A Normed Fractional Chain Rule Counterpart To The Faà di Bruno Identity

SEAN DOUGLAS

Abstract. In this paper, we establish a normed fractional Faà di Bruno inequality within the framework of Lebesgue spaces. This extends the classical fractional chain rule from the range $0 < s < 1$ to an arbitrarily large value of s.

1. Introduction

The Faà di Bruno formula provides an identity for the higher-order derivative of a composition. In our goal to establish a normed fractional version of this identity, the specific coefficients are not of interest. Therefore, we seek a suitable fractional analogue of the following expression

$$\left|\frac{d^m}{dx^m}F(u)\right| \lesssim_m \sum_{i=1}^{m} |F^{(i)}(u)| \sum_{\kappa_1^i+\cdots+\kappa_i^i=m} |u^{(\kappa_1^i)}|\cdots|u^{(\kappa_i^i)}|. \tag{1.1}$$

This paper builds on the fractional chain rule (abbreviated by FCR) initially developed by Christ and Weinstein [1] in the study of generalized Korteweg-de Vries (gKdV) equations. Specifically, these authors used fractional chain and Leibniz rules to provide estimates for the solution of a gKdV represented as an integral equation that enabled the study of the long-time behavior of the solution. The FCR of Christ and Weinstein states that if $0 < s < 1$, $p, p_1, p_2 > 0$, $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$, and F' is convex, then

$$\|D^s(F\circ u)\|_{L^p} \lesssim \|F'\circ u\|_{L^{p_1}}\|D^s u\|_{L^{p_2}}.$$

Fractional chain rules are also useful in establishing the well-posedness of wave equations. A more general smoothness condition, similar to that presented in Theorem 1.1, was developed by Kato [7] and Staffilani [10] in their study of the well-posedness of gKdV and Schrödinger equations.

A more recent advancement is the extension of the fractional chain rule to weighted Triebel-Lizorkin spaces, which, in particular, implies a FCR for Lebesgue spaces when the integrability index $p \leqslant 1$ [2]. This extension to $p \leqslant 1$, and the inclusion of weights for the FCR parallels the result of Naibo and Thomson [9] on the related fractional Leibniz rule for weighted Triebel-Lizorkin spaces.

The focus of this paper is on extending the range of the smoothness index $s > 0$. A FCR was extended to the range $0 < s < \min\{2, \rho\}$, $\rho \geqslant 1$ by Ginibre, Ozawa, and Velo for homogeneous Besov spaces, $\dot{B}^s_{p,q}$, with $p > 1$ and q depending on ρ, when the exterior function F behaves like $|x|^{\rho-1}x$ [6]. The key insight was to use the following representation of the Besov norm,

$$\|u\|_{\dot{B}^s_{p,q}} \sim \left(\int_0^\infty \lambda^{-sq-1}\sup_{|y|<\lambda}\|u(\cdot-y)+u(\cdot+y)-2u\|_{L^p}^q\,d\lambda\right)^{\frac{1}{q}}.$$

Date: September 1, 2024.

Under the same conditions on the exterior function F, a similar result was later extended to Lebesgue spaces by Fujiwara [3]. He employed a technique also involving a centered second order finite difference, but in the context of Littlewood-Paley operators,

$$\Delta_j u(x-y) + \Delta_j u(x+y) - 2\Delta_j u(x).$$

The use of finite differences involving Littlewood-Paley operators is an idea further developed in this paper.

As touched upon in [3] one way to establish a fractional Faà di Bruno inequality is by combining the the classical fractional Leibniz rule of Kato and Ponce [8] , with the FCR for $0 < s < 1$. To illustrate, let $2 < s < 3$, $p, p_i^k > 1$ be related in the sense of Hölder, that is $\frac{1}{p} = \sum_{i=0} \frac{1}{p_i^k}$ for each $1 \leqslant k \leqslant 4$, for simplicity suppose all integrals are over $\mathbb{R}$, and assume F and its derivatives satisfy the appropriate smoothness conditions (such as being convex), then

$$\begin{aligned}
\|D^s F(u)\|_{L^p} &\sim \left\| D^{s-2}\big(F''(u)u'u' + F'(u)u''\big)\right\|_{L^p} \\
&\lesssim \|F'''(u)\|_{L^{p_0^1}} \|D^{s-2}u\|_{L^{p_1^1}} \|u'\|_{L^{p_2^1}} \|u'\|_{L^{p_3^1}} \\
&\quad + \|F''(u)\|_{L^{p_0^2}} \|D^{s-1}u\|_{L^{p_1^2}} \|u'\|_{L^{p_2^2}} \\
&\quad + \|F''(u)\|_{L^{p_0^3}} \|D^{s-2}u\|_{L^{p_1^3}} \|u''\|_{L^{p_2^3}} \\
&\quad + \|F'(u)\|_{L^{p_0^4}} \|D^{s}u\|_{L^{p_1^4}}
\end{aligned} \tag{1.2}$$

by a direct application of the fractional Leibniz rule, FCR, and the fact that $\frac{|\cdot|^2}{(\cdot)^2}$, $\frac{(\cdot)}{|\cdot|}$ are L^p Fourier multipliers for $p > 1$. This approach, however, has two drawbacks. The first is that the appearance of integer order derivatives on u is unavoidable. For example, consider the first summand in (1.2); ideally, we would like to replace it with something like

$$\|F'''(u)\|_{L^{p_0^1}} \|D^{s_1}u\|_{L^{p_1^1}} \|D^{s_2}u\|_{L^{p_2^1}} \|D^{s_3}u\|_{L^{p_3^1}},$$

where $0 < s_1, s_2, s_3 < 1$ and $s_1 + s_2 + s_3 = s$.

The second drawback is there are two separate summands associated with $F''(u)$. Ideally, we would like to replace the expression $\|F''(u)\|_{L^{p_0^2}} \|D^{s-1}u\|_{L^{p_1^2}} \|u'\|_{L^{p_2^2}} + \|F''(u)\|_{L^{p_0^3}} \|D^{s-2}u\|_{L^{p_1^3}} \|u''\|_{L^{p_2^3}}$ with a single summand such as

$$\|F''(u)\|_{L^{p_0^2}} \|D^{\dot{s}_1}u\|_{L^{p_1^2}} \|D^{\dot{s}_2}u\|_{L^{p_2^2}},$$

where $0 < \dot{s}_1 < 1$, $1 < \dot{s}_2 < 2$, and $\dot{s}_1 + \dot{s}_2 = s$. Both of these issues are addressed in Theorem 1.1.

We denote the space of Schwartz functions, which are smooth functions that decay rapidly on $\mathbb{R}^n$, by $\mathcal{S}$. The dual space of $\mathcal{S}$, consisting of tempered distributions, is denoted by $\mathcal{S}'$. Since we are using the homogeneous fractional derivative we need the space $\mathcal{S}_0$, the subspace of Schwartz functions with all moments vanishing; specifically, $f \in \mathcal{S}_0$ if and only if $f \in \mathcal{S}$ and $\int_{\mathbb{R}^n} x^\alpha f(x)\,dx = 0$ for every multi-index α. The dual of $\mathcal{S}_0$ is $\mathcal{S}'/\mathcal{P}$, the space of tempered distributions modulo polynomials, where two distributions are considered equivalent if their difference is a polynomial. It is known that $\mathcal{S}_0$ is dense in $\mathcal{S}'/\mathcal{P}$, and $\mathcal{S}$ is dense in $\mathcal{S}'$. For further details on tempered distributions, refer to [[4], Chap. 1].

For any $f \in L^1(\mathbb{R}^n)$, the Fourier transform and its inverse are respectively defined by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(y) e^{-2\pi i y\cdot\xi}\,dy \quad \text{and} \quad \mathscr{F}^{-1}(f)(\xi) = \check{f}(\xi) = \int_{\mathbb{R}^n} f(y) e^{2\pi i y\cdot\xi}\,dy.$$

The homogeneous fractional Laplacian of order $s > 0$ for $u \in \mathcal{S}'/\mathcal{P}$, denoted by $D^s u \in \mathcal{S}'/\mathcal{P}$, are defined as

$$\langle D^s u, \varphi\rangle := \langle u, (|\cdot|^s \check{\varphi})\widehat{\ }\rangle, \quad \text{for } \varphi \in \mathcal{S}_0,$$

note that $(|\cdot|^s \check{\varphi})\widehat{\ } \in \mathcal{S}_0$. A subtle point here is that by defining the homogeneous fractional derivative by its action on $\mathcal{S}_0$, rather than $\mathcal{S}$, we avoid the need to potentially mod out by a polynomial. We now state our main result.

Theorem 1.1. *Let $s > 0$ and $\mathfrak{m}$ be the smallest integer greater than s. Let $p, p_i^k > 1$ be related in the sense of Hölder, that is*

$$\frac{1}{p} = \sum_{i=0}^{k} \frac{1}{p_i^k}$$

for every $1 \leqslant k \leqslant \mathfrak{m}$. Let u be a tempered locally integrable function on $\mathbb{R}^n$. Let $F : \mathbb{C} \to \mathbb{C}$, $G : \mathbb{C} \to [0, \infty)$ and suppose $F \circ u$ is a tempered locally integrable function. Furthermore, suppose that for $x, y \in \mathbb{C}$ that

$$|F^{(\mathfrak{m})}(tx + (1-t)y)| \leqslant \mu(t)\big(G(x) + G(y)\big)$$

for all $t \in [0,1]$, and $\mu \in L^1([0,1])$. Let $C = C_{s,n,\mu,p_1^1,\ldots,p_{\mathfrak{m}}^{\mathfrak{m}}}$. Suppose the right hand side of (1.3) is finite, then $D^s(F(u))$ defined by its action on $\mathcal{S}_0$ coincides with a L^p function and satisfies the following estimate:

$$\begin{aligned}\|D^s(F(u))\|_{L^p} \leqslant C\Big(&\|G(u)\|_{L^{p_0^{\mathfrak{m}}}} \left\|D^{\dot{s}_1} u\right\|_{L^{p_1^{\mathfrak{m}}}} \cdots \left\|D^{\dot{s}_{\mathfrak{m}}} u\right\|_{L^{p_{\mathfrak{m}}^{\mathfrak{m}}}} \\ &+ \sum_{i=1}^{\mathfrak{m}-1} \left\|F^{(i)}(u)\right\|_{L^{p_0^i}} \sum_{\kappa_1^i + \cdots + \kappa_i^i = \mathfrak{m}} \left\|D^{s_{\kappa_1^i}} u\right\|_{L^{p_1^i}} \cdots \left\|D^{s_{\kappa_i^i}} u\right\|_{L^{p_i^i}}\Big)\end{aligned} \tag{1.3}$$

where $0 < \dot{s}_r < 1$, $\sum_{r=1}^{\mathfrak{m}} \dot{s}_r = s$, $\kappa_r^i \in \mathbb{N}$, $\kappa_r^i - 1 < s_{\kappa_r^i} < \kappa_r^i$ for $1 \leqslant r \leqslant i$ and $\sum_{r=1}^{i} s_{\kappa_r^i} = s$.

The key ingredients in the proof of Theorem (1.1) are Littlewood-Paley theory, elementary combinatorial arguments and natural extensions of the techniques used to prove the classical fractional chain rule. More specifically, we use forward finite differences to obtain the appropriate derivatives. The main challenge lies in balancing the use of finite differences while simultaneously being able to bound by the Hardy-Littlewood maximal operator. To illustrate with a simplified example, after a decomposition, we obtain sums over l_1, l_2, and j of the form:

$$\int 2^{jn} |\psi(2^j y)| |\Delta_{l_1} u(x + k_1 y)| |\Delta_{l_2} u(x + k_2 y)| \, dy.$$

With careful management of the finite differences, we can set $k_1 = k_2$ in the expression above, which then transforms the above integral into convolution with an L^1 dilate of a Schwartz function. This makes it bounded by the maximal operator. Our decomposition in Section 4 and the lemmas involving simple combinatorial arguments ensure that this is always the case.

2. Notation

All functions are assumed to be defined on $\mathbb{R}^n$, so we omit $\mathbb{R}^n$ in the notation for simplicity. $B(0,k)$ denotes a ball centered at the origin with radius k. We denote by M the uncentered Hardy-Littlewood maximal function with respect to cubes. The $L^p(\ell_2)$ norm is defined by

$$\|f_j\|_{L^p(\ell_2)} = \Big\|\Big(\sum_{j\in\mathbb{Z}} |f_j|^2\Big)^{\frac{1}{2}}\Big\|_{L^p} = \Big(\int_{\mathbb{R}^n} \Big(\sum_{j\in\mathbb{Z}} |f_j(y)|^2\Big)^{\frac{p}{2}} dy\Big)^{\frac{1}{p}}$$

for $p > 1$.

For real numbers A, B we use $A \lesssim B$ to mean $A \leqslant CB$ for some positive constant C. The dependence of C on other variables and parameters will often be suppressed and clear from the context. The notation $A \sim B$ means $A \lesssim B$ and $B \lesssim A$.

Let $\widehat{\phi}(\xi)$ be a positive radially decreasing $C^\infty(\mathbb{R}^n)$ function on $\mathbb{R}^n$ supported in $B(0,2)$ and equal to one on the unit ball. Let $\widehat{\psi}(\xi) = \widehat{\phi}(\xi) - \widehat{\phi}(2\xi)$, which is non-negative and supported in the annulus $\frac{1}{2} \leqslant |\xi| \leqslant 2$. Note that ψ is radial and has average value zero. The Littlewood-Paley operator Δ_j is defined to be convolution with $2^{jn}\psi(2^j\cdot)$ for $j \in \mathbb{Z}$. For ease of notation denote

$$\psi_j := 2^{jn}\psi(2^j\cdot).$$

If a Littlewood-Paley operator is defined with a Shwartz function, φ, different than ψ then it will be denoted by Δ_j^φ.

Let $k, m \in \mathbb{N}_0$ and $l \in \mathbb{Z}$, and $u : \mathbb{R}^n \to \mathbb{C}$. Define $a_k(x,y) := u(x+ky)$ and

$$a_{k,l}(x,y) := \Delta_l u(x+ky).$$

We will often omit the variables x and y for simplicity. Note that $a_0 = a_0(x,y) = u(x)$ is independent of y. Let $k \geqslant m$, a mth order forward finite difference is denoted by $b_k^m := \sum_{i=0}^m \binom{m}{i}(-1)^i a_{k-i}$ and

$$b_{k,l}^m := \sum_{i=0}^m \binom{m}{i}(-1)^i a_{k-i,l},$$

which is implicitly a function of x and y.

We will frequently use the Cauchy-Schwarz inequality for multiple factors, that is for $d_{1,j}, d_{2,j}, \dots, d_{m,j} \in \mathbb{R}$

$$\sum_{j\in\mathbb{Z}} \left(\prod_{i=1}^m d_{i,j}\right)^2 \leqslant \prod_{i=1}^m \sum_{j\in\mathbb{Z}} d_{i,j}^2.$$

Additionally, we will use the vector-valued Fefferman-Stein inequality,

$$\|M(f_j)\|_{L^q(\ell_2)} \lesssim \|f_j\|_{L^q(\ell_2)}, \quad \text{for } 1 < q < \infty.$$

3. Preliminaries

For Lemmas 3.1, 3.2 and 3.3 the frequency isolator index l does not play a role, so it will be omitted from the proofs. In fact, the following three lemmas are purely combinatorial in nature, so, for simplicity, we may assume that $\{a_i\}_{i\in\mathbb{N}_0}$ is a finite sequence.

Lemma 3.1. *Let $k, m \in \mathbb{N}_0$, $l \in \mathbb{Z}$ and $k \geqslant m$, then*

$$b_{k,l}^m = b_{k,l}^{m-1} - b_{k-1,l}^{m-1}.$$

Proof. Recall the identity

$$\binom{m+1}{k} = \binom{m}{k} + \binom{m}{k-1} \tag{3.1}$$

for $k > 0$. Observe,

$$\begin{aligned} b_k^{m-1} - b_{k-1}^{m-1} &= \sum_{i=0}^{m-1}\binom{m-1}{i}(-1)^i a_{k-i} - \sum_{i=0}^{m-1}\binom{m-1}{i}(-1)^i a_{k-1-i} \\ &= \sum_{i=0}^{m-1}\binom{m-1}{i}(-1)^i a_{k-i} + \sum_{i=1}^{m}\binom{m-1}{i-1}(-1)^i a_{k-i} \\ &= \binom{m-1}{0} a_k + \binom{m-1}{m-1}(-1)^m a_{k-m} \end{aligned}$$

$$+\sum_{i=1}^{m-1}\left[\binom{m-1}{i}+\binom{m-1}{i-1}\right](-1)^i a_{k-i}$$

$$=a_k+(-1)^m a_{k-m}+\sum_{i=1}^{m-1}\binom{m}{i}(-1)^i a_{k-i} \tag{3.2}$$

$$=\sum_{i=0}^{m}\binom{m}{i}(-1)^i a_{k-i}$$

$$=b_k^m$$

where in (3.2) we applied (3.1). □

Lemma 3.2. *Let $k \in \mathbb{N}$, $l \in \mathbb{Z}$, then*

$$a_{k,l}-a_{0,l}=\sum_{i=0}^{k-1}\binom{k}{k-i}b_{k-i,l}^{k-i}.$$

Proof. Observe,

$$\sum_{i=0}^{k-1}\binom{k}{k-i}b_{k-i}^{k-i}=\sum_{i=0}^{k-1}\sum_{m=0}^{k-i}\binom{k}{k-i}\binom{k-i}{k-i-m}(-1)^m a_{k-i-m}$$

$$=\sum_{i=0}^{k-1}\sum_{m=0}^{k-i}\frac{k!}{i!m!(k-i-m)!}(-1)^m a_{k-i-m}$$

$$=\sum_{i=0}^{k-1}\frac{k!}{i!(k-i)!}(-1)^{k-i}a_0+\sum_{t=0}^{k-1}\sum_{m=0}^{t}\frac{k!}{m!(t-m)!(k-t)!}(-1)^{t-m}a_{k-t} \tag{3.3}$$

$$=\sum_{i=0}^{k-1}\frac{k!}{i!(k-i)!}(-1)^{k-i}a_0+\sum_{t=0}^{k-1}\frac{k!}{t!(k-t)!}a_{k-t}\sum_{m=0}^{t}\frac{t!}{m!(t-m)!}(-1)^{t-m}$$

$$=a_k-a_0 \tag{3.4}$$

where in (3.3) we applied a change of variables, and (3.4) is a result of

$$\sum_{m=0}^{t}\frac{t!}{m!(t-m)!}(-1)^{t-m}=0$$

if $t \neq 0$. This follows from the inherent symmetry in a row of Pascal's triangle. □

The following lemma is the critical observation in Section 6.3 that allows us to attain finite differences of the proper orders.

Lemma 3.3. *Let $m, d, n_1, \cdots, n_d \in \mathbb{N}$, $d < m$ and $n_1 + \cdots + n_d < m$ then*

$$\sum_{i=0}^{m-1}(-1)^i\binom{m}{i}\binom{m-i}{n_1}\cdots\binom{m-i}{n_d}=0,$$

where it is understood that $\binom{a}{b}=0$ if $b>a$.

Proof. We proceed by induction on m,

$$\sum_{i=0}^{m}(-1)^i\binom{m+1}{i}\binom{m+1-i}{n_1}\cdots\binom{m+1-i}{n_d}$$

$$=\binom{m+1}{n_1}\cdots\binom{m+1}{n_d}+\sum_{i=1}^{m}(-1)^i\left[\binom{m}{i}+\binom{m}{i-1}\right]\binom{m+1-i}{n_1}\cdots\binom{m+1-i}{n_d}$$

$$=\sum_{i=0}^{m}(-1)^i\binom{m}{i}\binom{m+1-i}{n_1}\cdots\binom{m+1-i}{n_d}$$

$$+\sum_{i=1}^{m}(-1)^i\binom{m}{i-1}\binom{m-(i-1)}{n_1}\cdots\binom{m-(i-1)}{n_d}$$

$$\begin{aligned}&=\sum_{i=0}^{m}(-1)^i\binom{m}{i}\left[\binom{m-i}{n_1}+\binom{m-i}{n_1-1}\right]\cdots\left[\binom{m-i}{n_d}+\binom{m-i}{n_d-1}\right]\\&\qquad-\sum_{i=0}^{m}(-1)^i\binom{m}{i}\binom{m-i}{n_1}\cdots\binom{m-i}{n_d},\end{aligned}\tag{3.5}$$

which is equal to a sum of terms of the form

$$\sum_{i=0}^{m}(-1)^i\binom{m}{i}\binom{m-i}{\eta_1}\cdots\binom{m-i}{\eta_d}\tag{3.6}$$

where η_r for $1\leqslant r\leqslant d$ is n_r or n_r-1. Note that due to the difference in (3.5) that $\eta_r=n_r-1$ for at least one r between 1 and d; this will allow us to apply the induction hypothesis. It is possible that $\eta_r=0$, so reindex (3.6) to

$$\sum_{i=0}^{m-1}(-1)^i\binom{m}{i}\binom{m-i}{\eta_1}\cdots\binom{m-i}{\eta_q}\tag{3.7}$$

such that $q\leqslant d$ and $\eta_r>0$ for $1\leqslant r\leqslant q$. Observe that

$$\sum_{r=1}^{q}\eta_r\leqslant\Big(\sum_{r=1}^{d}n_r\Big)-1<m.$$

Hence, if $q<m$ then we can apply the induction hypothesis to obtain that (3.7) is zero. Note, it must be the case that $q<m$. To see this suppose $q=m$, then $d=m$. Hence $n_1+\cdots+n_m<m+1$ implies that $n_1=\ldots=n_m=1$, since $n_r>0$ for $1\leqslant r\leqslant d$. It follows $\eta_r=0$ for at least one $1\leqslant r\leqslant m$ implying $q<m$, a contradiction. □

Lemma 3.4. *Let $k\geqslant m$ and $|y|\leqslant 2^{-l}$ then a m order finite difference admits the following bound*

$$|b^m_{k,l}(x,y)|\lesssim 2^{ml}|y|^mM(\Delta_l u)(x).\tag{3.8}$$

Furthermore,

$$|a_{k,l}-a_{0,l}|\lesssim 2^l|y|M(\Delta_l(u))(x).$$

Proof. This proof will be by induction on m. Let ψ^* be a Schwartz function such that its Fourier transform is 1 on $\frac{1}{2}\leqslant|\xi|\leqslant 2$, and is supported in $\frac{1}{4}\leqslant|\xi|\leqslant 4$. First we will show that

$$\begin{aligned}b^m_{k,l}&=-2^{ml}\int_{\mathbb{R}^n}\Delta_l u(z)\sum_{r_1=1}^{n}\cdots\sum_{r_m=1}^{n}\int_0^1\cdots\int_0^1\\&\qquad(\partial_{y_{r_1}}\cdots\partial_{y_{r_m}}\psi^*)_l(x+ky-\sum_{i=1}^{m}t_iy-z)y_{r_1}\cdots y_{r_m}dt_1\cdots dt_mdz.\end{aligned}\tag{3.9}$$

For $m=1$ observe,

$$b^1_{k,l}=a_{k,l}-a_{k-1,l}$$

$$
\begin{aligned}
&= \Delta_l^{\psi*}\Delta_l u(x+ky) - \Delta_l^{\psi*}\Delta_l u(x+(k-1)y)\\
&= \int_{\mathbb{R}^n} \Delta_l u(z)\Big(\psi_l^*(x+ky-z) - \psi_l^*(x+(k-1)y-z)\Big)dz\\
&= -2^l \int_{\mathbb{R}^n} \Delta_l u(z) \int_0^1 \nabla\psi_j^*(x+ky-t_1y-z)\cdot y dt_1 dz \qquad (3.10)\\
&= -2^l \int_{\mathbb{R}^n} \Delta_l u(z) \sum_{r_1=1}^{n}\int_0^1 \partial_{y_{r_1}}\psi_j^*(x+ky-t_1y-z) y_{r_1} dt_1 dz
\end{aligned}
$$

where in (3.10) we applied the Fundamental Theorem of Calculus. Now suppose (3.9) holds. Let $\partial_{\vec{r}}$ denote $\partial_{y_{r_1}}\cdots\partial_{y_{r_m}}$. Using Lemma 3.1 we obtain

$$
\begin{aligned}
b_k^{m+1} =& b_k^m - b_{k-1}^m\\
=& -2^{ml}\int_{\mathbb{R}^n}\Delta_l u(z)\sum_{r_1=1}^{n}\cdots\sum_{r_m=1}^{n}\int_0^1\cdots\int_0^1(\partial_{\vec{r}}\psi)_l(x+ky-\sum_{i=1}^{m}t_iy-z)\\
&\times y_{r_1}\cdots y_{r_m}dt_1\cdots dt_m dz\\
&+2^{ml}\int_{\mathbb{R}^n}\Delta_l u(z)\sum_{r_1=1}^{n}\cdots\sum_{r_m=1}^{n}\int_0^1\cdots\int_0^1(\partial_{\vec{r}}\psi)_l(x+(k-1)y-\sum_{i=1}^{m}t_iy-z)\\
&\times y_{r_1}\cdots y_{r_m}dt_1\cdots dt_m dz\\
=&2^{ml}\int_{\mathbb{R}^n}\Delta_l u(z)\sum_{r_1=1}^{n}\cdots\sum_{r_m=1}^{n}\int_0^1\cdots\int_0^1 y_{r_1}\cdots y_{r_m}\\
&\times\Big[(\partial_{\vec{r}}\psi)_l(x+(k-1)y-\sum_{i=1}^{m}t_iy-z)-(\partial_{\vec{r}}\psi)_l(x+ky-\sum_{i=1}^{m}t_iy-z)\Big]dt_1\cdots dt_m dz\\
=& 2^{ml}\int_{\mathbb{R}^n}\Delta_l u(z)\sum_{r_1=1}^{n}\cdots\sum_{r_m=1}^{n}\int_0^1\cdots\int_0^1 y_{r_1}\cdots y_{r_m}\\
&\times\Big[2^l\int_0^1(\nabla\partial_{\vec{r}}\psi)_l(x+ky-\sum_{i=1}^{m+1}t_iy-z)\cdot -y dt_{m+1}\Big]dt_1\cdots dt_m dz\\
=& -2^{(m+1)l}\int_{\mathbb{R}^n}\Delta_l u(z)\sum_{r_1=1}^{n}\cdots\sum_{r_{m+1}=1}^{n}\int_0^1\cdots\int_0^1(\partial_{\vec{r}}\psi)_l(x+ky-\sum_{i=1}^{m+1}t_iy-z)\\
&\times y_{r_1}\cdots y_{r_{m+1}}dt_1\cdots dt_{m+1}dz
\end{aligned}
$$

which establishes (3.9). To obtain the estimate in (3.8) we will use the following simple inequality for $v_1, v_2 \in \mathbb{R}^n$

$$
\frac{1}{1+|v_2+v_1|} \leqslant \frac{1+|v_1|}{1+|v_2|}. \qquad (3.11)
$$

From (3.9) it follows the absolute value of $b_{k,l}^m$ is bounded by a constant multiple of

$$
\begin{aligned}
&2^{ml}\sum_{r_1=1}^{n}\cdots\sum_{r_m=1}^{n}|y_{r_1}\cdots y_{r_m}|\int_{\mathbb{R}^n}|\Delta_l u(z)|\\
&\qquad\times\int_0^1\cdots\int_0^1\frac{2^{ln}C_{\vec{r},\psi*}}{(1+2^l|x+ky-\sum_{i=1}^{m}t_iy-z|)^N}dt_1\cdots dt_m dz
\end{aligned}
$$

$$\lesssim_{n,m} 2^{ml}|y|^m \int_{\mathbb{R}^n} |\Delta_l u(z)| \int_0^1 \cdots \int_0^1 \frac{2^{ln}(1+2^l|y||k-\sum_{i=1}^m t_i|)^N}{(1+2^l|x-z|)^N} dt_1 \cdots dt_m dz \tag{3.12}$$

$$\lesssim_{k,m} 2^{ml}|y|^m \int_{\mathbb{R}^n} |\Delta_l u(z)| \frac{2^{ln}}{(1+2^l|x-z|)^N} dz \tag{3.13}$$

$$\lesssim 2^{ml}|y|^m M(\Delta_l u)(x) \tag{3.14}$$

where $N > n$, in (3.12) we applied (3.11) and in (3.13) we used the hypothesis that $|y| < 2^{-l}$. This ends the proof of (3.8).

The second statement now follows by a simple telescoping argument,

$$|a_{k,l} - a_{0,l}| = \Big| \sum_{i=0}^{k-1} a_{i+1,l} - a_{i,l} \Big| \leqslant \sum_{i=0}^{k-1} |b^1_{i+1,l}| \lesssim_k 2^l |y| M(\Delta_l(u))(x).$$

□

The inequality in (3.15) for $0 \leqslant s < 1$ is given in [[11], Lemma 4.2], here we provide a straight forward extension to larger s.

Lemma 3.5. *Let $\{d_l\}_{l\in\mathbb{Z}}$ be a sequence of non-negative real numbers, then*

$$\Big(\sum_{j\in\mathbb{Z}} \Big(2^{js} \sum_{l\leqslant j} 2^{m(l-j)} d_l\Big)^2\Big)^{\frac{1}{2}} \lesssim \Big(\sum_{j\in\mathbb{Z}} (2^{js} d_j)^2\Big)^{\frac{1}{2}} \quad \textit{if } 0 \leqslant s < m \tag{3.15}$$

$$\Big(\sum_{j\in\mathbb{Z}} \Big(2^{js} \sum_{j<l} d_l\Big)^2\Big)^{\frac{1}{2}} \lesssim \Big(\sum_{j\in\mathbb{Z}} (2^{js} d_j)^2\Big)^{\frac{1}{2}} \quad \textit{if } 0 \leqslant s. \tag{3.16}$$

Proof. Observe,

$$\begin{aligned}
\Big(\sum_{j\in\mathbb{Z}} \Big(2^{js} \sum_{l\leqslant j} 2^{m(l-j)} d_l\Big)^2\Big)^{\frac{1}{2}} &= \Big(\sum_{j\in\mathbb{Z}} \Big(\sum_{l=0}^{\infty} 2^{js} 2^{-lm} d_{j-l}\Big)^2\Big)^{\frac{1}{2}} \\
&= \Big(\sum_{j\in\mathbb{Z}} \Big(\sum_{l=0}^{\infty} 2^{l(s-m)/2} 2^{l(s-m)/2} 2^{(j-l)s} d_{j-l}\Big)^2\Big)^{\frac{1}{2}} \\
&\lesssim \Big(\sum_{j\in\mathbb{Z}} \sum_{l=0}^{\infty} (2^{l(s-m)/2} 2^{(j-l)s} d_{j-l})^2\Big)^{\frac{1}{2}} \\
&= \Big(\sum_{l=0}^{\infty} 2^{l(s-m)} \sum_{j\in\mathbb{Z}} (2^{(j-l)s} d_{j-l})^2\Big)^{\frac{1}{2}} \\
&\lesssim \Big(\sum_{j\in\mathbb{Z}} (2^{js} d_j)^2\Big)^{\frac{1}{2}}.
\end{aligned}$$

where in the first inequality we applied the Cauchy-Schwarz inequality. The proof of (3.16) is similar and omitted. □

The following is the celebrated Peetre's lemma. For our purposes we only need the below statement for $t = 1$.

Lemma 3.6 ([12], Theorem 2.10). *Let $0 < t < \infty$, and u be a function on $\mathbb{R}^n$ whose distributional Fourier transform satisfies* $\mathrm{supp}(\widehat{u}) \subset B(0,k)$*, then*

$$\sup_{y \in \mathbb{R}^n} \frac{|u(x-y)|}{(1+k|y|)^{\frac{n}{t}}} \leqslant C_{n,t} M_t(u)(x)$$

where the constant is independent of k.

4. Initial Decomposition for Theorem 1.1

Let $s < \mathfrak{m}$ with $\mathfrak{m}$ being the smallest integer bigger than s. Let

$$\widehat{\Psi}(\xi) = \sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i} (-1)^i \widehat{\psi}\big((\mathfrak{m}-i)\xi\big). \tag{4.1}$$

So for example, if $\mathfrak{m} = 3$ then $\widehat{\Psi}(\xi) = \widehat{\psi}(3\xi) - 3\widehat{\psi}(2\xi) + 3\widehat{\psi}(\xi)$. It is easy to see via induction that $\sum_{j \in \mathbb{Z}} \widehat{\Psi}(2^{-j}\xi)$ for $\xi \neq 0$ equals either 1 or -1 depending on if $\mathfrak{m}$ is odd or even. Since Ψ is a Schwartz function whose Fourier transform is supported in an annulus and forms a partition of unity when summed dyadically (multiplying by -1 if necessary), we can apply the lower Littlewood-Paley inequality [[5], Theorem 4.5.6]. It follows that

$$\begin{aligned} \|D^s(F(u))\|_{L^p} &\lesssim \Big\| \Big(\sum_{j \in \mathbb{Z}} |\Delta_j^{\Psi}(D^s(F(u))|^2 \Big)^{\frac{1}{2}} \Big\|_{L^p} \\ &\lesssim \Big\| \Big(\sum_{j \in \mathbb{Z}} |2^{js} \Delta_j^{\Psi}(F(u))|^2 \Big)^{\frac{1}{2}} \Big\|_{L^p} \end{aligned} \tag{4.2}$$

where the last inequality is the lifting property of Triebel-Lizorkin spaces, which is generally expressed as $\|f\|_{\dot{F}^s_{p,q}} \sim \|D^s f\|_{\dot{F}^0_{p,q}}$. Recall that $D^s(F(u))$, as a tempered distribution, is defined by its action on $\varphi \in \mathcal{S}_0$. Therefore, there is no need to potentially mod out by a polynomial when applying the lower Littlewood-Paley inequality.

For simplicity let $h := F \circ u$, then $\Delta_j^{\Psi}(F(u))$ can be express as

$$\Delta_j^{\Psi}(F(u))(x) = \sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i} (-1)^i \mathscr{F}^{-1}\big(\widehat{\psi}\big((\mathfrak{m}-i)2^{-j}\,\cdot\,\big)\widehat{h}\big)(x). \tag{4.3}$$

Consider a summand of the the above sum, namely for a fixed k observe

$$\begin{aligned} \mathscr{F}^{-1}\big(\widehat{\psi}\big(k2^{-j}\,\cdot\,\big)\widehat{h}\big)(x) &= \int_{\mathbb{R}^n} \frac{1}{(k2^{-j})^n} \psi\big(k^{-1}2^j y\big) h(x-y)\,dy \\ &= \int_{\mathbb{R}^n} 2^{jn} \psi(2^j y) h(x+ky)\,dy \end{aligned}$$

where in the last line we used a change and variables and the fact that ψ is radial. It follows that (4.3) can be written as

$$\Delta_j^{\Psi}(F(u))(x) = \sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i} (-1)^i \int_{\mathbb{R}^n} \psi_j(y) F\big(u(x+(\mathfrak{m}-i)y)\big)\,dy. \tag{4.4}$$

Recall that for a fixed k, $a_k = a_k(x,y) = u(x+ky)$, note that a_0 is independent of y. Focusing on the integral within a summand of (4.4) we proceed by repeatedly adding and subtracting expressions to apply the Fundamental Theorem of Calculus (FTC) as follows

$$\int \psi_j(y) F\big(u(x+ky)\big)\,dy = \int \psi_j(y) F\big(u(x+ky)\big)\,dy - F(u(x)) \int \psi_j(y)\,dy$$

$$= \int \psi_j(y) \int_0^1 F'(t_1 a_k + (1-t_1)a_0)(a_k - a_0) dt_1 dy$$

$$= \int \psi_j(y)(a_k - a_0) \int_0^1 F'(a_0 + (a_k - a_0)t_1) dt_1 dy. \tag{4.5}$$

Adding and subtracting $F'(a_0)$ from the inner integral in (4.5) and applying the FTC we obtain

$$\begin{aligned} &F'(a_0) \int \psi_j(y)(a_k - a_0) \\ &+ \int \psi_j(y)(a_k - a_0)^2 \int_0^1 \int_0^1 t_1 F''(a_0 + (a_k - a_0)t_1 t_2) dt_2 dt_1 dy. \end{aligned} \tag{4.6}$$

Now add and subtract $F''(a_0)$ to the inner integral in (4.6) and apply the FTC to obtain

$$\begin{aligned} &\int \psi_j(y)(a_k - a_0)^2 \int_0^1 \int_0^1 t_1 F''(a_0) dt_2 dt_1 dy \\ &+ \int \psi_j(y)(a_k - a_0)^3 \int_0^1 \int_0^1 \int_0^1 t_1^2 t_2 F'''(a_0 + (a_k - a_0)t_1 t_2 t_3) dt_3 dt_2 dt_1 dy \end{aligned} \tag{4.7}$$

Once more, add and subtract $F'''(a_0)$ to the inner integral in (4.7) and apply the FTC to obtain

$$\begin{aligned} &\int \psi_j(y)(a_k - a_0)^3 \int_0^1 \int_0^1 \int_0^1 t_1^2 t_2 F'''(a_0) dt_3 dt_2 dt_1 dy \\ &+ \int \psi_j(y)(a_k - a_0)^4 \int_0^1 \int_0^1 \int_0^1 \int_0^1 t_1^3 t_2^2 t_3 F^{(4)}(a_0 + (a_k - a_0)t_1 t_2 t_3 t_4) dt_4 dt_3 dt_2 dt_1 dy. \end{aligned}$$

Continuing until the $\mathfrak{m}$th derivative is reached it follows

$$\int \psi_j(y) F\big(u(x+ky)\big) dy = \sum_{d=1}^{\mathfrak{m}-1} C_d F^{(d)}(a_0) \int \psi_j(y)(a_k - a_0)^d dy \tag{4.8}$$

$$\begin{aligned} &+ \int \psi_j(y)(a_k - a_0)^{\mathfrak{m}} \int_0^1 \cdots \int_0^1 t_1^{\mathfrak{m}-1} t_2^{\mathfrak{m}-2} \cdots t_{\mathfrak{m}-1} \\ &\times F^{(\mathfrak{m})}(a_0 + (a_k - a_0)t_1 t_2 \cdots t_{\mathfrak{m}}) dt_{\mathfrak{m}} \cdots dt_2 dt_1 dy \end{aligned} \tag{4.9}$$

where C_d only depends on d. Combining (4.4), (4.8) and (4.9) we obtain

$$\begin{aligned} &\Delta_j^\Psi(F(u))(x) \\ &= \sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i} (-1)^i \sum_{d=1}^{\mathfrak{m}-1} C_d F^{(d)}(a_0) \int \psi_j(y)(a_{\mathfrak{m}-i} - a_0)^d dy \end{aligned} \tag{4.10}$$

$$\begin{aligned} &+ \sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i} (-1)^i \int \psi_j(y)(a_{\mathfrak{m}-i} - a_0)^{\mathfrak{m}} \\ &\times \int_0^1 \cdots \int_0^1 t_1^{\mathfrak{m}-1} t_2^{\mathfrak{m}-2} \cdots t_{\mathfrak{m}-1} \\ &\times F^{(\mathfrak{m})}(a_0 + (a_{\mathfrak{m}-i} - a_0)t_1 t_2 \cdots t_{\mathfrak{m}}) dt_{\mathfrak{m}} \cdots dt_1 dy. \end{aligned} \tag{4.11}$$

The decomposition so far has focused on representing

$$\int \psi_j(y) F\big(u(x + (\mathfrak{m} - i)y)\big) dy$$

in terms of derivatives of F. We now turn our attention to further decomposing $(a_{\mathfrak{m}-i}-a_0)^d = (u(x+(\mathfrak{m}-i)y)-u(x))^d$ in terms of frequencies. Notice both (4.10) and (4.11) contain $(a_{\mathfrak{m}-i}-a_0)$ raised to a power. Write $(a_{\mathfrak{m}-i}-a_0)^d$ for $1 \leqslant d \leqslant \mathfrak{m}$ as

$$\sum_{l_1\in\mathbb{Z}} \cdots \sum_{l_d\in\mathbb{Z}} \prod_{r=1}^{d} (a_{\mathfrak{m}-i,l_r} - a_{0,l_r}). \tag{4.12}$$

For simplicity of notation let $\beta_{\mathfrak{m}-i} = \int_0^1 \cdots \int_0^1 t_1^{\mathfrak{m}-1} t_2^{\mathfrak{m}-2} \cdots t_{\mathfrak{m}-1} F^{(\mathfrak{m})}(a_0+(a_{\mathfrak{m}-i}-a_0)t_1t_2\cdots t_{\mathfrak{m}})dt_{\mathfrak{m}}\cdots dt_2dt_1$ from (4.11). From (4.10) and (4.11) it follows that $|\Delta_j^{\Psi}(F(u))(x)|$ is bounded by

$$\sum_{d=1}^{\mathfrak{m}-1} |C_d||F^{(d)}(a_0)| \sum_{l_1\in\mathbb{Z}} \cdots \sum_{l_d\in\mathbb{Z}} \Big| \int \sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i} (-1)^i \psi_j(y) \prod_{r=1}^{d} (a_{\mathfrak{m}-i,l_r} - a_{0,l_r}) dy \Big| \tag{4.13}$$

$$+ \sum_{l_1\in\mathbb{Z}} \cdots \sum_{l_{\mathfrak{m}}\in\mathbb{Z}} \Big| \int \sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i} (-1)^i \psi_j(y) \prod_{r=1}^{\mathfrak{m}} (a_{\mathfrak{m}-i,l_r} - a_{0,l_r}) \beta_{\mathfrak{m}-i} dy \Big|. \tag{4.14}$$

Further decomposing the sum $\sum_{l_1\in\mathbb{Z}} \cdots \sum_{l_d\in\mathbb{Z}}$ for $1 \leqslant d \leqslant \mathfrak{m}$, break up each sum at j, that is $\sum_{l_d\in\mathbb{Z}} = \sum_{l_d\leqslant j} + \sum_{l_d>j}$. Then $\sum_{l_1\in\mathbb{Z}} \cdots \sum_{l_d\in\mathbb{Z}}$ can be written as 2^d sums, by symmetry it is sufficient to focus on only a sum of the form

$$\sum_{l_d>j} \cdots \sum_{l_{\alpha+1}>j} \sum_{l_\alpha\leqslant j} \cdots \sum_{l_1\leqslant j} \tag{4.15}$$

for some $0 \leqslant \alpha \leqslant d$. The low frequency sums can be further bounded by

$$\sum_{l_\alpha\leqslant j} \cdots \sum_{l_1\leqslant j} \leqslant \sum_{\Lambda\in S_\alpha} \sum_{\substack{(l_1,\cdots,l_\alpha)\in\mathbb{Z}^\alpha \\ l_{\Lambda(\alpha)}\leqslant\ldots\leqslant l_{\Lambda(1)}\leqslant j}}$$

where Λ is a permutation of $\{1,\ldots,\alpha\}$. Note, $\sum_{l_\alpha\leqslant j} \cdots \sum_{l_1\leqslant j}$ is not equal to the sum of these $\alpha!$ permutation sums due to the diagonal terms, but the absolute value in (4.10) and (4.11) allows us to use them as bound. By symmetry it suffices to consider the following expression

$$\sum_{l_d>j} \cdots \sum_{l_{\alpha+1}>j} \sum_{\substack{(l_1,\cdots,l_\alpha)\in\mathbb{Z}^\alpha \\ l_\alpha\leqslant\ldots\leqslant l_1\leqslant j}} \Big| \int \Big|.$$

Lastly, partition the integral into $\alpha+1$ pieces, that is

$$\int_{\mathbb{R}^n} = \int_{|y|<2^{-l_1}} + \int_{2^{-l_1}<|y|<2^{-l_2}} + \cdots + \int_{2^{-l_{\alpha-1}}<|y|<2^{-l_\alpha}} + \int_{|y|>2^{-l_\alpha}}.$$

Thus, we have reduced estimating $|\Delta_j F(u)|$ to bounding the following expressions which respectfully correspond to (4.13) and (4.14),

$$\begin{aligned} &\sum_{d=1}^{\mathfrak{m}-1} |C_d||F^{(d)}(a_0)| \sum_{l_d>j} \cdots \sum_{l_{\alpha+1}>j} \sum_{\substack{(l_1,\cdots,l_\alpha)\in\mathbb{Z}^\alpha \\ l_\alpha\leqslant\ldots\leqslant l_1\leqslant j}} \Big| \int_{2^{\eta_1}<|y|<2^{\eta_2}} \\ &\qquad \sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i} (-1)^i \psi_j(y) \prod_{r=1}^{d} (a_{\mathfrak{m}-i,l_r} - a_{0,l_r}) dy \Big| \end{aligned} \tag{4.16}$$

and

$$\begin{aligned}\sum_{l_{\mathfrak{m}}>j}\cdots\sum_{l_{\alpha+1}>j}\sum_{\substack{(l_1,\cdots,l_\alpha)\in\mathbb{Z}^\alpha\\ l_\alpha\leqslant\ldots\leqslant l_1\leqslant j}}\Big|\int_{2^{\eta_1}<|y|<2^{\eta_2}}\sum_{i=0}^{\mathfrak{m}-1}\binom{\mathfrak{m}}{i}(-1)^i\psi_j(y)\\ \times\prod_{r=1}^{\mathfrak{m}}(a_{\mathfrak{m}-i,l_r}-a_{0,l_r})\int_0^1\cdots\int_0^1 t_1^{\mathfrak{m}-1}t_2^{\mathfrak{m}-2}\cdots t_{\mathfrak{m}-1}\\ \times F^{(\mathfrak{m})}(a_0+(a_{\mathfrak{m}-i}-a_0)t_1t_2\cdots t_{\mathfrak{m}})dt_{\mathfrak{m}}\cdots dt_2dt_1dy\Big|\end{aligned}\tag{4.17}$$

where $\eta_1<\eta_2$ are consecutively from the ordered set $\{-\infty,-l_1,\ldots,-l_\alpha,\infty\}$.

5. Techniques

The summation indices, l_r, are categorized into three classes. Let $q\in\mathbb{Z}$. A high index, l_r, means that the sum is over $l_r>q$ for some q. A low-greater index implies the sum is over $l_r\leqslant q$ and the integral is taken over a set where $|y|>2^{-l_r}$. A low-less index means the sum is over $l_r\leqslant q$ and the integral is over a set where $|y|\leqslant 2^{-l_r}$.

For example, consider (4.16) with $\mathfrak{m}=6$, $d=4$, $\eta_1=-l_1$, and $\eta_2=-l_2$, which is expressed as

$$|F^{(d)}(a_0)|\sum_{l_4>j}\sum_{l_3>j}\sum_{l_1\leqslant j}\sum_{l_2\leqslant l_1}\left|\int_{2^{-l_1}<|y|<2^{-l_2}}\sum_{i=0}^{5}\binom{6}{i}(-1)^i\psi_j(y)\prod_{r=1}^{4}(a_{6-i,l_r}-a_{0,l_r})\,dy\right|.$$

Here, l_4 and l_3 are high indices, l_2 is a low-less index, and l_1 is a low-greater index. In this section, we discuss the three methods used to bound the sum corresponding to each type of index. The aim is to demonstrate these methods in simpler cases to avoid excessive complexity with indices and notation. In practice, these methods will be used in combination with additional tools, such as Hölder's inequality, the Cauchy-Schwarz inequality. In the next section, we will explore how these methods are combined and work together.

5.1. High indexed sum. A high indexed sum is simply bounded by the maximal function,

$$\sum_{l>j}\int|\psi_j(y)\Delta_l u(x+ky)|dy\lesssim M\Big(\sum_{l>j}\Delta_l u\Big)(x).$$

Let $s_\star>0$, then multiplying by $2^{js_\star}$ and applying the $L^q(\ell_2)$ norm for $q>1$ we obtain

$$\Big\|2^{js_\star}M\Big(\sum_{l>j}\Delta_l u\Big)\Big\|_{L^q(\ell_2)}\lesssim\Big\|2^{js_\star}\Delta_j u\Big\|_{L^q(\ell_2)}\lesssim\|D^{s_\star}u\|_{L^q}\tag{5.1}$$

where in the first inequality we used Lemma 3.5 and the Fefferman-Stein inequality. The benefit of a high indexed sum is that it can "absorb" arbitrarily large $s_\star$. This will allow a high index sum take the highest order derivative in Section 6.1.

5.2. Low-greater indexed sum. Let $N>>n$, and $m>s_\star>0$. Observe,

$$\begin{aligned}\sum_{l\leqslant j}\int_{|y|>2^{-l}}|\psi_j(y)\Delta_l u(x+ky)|dy&=\sum_{l\leqslant j}\int_{|y|>2^{-l}}|\psi_j(y)|\frac{|\Delta_l u(x+ky)|}{(1+2^l|ky|)^n}(1+2^l|ky|)^n dy\\ &\lesssim\sum_{l\leqslant j}M(\Delta_l u)(x)\int_{|y|>2^{-l}}|\psi_j(y)|(1+2^l|ky|)^n dy\\ &\lesssim\sum_{l\leqslant j}M(\Delta_l u)(x)\int_{|y|>2^{-l}}\frac{2^{jn}(1+2^j|ky|)^n}{(1+2^j|y|)^{N+m}}dy\end{aligned}\tag{5.2}$$

$$\leqslant \sum_{l\leqslant j} 2^{m(l-j)} M(\Delta_l u)(x) \int_{|y|>2^{-l}} \frac{2^{jn}(1+2^j k|y|)^n}{(1+2^j|y|)^{N+m}} dy$$

$$\lesssim \sum_{l\leqslant j} 2^{m(l-j)} M(\Delta_j u)(x) \tag{5.3}$$

where in (5.2) we used Lemma 3.6 and $j \geqslant l$. Multiplying (5.3) by $2^{js_\star}$ and applying the $L^q(\ell_2)$ norm, Lemma 3.5, and the Fefferman-Stein inequality we obtain the bound

$$\|D^{s_\star} u\|_{L^q}. \tag{5.4}$$

Low-greater indexed sums also have the benefit of being able to absorb a higher order derivative. Given a $s_\star$ we are free to choose any $m > s_\star$, due to the rapid decay of ψ, to obtain the final bound in (5.4).

5.3. Low-lower indexed sum. A low-lower indexed sum is more restrictive than both of the previous cases, this is due to the fact that we can not put an arbitrarily large derivative on a low-lower term and still achieve summablity. The decomposition's, (4.16) and (4.17), have been done in such way that a low-lower indexed sum will always be able to be paired with a finite difference as we will see in Section 6.3. The following illustrates the technique,

$$\sum_{l\leqslant j} \int_{|y|<2^{-l}} |\psi_j(y)||b^3_{5,l}| dy$$

$$= \sum_{l\leqslant j} \int_{|y|<2^{-l}} |\psi_j(y)||\Delta_l u(x+5y) - 3\Delta_l u(x+4y)$$

$$+ 3\Delta_l u(x+3y) - \Delta_l u(x+2y)| dy$$

$$\lesssim \sum_{l\leqslant j} 2^{3l} M(\Delta_l u)(x) \tag{5.5}$$

where in (5.5) we applied Lemma 3.4. Multiplying (5.5) by $2^{js_\star}$ and applying the $L^q(\ell_2)$ norm, Lemma 3.5, and the Fefferman-Stein inequality we obtain the bound $\|D^{s_\star} u\|_{L^q}$.

6. Bounding (4.16)

Recall (4.16) is given by

$$\sum_{d=1}^{\mathfrak{m}-1} |C_d||F^{(d)}(a_0)| \sum_{l_d>j} \cdots \sum_{l_{\alpha+1}>j} \sum_{\substack{(l_1,\cdots,l_\alpha)\in\mathbb{Z}^\alpha \\ l_\alpha\leqslant\ldots\leqslant l_1\leqslant j}} \Big| \int_{2^{\eta_1}<|y|<2^{\eta_2}}$$

$$\times \sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i} (-1)^i \psi_j(y) \prod_{r=1}^{d} (a_{\mathfrak{m}-i,l_r} - a_{0,l_r}) dy \Big|$$

We consider three cases:

- Case 1: $\alpha < \mathfrak{m}$, when there is at least one high indexed sum.
- Case 2: $\alpha = \mathfrak{m}$ and $\eta_2 \neq -l_1$, when there is at least one low-greater indexed sum.
- Case 3: $\alpha = \mathfrak{m}$ and $\eta_2 = -l_1$, when *all* the sums are low-less indexed.

In Case 1 and 2, where we have at least one high or low-greater indexed sum the key idea is to let the sum associated with corresponding index take the the highest order derivative. However, in Case 3, where every sum is low-less indexed this is not possible, and we must rely on the cancellation provided by combinatorial arguments.

For Cases 1 and 2, it is sufficient to illustrate the techniques using a specific example, such as $\mathfrak{m} = 6$ and $d = 4$, to avoid the proof becoming obscured by notation and since the sums split. These techniques easily extend to different values of $\mathfrak{m}$ and d. In Case 3, however, we will prove the result for general $\mathfrak{m}$ and d, as this case depends on combinatorial considerations that are specific to those values.

6.1. Case 1: $\alpha < \mathfrak{m}$**.** Let $\mathfrak{m} = 6$ and $d = 4$. In the classical Fa di Bruno setting (see (1.1)), this corresponds to the expressions $F^{(4)}(u)u'u'u'u'''$, the 1-1-1-3 derivative and $F^{(4)}(u)u'u'u''u''$, the 1-1-2-2 derivative, where the sum of the derivatives on the u's adds up to 6. In Case 1 and Case 2 we will bound using the fractional version of the 1-1-1-3 derivative, that is $0 < s_1, s_2, s_3 < 1$, $2 < s_4 < 3$ and $s_1 + s_2 + s_3 + s_4 = s$. Here the higher frequency sum will take on the higher order derivative, s_4, corresponding to 3. The general situation for Case 1 and 2 would be bounded using the fractional version of the 1-1-$\cdots$-1-$\mathfrak{m} - d + 1$ derivative. The only finite differences that will be used in Case 1 and 2 are first order.

To provide a clear illustration that incorporates all three techniques discussed in Section 5, consider the case where $\alpha = 2$, and $\eta_1 = -l_1, \eta_2 = -l_2$. Observe,

$$
\begin{aligned}
&|F^{(d)}(a_0)| \sum_{l_4>j}\sum_{l_3>j}\sum_{l_1\leqslant j}\sum_{l_2\leqslant l_1}\Big|\int_{2^{-l_1}<|y|<2^{-l_2}}\sum_{i=0}^{5}\binom{6}{i}(-1)^i\psi_j(y)\prod_{r=1}^{4}(a_{6-i,l_r}-a_{0,l_r})dy\Big| \\
&\lesssim |F^{(d)}(a_0)|\sum_{i=0}^{5}\sum_{l_1\leqslant j}\sum_{l_2\leqslant l_1}\int_{2^{-l_1}<|y|<2^{-l_2}}\sum_{l_4>j}\sum_{l_3>j}\Big|\psi_j(y)\prod_{r=1}^{4}(a_{6-i,l_r}-a_{0,l_r})\Big|dy \\
&\lesssim |F^{(d)}(a_0)|\sum_{i=0}^{5}\sum_{l_1\leqslant j}\sum_{l_2\leqslant l_1}2^{l_2-j}M(\Delta_{l_2}u)(x)\int_{2^{-l_1}<|y|<2^{-l_2}}|2^jy||\psi_j(y)||a_{6-i,l_1}-a_{0,l_1}| \\
&\qquad\times\sum_{l_4>j}\sum_{l_3>j}\Big|\prod_{r=3}^{4}(a_{6-i,l_r}-a_{0,l_r})dy\Big|
\end{aligned}
\tag{6.1}
$$

where in (6.1) we applied Lemma 3.4 to $|a_{6-i,l_2} - a_{0,l_2}|$. In general all the low-less indexed sums would be handled in this way. Focusing on the low-greater indexed sum continuing from (6.1) we obtain

$$
\begin{aligned}
&\lesssim |F^{(d)}(a_0)|\sum_{i=0}^{5}\sum_{l_1\leqslant j}\sum_{l_2\leqslant l_1}2^{l_2-j}M(\Delta_{l_2}u)(x)\int_{2^{-l_1}<|y|<2^{-l_2}}\frac{2^{jn}|2^jy|}{(1+|2^jy|)^{N+1}} \\
&\qquad\times\Big[\frac{|a_{6-i,l_1}|}{(1+|2^{l_1}(6-i)y|)^n}(1+2^{l_1}|(6-i)y|)^n+|a_{0,l_1}|\Big]\sum_{l_4>j}\sum_{l_3>j}\Big|\prod_{r=3}^{4}(a_{6-i,l_r}-a_{0,l_r})dy\Big|.
\end{aligned}
\tag{6.2}
$$

The expression in square brackets above is bounded by

$$
M(\Delta_{l_1}u)(x)\big((1+2^{l_1}|(6-i)y|)^n+1\big)\leqslant M(\Delta_{l_1}u)(x)2(1+2^j|(6-i)y|)^n \tag{6.3}
$$

by Lemma 3.6 and since $l_1 \leqslant j$. Also, note that

$$
\frac{1}{(1+2^j|y|)} < 2^{l_1-j}
$$

since $|y| > 2^{-l_1}$. Applying these estimates to (6.2) we obtain

$$\begin{aligned}(6.4)\quad &\lesssim |F^{(d)}(a_0)| \sum_{i=0}^{5} \sum_{l_1 \leqslant j} \sum_{l_2 \leqslant l_1} 2^{l_2-j} M(\Delta_{l_2} u)(x) 2^{l_1-j} M(\Delta_{l_1} u)(x) \\ &\quad \times \int \frac{2^{jn} 2(1+2^j|(6-i)y|)^n |2^j y|}{(1+|2^j y|)^N} \sum_{l_4>j} \sum_{l_3>j} \Big| \prod_{r=3}^{4} (a_{6-i,l_r} - a_{0,l_r}) dy \Big|.\end{aligned}$$

In the case where there are multiple low-greater indices, say q of them, choose N sufficiently large (specifically, $N > n + 2\mathfrak{m}n$) and bound $|\psi_j(y)|$ by

$$\frac{2^{jn}}{(1+2^j|y|)^{N+q}}.$$

This ensures the integral converges, and every low-greater indexed sum is paired with a 2^{l_r-j}.

Lastly, we address the high indexed sums. First pointwise bound the product as follows

$$\begin{aligned}&\prod_{r=3}^{4} |(a_{6-i,l_r} - a_{0,l_r})| \\ &\leqslant |a_{6-i,l_3}||a_{6-i,l_4}| + |a_{6-i,l_3}||a_{0,l_4}| + |a_{6-i,l_4}||a_{0,l_3}| + |a_{0,l_3}||a_{0,l_4}| \\ &\lesssim |a_{6-i,l_3}||a_{6-i,l_4}| + |a_{6-i,l_3}| M(\Delta_{l_4} u)(x) \\ &\quad + |a_{6-i,l_4}| M(\Delta_{l_3} u)(x) + M(\Delta_{l_3} u)(x) M(\Delta_{l_4} u)(x).\end{aligned}$$

In the general case after multiplying out $\prod_{r=1}^{d} |a_{\mathfrak{m}-i,l_r} - a_{0,l_r}|$ simply bound the $|a_{0,l_r}|$ terms by $M(\Delta_{l_r} u)(x)$. Applying this estimate to (6.4), and using that convolution with a L^1 dilate of a Shwartz function is bounded by the Maximal function we obtain

$$\begin{aligned}(6.5)\quad &\lesssim |F^{(d)}(a_0)| \sum_{l_1 \leqslant j} 2^{l_1-j} M(\Delta_{l_1} u)(x) \sum_{l_2 \leqslant j} 2^{l_2-j} M(\Delta_{l_2} u)(x) \\ &\quad \times \Big[M\Big(\sum_{l_4>j} |\Delta_{l_4} u| \sum_{l_3>j} |\Delta_{l_3} u| \Big)(x) + \sum_{l_4>j} M(\Delta_{l_4} u)(x) M\Big(\sum_{l_3>j} |\Delta_{l_3} u| \Big)(x) \\ &\quad + \sum_{l_3>j} M(\Delta_{l_3} u)(x) M\Big(\sum_{l_4>j} |\Delta_{l_4} u| \Big)(x) + \sum_{l_4>j} M(\Delta_{l_4} u)(x) \sum_{l_3>j} M(\Delta_{l_3} u)(x) \Big].\end{aligned}$$

6.1.1. *Applying the $L^p(\ell_2)$ norm.* Recall $0 < s_1, s_2, s_3 < 1$, $2 < s_4 < 3$ and $s_1 + s_2 + s_3 + s_4 = s$. In general we would have $0 < s_1, \dots, s_{d-1} < 1$, $\mathfrak{m} - d < s_d < \mathfrak{m} - d + 1$ and $s_1 + \cdots + s_d = s$. The largest derivative s_4, will be put on a higher indexed sum. Let p^* be such that $\frac{1}{p_3} + \frac{1}{p_4} = \frac{1}{p^*}$, and note $p^* > 1$. We will only focus on bounding one of four terms in (6.5), namely

$$M\Big(\sum_{l_4>j} |\Delta_{l_4} u| \sum_{l_3>j} |\Delta_{l_3} u| \Big),$$

as all four can be handled similarly by the following method of using the Cauchy-Schwarz inequality for multiple factors, Hölder's inequality, Lemma 3.5, and the Fefferman-Stein inequality. Multiplying (6.5) by 2^{js} and applying the $L^p(\ell_2)$ norm we obtain

$$\Big\| |F^{(d)}(a_0)| 2^{js} \sum_{l_1 \leqslant j} 2^{l_1-j} M(\Delta_{l_1} u) \sum_{l_2 \leqslant j} 2^{l_2-j} M(\Delta_{l_2} u) M\Big(\sum_{l_4>j} |\Delta_{l_4} u| \sum_{l_3>j} |\Delta_{l_3} u| \Big) \Big\|_{L^p(\ell_2)}$$

(6.6)

$$\leqslant \left\| F^{(d)}(u) \right\|_{L^{p_0}} \left\| \Big(\sum_{j\in\mathbb{Z}} \big(2^{js_1} \sum_{l_1 \leqslant j} 2^{l_1-j} M(\Delta_{l_1} u) \big)^2 \Big)^{\frac{1}{2}} \right\|_{L^{p_1}} \left\| \Big(\sum_{j\in\mathbb{Z}} \big(2^{js_2} \sum_{l_2 \leqslant j} 2^{l_2-j} M(\Delta_{l_2} u) \big)^2 \Big)^{\frac{1}{2}} \right\|_{L^{p_2}}$$
$$\times \left\| \Big(\sum_{j\in\mathbb{Z}} \Big(2^{j(s_3+s_4)} M\Big(\sum_{l_4>j} |\Delta_{l_4} u| \sum_{l_3>j} |\Delta_{l_3} u| \Big) \Big)^2 \Big)^{\frac{1}{2}} \right\|_{L^{p^*}}$$

(6.7)

$$\lesssim \left\| F^{(d)}(u) \right\|_{L^{p_0}} \| D^{s_1} u \|_{L^{p_1}} \| D^{s_2} u \|_{L^{p_2}} \left\| \Big(\sum_{j\in\mathbb{Z}} \Big(2^{j(s_3+s_4)} M\Big(\sum_{l_4>j} |\Delta_{l_4} u| \sum_{l_3>j} |\Delta_{l_3} u| \Big) \Big)^2 \Big)^{\frac{1}{2}} \right\|_{L^{p^*}}$$

(6.8)

$$\lesssim \left\| F^{(d)}(u) \right\|_{L^{p_0}} \| D^{s_1} u \|_{L^{p_1}} \| D^{s_2} u \|_{L^{p_2}} \| D^{s_3} u \|_{L^{p_3}} \| \| D^{s_4} u \|_{L^{p_4}}.$$

In (6.6), we applied the Cauchy-Schwarz inequality for multiple factors and then Hölder's inequality. In (6.7), we used Lemma 3.5 followed by the Fefferman-Stein inequality. In (6.8), we applied the Fefferman-Stein inequality first, then Hölder's inequality, and finally Lemma 3.5.

6.2. Case 2: $\alpha = \mathfrak{m}$ and $\eta_2 \neq -l_1$. For simplicity, and because the sums split, it suffices to again consider the case where $\mathfrak{m} = 6$ and $d = 4$. In this case there are no high indexed sums. However, we do have a low-greater index, allowing us to assign the highest-order derivative to the corresponding sum. Therefore, we will decompose s as in Case 1, with $0 < s_1, s_2, s_3 < 1$, $2 < s_4 < 3$, and $s_1 + s_2 + s_3 + s_4 = s$. We will only consider the case where we have two low-greater indices l_1, l_2. Observe,

$$|F^{(d)}(a_0)| \sum_{l_1 \leqslant j} \sum_{l_2 \leqslant l_1} \sum_{l_3 \leqslant l_2} \sum_{l_4 \leqslant l_3} \Big| \int_{2^{-l_2} < |y| < 2^{-l_3}} \sum_{i=0}^{5} \binom{6}{i} (-1)^i \psi_j(y) \prod_{r=1}^{4} (a_{6-i,l_r} - a_{0,l_r}) dy \Big|$$

(6.9)

$$\lesssim |F^{(d)}(a_0)| \sum_{i=0}^{5} \sum_{l_1 \leqslant j} \sum_{l_2 \leqslant l_1} \sum_{l_3 \leqslant l_2} \sum_{l_4 \leqslant l_3} 2^{l_4-j} M(\Delta_{l_4} u)(x) 2^{l_3-j} M(\Delta_{l_3} u)(x)$$
$$\times \int_{2^{-l_2} < |y| < 2^{-l_3}} |2^j y| |2^j y| |\psi_j(y)| \prod_{r=1}^{2} |(a_{6-i,l_r} - a_{0,l_r})| dy$$

(6.10)

$$\lesssim |F^{(d)}(a_0)| \sum_{i=0}^{5} \sum_{l_1 \leqslant j} \sum_{l_2 \leqslant l_1} \sum_{l_3 \leqslant l_2} \sum_{l_4 \leqslant l_3} 2^{l_4-j} M(\Delta_{l_4} u)(x) 2^{l_3-j} M(\Delta_{l_3} u)(x) M(\Delta_{l_2} u)(x) M(\Delta_{l_1} u)(x)$$
$$\times \int_{2^{-l_2} < |y| < 2^{-l_3}} |2^j y| |2^j y| |\psi_j(y)| 4(1 + 2^j |(6-i) y|)^{2n} dy$$

(6.11)

$$\lesssim |F^{(d)}(a_0)| \sum_{i=0}^{5} \sum_{l_1 \leqslant j} \sum_{l_2 \leqslant l_1} \sum_{l_3 \leqslant l_2} \sum_{l_4 \leqslant l_3} 2^{l_4-j} M(\Delta_{l_4} u)(x) 2^{l_3-j} M(\Delta_{l_3} u)(x) M(\Delta_{l_2} u)(x) M(\Delta_{l_1} u)(x)$$
$$\times \int_{2^{-l_2} < |y| < 2^{-l_3}} \frac{2^{jn} |2^j y|^2 4(1 + 2^j |(6-i) y|)^{2n} |2^j y|^2}{(1 + 2^j |y|)^{N+1+3}} dy$$

$$\lesssim |F^{(d)}(a_0)| \sum_{i=0}^{5} \sum_{l_1 \leqslant j} \sum_{l_2 \leqslant l_1} \sum_{l_3 \leqslant l_2} \sum_{l_4 \leqslant l_3} 2^{l_4-j} M(\Delta_{l_4}u)(x) 2^{l_3-j} M(\Delta_{l_3}u)(x) M(\Delta_{l_2}u)(x) M(\Delta_{l_1}u)(x)$$
$$\times \int_{2^{-l_2}<|y|<2^{-l_3}} \frac{2^{jn}4(1+2^j|(6-i)y|)^{2n}|2^jy|^2}{(1+2^j|y|)^N} 2^{l_2-j}2^{3(l_1-j)}dy$$

(6.12)

$$\lesssim |F^{(d)}(a_0)| \sum_{l_1 \leqslant j} 2^{3(l_1-j)} M(\Delta_{l_1}u)(x) \sum_{l_2 \leqslant j} 2^{l_2-j} M(\Delta_{l_2}u)(x)$$
$$\times \sum_{l_3 \leqslant j} 2^{l_3-j} M(\Delta_{l_3}u)(x) \sum_{l_4 \leqslant j} 2^{l_4-j} M(\Delta_{l_4}u)(x).$$

In (6.9), we applied Lemma 3.4 to the low-less indexed terms. In (6.10), we used the estimate given in (6.3) for the less-greater indices. In (6.11) we chose the exponent on $(1+2^j|y|)$ to be $N+1+3$, where N is large enough to ensure the convergence of the integral and the $1+3$ is used to obtain $2^{l_2-j}2^{3(l_1-j)}$. In general, we set the exponent to be $N+(q-1)+(\mathfrak{m}-d+1)$ where $N > n+2n\mathfrak{m}$, and q denotes the number of low-greater indices.

Finally, multiplying (6.12) by 2^{js} and applying the $L^p(\ell_2)$ norm, we proceed as in subsection 6.1.1, using the Cauchy-Schwarz inequality for multiple factors, Hölder's inequality, Lemma 3.5, and the Fefferman-Stein inequality. This yields the desired estimate

$$\left\|F^{(d)}(u)\right\|_{L^{p_0}} \|D^{s_1}u\|_{L^{p_1}} \|D^{s_2}u\|_{L^{p_2}} \|D^{s_3}u\|_{L^{p_3}} \| \|D^{s_4}u\|_{L^{p_4}}.$$

6.3. Case 3: $\alpha = \mathfrak{m}$ and $\eta_2 = -l_1$. In the final case there are no high or low-greater indexed terms to absorb the highest order derivative. We will use combinatorial arguments to eliminate all terms except for the desired finite differences. In this case we seek to bound an expression of the form,

$$\sum_{d=1}^{\mathfrak{m}-1} |C_d||F^{(d)}(a_0)| \sum_{l_1 \leqslant j} \sum_{l_2 \leqslant l_1} \sum_{l_3 \leqslant l_2} \cdots \sum_{l_d \leqslant l_{d-1}} \Big| \int_{|y|<2^{-l_1}} \psi_j(y) \times \sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i} (-1)^i \prod_{r=1}^{d} (a_{\mathfrak{m}-i,l_r} - a_{0,l_r}) dy \Big| \tag{6.13}$$

To begin we will focus on the sum in the integrand,

$$\sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i} (-1)^i \prod_{r=1}^{d} (a_{\mathfrak{m}-i,l_r} - a_{0,l_r})$$

$$= \sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i} (-1)^i \underbrace{\prod_{r=1}^{d} \sum_{\nu=0}^{\mathfrak{m}-i-1} \binom{\mathfrak{m}-i}{\mathfrak{m}-i-\nu} b^{\mathfrak{m}-i-\nu}_{\mathfrak{m}-i-\nu,l_r}}_{\text{multiply out}} \tag{6.14}$$

$$= \sum_{(\kappa_1,\ldots,\kappa_d)\in\{1,\ldots,\mathfrak{m}\}^d} \sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i} (-1)^i \binom{\mathfrak{m}-i}{\kappa_1} \cdots \binom{\mathfrak{m}-i}{\kappa_d} b^{\kappa_1}_{\kappa_1,l_1} \cdots b^{\kappa_d}_{\kappa_d,l_d} \tag{6.15}$$

where $\binom{a}{b} = 0$ for $b > a$, and in (6.14) we applied Lemma 3.2. Using Lemma 3.3 we obtain that (6.15) is equal to

$$\sum_{\substack{(\kappa_1,\ldots,\kappa_d)\in\{1,\ldots,\mathfrak{m}\}^d \\ \kappa_1+\cdots+\kappa_d\geqslant\mathfrak{m}}} \sum_{i=0}^{\mathfrak{m}-1} \binom{\mathfrak{m}}{i}(-1)^i \binom{\mathfrak{m}-i}{\kappa_1}\cdots\binom{\mathfrak{m}-i}{\kappa_d} b^{\kappa_1}_{\kappa_1,l_1}\cdots b^{\kappa_d}_{\kappa_d,l_d}. \tag{6.16}$$

For each $(\kappa_1,\ldots,\kappa_d)\in\{1,\ldots,\mathfrak{m}\}^d$ such that $\kappa_1+\cdots+\kappa_d\geqslant\mathfrak{m}$ associate a $(\tau_1,\ldots,\tau_d)\in\{1,\ldots,\mathfrak{m}\}^d$ in the following way: select the smallest index r for which $\kappa_r > 1$ and replace κ_r with $\kappa_r - 1$. Repeat this process until $(\kappa_1,\ldots,\kappa_d)$ becomes $(\tau_1,\ldots,\tau_d)$ where $\tau_1+\cdots+\tau_d = \mathfrak{m}$. By repeatedly applying Lemma 3.1, $b^{\kappa_1}_{\kappa_1,l_1}\cdots b^{\kappa_d}_{\kappa_d,l_d}$ can be written as a linear combination of terms of the form

$$b^{\tau_1}_{\sigma_1,l_1}\cdots b^{\tau_d}_{\sigma_d,l_d} \tag{6.17}$$

where $\sigma_r \geqslant \tau_r$ and $\tau_1+\cdots+\tau_d = \mathfrak{m}$. Thus (6.16) can be written as a linear combination of terms of the form in (6.17) where the coefficients of the linear combination depend on $\mathfrak{m}, d$. Hence we have reduced bounding (6.13) to bounding a term of following form

$$|F^{(d)}(a_0)| \sum_{l_1\leqslant j}\sum_{l_2\leqslant l_1}\sum_{l_3\leqslant l_2}\cdots\sum_{l_d\leqslant l_{d-1}} \int_{|y|<2^{-l_1}} |\psi_j(y) b^{\tau_1}_{\sigma_1,l_1}\cdots b^{\tau_d}_{\sigma_d,l_d}|dy.$$

By Lemma 3.4 the above expression is bounded by a constant multiple of

$$|F^{(d)}(a_0)| \sum_{l_1\leqslant j}\sum_{l_2\leqslant j}\sum_{l_3\leqslant j}\cdots\sum_{l_d\leqslant j} 2^{\tau_1(l_1-j)}M(\Delta_{l_1}u)(x)\cdots 2^{\tau_d(l_d-j)}M(\Delta_{l_d}u)(x). \tag{6.18}$$

Now select s_i such that $\tau_i - 1 < s_i < \tau_i$ and $\sum_{i=1}^d s_i = s$. Multiplying (6.18) by 2^{js} and applying the $L^p(\ell_2)$ norm, we proceed as in subsection 6.1.1, using the Cauchy-Schwarz inequality for multiple factors, Hölder's inequality, Lemma 3.5, and the Fefferman-Stein inequality. This yields the desired estimate

$$\left\|F^{(d)}(u)\right\|_{L^{p_0}} \|D^{s_1}u\|_{L^{p_1}}\cdots\|D^{s_d}u\|_{L^{p_d}}.$$

□

7. Bounding (4.17)

Split s such that $0 < s_i < 1$ and $\sum_{i=1}^{\mathfrak{m}} s_i = s$. Bringing the absolute value inside the integral, (4.17) is bounded by constant multiple (depending on $\mathfrak{m}$) of

$$\begin{aligned}&\sum_{i=0}^{\mathfrak{m}-1}\sum_{l_{\mathfrak{m}}>j}\cdots\sum_{l_{\alpha+1}>j}\sum_{\substack{(l_1,\cdots,l_\alpha)\in\mathbb{Z}^\alpha \\ l_\alpha\leqslant\ldots\leqslant l_1\leqslant j}} \int_{2^{\eta_1}<|y|<2^{\eta_2}} |\psi_j(y)| \prod_{r=1}^{\mathfrak{m}} |a_{\mathfrak{m}-i,l_r} - a_{0,l_r}| \\ &\times \int_0^1\cdots\int_0^1 t_1^{\mathfrak{m}-1}t_2^{\mathfrak{m}-2}\cdots t_{\mathfrak{m}-1}|F^{(\mathfrak{m})}(a_0 + (a_{\mathfrak{m}-i}-a_0)t_1t_2\cdots t_{\mathfrak{m}})|dt_{\mathfrak{m}}\cdots dt_2dt_1dy\end{aligned} \tag{7.1}$$

By the smoothness condition on $F^{(\mathfrak{m})}$ the second line of (7.1) is bounded by

$$\begin{aligned}&\Big(G(a_0) + G(a_{\mathfrak{m}-i})\Big)\int_0^1\cdots\int_0^1 \mu(t_1\cdots t_{\mathfrak{m}})t_1^{\mathfrak{m}-1}t_2^{\mathfrak{m}-2}\cdots t_{\mathfrak{m}-1}dt_{\mathfrak{m}}\cdots dt_2dt_1 \\ &\lesssim_{\mathfrak{m},\mu} \Big(G(a_0) + G(a_{\mathfrak{m}-i})\Big).\end{aligned}$$

Which follows from a simple change of variables namely $t'_m = t_1 \cdots t_{m-1} t_m$. Applying this estimate to (7.1) yields

$$\sum_{i=0}^{\mathfrak{m}-1} \sum_{l_{\mathfrak{m}}>j} \cdots \sum_{l_{\alpha+1}>j} \sum_{\substack{(l_1,\cdots,l_\alpha)\in\mathbb{Z}^\alpha \\ l_\alpha\leqslant\ldots\leqslant l_1\leqslant j}} \int_{2^{\eta_1}<|y|<2^{\eta_2}} \Big(G(a_0)+G(a_{\mathfrak{m}-i})\Big)|\psi_j(y)| \prod_{r=1}^{\mathfrak{m}} |a_{\mathfrak{m}-i,l_r} - a_{0,l_r}|dy \tag{7.2}$$

$$= \sum_{i=0}^{\mathfrak{m}-1} \sum_{l_{\mathfrak{m}}>j} \cdots \sum_{l_{\alpha+1}>j} \sum_{\substack{(l_1,\cdots,l_\alpha)\in\mathbb{Z}^\alpha \\ l_\alpha\leqslant\ldots\leqslant l_1\leqslant j}} \int_{2^{\eta_1}<|y|<2^{\eta_2}} G(a_0)|\psi_j(y)| \prod_{r=1}^{\mathfrak{m}} |a_{\mathfrak{m}-i,l_r} - a_{0,l_r}|dy \tag{7.3}$$

$$+ \sum_{i=0}^{\mathfrak{m}-1} \sum_{l_{\mathfrak{m}}>j} \cdots \sum_{l_{\alpha+1}>j} \sum_{\substack{(l_1,\cdots,l_\alpha)\in\mathbb{Z}^\alpha \\ l_\alpha\leqslant\ldots\leqslant l_1\leqslant j}} \int_{2^{\eta_1}<|y|<2^{\eta_2}} G(a_{\mathfrak{m}-i})|\psi_j(y)| \prod_{r=1}^{\mathfrak{m}} |a_{\mathfrak{m}-i,l_r} - a_{0,l_r}|dy \tag{7.4}$$

Bounding (7.2) follows a proof very similar to that used for (4.16) in the previous section. Additionally, we do not need to consider the combinatorial arguments from subsection 6.3 because we can directly apply Lemma 3.4 to $|a_{\mathfrak{m}-i,l_r} - a_{0,l_r}|$ when l_r is a low-less index, since $s_r < 1$. The only difference from the proof in Section 6 is the addition of $(G(a_0) + G(a_{\mathfrak{m}-i}))$ to the integrand. The expression in (7.3) can be bounded using the same argument from Section 6, with $F^{(d)}(u(x))$ replaced by $G(a_0) = G(u(x))$.

The effects of the addition of $G(a_{\mathfrak{m}-i}) = G(u(x + (\mathfrak{m} - i)y))$ in (7.3) are minimal since all integrability indices involved are greater than 1, allowing us to use the boundedness of the maximal operator. To illustrate first suppose that we have all low-less or low-greater indices then using Lemma 3.4 and Lemma 3.6 we can pull all the Littlewood-Paley operators (by bounding them with $M(\Delta_{l_r}u)(x)$ as seen in (5.2) and (5.3)) out of the integral leaving

$$\int \frac{2^{jn}}{(1+2^j|y|)^N} G(u(x+(\mathfrak{m}-i)y))dy \lesssim_{\mathfrak{m},i} M(G\circ u)(x).$$

The situation for when there are high index sums is similar since the integral is a convolution. To illustrate, suppose all low-less or low-greater indices have been extracted from the integral using Lemma 3.4 or Lemma 3.6. This leaves an integral of the form

$$\int \frac{2^{jn}}{(1+2^j|y|)^N} G(a_{\mathfrak{m}-i}) \sum_{l_{\mathfrak{m}}>j} \cdots \sum_{l_{\alpha+1}>j} \prod_{r=\alpha+1}^{\mathfrak{m}} |a_{\mathfrak{m}-i,l_r}|dy,$$

which is bounded by

$$M\Big(G\circ u \sum_{l_{\mathfrak{m}}>j} \cdots \sum_{l_{\alpha+1}>j} \prod_{r=\alpha+1}^{\mathfrak{m}} |\Delta_{l_r}u|\Big).$$

The proof then proceeds in the same manner as in Subsection 6.1.1. Specifically, by multiplying (6.18) by 2^{js} and applying the $L^p(\ell_2)$ norm, then the Cauchy-Schwarz inequality for multiple factors, Hölder's inequality, Lemma 3.5, and the Fefferman-Stein inequality. □

References

[1] Christ, Michael; Weinstein, Michael. *Dispersion of small amplitude solutions of the generalized Korteweg-de Vries equation.* J. Funct. Anal. 100 (1991), no. 1, 87–109.

[2] Douglas, Sean. *Chain rule for weighted Triebel-Lizorkin spaces.* J. Funct. Anal. *Revised version submitted.*

[3] Fujiwara, Kazumasa. *Remark on the chain rule of fractional derivative in the Sobolev framework.* Math. Inequal. Appl. 24 (2021), no.4, 1113–1124.

[4] Grafakos, Loukas. *Modern Fourier analysis.* Third edition. Graduate Texts in Mathematics, 250. Springer, New York, 2014.

[5] Grafakos, Loukas. *Fundamentals of Fourier Analysis*. Graduate Texts in Mathematics, 302. Springer, New York, 2024.

[6] Ginibre, Jean; Ozawa, Tohru; Velo, Giorgio. *On the existence of the wave operators for a class of nonlinear Schrödinger equations* Ann. Inst. H. Poincaré Phys. Théor. 60 (1994), no.2, 211–239

[7] Kato, Tosio. *On nonlinear Schrödinger equations. II. Hs-solutions and unconditional well-posedness*. J. Anal. Math. 67 (1995), 281–306.

[8] Kato, Tosio; Ponce, Gustavo. *Commutator estimates and the Euler and Navier-Stokes equations*. Comm. Pure Appl. Math. 41 (1988), no. 7, 891–907.

[9] Naibo, Virginia; Thomson, Alexander. *Coifman-Meyer multipliers: Leibniz-type rules and applications to scattering of solutions to PDEs*. Trans. Amer. Math. Soc. 372 (2019), no. 8, 5453–5481.

[10] Staffilani, Gigliola. *On the generalized Korteweg-de Vries-type equations*. Differential Integral Equations 10 (1997), no.4, 777–796.

[11] Taylor, Michael. *Tools for PDE: Pseudodifferential Operators, Paradifferential Operators, and Layer Potentials*. Math. Surveys Monogr., 81. AMS, Rhode Island, 2000.

[12] Yamazaki, Masao. *A quasihomogeneous version of paradifferential operators. I. Boundedness on spaces of Besov type*. J. Fac. Sci. Univ. Tokyo Sect. IA Math. 33 (1986), no.1, 131–174.

Department of Mathematics, University of British Columbia
Email address: spdthc@gmail.com