CHAIN RULE FOR WEIGHTED TRIEBEL-LIZORKIN SPACES

SEAN DOUGLAS

Abstract. In this paper we establish a fractional chain rule in the setting of weighted Triebel-Lizorkin spaces under various smoothness conditions. Notably, this provides a fractional chain rule for weighted Lebesgue spaces, $L^p(w)$, for $p \leqslant 1$. Additionally, an explicit relationship between the smoothness index, integrablity index, and the choice of weights is determined.

1. Introduction

In this paper we consider a fractional chain rule that forms a good substitute for the identity

$$\frac{d}{dx}F(u) = F'(u)u'$$

in the framework of weighted Triebel-Lizorkin (TL) spaces. This includes both homogeneous and inhomogeneous weighted TL spaces. The early use of a fractional chain rule goes back to Christ and Weinstein [3] 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.

A fractional chain rule also plays an important role in establishing well-posedness of wave equations [9], [8]. In particular, a more general smoothness condition, similar to that given subsequently in Theorem 1.1, was developed by Kato and Staffilani [11], [13] in the study of well-posedness of NLS equations. More recently a weighted fractional chain rule was used to establish Morawetz type estimates in order to determine the local well-posedness of a semi-linear wave equation [10]. A fractional chain rule has also been obtained in the more general setting of Besov spaces, however the integrablity index is greater than or equal to 1 [8].

A related inequality to a fractional chain rule is a fractional Leibniz rule, which has a multitude of generalizations and extensions. A fractional Leibniz rule, also known as a Kato-Ponce inequality [12] are estimates of the form

$$\|D^s(fg)\|_{L^p} \leqslant C_{n,s,p_1,p_2}\left(\|D^s f\|_{L^{p_1}}\|g\|_{L^{p_2}} + \|f\|_{L^{p_1}}\|D^s g\|_{L^{p_2}}\right), \tag{1.1}$$

where f, g are Schwartz functions and $p^{-1} = p_1^{-1} + p_2^{-1}$. In [1] Bernicot, Naibo, Maldonado, and Moen extended (1.1) to include $p \leqslant 1$, but required that $s > n$. Grafakos and Oh [5] obtained (1.1) for $\frac{1}{2} < p < \infty$ with a sharp lower bound on the smoothness index $s > n(1/\min(p,1) - 1)$. This same lower bound is also shown to be sharp for a fractional chain rule and is discussed in Section 7.

The aim of this article is to provide a weighted fractional chain rule for $p \leqslant 1$, similar to the extension in the range of indices of the Kato-Ponce inequality. The main results in this article are in the spirit of Christ and Weinstein's work. However, their proof heavily relied on the boundedness of the Hardy-Littlewood maximal operator, which is not available for $p \leqslant 1$. To overcome this obstacle we adapt a similar technique used by Naibo and Thomson [15] to obtain a weighted Kato-Ponce inequality for Triebel-Lizorkin spaces. Specifically, we use a Peetre type lemma from [16], which allows us to circumvent this obstacle.

All functions are defined over $\mathbb{R}^n$, so $\mathbb{R}^n$ is suppressed from in the notation. The space of Schwartz functions, smooth rapidly decaying functions on $\mathbb{R}^n$, is denoted by $\mathcal{S}$. The dual space of $\mathcal{S}$, the space of tempered distributions, is denoted by $\mathcal{S}'$. For the homogeneous setting we need $\mathcal{S}_0$, the space of Schwartz functions that have vanishing moments of all orders; that is $f \in \mathcal{S}_0$ if and only if f is a Schwartz function and $\int_{\mathbb{R}^n} x^\alpha f(x)dx = 0$ where α is a multi-index. The dual space of $\mathcal{S}_0$ is the space of tempered distributions

Date: January 1, 2024.
The author acknowledges the support of the University of Missouri Research Council.

modulo polynomials and is denoted by $\mathcal{S}'/\mathcal{P}$. This is the space of tempered distributions in which two elements are identified if their difference is a polynomial. It is a known fact that $\mathcal{S}_0$ is dense in $\mathcal{S}'/\mathcal{P}$ and $\mathcal{S}$ is dense in $\mathcal{S}'$. For more information about tempered distributions see [[7], Chap. 1].

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

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(y)e^{-2\pi i y\cdot\xi}dy \qquad \check{f}(\xi) = \int_{\mathbb{R}^n} f(y)e^{2\pi i y\cdot\xi}dy.$$

The homogeneous fractional Laplacian for $u \in \mathcal{S}'/\mathcal{P}$ and $s > 0$ is denoted by $D^s u \in \mathcal{S}'/\mathcal{P}$ and defined by $\langle D^s u, \varphi\rangle := \langle u, (|\cdot|^s \check{\varphi})\widehat{\ }\rangle$ for $\varphi \in \mathcal{S}_0$ (it turns out $(|\cdot|^s\check{\varphi})\widehat{\ } \in \mathcal{S}_0$). The inhomogeneous fractional Laplacian for $u \in \mathcal{S}'$ and $s > 0$ is denoted by $J^s u \in \mathcal{S}'$ and defined by $\langle J^s u, \varphi\rangle := \langle u, (1+|\cdot|^2)^{\frac{s}{2}}\check{\varphi})\widehat{\ }\rangle$ for $\varphi \in \mathcal{S}$.

Let $\widehat{\Phi}(\xi)$ be a positive radially decreasing $C^\infty(\mathbb{R}^n)$ function on $\mathbb{R}^n$ supported in twice the unit ball 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 $\frac{1}{2} \leqslant |\xi| \leqslant 2$. 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

$$\Psi_j := 2^{jn}\Psi(2^j\cdot).$$

The weighted homogeneous Triebel-Lizorkin space is $u \in \mathcal{S}'/\mathcal{P}$ such that

$$\|u\|_{\dot{F}^s_{p,q}(w)} := \left\|\Big(\sum_{j\in\mathbb{Z}} |2^{js}\Delta_j u|^q\Big)^{\frac{1}{q}}\right\|_{L^p(w)} < \infty,$$

for $w \in A_\infty$ and $0 < s,p,q < \infty$. The weighted inhomogeneous Triebel-Lizorkin space is $u \in \mathcal{S}'$ such that

$$\|u\|_{F^s_{p,q}(w)} := \|S_0 u\|_{L^p(w)} + \left\|\left(\sum_{j\geqslant 1}\Big(2^{js}|\Delta_j u|\Big)^q\right)^{\frac{1}{q}}\right\|_{L^p(w)} < \infty,$$

for $w \in A_\infty$ and $0 < s,p,q < \infty$. The Hardy space $H^p(w)$ for $0 < p < \infty$, and $w \in A_\infty$ is defined to be $u \in \mathcal{S}'$, such that

$$\|u\|_{H^p(w)} := \left\|\sup_{0<t<\infty} |t^{-n}\Phi(t^{-1}\cdot) * u|\right\|_{L^p(w)} < \infty.$$

The local Hardy space $h^p(w)$ for $0 < p < \infty$, and $w \in A_\infty$ is defined to be $u \in \mathcal{S}'$, such that

$$\|u\|_{h^p(w)} := \left\|\sup_{0<t<1} |t^{-n}\Phi(t^{-1}\cdot) * u|\right\|_{L^p(w)} < \infty.$$

The lifting property of Triebel-Lizorkin spaces states for $u \in S'$ and $w \in A_\infty$ that $\|J^s u\|_{F^0_{p,q}(w)} \sim \|u\|_{F^s_{p,q}(w)}$ for $0 \leqslant s < \infty$ and $0 < p,q < \infty$. For homogeneous spaces the analogous lifting property for $u \in S'/\mathcal{P}$ and $w \in A_\infty$ is $\|D^s u\|_{\dot{F}^0_{p,q}(w)} \sim \|u\|_{\dot{F}^s_{p,q}(w)}$ for $0 \leqslant s < \infty$ and $0 < p,q < \infty$. Furthermore, there is the following relationship between Hardy and TL norms; $\|\cdot\|_{F^0_{p,2}(w)} \sim \|\cdot\|_{h^p(w)}$ and $\|\cdot\|_{\dot{F}^0_{p,2}(w)} \sim \|\cdot\|_{H^p(w)}$ for $w \in A_\infty$ and $0 < p < \infty$. A subtle point is that $\|\cdot\|_{\dot{F}^0_{p,2}(w)} \sim \|\cdot\|_{H^p(w)}$ holds for $u \in S'/\mathcal{P}$, that is u is a tempered distribution defined by its action on $\mathcal{S}_0$. If $1 < p < \infty$ and $w \in A_p$ then $\|\cdot\|_{L^p(w)} \sim \|\cdot\|_{H^p(w)} \sim \|\cdot\|_{h^p(w)}$. Note that if g is a function on $\mathbb{R}^n$ then $\|g\|_{L^p(w)} \lesssim \|g\|_{h^p(w)} \lesssim \|g\|_{H^p(w)}$ for $0 < p < \infty$ and $w \in A_\infty$. For more information on weighted function spaces see [2].

For a Muckenhoupt weight w we define $\tau_w = \inf\{\tau \in (1,\infty) : w \in A_\tau\}$. Further definitions and notation are defined in the next section. Now that all the relevant definitions have been discussed we state our main result and its consequences.

Theorem 1.1. *Let* $0 < s < 1$, $0 < p < \infty$, $1 < p_1 \leqslant \infty$, $0 < p_2 < \infty$, $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$, $w_1 \in A_{p_1}$, $w_2 \in A_\infty$, $w = w_1^{\frac{p}{p_1}} w_2^{\frac{p}{p_2}}$, *and* $0 \leqslant n\Big(\max\big(\frac{\tau_{w_2}}{p_2}, \frac{1}{q}\big) - \frac{p_1 - \tau_{w_1}}{p_1}\Big) < s$ *or* $\frac{p}{\tau_w} > 1$ *(in the case* $p_1 = \infty$ *assume* $0 \leqslant n\Big(\max\big(\frac{\tau_{w_2}}{p_2}, \frac{1}{q}\big) - 1\Big) < s$*) and* $u \in \mathcal{S}_0$. *Let* $F : \mathbb{C} \to \mathbb{C}$ *and* $G : \mathbb{C} \to [0,\infty)$ *such that* $F(0) = 0$ *and* $G(u) \in L^{p_1}(w_1)$. *Furthermore, suppose that for* $x, y \in \mathbb{C}$ *that*

$$|F(x) - F(y)| \leqslant [G(x) + G(y)]|x - y|,$$

then

$$\|F(u)\|_{\dot{F}^s_{p,q}(w)} \leqslant C_{n,s,q,p_1,p_2,w_1,w_2}\|G(u)\|_{L^{p_1}(w_1)}\|u\|_{\dot{F}^s_{p_2,q}(w_2)}. \tag{1.2}$$

If $u \in \mathcal{S}$ *then*

$$\|F(u)\|_{F^s_{p,q}(w)} \leqslant C_{n,s,q,p_1,p_2,w_1,w_2}\Big(\|F(u)\|_{L^p(w)} + \|G(u)\|_{L^{p_1}(w_1)}\|u\|_{F^s_{p_2,q}(w_2)}\Big). \tag{1.3}$$

Moreover, if $G(0) = 0$ *then we obtain*

$$\|F(u)\|_{F^s_{p,q}(w)} \leqslant C_{n,s,q,p_1,p_2,w_1,w_2}\|G(u)\|_{L^{p_1}(w_1)}\|u\|_{F^s_{p_2,q}(w_2)}. \tag{1.4}$$

Remark 1.1.1. *The hypothesis* $F(0) = 0$ *is not necessary for* (1.2) *since* $\Delta_j(c) = 0$ *in the sense of distributions for any constant c. Furthermore, the assumption* $u \in \mathcal{S}_0$ *(or* $u \in \mathcal{S}$*) is also inessential and only serves to guarantee* $F(u)$ *is a well defined tempered distribution, so we may take the Triebel-Lizorkin norm of it. This follows from the fact* $|F(u)| \leqslant C[G(u) + G(0)]|u|$ *and* $\|G(u)\|_{L^{p_1}(w_1)} < \infty$*, so* $F(u)$ *is a* $L^{p_1}(w_1)$ *function, hence a tempered distribution. If we knew that* $G(0) = 0$ *for example, then we could allow for* $u \in \mathcal{S}'/\mathcal{P} \cap L^\infty$ *(or* $u \in \mathcal{S}' \cap L^\infty$*) and still obtain that* $F(u)$ *is a well defined tempered distribution. This is a detail overlooked in the existing literature.*

A weighted fractional chain rule for Lebesgue spaces is given in [10] with a similar smoothness condition of Theorem 1.1. However, in [10] it is assumed that $1 < p_1, p_2, p < \infty$ and $w_1 \in A_{p_1}$, $w_2 \in A_{p_2}$ and importantly $w := w_1^{\frac{p}{p_1}} w_2^{\frac{p}{p_2}} \in A_p$. Nevertheless, Theorem 1.1 can hold even when $w \notin A_p$. For example suppose that $p > 1$ and $q = 2$ so that the weighted Triebel-Lizorkin norm reduces to the weighted Hardy norm, which controls the Lebesgue norm. It is well known that $|\cdot|^a \in A_\rho$ if and only if $\rho \in (-n, n(\rho - 1))$. Let $w_1 = |x|^{n(\tau_{w_1}-1)}, w_2 = |x|^{n(\tau_{w_2}-1)}$, then

$$w(x) = |x|^{n(\tau_{w_1}-1)\frac{p}{p_1}}|x|^{n(\tau_{w_2}-1)\frac{p}{p_2}} = |x|^{n(p[\frac{\tau_{w_1}}{p_1}+\frac{\tau_{w_2}}{p_2}]-1)}.$$

Choosing for example $p_1 = 1.5, \tau_{w_1} = 1.1, p_2 = 4, \tau_{w_2} = 2$ gives $\frac{\tau_{w_1}}{p_1} + \frac{\tau_{w_2}}{p_2} > 1$, hence $w \notin A_p$. But if $D^s(F(u))$ is a function then (1.2) implies

$$\|D^sF(u)\|_{L^p(w)} \lesssim \|D^sF(u)\|_{H^p(w)} \lesssim \|F(u)\|_{\dot{F}^s_{p,2}(w)} \lesssim \|G(u)\|_{L^{p_1}(w_1)}\|D^su\|_{L^{p_2}(w_2)}$$

for sufficiently large $s < 1$.

We now consider a special case of Theorem 1.1. Let $p \leqslant 1$, $p_2 \leqslant q = 2$ and $w_1, w_2 \in A_1$. Then the lower bound on the smoothness index becomes $0 \leqslant n(\frac{1}{p_2} - \frac{1}{p_1'}) = n(\frac{1}{p} - 1) < s$. This is the same lower bound for s in the unweighted Kato-Ponce inequality [5]. This discussion in conjunction with the reduction of the Triebel-Lizorkin norm to the Hardy space norm when $q = 2$ provides the following corollary.

Corollary 1.1.1. *Let* $0 < s < 1$, $0 < p \leqslant 1$, $1 < p_1 \leqslant \infty$, $0 < p_2 \leqslant 2$, $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$, $w_1, w_2 \in A_1$, $w = w_1^{\frac{p}{p_1}} w_1^{\frac{p}{p_2}}$ *and* $n(1/p - 1) < s$ *and* $u \in \mathcal{S}_0$ *(or* $u \in \mathcal{S}$*). Let* $F : \mathbb{C} \to \mathbb{C}$ *and* $G : \mathbb{C} \to [0, \infty)$ *such that* $F(0) = 0$ *and* $G(u) \in L^{p_1}(w_1)$*. Furthermore, suppose that for* $x, y \in \mathbb{C}$ *that*

$$|F(x) - F(y)| \leqslant [G(x) + G(y)]|x - y|,$$

then

$$\|D^sF(u)\|_{H^p(w)} \leqslant C_{n,s,p_1,p_2,w_1,w_2}\|G(u)\|_{L^{p_1}(w_1)}\|D^su\|_{H^{p_2}(w_2)}. \tag{1.5}$$

Furthermore, if $0 < s \leqslant n(1/p - 1)$ *then* (1.5) *can fail. If* $u \in \mathcal{S}$ *then*

$$\|J^sF(u)\|_{h^p(w)} \leqslant C_{n,s,p_1,p_2,w_1,w_2}\Big(\|F(u)\|_{L^p(w)} + \|G(u)\|_{L^{p_1}(w_1)}\|J^su\|_{h^{p_2}(w_2)}\Big).$$

Remark 1.1.2. *Corollary 1.1.1 implies a fractional chain rule for Lebesgue spaces for* $p \leqslant 1$ *when the tempered distributions* $D^s(F(u))$ *and* $J^s(F(u))$*, defined by their actions of* $\mathcal{S}_0$ *and* $\mathcal{S}$ *respectively, correspond to functions. Section 7 addresses the sharpness of the lower bound on s in* (1.5) *for the unweighted case.*

Assuming a stronger smoothness condition like that in the original fractional chain rule of Christ and Weinstein [3] we can allow for more general weights.

Corollary 1.1.2. *Let* $0 < s < 1$, $0 < p < \infty$, $1 < p_1 \leqslant \infty$, $0 < p_2 < \infty$, $p_2 \leqslant q < \infty$, $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$, $w_1 \in A_{p_1}$, $w_2 \in A_\infty$ $w = w_1^{\frac{p}{p_1}} w_2^{\frac{p}{p_2}}$, $s > n(\max\left(\frac{\tau_{w_2}}{p_2}, \frac{1}{q}, 1\right) - 1)$ *and* $u \in \mathcal{S}_0(\mathbb{R}^n)$. *Furthermore, suppose that* $F \in C^1(\mathbb{C})$, $F(0) = 0$ *and* $F'(u) \in L^{p_1}(w_1)$ *then*

$$\|F(u)\|_{\dot{F}^s_{p,q}(w)} \leqslant C_{n,s,q,p_1,p_2,w_1,w_2} \left\|F'(u)\right\|_{L^{p_1}(w_1)} \|u\|_{\dot{F}^s_{p_2,q}(w_2)}.$$

Using the fact $F(0) = 0$ and the fundamental theorem of calculus it can easily be seen that $F(u)$ is in $L^{p_1}(w_1)$, hence is a well defined tempered distribution. It is worth pointing out that in Corollary 1.1.2 for the case $p_1 = \infty$ the lower bound for the smoothness index becomes the same lower bound as the Muckenhoupt weighted Kato-Ponce inequality for Triebel-Lizorkin spaces [15].

2. Notation

Much of the notation is adopted from [15]. We denote by M the uncentered Hardy-Littlewood maximal function with respect to cubes. For a locally integrable function g and $t > 0$, the maximal operator M_t is given by $M_t(g) := M(|g|^t)^{\frac{1}{t}}$. 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 $B \lesssim A$. For a set $E \subset \mathbb{R}^n$ we denote χ_E to be the characteristic function of E.

Let $\widehat{\Psi^\star}(\xi)$ be a Schwartz function supported in $4^{-1} \leqslant |\xi| \leqslant 4$, 1 on $4^{-1} \leqslant |\xi| \leqslant 4$ and denote $\Delta_j^\star$ to be convolution with $2^{jn}\Psi^\star(2^j\cdot)$. Notice that $\widehat{\Psi}\widehat{\Psi^\star} = \widehat{\Psi}$ and $\Delta_j^\star \Delta_j = \Delta_j$. For $f \in \mathcal{S}_0$ we have the identity $\sum_{j\in\mathbb{Z}} \Delta_j f = f$. Also note that $\int \Psi = 0$, in fact Ψ has vanishing moments of all orders. The operator S_0 is defined to be convolution with Φ, and the operator $\Delta_{\geqslant k}$ is defined to be $\sum_{j\geqslant k} \Delta_j$.

A Muckenhoupt weight or A_p weight is a non-negative locally integrable function w on $\mathbb{R}^n$ such that for $1 < p < \infty$ and for all cubes Q in $\mathbb{R}^n$ with sides parallel to the axes, we have

$$[w]_{A_p} := \sup_Q \left(\frac{1}{|Q|}\int_Q w(x)\,dx\right)\left(\frac{1}{|Q|}\int_Q w(x)^{-\frac{1}{p-1}}\,dx\right)^{p-1} < \infty.$$

We say $w \in A_1$ if

$$[w]_{A_1} := \sup_Q \left(\frac{1}{|Q|}\int_Q w(x)\,dx\right)\left\|w^{-1}\right\|_{L^\infty} < \infty.$$

The space $L^p(w)$, is defined as the set of Lebesgue measurable functions on $\mathbb{R}^n$ such that

$$\|f\|_{L^p(w)} := \left(\int_{\mathbb{R}^n} |f(x)|^p w(x)\,dx\right)^{1/p} < \infty.$$

For $p = \infty$, Lebesgue measure and wdx are mutually absolutely continuous, thus the essential supremum with respect to wdx and Lebesgue measure are the same, hence $\|\cdot\|_{L^\infty} = \|\cdot\|_{L^\infty(w)}$. Muckenhoupt's Theorem states that if $1 < p < \infty$ and $f \in L^p(w)$ then, $w \in A_p$ if and only if $\|M(f)\|_{L^p(w)} \leqslant C_{p,n,[w]_{A_p}}\|f\|_{L^p(w)}$. Muckenhoupt weights are nested, that is if $1 \leqslant p_1 < p_2 < \infty$ then $A_{p_1} \subset A_{p_2}$. As stated in the introduction for $w \in A_\infty := \bigcup_{1\leqslant p} A_p$ we define $\tau_w = \inf\{\tau \in (1,\infty) : w \in A_\tau\}$.

3. Preliminaries

The inequality in (3.1) is an extension of a lemma from Taylor's book [[14], Lemma 4.2], which was given for $q = 2$ and $a = -\infty$. A proof of the inequality in (3.2) can be found in [[15], Lemma A.3] with $\tau = \theta - s$.

Proposition 3.1. *Let* $\{a_k\}_{k\in\mathbb{Z}}$ *be a sequence of non-negative real numbers,* $0 < q < \infty$ *and* $a \in \{0, -\infty\}$ *then*

$$\left(\sum_{j>a}\left(2^{js}\sum_{a<k<j} 2^{k-j}a_k\right)^q\right)^{\frac{1}{q}} \lesssim \left(\sum_{j>a}(2^{js}a_j)^q\right)^{\frac{1}{q}} \quad \text{if } 0 \leqslant s < 1 \tag{3.1}$$

$$\left(\sum_{j>a}\left(2^{js}\sum_{j\leqslant k} 2^{(k-j)\theta}a_k\right)^q\right)^{\frac{1}{q}} \lesssim \left(\sum_{j>a}(2^{js}a_j)^q\right)^{\frac{1}{q}} \quad \text{if } 0 \leqslant \theta < s. \tag{3.2}$$

Proof. We only consider the case $a = 0$, the proof for $a = -\infty$ follows by a similar argument. Suppose that $q \leqslant 1$, then

$$\begin{aligned}\left(\sum_{j\geqslant 1}\Big(2^{js}\sum_{1\leqslant k<j}2^{k-j}a_k\Big)^q\right)^{\frac{1}{q}} &= \left(\sum_{j\geqslant 1}\Big(\sum_{l=1}^{j-1}2^{js}2^{-l}a_{j-l}\Big)^q\right)^{\frac{1}{q}}\\ &\leqslant \left(\sum_{j\geqslant 1}\sum_{l=1}^{j-1}2^{jsq}2^{-lq}a_{j-l}^q\right)^{\frac{1}{q}}\\ &= \left(\sum_{l=1}^{\infty}2^{lq(s-1)}\sum_{j\in\mathbb{Z}}2^{(j-l)sq}a_{j-l}^q\chi_{\mathbb{N}}(j-l)\right)^{\frac{1}{q}}\\ &\lesssim \Big(\sum_{j\geqslant 1}(2^{js}a_j)^q\Big)^{\frac{1}{q}}.\end{aligned}$$

Now suppose $q > 1$, then

$$\begin{aligned}\left(\sum_{j\geqslant 1}\Big(2^{js}\sum_{1\leqslant k<j}2^{k-j}a_k\Big)^q\right)^{\frac{1}{q}} &= \left(\sum_{j\geqslant 1}\Big(\sum_{l=1}^{j-1}2^{js}2^{-l}a_{j-l}\Big)^q\right)^{\frac{1}{q}}\\ &= \left(\sum_{j\geqslant 1}\Big(\sum_{l=1}^{j-1}2^{l(s-1)/2}2^{l(s-1)/2}2^{(j-l)s}a_{j-l}\Big)^q\right)^{\frac{1}{q}}\\ &\lesssim \Big(\sum_{j\geqslant 1}\sum_{l=1}^{j-1}(2^{l(s-1)/2}2^{(j-l)s}a_{j-l})^q\Big)^{\frac{1}{q}}\\ &\lesssim \Big(\sum_{l=1}^{\infty}2^{lq(s-1)/2}\sum_{j\geqslant 1}(2^{(j-l)s}a_{j-l}\chi_{\mathbb{N}}(j-l))^q\Big)^{\frac{1}{q}}\\ &\lesssim \Big(\sum_{j\geqslant 1}(2^{js}a_j)^q\Big)^{\frac{1}{q}}.\end{aligned}$$

where in the first inequality we applied Hölder's inequality.

□

We will regularly use the fact that a locally integrable function convolved with a L^1 dilate of a radially decreasing integrable function is pointwise bounded by the Hardy-Littlewood maximal operator [[6], Theorem 2.1.10]. A direct application of this fact provides the following result.

Lemma 3.2 ([16], Lemma 2.6)**.** *Let $0 < t < \infty$, $f \in L^1_{loc}(\mathbb{R}^n)$, $A > 0$ and $a > \frac{n}{t}$ then for $x \in \mathbb{R}^n$*

$$\big\|(1 + |A\cdot|)^{-a}f(x - \cdot)\big\|_{L^t} \leqslant C_{n,t}A^{-\frac{n}{t}}\mathcal{M}_t(f)(x).$$

The following is the celebrated Peetre's lemma, an essential estimate when dealing with function spaces.

Lemma 3.3 ([16], Theorem 2.10)**.** *Let $0 < t < \infty$, and u be a function on $\mathbb{R}^n$ whose distributional Fourier transform satisfies $\operatorname{supp}(\widehat{u}) \subset B(0,k)$, then*

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

where the constant is independent of k.

The proceeding lemma is a slight modification of a corollary from [[16], Corollary 2.12]. As pointed out in the appendix of [15] one can allow for $A > 0$, rather than $A \geqslant 1$ as originally stated in the corollary. Also there is a typo, $A^{-\frac{n}{t}}$ should be replaced by $A^{-\frac{n}{\eta'}}$. The proof is essentially the same as that given in [16], with the only addition being the function h. The proof is provided for the reader's convenience.

Lemma 3.4. *Suppose $1 \leqslant \eta \leqslant \infty$ and $0 < t \leqslant \eta'$ where η and η' are related by $\frac{1}{\eta} + \frac{1}{\eta'} = 1$. Let $A > 0$, $R \geqslant 1$, $h \geqslant 0$ and $a > \frac{n}{t}$. If u and φ are functions on $\mathbb{R}^n$ and the distributional Fourier transform of u satisfies* $\operatorname{supp}(\widehat{u}) \subset B(0, AR)$, *then*

$$\int_{\mathbb{R}^n} h(y)|u(y)||\varphi(x-y)|\, dy \lesssim R^{n(\frac{1}{t}-\frac{1}{\eta'})} A^{-\frac{n}{\eta'}} \|h(x-\cdot)(1+|A\cdot|)^a \varphi\|_{L^\eta}\, \mathcal{M}_t(u)(x)$$

where the implicit constant is independent of A, R, φ and u.

Proof. Observe,

$$\begin{aligned}
&\int_{\mathbb{R}^n} h(y)|u(y)||\varphi(x-y)|\, dy \\
&= \int \frac{|u(x-y)|}{(1+|Ay|)^a} |h(x-y)|(1+|Ay|)^a \varphi(y)\, dy \\
&\leqslant \left\| \frac{|u(x-\cdot)|}{(1+|A\cdot|)^a} \right\|_{L^{\eta'}} \|h(x-\cdot)(1+|A\cdot|)^a\varphi\|_{L^\eta} \\
&\leqslant \|h(x-\cdot)(1+|A\cdot|)^a\varphi\|_{L^\eta} \left\| \frac{|u(x-\cdot)|}{(1+|A\cdot|)^a} \right\|_{L^t}^{\frac{t}{\eta'}} \Big(\sup_{y\in\mathbb{R}^n} \frac{|u(x-y)|}{(1+|Ay|)^a} \Big)^{1-\frac{t}{\eta'}} \\
&\lesssim \|h(x-\cdot)(1+|A\cdot|)^a\varphi\|_{L^\eta} \left(A^{-\frac{n}{t}} \mathcal{M}_t(u)(x)\right)^{\frac{t}{\eta'}} \Big(\sup_{y\in\mathbb{R}^n} \frac{R^{\frac{n}{t}}|u(x-y)|}{(R+|RAy|)^{\frac{n}{t}}} \Big)^{1-\frac{t}{\eta'}} \qquad (3.3) \\
&\leqslant \|h(x-\cdot)(1+|A\cdot|)^a\varphi\|_{L^\eta} A^{-\frac{n}{\eta'}} R^{\frac{n}{t}(1-\frac{t}{\eta'})} \left(\mathcal{M}_t(u)(x)\right)^{\frac{t}{\eta'}} \Big(\sup_{y\in\mathbb{R}^n} \frac{|u(x-y)|}{(1+|ARy|)^{\frac{n}{t}}} \Big)^{1-\frac{t}{\eta'}} \\
&\lesssim \|h(x-\cdot)(1+|A\cdot|)^a\varphi\|_{L^\eta} A^{-\frac{n}{\eta'}} R^{\frac{n}{t}(1-\frac{t}{\eta'})} \left(\mathcal{M}_t(u)(x)\right)^{\frac{t}{\eta'}} \Big(\mathcal{M}_t(u)(x)\Big)^{1-\frac{t}{\eta'}} \qquad (3.4) \\
&= R^{n(\frac{1}{t}-\frac{1}{\eta'})} A^{-\frac{n}{\eta'}} \|h(x-\cdot)(1+|A\cdot|)^a\varphi\|_{L^\eta}\, \mathcal{M}_t(u)(x)
\end{aligned}$$

where in (3.3) we applied Lemma 3.2, and in (3.4) we applied Lemma 3.3 □

In [15] Naibo and Thomson only used the case $\eta = \infty$, our use however, will necessitate an optimal value of η.

The following weighted vector-valued Fefferman-Stein inequality will be of great use: If $0 < p < \infty$, $0 < q < \infty$, $w \in A_\infty$, and $0 < t < \min(p/\tau_w, q)$, then for all sequences $\{f_j\}_{j=1}^\infty$ of locally integrable functions defined on $\mathbb{R}^n$, we have

$$\left\| \left(\sum_j |M_t(f_j)|^q \right)^{\frac{1}{q}} \right\|_{L^p(w)} \lesssim \left\| \left(\sum_j |f_j|^q \right)^{\frac{1}{q}} \right\|_{L^p(w)}, \qquad (3.5)$$

where the implicit constant depends on t, p, q, and w.

4. Proof of Theorem 1.1

The proof is inspired by the method in [[14], Proposition 5.1] for L^p spaces with $p > 1$. However, our argument is more delicate and relies on the use of Lemma 3.3 and Lemma 3.4. First suppose $p_1 < \infty$.

To bound

$$\|F(u)\|_{\dot{F}^s_{p,q}(w)} = \left\| \left(\sum_{j\in\mathbb{Z}} \left(2^{js}|\Delta_j(F(u))|\right)^q \right)^{\frac{1}{q}} \right\|_{L^p(w)} \qquad (4.1)$$

we begin by estimating the Littlewood-Paley operator of the composition. Recall $\int \Psi = 0$, and observe

$$
\begin{aligned}
|\Delta_j F(u)(x)| &= \left| \int_{\mathbb{R}^n} F(u)(y)\Psi_j(x-y)\,dy - F(u)(x)\int_{\mathbb{R}^n} \Psi_j(x-y)\,dy \right| \\
&\leqslant \int_{\mathbb{R}^n} |F(u)(x) - F(u)(y)|\,|\Psi_j(x-y)\,dy| \\
&\lesssim G(u)(x)\int_{\mathbb{R}^n} |u(x)-u(y)|\,|\Psi_j(x-y)|\,dy \qquad (4.2)\\
&\qquad + \int_{\mathbb{R}^n} G(u)(y)|u(x)-u(y)|\,|\Psi_j(x-y)|\,dy. \qquad (4.3)
\end{aligned}
$$

We will first bound (4.3) by making several further decompositions,

$$
\begin{aligned}
&\int_{\mathbb{R}^n} G(u)(y)|u(x)-u(y)|\,|\Psi_j(x-y)|\,dy \\
&\leqslant \sum_{k<j} \int_{\mathbb{R}^n} G(u)(y)|\Delta_k u(x) - \Delta_k u(y)|\,|\Psi_j(x-y)|\,dy + \qquad (4.4)\\
&\qquad \sum_{j\leqslant k} \int_{\mathbb{R}^n} G(u)(y)|\Delta_k u(x) - \Delta_k u(y)|\,|\Psi_j(x-y)|\,dy. \qquad (4.5)
\end{aligned}
$$

4.1. Bound for (4.4)

We estimate (4.4) by further expressing it as

$$
\begin{aligned}
&\sum_{k<j} \int_{|x-y|\leqslant 2^{-k}} G(u)(y)|\Delta_k u(y) - \Delta_k u(x)||\Psi_j(x-y)|dy \qquad (4.6)\\
&+ \sum_{k<j} \int_{|x-y|> 2^{-k}} G(u)(y)|\Delta_k u(y) - \Delta_k u(x)||\Psi_j(x-y)|dy. \qquad (4.7)
\end{aligned}
$$

Fix $t > 0$, and choose $N \in \mathbb{N}$ so that $N - \frac{n}{t} > n$. To estimate (4.6) observe that

$$
\begin{aligned}
|\Delta_k u(y) - \Delta_k u(x)| &= |\Delta_k^\star \Delta_k u(y) - \Delta_k^\star \Delta_k u(x)| \\
&\leqslant \int_{\mathbb{R}^n} |\Delta_k u(z)||\Psi_k^\star(x-z) - \Psi_k^\star(y-z)|dz \\
&= \int_{\mathbb{R}^n} |\Delta_k u(z)| \left| \int_0^1 2^k (\nabla \Psi^\star)_k(\theta y + (1-\theta)x - z)\cdot (y-x)d\theta \right| dz \\
&\lesssim 2^k|y-x| \int_{\mathbb{R}^n}\int_0^1 |\Delta_k u(z)| \frac{2^{kn}}{(1+2^k|\theta(y-x)+x-z|)^N} d\theta dz \\
&\lesssim 2^k|y-x| \int_{\mathbb{R}^n} |\Delta_k u(z)| \frac{2^{kn}(1+2^k|y-x|)^N}{(1+2^k|x-z|)^N} dz, \qquad (4.8)
\end{aligned}
$$

where in (4.8) we used that $\theta \leqslant 1$ and the following simple inequality for $v_1, v_2 \in \mathbb{R}^n$

$$
\frac{1}{1+|v_2+v_1|} \leqslant \frac{1+|v_1|}{1+|v_2|}.
$$

Recalling that in (4.6) we are assuming that $|x-y| \leqslant 2^{-k}$ we obtain (4.8) is bounded by a constant multiple of

$$
\begin{aligned}
&2^k|y-x| \int_{\mathbb{R}^n} |\Delta_k u(x-z)| \frac{2^{kn}}{(1+2^k|z|)^N} dz \\
&= 2^k|y-x| \int_{\mathbb{R}^n} \frac{|\Delta_k u(x-z)|}{(1+2^k|z|)^{\frac{n}{t}}} (1+2^k|z|)^{\frac{n}{t}} \frac{2^{kn}}{(1+2^k|z|)^N} dz
\end{aligned}
$$

$$= 2^k|y-x|M_t(\Delta_k u)(x)\int_{\mathbb{R}^n} 2^{kn}\frac{(1+2^k|z|)^{\frac{n}{t}}}{(1+2^k|z|)^N}dz \tag{4.9}$$
$$\lesssim 2^k|y-x|M_t(\Delta_k u)(x)$$

where in (4.9) we used Lemma 3.3. Thus (4.6) can be written as

$$= \sum_{k<j}\int_{|x-y|\leqslant 2^{-k}} G(u)(y)2^k|y-x|M_t(\Delta_k u)(x)|\Psi_j(x-y)|dy$$
$$\leqslant \sum_{k<j} 2^{k-j}M_t(\Delta_k u)(x)\int_{\mathbb{R}^n} G(u)(y)|2^jy-2^jx||\Psi_j(x-y)|dy$$
$$\lesssim M(G(u))(x)\sum_{k<j}2^{k-j}M_t(\Delta_k u)(x). \tag{4.10}$$

Now to bound (4.7)

$$\sum_{k<j}\int_{|x-y|>2^{-k}} G(u)(y)|\Delta_k u(y)-\Delta_k u(x)||\Psi_j(x-y)|dy$$
$$\leqslant \sum_{k<j}|\Delta_k u(x)|\int_{|x-y|>2^{-k}} G(u)(y)|\Psi_j(x-y)|dy \tag{4.11}$$
$$+\sum_{k<j}\int_{|x-y|>2^{-k}} G(u)(y)|\Delta_k u(y)||\Psi_j(x-y)|dy. \tag{4.12}$$

The sum in (4.11) can be estimated by

$$\sum_{k<j} M_t(\Delta_k u)(x)\int_{|x-y|>2^{-k}} G(u)(y)\frac{2^{jn}}{(1+2^j|x-y|)^{n+1}}(1+2^j|x-y|)^{-1}dy$$
$$\lesssim M(G(u))(x)\sum_{k<j}2^{k-j}M_t(\Delta_k u)(x).$$

The second sum in (4.12) is estimated similarly,

$$\sum_{k<j}\int_{|x-y|>2^{-k}} G(u)(y)|\Delta_k u(y)||\Psi_j(x-y)|dy$$
$$\lesssim \sum_{k<j}\int_{|x-y|>2^{-k}} G(u)(y)|\Delta_k u(y)|\frac{2^{jn}}{(1+2^j|x-y|)^N}(1+2^j|x-y|)^{-1}dy$$
$$\leqslant \sum_{k<j}2^{k-j}\int_{|x-y|>2^{-k}} G(u)(x-y)\frac{|\Delta_k u(x-y)|}{(1+2^k|y|)^{\frac{n}{t}}}(1+2^k|y|)^{\frac{n}{t}}\frac{2^{jn}}{(1+2^j|y|)^N}dy$$
$$\lesssim \sum_{k<j}2^{k-j}M_t(\Delta_k u)(x)\int_{\mathbb{R}^n} G(u)(x-y)(1+2^k|y|)^{\frac{n}{t}}\frac{2^{jn}}{(1+2^j|y|)^N}dy \tag{4.13}$$
$$\lesssim M(G(u))(x)\sum_{k<j}2^{k-j}M_t(\Delta_k u)(x) \tag{4.14}$$

where in (4.13) we used Lemma 3.3 and in (4.14) we used $2^k<2^j$. So (4.4) is pointwise bounded above by $M(G(u)(x))\sum_{k<j}2^{k-j}M_t(\Delta_k u)(x)$.

4.2. Bound for (4.5)

To estimate (4.5) we write

$$\sum_{j\leqslant k}\int_{\mathbb{R}^n} G(u)(y)|\Delta_k u(y)-\Delta_k u(x)||\Psi_j(x-y)|dy$$

$$\leqslant \sum_{j\leqslant k} |\Delta_k u(x)| \int_{\mathbb{R}^n} G(u)(y)|\Psi_j(x-y)|dy \tag{4.15}$$

$$+ \sum_{j\leqslant k} \int_{\mathbb{R}^n} G(u)(y)|\Delta_k u(y)||\Psi_j(x-y)|dy. \tag{4.16}$$

To bound (4.15) observe,

$$\begin{aligned}
&\sum_{j\leqslant k} |\Delta_k^\star \Delta_k u(x)| \int_{\mathbb{R}^n} G(u)(y)|\Psi_j(x-y)|dy \\
&\sum_{j\leqslant k} M_t(\Delta_k u)(x) \int_{\mathbb{R}^n} G(u)(y)|\Psi_j(x-y)|dy \\
&\lesssim M(G(u))(x) \sum_{j\leqslant k} M_t(\Delta_k u)(x).
\end{aligned}$$

4.3. Bound for the bad term (4.16)

The term in (4.16) i.e.

$$\sum_{j\leqslant k} \int_{\mathbb{R}^n} G(u)(y)|\Delta_k u(y)||\Psi_j(x-y)|\, dy$$

is the most delicate. If $\frac{p}{\tau_w} > 1$ then here we would pointwise bound by $\sum_{j\leqslant k} M(G(u)|\Delta_k u|)$, however this method breaks down for $p/\tau_w \leqslant 1$. Furthermore, a direct application of Lemma 3.3 like in (4.11) will not work here since $j \leqslant k$.

Applying Lemma 3.4 with $R = 2^{k-j} \geqslant 1$, $A = 2^j$, $\varphi = \Psi_j$, $u_k = \Delta_k u$ and $1 \leqslant \eta \leqslant \infty$, $0 < r \leqslant \eta'$, $a > n/r$, where r and η are to be determined. A summand of (4.16) is bounded by a constant multiple of

$$\begin{aligned}
&2^{(k-j)n(1/r-1/\eta')} M_r(\Delta_k u)(x) 2^{-j\frac{n}{\eta'}} \left\| G(u)(x-\cdot)(1+|2^j\cdot|)^a \Psi_j \right\|_{L^\eta} \\
&= 2^{(k-j)n(1/r-1/\eta')} M_r(\Delta_k u)(x) \left(2^{jn} \int_{\mathbb{R}^n} (1+|2^j y|)^{\eta a} |\Psi(2^j y)|^\eta (G(u)(x-y))^\eta \right)^{\frac{1}{\eta}} \\
&\lesssim 2^{(k-j)n(1/r-1/\eta')} M_r(\Delta_k u)(x) M_\eta(G(u))(x).
\end{aligned}$$

4.4. Applying the Triebel-Lizorkin norm

We have estimated (4.3) above by a constant multiple of

$$M(G(u))(x) \sum_{k<j} 2^{k-j} M_t(\Delta_k u)(x) \tag{4.17}$$

$$+ M(G(u))(x) \sum_{j\leqslant k} M_t(\Delta_k u)(x) \tag{4.18}$$

$$+ M_\eta(G(u))(x) \sum_{j\leqslant k} 2^{(k-j)n(1/r-1/\eta')} M_r(\Delta_k u)(x). \tag{4.19}$$

We now apply the homogeneous Triebel-Lizorkin norm to (4.17), (4.18), (4.19). Starting with (4.17) and choosing $0 < t < \min(p_2/\tau_{w_2}, q)$ we obtain

$$\left\| M(G(u)) 2^{js} \sum_{k<j} 2^{k-j} M_t(\Delta_k u) \right\|_{L^p(\ell^q)(w)}$$

$$\leqslant \|M(G(u))\|_{L^{p_1}(w_1)} \left\| \left(\sum_{j\in\mathbb{Z}} \Big(2^{js} \sum_{k<j} 2^{k-j} M_t(\Delta_k u) \Big)^q \right)^{\frac{1}{q}} \right\|_{L^{p_2}(w_2)} \tag{4.20}$$

$$\lesssim \|G(u)\|_{L^{p_1}(w_1)} \left\| \left(\sum_{j\in\mathbb{Z}} \Big(2^{js} M_t(\Delta_j u) \Big)^q \right)^{\frac{1}{q}} \right\|_{L^{p_2}(w_2)} \tag{4.21}$$

$$\lesssim \|G(u)\|_{L^{p_1}(w_1)} \left\| \left(\sum_{j\in\mathbb{Z}} \Big(2^{js} |\Delta_j u| \Big)^q \right)^{\frac{1}{q}} \right\|_{L^{p_2}(w_2)} \tag{4.22}$$

$$\lesssim \|G(u)\|_{L^{p_1}(w_1)} \|u\|_{\dot{F}^s_{p_2,q}(w_2)}$$

where in (4.20) we applied Hölder's inequality, in (4.21) we applied (3.1), and in (4.22) we applied the Fefferman-Stein inequality. We obtain the same normed estimate for (4.18) by a similar method used to bound (4.17); namely Hölder's inequality, Lemma 3.1 with $\theta = 0$, followed the Fefferman-Stein inequality.

4.5. Bounding the worse term (4.19)

To bound (4.19), first choose r and η. Recall by hypothesis

$$0 \leqslant n\Big(\max\big(\frac{\tau_{w_2}}{p_2}, \frac{1}{q}\big) - \frac{p_1 - \tau_{w_1}}{p_1} \Big) < s.$$

Note that

$$\Big(\frac{p_1}{\tau_{w_1}}\Big)' = \frac{p_1}{p_1 - \tau_{w_1}},$$

also since $p_1 > 1$ and $w_1 \in A_{p_1}$ we have $\tau_{w_1} < p_1$, by the self improving property of Muckenhoupt weights. Pick $\epsilon > 0$ small enough so that

$$0 < n\Big(\frac{1}{\min\big(\frac{p_2}{\tau_{w_2}}, q\big) - \epsilon} - \frac{1}{\big(\frac{p_1}{\tau_{w_1}} - \epsilon\big)'} \Big) < s,$$

and let $\eta = \frac{p_1}{\tau_{w_1}} - \epsilon > 1$ and $r = \min\big(\frac{p_2}{\tau_{w_2}}, q\big) - \epsilon > 0$. Note that $0 < r < \eta'$ which is required for Lemma 3.4. Observe,

$$\left\| M_\eta(G(u))(x) 2^{js} \sum_{j\leqslant k} 2^{(k-j)n(1/r - 1/\eta')} M_r(\Delta_k u)(x) \right\|_{L^p(\ell^q)(w)}$$

$$\leqslant \big\|M_\eta(G(u))\big\|_{L^{p_1}(w_1)} \left\| \left(\sum_{j\in\mathbb{Z}} \Big(2^{js} \sum_{j\leqslant k} 2^{(k-j)n(1/r-1/\eta')} M_r(\Delta_k u)(x) \Big)^q \right)^{\frac{1}{q}} \right\|_{L^{p_2}(w_2)} \tag{4.23}$$

$$\lesssim \big\|M_\eta(G(u))\big\|_{L^{p_1}(w_1)} \left\| \left(\sum_{j\in\mathbb{Z}} \Big(2^{js} M_r(\Delta_j u) \Big)^q \right)^{\frac{1}{q}} \right\|_{L^{p_2}(w_2)} \tag{4.24}$$

$$\lesssim \|G(u)\|_{L^{p_1}(w_1)} \left\| \left(\sum_{j\in\mathbb{Z}} \Big(2^{js} |\Delta_j u| \Big)^q \right)^{\frac{1}{q}} \right\|_{L^{p_2}(w_2)} \tag{4.25}$$

$$\lesssim \|G(u)\|_{L^{p_1}(w_1)} \|u\|_{\dot{F}^s_{p_2,q}(w_2)}$$

where in (4.23) we applied Hölder's inequality, in (4.24) we applied (3.2) with $0 < \theta = n(1/r - 1/\eta') < s$, and in (4.25) we applied the Fefferman-Stein inequality and the boundedness of the Hardy-Littlewood maximal operator. Notice we may apply the Fefferman-Stein inequality as $r < \min(p_2/\tau_{w_2}, q)$. Furthermore, $\eta < p_1/\tau_{w_1}$ so M_η maps $L^{p_1}(w_1)$ to $L^{p_1}(w_1)$. This finishes the desired estimate for (4.3).

4.6. Bound for (4.2)

To finish the proof we turn our attention to (4.2) i.e.

$$G(u)(x)\int_{\mathbb{R}^n} |u(x)-u(y)|\,|\Psi_j(x-y)|\,dy.$$

From the proof of (4.3) we see if we replace $G(u)(y)$ by 1, then (4.2) is bounded above by a constant multiple of

$$\begin{aligned} &G(u)(x)\sum_{k<j} 2^{k-j} M_t(\Delta_k u)(x) \\ &+ G(u)(x)\sum_{j\leqslant k} M_t(\Delta_k u)(x) \\ &+ G(u)(x)\sum_{j\leqslant k} 2^{(k-j)n(1/r-1/\eta')} M_r(\Delta_k u)(x) 2^{-j\frac{n}{\eta'}} \left\|(1+|2^j\cdot|)^a \Psi_j\right\|_{L^\eta}. \end{aligned}$$

which all give the desired bound by the same proof used for (4.17), (4.18), and (4.19) respectively; that is Hölder's inequality, Lemma 3.1, followed by the Fefferman-Stein inequality.

All that remains is the case $p_1 = \infty$. The only part of the proof that this effects is the estimate for the bad term (4.16) i.e.

$$2^{(k-j)n(1/r-1/\eta')} M_r(\Delta_k u)(x) 2^{-j\frac{n}{\eta'}} \left\|G(u)(x-\cdot)(1+|2^j\cdot|)^a \Psi_j\right\|_{L^\eta}.$$

Recall the smoothness condition in this case is $0 \leqslant n\Big(\max\big(\frac{\tau_{w_2}}{p_2}, \frac{1}{q}\big) - 1\Big) < s$. Choosing $\eta = \infty$ and r close enough to $\max\big(\frac{\tau_{w_2}}{p_2}, \frac{1}{q}\big)$ gives $0 < \theta = n(\frac{1}{r} - 1) < s$. We can then applying Lemma 3.1 and the Fefferman-Stein inequality analogously to subsection 4.6. ☐

5. Inhomogeneous Triebel-Lizorkin chain rule

Most of the proof reduces to the homogeneous version. We need to bound

$$\|F(u)\|_{F^s_{p,q}(w)} = \|S_0(F(u))\|_{L^p(w)} + \left\|\left(\sum_{j\geqslant 1}\Big(2^{js}|\Delta_j(F(u))|\Big)^q\right)^{\frac{1}{q}}\right\|_{L^p(w)}. \tag{5.1}$$

Using the estimates (4.2) and (4.3) along with the identity $g = S_0 g + \Delta_{k\geqslant 1} g$ for a tempered distribution g, we obtain $|\Delta_j F(u)|$ is bounded above by a constant multiple of

$$G(u)(x)\int_{\mathbb{R}^n} |S_0u(x)-S_0u(y)|\,|\Psi_j(x-y)|\,dy \tag{5.2}$$

$$+\sum_{k\geqslant 1} G(u)(x)\int_{\mathbb{R}^n} |\Delta_k u(x)-\Delta_k u(y)|\,|\Psi_j(x-y)|\,dy \tag{5.3}$$

$$+\int_{\mathbb{R}^n} G(u)(y)|S_0u(x)-S_0u(y)|\,|\Psi_j(x-y)|\,dy \tag{5.4}$$

$$+\sum_{k\geqslant 1}\int_{\mathbb{R}^n} G(u)(y)|\Delta_k u(x)-\Delta_k u(y)|\,|\Psi_j(x-y)|\,dy. \tag{5.5}$$

Bounding (5.5) is very similar to bounding (4.3); that is we break up the sum of $k \in \mathbb{N}$ into two parts $1 \leqslant k < j$ and $j \leqslant k$. The proof then follows analogously to treatment of (4.4) and (4.5) with the only difference coming from the Lemma 3.1, where we use $a = 0$ rather than $a = -\infty$, which does not affect the proof.

Since the only property of the operator Δ_k used is that it is convolution with a L^1 dilate of a Schwartz function (5.4) can be obtained from the same argument in subsection 4.1 where $k = 0$. From this we obtain (5.4) is pointwise bounded by a constant multiple of

$$M(G(u))(x)2^{-j}M_t(S_0u)(x).$$

Since $s < 1$ applying the TL norm we obtain the desired bound.

The bounds for (5.2) and (5.3) follow from the the proof for (5.4) and (5.5) by setting $G(u)(y) = 1$. This just leaves the $S_0(F(u))$ term. Observe,

$$\begin{aligned}|S_0(F(u))(x)| &\leqslant \int |\Phi(x-y)||F(u)(y)|dy\\ &\leqslant |G(u)(x)|\int |\Phi(x-y)||u(x)-u(y)|dy \qquad (5.6)\\ &\quad + \int |\Phi(x-y)||G(u)(y)||u(x)-u(y)|dy \qquad (5.7)\\ &\quad + |F(u)(x)|\int |\Phi(x-y)|dy. \qquad (5.8)\end{aligned}$$

It is clear that (5.8) will give the bound $\|F(u)\|_{L^p(w)}$. Furthermore, the bound for (5.6) will follow from our bound for (5.7) by setting $G(u(y)) = 1$, hence we focus on (5.7). Using the identity $g = S_0g + \Delta_{\geqslant 1}g$ for a tempered distribution g, (5.7) is further decomposed by

$$\int |\Phi(x-y)||G(u)(y)||S_0u(x) - S_0u(y)|dy \qquad (5.9)$$

$$+ \sum_{k\geqslant 1}\int |\Phi(x-y)||G(u)(y)||\Delta_k u(x) - \Delta_k u(y)|dy. \qquad (5.10)$$

Using Lemma 3.3 and the triangle inequality (5.9) is further bound by

$$|S_0u(x)|\int G(u)(y)|\Phi(x-y)|dy + M_t(S_0(u))(x)\int G(u)(x-y)(1+2|y|)^{\frac{n}{t}}|\Phi(y)|dy$$

for t arbitrarily small. Hence the $L^p(w)$ norm of (5.9) is bounded by a constant multiple of $\|G(u)\|_{L^{p_1}w_1}\|u\|_{F^s_{p,q}}$.

By the triangle inequality we obtain (5.10) is bounded by

$$\begin{aligned}&\sum_{k\geqslant 1}|\Delta_k u(x)|\int |\Phi(x-y)||G(u)(y)|dy + \sum_{k\geqslant 1}\int |\Delta_k u(y)||\Phi(x-y)||G(u)(y)|dy\\ &\lesssim M(G(u))(x)\sum_{k\geqslant 1}|\Delta_k u(x)| + M_\eta(G(u))(x)\sum_{k\geqslant 1}2^{kn(\frac{1}{r}-\frac{1}{\eta'})}M_r(\Delta_k u)(x) \qquad (5.11)\\ &\leqslant \Big(M(G(u))(x) + M_\eta(G(u))(x)\Big)\sum_{k\geqslant 1}2^{kn(\frac{1}{r}-\frac{1}{\eta'})}M_r(\Delta_k u)(x)\\ &\lesssim \Big(M(G(u))(x) + M_\eta(G(u))(x)\Big)\Big(\sum_{j\geqslant 1}\Big(2^{js}\sum_{k\geqslant j}2^{kn(\frac{1}{r}-\frac{1}{\eta'})}M_r(\Delta_k u)(x)\Big)^q\Big)^{\frac{1}{q}}\end{aligned}$$

where in (5.11) we applied Lemma 3.4. The estimate now follows by the same method used in subsection 4.5, with the same choices of r and η. Hence, the $L^p(w)$ norm of (5.10) is bounded by a constant multiple of $\|G(u)\|_{L^{p_1}w_1}\|u\|_{F^s_{p,q}}$.

Lastly if $G(0) = 0$ then we have the bound $|S_0(F(u))(x)| \leqslant \int |\Phi(x-y)|G(u)(y)|u(y)|dy$ which is the same term as in (5.7), and therefore we may drop the $\|F(u)\|_{L^p(w)}$ term in (1.3) to obtain (1.4). ☐

6. Proof of Corollary 1.1.2

The proof of Theorem 1.1.2 is similar to the proof of Theorem 1.1. The improved smoothness condition greatly simplifies the decomposition and avoids the bad term (4.16). The following estimate was shown in

[3],

$$\begin{aligned}|\Delta_j(F(u))(x)| &= \left|\int_{\mathbb{R}^n} F(u)(y)\Psi_j(x-y)\,dy - F(u)(x)\int_{\mathbb{R}^n}\Psi_j(x-y)\,dy\right|\\ &\leqslant \int_{\mathbb{R}^n} |F(u)(x)-F(u)(y)|\,|\Psi_j(x-y)\,dy|\\ &\leqslant \int_{\mathbb{R}^n}\int_0^1 \left|F'(tu(y)+(1-t)u(x))\right||u(x)-u(y)|\,|\Psi_j(x-y)\,dtdy|\\ &\leqslant 2M(F'(u))(x)\int_{\mathbb{R}^n}|u(x)-u(y)|\,|\Psi_j(x-y)|\,dy \qquad (6.1)\end{aligned}$$

where (6.1) follows from a substitution.

By replacing $G(u)$ by $M(F'(u))$ in subsection 4.6 we obtain $|\Delta_j(F(u))|$ is bounded above by a constant multiple of

$$M(F'(u))(x)\sum_{k<j} 2^{k-j} M_t(\Delta_k u)(x) \qquad (6.2)$$

$$+\, M(F'(u))(x)\sum_{j\leqslant k} M_t(\Delta_k u)(x) \qquad (6.3)$$

$$+\, M(F'(u))(x)\sum_{j\leqslant k} 2^{(k-j)n(1/r-1/\eta')} M_r(\Delta_k u)(x) 2^{-j\frac{n}{\eta'}} \left\|(1+|2^j\cdot|)^a\Psi_j\right\|_{L^\eta}. \qquad (6.4)$$

where t, r and η are to be determined. Picking $0 < t < \min(p_2/\tau_{w_2}, q)$ we have (6.2) and (6.3) follow by the same method as (4.17) and (4.18). To bound (6.4) let $\eta = \infty$, so $\eta' = 1$. Recall by assumption $s > n(1/\min(p_2/\tau_{w_2}, q, 1) - 1)$, so choose $0 < r < \min(p_2/\tau_{w_2}, q, 1)$ close enough to $\min(p_2/\tau_{w_2}, q, 1)$ so that $n(1/r-1) < s$. Applying (3.2) with $\theta = n(1/r-1)$ and noting $0 < r < \min(p_2/\tau_{w_2}, q)$ we can apply Hölder's inequality, Lemma 3.1, and the Fefferman-Stein inequality to obtain the desired result.

□

7. Sharpness

We now prove the lower bound on s is sharp for (1.5) where $w_1 = w_2 = 1$. For this we need the following lemma.

Lemma 7.1. *[Lemma 1, [5]] Let $f \in \mathcal{S}(\mathbb{R}^n)$ and $s > 0$, $s \notin 2\mathbb{N}$. If $f(x) \geqslant 0$ for all $x \in \mathbb{R}^n$ and $f \not\equiv 0$, then there exists a $R \gg 1$ such that*

$$|D^s f| \geqslant C_{n,s,f,R}|x|^{-n-s} \text{ for } |x| > R.$$

For $p < 1$ and $0 < s \leqslant n(1/p-1)$ we construct an example such that (1.5) fails. Let $F(z) = z^2$ then $G(z) = |z|$, that is

$$|F(z_1) - F(z_2)| \leqslant \int_0^1 2|tz_1 + (1-t)z_2||z_1 - z_2| \leqslant 2(|z_1| + |z_2|)|z_1 - z_2|.$$

Let $u = \Psi$. Note that $\widehat{\Psi}$ is compactly supported away from the origin so the RHS of (1.5) is finite. Since $F(\Psi) = \Psi^2 \in \mathcal{S}$ we have $D^s(\Psi^2)$ is a L^2 function. So $\|D^s(F(\Psi))\|_{L^p} \lesssim \|D^s(F(\Psi))\|_{H^p}$. By Lemma 7.1 and the fact $s \leqslant n(1/p-1)$ we obtain that $\|D^s(F(\Psi))\|_{L^p} = \infty$. Therefore the LHS of (1.5) is infinite.

References

[1] Bernicot, Frederic; Naibo, Virginia; Maldonado, Diego; Moen, Kabe. *Bilinear Sobolev-Poincaré inequalities and Leibniz-type rules.* J. Geom. Anal. 24 (2014), no. 2, 1144–1180.

[2] Bui, Huy-Qui. *Weighted Besov and Triebel spaces: interpolation by the real method.* Hiroshima Math. J. 12 (1982), no.3, 581–605.

[3] 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.

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

[5] Grafakos, Loukas; Oh, Seungly. *The Kato-Ponce inequality.* Comm. Partial Differential Equations 39 (2014), no. 6, 1128–1157.

[6] Grafakos, Loukas. *Classical Fourier analysis.* Third edition. Graduate Texts in Mathematics, 249. Springer, New York, 2014.

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

[8] 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

[9] Hidano, Kunio; Wang, Chengbo. *Fractional derivatives of composite functions and the Cauchy problem for the nonlinear half wave equation.* Selecta Math. 25 (2019), no.1, Paper No. 2, 28 pp.

[10] Hidano, Kunio; Jiang, Jin-Cheng; Lee, Sanghyuk; Wang, Chengbo. *Weighted fractional chain rule and nonlinear wave equations with minimal regularity.* Rev. Mat. Iberoam. 36 (2020), no.2, 341–356.

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

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

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

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

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

[16] 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 Missouri
Email address: spdthc@umsystem.edu