd_j=j-\alpha_1\\ d_1+\dots+d_j=\alpha_2}} C_{\alpha,d} \prod_{i=1}^{j} y^{(i)}(x)^{d_i} \tag{C.5}$$

holds. The precise values of the $C_{\alpha,d}$ are given by the multivariate Faà di Bruno formula: see [17, Theorem 4.2]. Similarly, for $1 \leqslant k \leqslant |\alpha|$ there exists a sequence of coefficients $(C_{k,e})_{e\in\mathcal{E}(\alpha,k)}$, depending only on α, such that

$$\partial_x^{\alpha_1}\partial_y^{\alpha_2}\big[(\partial_y G)(x,y)^{-1}\big] = \sum_{k=1}^{|\alpha|} (\partial_y G)(x,y)^{-k-1} \sum_{e\in\mathcal{E}(\alpha,k)} C_{k,e} \prod_{\beta\preceq\alpha} (\partial_x^{\beta_1}\partial_y^{\beta_2+1}G)(x,y)^{e_\beta} \tag{C.6}$$

where

$$\mathcal{E}(\alpha,k) := \Big\{e = (e_\beta)_{\beta\preceq\alpha} : e_\beta \in \mathbb{N}_0 \text{ for all } \beta \preceq \alpha \text{ and } \sum_{\beta\preceq\alpha} \beta_\ell \cdot e_\beta = \alpha_\ell \text{ for } \ell = 1,2,\ \sum_{\beta\preceq\alpha} e_\beta = k\Big\}$$

and the notation $\beta \preceq \alpha$ refers to those $\beta \in \mathbb{N}_0^2\backslash\{0\}$ which satisfy $\beta_\ell \leqslant \alpha_\ell$ for $\ell = 1, 2$. Once again, the precise values of the $C_{k,e}$ are given by the multivariate Faà di Bruno formula.

Both identities (C.5) and (C.6) play a rôle in the proof of Lemma C.1.

Proof. By scaling, it suffices to show the case $A = 1$. The proof of (C.2) proceeds by (strong) induction on j. By implicit differentiation,

$$y'(x) = Q(x, y(x)) \quad \text{where} \quad Q(x,y) := -(\partial_x G)(x,y) \cdot (\partial_y G)(x,y)^{-1} \quad \text{for } (x,y) \in \Omega \times I. \tag{C.7}$$

Thus, the $j = 1$ case is an immediate consequence of this identity together with the hypothesised bounds (C.1). Now let $j \geqslant 1$ and suppose $|y^{(i)}(x)| \lesssim_i M_1^i M_2^{-1}$ holds for all $1 \leqslant i \leqslant j$.

To bound the higher order derivative $y^{(j+1)}$ we make use of the differential identity (C.5), taking $H := Q$. In particular, (C.5) together with (C.7) directly imply that

$$y^{(j+1)}(x) = \sum_{\substack{\alpha\in\mathbb{N}_0^2\backslash\{0\}\\ \alpha_1,\alpha_2\leqslant j}} (\partial_x^{\alpha_1}\partial_y^{\alpha_2} Q)(x, y(x)) \sum_{\substack{d_1+\dots+jd_j=j-\alpha_1\\ d_1+\dots+d_j=\alpha_2}} C_{\alpha,d} \prod_{i=1}^{j} y^{(i)}(x)^{d_i}. \tag{C.8}$$

The bound (C.2) is now reduced to showing

$$|(\partial_x^{\alpha_1}\partial_y^{\alpha_2} Q)(x, y(x))| \lesssim_\alpha M_1 M_2^{-1} M_1^{\alpha_1} M_2^{\alpha_2}. \tag{C.9}$$

Indeed, once (C.9) is established, one may use this inequality to bound the derivatives of Q appearing on the right-hand side of (C.8) and the induction hypothesis to bound the $y^{(i)}(x)$ terms. Consequently, one deduces that

$$|y^{(j+1)}(x)| \lesssim_j M_1^{j+1} M_2^{-1}.$$

This closes the induction and completes the proof of (C.2).

Turning to the proof of (C.9), note that (C.6) and the hypothesised bounds (C.1) imply

$$\big|\partial_x^{\alpha_1}\partial_y^{\alpha_2}\big[(\partial_y G)(x,y)^{-1}\big]|_{y=y(x)}\big| \lesssim_\alpha M_2^{-1}M_1^{\alpha_1}M_2^{\alpha_2} \qquad \text{for all } \alpha \in \mathbb{N}_0^2\backslash\{0\}. \tag{C.10}$$

On the other hand, (C.1) immediately implies that

$$\big|\partial_x^{\alpha_1}\partial_y^{\alpha_2}(\partial_x G)(x,y)|_{y=y(x)}\big| \lesssim_\alpha M_1M_1^{\alpha_1}M_2^{\alpha_2} \qquad \text{for all } \alpha \in \mathbb{N}_0^2\backslash\{0\}. \tag{C.11}$$

Combining (C.10) and (C.11) with the Leibniz rule one obtains (C.9).

The bound (C.4) is a simple consequence of (C.2) and (C.3) via the formula (C.5). □

Lemma C.1 immediately implies the following multivariate extension. Let $\Omega \subseteq \mathbb{R}^n$ be an open set, $I \subseteq \mathbb{R}$ an open interval and $G\colon \Omega\times I \to \mathbb{C}$ a C^∞ mapping, for some $N \in \mathbb{N}$. Suppose $\partial_y G(\boldsymbol{x}, y)$ is non-vanishing on $\Omega \times I$ and $y\colon \Omega \to I$ is a C^∞ mapping such that

$$G(\boldsymbol{x}, y(\boldsymbol{x})) = 0 \qquad \text{for all } \boldsymbol{x} \in \Omega.$$

For $\mathbf{e} \in S^{n-1}$ let $\nabla_{\mathbf{e}}$ denote the directional derivative operator with respect to $\boldsymbol{x}$ in the direction of $\mathbf{e}$. Suppose A, M_1, $M_2 > 0$ are constants such that

$$\begin{cases} \big|(\partial_y G)(\boldsymbol{x}, y(\boldsymbol{x}))\big| & \geqslant & AM_2, \\ \big|(\nabla_{\mathbf{e}}^{\alpha_1}\partial_y^{\alpha_2} G)(\boldsymbol{x}, y(\boldsymbol{x}))\big| & \lesssim_N & AM_1^{\alpha_1}M_2^{\alpha_2} \end{cases} \qquad \text{for all } \alpha \in \mathbb{N}_0^2\backslash\{0\} \text{ and all } \boldsymbol{x} \in \Omega. \tag{C.12}$$

Then the function y satisfies

$$|\nabla_{\mathbf{e}}^N y(\boldsymbol{x})| \lesssim_N M_1^N M_2^{-1} \qquad \text{for all } \boldsymbol{x} \in \Omega \text{ and all } N \in \mathbb{N}_0. \tag{C.13}$$

Similarly, (C.4) has a multivariate extension. In particular, suppose, in addition to the above, that $H\colon \Omega \times I \to \mathbb{C}$ a C^∞ mapping and $B > 0$ is a constant such that

$$\big|(\nabla_{\mathbf{e}}^{\alpha_1}\partial_y^{\alpha_2} H)(\boldsymbol{x}, y(\boldsymbol{x}))\big| \lesssim_N BM_1^{\alpha_1}M_2^{\alpha_2} \qquad \text{for all } \alpha \in \mathbb{N}_0^2\backslash\{0\}. \tag{C.14}$$

Then it follows from (C.4) that

$$\big|\nabla_{\mathbf{e}}^N H(\boldsymbol{x}, y(\boldsymbol{x}))\big| \lesssim_N BM_1^N \qquad \text{for all } \boldsymbol{x} \in \Omega \text{ and all } N \in \mathbb{N}. \tag{C.15}$$

For the purposes of this paper, we are interested in the special case where G, $H\colon \mathbb{R}^n \times I \to \mathbb{R}$ are both linear in the $\boldsymbol{x}$ variable. Thus, G and H are of the form

$$G(\boldsymbol{x}, y) = \langle g(y), \boldsymbol{x}\rangle, \quad H(\boldsymbol{x}, y) = \langle h(y), \boldsymbol{x}\rangle,$$

for some C^∞ functions g, $h\colon I \to \mathbb{R}^n$. Furthermore, the conditions in (C.12) can be written as

$$\begin{cases} |\langle g' \circ y(\boldsymbol{x}), \boldsymbol{x}\rangle| & \geqslant & AM_2, \\ |\langle g^{(N)} \circ y(\boldsymbol{x}), \boldsymbol{x}\rangle| & \lesssim_N & AM_2^N, \\ |\langle g^{(N)} \circ y(\boldsymbol{x}), \mathbf{e}\rangle| & \lesssim_N & AM_1M_2^N \end{cases} \qquad \text{for all } N \in \mathbb{N} \text{ and all } \boldsymbol{x} \in \Omega \tag{C.16}$$

and the condition in (C.14) can be written as

$$\begin{cases} |\langle h^{(N)} \circ y(\boldsymbol{x}), \boldsymbol{x}\rangle| & \lesssim_N & BM_2^N \\ |\langle h^{(N)} \circ y(\boldsymbol{x}), \mathbf{e}\rangle| & \lesssim_N & BM_1M_2^N \end{cases} \qquad \text{for all } N \in \mathbb{N} \text{ and all } \boldsymbol{x} \in \Omega. \tag{C.17}$$

Example C.2 (Application to Lemma 8.7). *Let $\gamma \in \mathfrak{G}_3(\delta_0)$, and $\theta_2 : \widehat{\mathbb{R}}^3\backslash\{0\} \to I_0$ satisfying*

$$\langle \gamma'' \circ \theta_2(\xi), \xi\rangle = 0.$$

We apply the previous result with $g = \gamma''$ and $h = \gamma'$. If $B \leqslant A$ the conditions (C.16) and (C.17) read succinctly as

$$\begin{cases} |\langle \gamma^{(3)} \circ \theta_2(\xi), \xi \rangle| & \geqslant & AM_2, \\ |\langle \gamma^{(1+N)} \circ \theta_2(\xi), \xi \rangle| & \lesssim_N & BM_2^N, \\ |\langle \gamma^{(1+N)} \circ \theta_2(\xi), \mathbf{e} \rangle| & \lesssim_N & BM_1 M_2^N \end{cases} \qquad \text{for all } N \in \mathbb{N} \text{ and all } \xi \in \Omega \subset \hat{\mathbb{R}}^3 \backslash \{0\},$$

which imply

$$|\nabla_{\boldsymbol{e}}^N \theta_2(\xi)| \lesssim_N M_1^N M_2^{-1} \qquad \text{and} \qquad |\nabla_{\boldsymbol{e}}^N \langle \gamma' \circ \theta_2(\xi), \xi \rangle| \lesssim_N BM_1^N$$

for all $N \in \mathbb{N}$ and all $\xi \in \Omega \subset \hat{\mathbb{R}}^3 \backslash \{0\}$.

The application with respect to $\theta_1^{\pm} : \widehat{\mathbb{R}}^3 \backslash \{0\} \to I_0$ satisfying

$$\langle \gamma' \circ \theta_1^{\pm}(\xi), \xi \rangle = 0$$

is similar, with $g = \gamma'$ (we do not require to take an auxiliary h in this case).

Appendix D. Integration-by-parts

For $a \in C_c^\infty(\mathbb{R})$ supported in an interval $I \subset \mathbb{R}$ and $\phi \in C^\infty(I)$, define the oscillatory integral

$$\mathcal{I}[\phi, a] := \int_{\mathbb{R}} e^{i\phi(s)} a(s)\, \mathrm{d}s.$$

The following lemma is a standard application of integration-by-parts.

Lemma D.1 (Non-stationary phase)**.** *Let $R \geqslant 1$ and ϕ, a be as above. Suppose that for each $j \in \mathbb{N}_0$ there exist constants $C_j \geqslant 1$ such that the following conditions hold on the support of a:*

i) $|\phi'(s)| > 0$,
ii) $|\phi^{(j)}(s)| \leqslant C_j R^{-(j-1)} |\phi'(s)|^j$ for all $j \geqslant 2$,
iii) $|a^{(j)}(s)| \leqslant C_j R^{-j} |\phi'(s)|^j$ for all $j \geqslant 0$.

Then for all $N \in \mathbb{N}_0$ there exists some constant $C(N)$ such that

$$|\mathcal{I}[\phi, a]| \leqslant C(N) \cdot |\operatorname{supp} a| \cdot R^{-N}.$$

Moreover, $C(N)$ depends on $C_1, \dots, C_N$ but is otherwise independent of ϕ and a and, in particular, does not depend on r.

Proof. Taking $D := \phi'(s)^{-1} \partial_s$, repeated integration-by-parts implies that

$$\mathcal{I}[\phi, a] = (-i)^{-N} \int_{\mathbb{R}} e^{i\phi(s)} (D^*)^N a(s)\, \mathrm{d}s$$

where D^* is the 'adjoint' differential operator $D^* \colon a \mapsto -\partial_s\big[(\phi')^{-1} \cdot a\big]$. Thus, the proof boils down to establishing a pointwise estimate

$$|(D^*)^N a(s)| \leqslant C(N) \cdot R^{-N}$$

under the hypotheses of the lemma.

It is in fact convenient to prove a more general inequality

$$|\partial_s^j (D^*)^N a(s)| \leqslant C(j, N) \cdot R^{-N-j} \cdot |\phi'(s)|^j, \qquad \text{for all } j, N \in \mathbb{N}_0, \tag{D.1}$$

where the $C(j, N)$ again only the constants C_k for $1 \leqslant k \leqslant N + j$. The inequality (D.1) is amenable to induction on the parameter N. Indeed, if $N = 0$, then (D.1) reduces to hypothesis iii), which establishes the base case.

Assume the inequality (D.1) holds for some $N \geqslant 0$ and all j. By the Leibniz rule,

$$\partial_s^j (D^*)^{N+1} a(s) = \sum_{i=0}^{j+1} \binom{j+1}{i} \big[\partial_s^i (\phi')^{-1}\big](s) \cdot \big[\partial^{j+1-i} (D^*)^N a\big](s). \tag{D.2}$$

Using the induction hypothesis, one may immediately bound

$$\big|\big[\partial^{j+1-i}(D^*)^N a\big](s)\big| \leqslant C(j+1-i,N)\cdot R^{-N-1-j+i}\cdot|\phi'(s)|^{j+1-i}. \tag{D.3}$$

On the other hand, an induction argument shows that there exists a polynomial $\wp \in \mathbb{R}[X_0,\dots,X_i]$, with coefficients depending only on i, with the following properties:

a) $\wp$ is a linear combination of monomials $X_0^{\alpha_0}\cdots X_i^{\alpha_i}$ for multi-indices $(\alpha_0,\dots,\alpha_i)$ satisfying

$$0\cdot\alpha_0 + 1\cdot\alpha_1 + \cdots + i\cdot\alpha_i = \alpha_0+\alpha_1+\cdots+\alpha_i = i.$$

b) The identity

$$\big[\partial_s^K(\phi')^{-1}\big](s) = \frac{\wp\big(\phi'(s),\dots,\phi^{(i+1)}(s)\big)}{\phi'(s)^{i+1}} \qquad \text{holds for all } s\in I.$$

If $(\alpha_0,\dots,\alpha_i)$ is a monomial satisfying a), then hypothesis ii) of the lemma implies that

$$\prod_{k=0}^{i}|\phi^{(k+1)}(s)|^{\alpha_k} \lesssim R^{-i}\cdot|\phi'(s)|^{2i},$$

where the implied constant is here allowed to depend on the C_k for $1\leqslant k\leqslant i+1$. Consequently, from the formula in b) above one deduces that

$$|\big[\partial_s^i(\phi')^{-1}\big](s)| \lesssim R^{-i}\cdot|\phi'(s)|^{i-1}. \tag{D.4}$$

Substituting the bounds (D.3) and (D.4) into (D.2), the induction now closes provided $C(j,N)$ is appropriately defined. ☐

References

[1] David Beltran, Shaoming Guo, Jonathan Hickman, and Andreas Seeger, *Sobolev improving for averages over curves in* $\mathbb{R}^4$, Adv. Math. **393** (2021), Paper No. 108089, 85. MR 4340226

[2] David Beltran, Jonathan Hickman, and Christopher D. Sogge, *Sharp local smoothing estimates for Fourier integral operators* , Geometric aspects of harmonic analysis, Springer INdAM Ser., vol. 45, Springer, Cham, [2021] ©2021, pp. 29–105. MR 4390222

[3] Jonathan Bennett, Neal Bez, Taryn C. Flock, and Sanghyuk Lee, *Stability of the Brascamp-Lieb constant and applications* , Amer. J. Math. **140** (2018), no. 2, 543–569. MR 3783217

[4] Jonathan Bennett, Anthony Carbery, Michael Christ, and Terence Tao, *The Brascamp-Lieb inequalities: finiteness, structure and extremals* , Geom. Funct. Anal. **17** (2008), no. 5, 1343–1415. MR 2377493

[5] Jonathan Bennett, Anthony Carbery, and Terence Tao, *On the multilinear restriction and Kakeya conjectures* , Acta Math. **196** (2006), no. 2, 261–302. MR 2275834

[6] J. Bourgain, *Averages in the plane over convex curves and maximal operators* , J. Analyse Math. **47** (1986), 69–85. MR 874045

[7] Jean Bourgain, *Estimations de certaines fonctions maximales* , C. R. Acad. Sci. Paris Sér. I Math. **301** (1985), no. 10, 499–502. MR 812567

[8] Jean Bourgain and Ciprian Demeter, *The proof of the* l^2 *decoupling conjecture* , Ann. of Math. (2) **182** (2015), no. 1, 351–389. MR 3374964

[9] Jean Bourgain and Larry Guth, *Bounds on oscillatory integral operators based on multilinear estimates* , Geom. Funct. Anal. **21** (2011), no. 6, 1239–1295. MR 2860188

[10] Anthony Carbery and Andreas Seeger, *Homogeneous Fourier multipliers of Marcinkiewicz type* , Ark. Mat. **33** (1995), no. 1, 45–80. MR 1340270

[11] Antonio Córdoba, *Geometric Fourier analysis* , Ann. Inst. Fourier (Grenoble) **32** (1982), no. 3, vii, 215–226. MR 688026

[12] Philip T. Gressman, Shaoming Guo, Lillian B. Pierce, Joris Roos, and Po-Lam Yung, *Reversing a philosophy: from counting to square functions and decoupling* , J. Geom. Anal. **31** (2021), no. 7, 7075–7095. MR 4289255

[13] Larry Guth, Hong Wang, and Ruixiang Zhang, *A sharp square function estimate for the cone in* $\mathbf{R}^3$, Ann. of Math. (2) **192** (2020), no. 2, 551–581. MR 4151084

[14] Seheon Ham and Sanghyuk Lee, *Restriction estimates for space curves with respect to general measures* , Adv. Math. **254** (2014), 251–279. MR 3161099

[15] Jonathan Hickman, *Uniform L^p_x–$L^q_{x,r}$ improving for dilated averages over polynomial curves* , J. Funct. Anal. **270** (2016), no. 2, 560–608. MR 3425895

[16] Sanghyuk Lee and Ana Vargas, *On the cone multiplier in $\mathbb{R}^3$* , J. Funct. Anal. **263** (2012), no. 4, 925–940. MR 2927399

[17] Roy B. Leipnik and Charles E. M. Pearce, *The multivariate Faà di Bruno formula and multivariate Taylor expansions with explicit integral remainder term* , ANZIAM J. **48** (2007), no. 3, 327–341. MR 2326318

[18] Gerd Mockenhaupt, Andreas Seeger, and Christopher D. Sogge, *Wave front sets, local smoothing and Bourgain's circular maximal theorem* , Ann. of Math. (2) **136** (1992), no. 1, 207–218. MR 1173929

[19] Malabika Pramanik and Andreas Seeger, *L^p regularity of averages over curves and bounds for associated maximal operators* , Amer. J. Math. **129** (2007), no. 1, 61–103. MR 2288738

[20] José L. Rubio de Francia, *Estimates for some square functions of Littlewood-Paley type* , Publ. Sec. Mat. Univ. Autònoma Barcelona **27** (1983), no. 2, 81–108. MR 765844

[21] Christopher D. Sogge, *Propagation of singularities and maximal functions in the plane* , Invent. Math. **104** (1991), no. 2, 349–376. MR 1098614

[22] Elias M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals* , Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993, With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. MR 1232192

[23] Terence Tao, *Sharp bounds for multilinear curved Kakeya, restriction and oscillatory integral estimates away from the endpoint* , Mathematika **66** (2020), no. 2, 517–576. MR 4130338

[24] T. Wolff, *Local smoothing type estimates on L^p for large p* , Geom. Funct. Anal. **10** (2000), no. 5, 1237–1288. MR 1800068

[25] Ruixiang Zhang, *The endpoint perturbed Brascamp-Lieb inequalities with examples* , Anal. PDE **11** (2018), no. 3, 555–581. MR 3738255

David Beltran: Department of Mathematics, University of Wisconsin, 480 Lincoln Drive, Madison, WI, 53706, USA.

Current address: Departament d'Anàlisi Matemàtica, Universitat de València, Dr. Moliner 50, 46100 Burjassot, Spain

E-mail address: david.beltran@uv.es

Shaoming Guo: Department of Mathematics, University of Wisconsin, 480 Lincoln Drive, Madison, WI, 53706, USA.

E-mail address: shaomingguo@math.wisc.edu

Jonathan Hickman: School of Mathematics, James Clerk Maxwell Building, The King's Buildings, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, UK.

E-mail address: jonathan.hickman@ed.ac.uk

Andreas Seeger: Department of Mathematics, University of Wisconsin, 480 Lincoln Drive, Madison, WI, 53706, USA.

E-mail address: seeger@math.wisc.edu