DIV-CURL SYSTEMS (PART 2)

PO-LAM YUNG

Contents

Last time we saw some remarkable consequences of the following approximation lemma for functions in $\dot{W}^{1,n}(\mathbb{R}^n)$, which can be thought of as some remedy of the fact that the Sobolev space $\dot{W}^{1,n}(\mathbb{R}^n)$ fails to embed into L^∞:

Lemma 1. *For any $\delta > 0$ there exists C_δ such that for any $f \in \dot{W}^{1,n}(\mathbb{R}^n)$ there exists $F \in \dot{W}^{1,n} \cap L^\infty(\mathbb{R}^n)$ satisfying*

$$\begin{cases} \sum_{i=2}^n \|\partial_i (f-F)\|_{L^n} \leq \delta \|\nabla f\|_{L^n} \\ \|\nabla F\|_{L^n} + \|F\|_{L^\infty} \leq C_\delta \|\nabla f\|_{L^n}. \end{cases}$$

We saw that it follows from the following non-linear[1] approximation lemma:

Lemma 2. *There exists $c_n < 1$ such that for $\delta > 0$, there exists C_δ such that for any $f \in \dot{W}^{1,n}(\mathbb{R}^n)$ with $\|\nabla f\|_{L^n} \leq c_n$, there exists $F \in \dot{W}^{1,n} \cap L^\infty(\mathbb{R}^n)$ such that*

$$\begin{cases} \|F\|_{L^\infty} \leq C_\delta \\ \|\nabla F\|_{L^n} \leq C_\delta \|\nabla f\|_{L^n} \\ \sum_{i=2}^n \|\partial_i (f-F)\|_{L^n} \leq \delta \|\nabla f\|_{L^n} + C_\delta \|\nabla f\|_{L^n}^2. \end{cases}$$

Today our goal is to give a complete proof of this second approximation lemma, following Bourgain-Brezis [1]. We shall not take the shortest possible route; rather, we shall try to explain some motivations behind the construction of F, and have some trial and errors on some model constructions before we carry out the actual one.

[1] Note the square of $\|\nabla f\|_{L^n}$ on the right hand side of the last assertion.

1. Motivations and a basic construction

The failure of $\dot{W}^{1,n}$ into L^∞ is the major obstacle that we need to get around. To tackle this we use Bernstein's inequality.

More precisely, let P_j be the Littlewood-Paley projection adapted to frequency 2^j; in fact we shall take a smooth function χ with support in $\{\frac{1}{2} \leq |\xi| \leq 2\}$, such that $\sum_{j=-\infty}^{\infty} \chi(2^{-j}\xi) = 1$ for all $\xi \neq 0$, and define

$$\widehat{P_j f}(\xi) = \chi(2^{-j}\xi)\widehat{f}(\xi).$$

Lemma 3 (Bernstein's inequality)**.** *There exists a constant C_n such that for all $f \in \dot{W}^{1,n}(\mathbb{R}^n)$,*

$$\|P_j f\|_{L^\infty} \leq C_n \|\nabla f\|_{L^n}$$

uniformly for all j.

Proof. One observes that

$$\widehat{P_j f}(\xi) = 2^{-j} \sum_{i=1}^{n} \frac{2^{-j}\xi_i}{(2^{-j}|\xi|)^2} \chi(2^{-j}\xi) \cdot \xi_i \widehat{f}(\xi),$$

from which it follows that there exist Schwartz functions $K^{(i)}$ such that

$$P_j f = \sum_{k=1}^{n} K_j^{(i)} * \partial_i f \tag{1}$$

for all j, where

$$K_j^{(i)}(x) = 2^{j(n-1)} K^{(i)}(2^j x).$$

One only needs to apply Holder's inequality to (1), noting that

$$\|K_j^{(i)}\|_{L^{\frac{n}{n-1}}} = c$$

independent of j. □

Hence if $f = P_j f$ for some j, then one can take $F = P_j f$ and that will be an L^∞ approximation of f verifying the conclusions of Lemma 2.

Now in general,

$$f = \sum_j P_j f,$$

and we cannot sum up $P_j f$ in L^∞ even though each piece is bounded. The general construction will rely on the following algebraic identity, which reads:

Lemma 4. *For any sequence $\{a_i\}$,*

$$1 = \sum_{j=1}^{N} a_j \prod_{1 \leq j' < j} (1 - a_{j'}) + \prod_{j=1}^{N} (1 - a_j).$$

Proof. This is just saying that

$$\begin{aligned}1 &= a_1 + (1 - a_1)\\ &= a_1 + a_2(1 - a_1) + (1 - a_1)(1 - a_2)\\ &= a_1 + a_2(1 - a_1) + a_3(1 - a_1)(1 - a_2) + (1 - a_1)(1 - a_2)(1 - a_3)\\ &= \dots\end{aligned}$$

□

Note that by renaming the indices, one can also write

$$1 = \sum_{j=1}^{N} a_j \prod_{j<j'\leq N} (1 - a_{j'}) + \prod_{j=1}^{N}(1 - a_j).$$

Hence if we have a sequence $\{a_i\}$ of numbers, all of which are non-negative and bounded by 1, then

$$\sum_{j=-\infty}^{\infty} a_j \prod_{j'>j} (1 - a_{j'}) \in [0, 1].$$

Now recall f can be written $f = \sum_j f_j$, if we take f_j to be $P_j f$. In view of the above, to approximate

$$f = \sum_j f_j, \tag{2}$$

one would take

$$F = \sum_j f_j \prod_{j'>j} (1 - G_{j'}) \tag{3}$$

where G_j are some non-negative functions such that

$$|f_j| \leq G_j \leq 1 \quad \text{pointwisely for all } j. \tag{4}$$

Then at least $\|F\|_{L^\infty} \leq 1$; in fact

$$|F(x)| \leq \sum_j |f_j| \prod_{j'>j} (1 - G_{j'}) \leq \sum_j G_j \prod_{j'>j} (1 - G_{j'}) \leq 1.$$

Now one would first ask whether this could be any sensible approximation of f. To understand this, write $f = \sum_j f_j$. If we think of each f_j as f_j multiplied by 1, then in constructing F we are replacing this 1 by the product over all $j' > j$ above. In fact,

$$f - F = \sum_j f_j \left(1 - \prod_{j'>j} (1 - G_{j'})\right).$$

Using Lemma 4 to expand the latter bracket and rearranging the resulting sum, we get

$$f - F = \sum_j G_j H_j, \tag{5}$$

where

$$H_j = \sum_{j'<j} f_{j'} \prod_{j'<j''<j} (1 - G_{j''}). \tag{6}$$

One would then estimate $\partial_i(f - F)$; one gets

$$\partial_i(f - F) = \sum_j (\partial_i G_j) H_j + \sum_j G_j (\partial_i H_j)$$

by Leibniz rule. Now

$$|H_j| \leq 1 \quad \text{pointwisely for all } j; \tag{7}$$

this is a consequence of the remark after Lemma 4, the reasoning of which is similar to why $|F|$ is bounded by 1. Also,

$$|\partial_i H_j| \leq \sum_{j'<j} (|\partial_i f_{j'}| + |\partial_i G_{j'}|). \tag{8}$$

This is because if one computes $\partial_i H_j$, either the derivative hits $f_{j'}$, in which case we get the first sum above, or the derivative hits $G_{j'}$ for some $j' < j$, and the coefficient of $\partial_i G_{j'}$ in $\partial_i H_j$ is

$$-H_{j'} \prod_{j'<j''<j} (1 - G_{j''}),$$

which is also bounded by 1. In fact,

$$\partial_i H_j = \sum_{j'<j} \left((\partial_i f_{j'}) - (\partial_i G_{j'}) H_{j'} \right) \prod_{j'<j''<j} (1 - G_{j''}).$$

It follows that

$$|\partial_i(f - F)| \leq \sum_j |\partial_i G_j| + \sum_j G_j \sum_{j'<j} (|\partial_i f_{j'}| + |\partial_i G_{j'}|); \tag{9}$$

we shall hope to estimate this in $L^n(\mathbb{R}^n)$ norm.

Equations (2), (3), (4), (5), (6), (7), (8) and (9) will form a basic paradigm of all our constructions below. For given f_j, we shall just take different choices of G_j, as long as (4) is satisfied. For instance, as a very naive attempt, one could try taking $G_j = |P_j f|$ when $f_j = P_j f$; then we shall then need to estimate $\partial_i(f - F)$ in $L^n(\mathbb{R}^n)$, and using (9), what we need to bound first is $\left\| \sum_j |\partial_i G_j| \right\|_{L^n}$, which basically requires one to bound

$$\left\| \sum_j 2^j |P_j f| \right\|_{L^n}. \tag{10}$$

It is well-known that this is not bounded by any multiple of $\|\nabla f\|_{L^n}$; in fact it is only the square function in j of $2^j |P_j f|$ whose L^n norm that is comparable to $\|\nabla f\|_{L^n}$, and there is no hope to gain any small factor like δ in any direction anyway since different directions are not distinguished in this naive construction. We need two ideas that deals with these two issues separately.

1.1. **Controlling the low frequencies by the high frequencies.** First, if we only need to bound

$$\left\| \sum_j 2^j |P_j f| \chi_{\{2^j|P_jf|>C\sum_{k<j} 2^k|P_kf|\}} \right\|_{L^n} \tag{11}$$

instead of (10), where $\chi_{\{\dots\}}$ denotes the characteristic function of the set where the low frequencies are controlled by the high frequencies, then we are in a better shape, since pointwisely

$$\sum_j 2^j |P_j f| \chi_{\{2^j|P_jf|>C\sum_{k<j} 2^k|P_kf|\}} \leq (C+1) \sup_j 2^j |P_j f|, \tag{12}$$

and the right hand side of this can be estimated in L^n using

$$\left\| \sup_j 2^j |P_j f| \right\|_{L^n} \leq C \|\nabla f\|_{L^n}. \tag{13}$$

(To see (12), first fix x, and for any j_0, look at the partial sum over all $j \leq j_0$ of the sum to be estimated; we need only consider j_0 that satisfies $2^{j_0}|P_{j_0}f|(x) > C\sum_{k<j_0} 2^k|P_kf|(x)$, since otherwise the characteristic function in the last term of this partial sum is zero at x, and we are reduced to a previous partial sum. Now assume j_0 is as such. Then

$$\begin{aligned} \sum_{j\leq j_0} 2^j |P_j f|(x) &= 2^{j_0}|P_{j_0}f|(x) + \sum_{k<j_0} 2^k |P_k f|(x) \\ &< (C+1) 2^{j_0} |P_{j_0} f|(x) \\ &\leq (C+1) \sup_j 2^j |P_j f|(x). \end{aligned}$$

Letting $j_0 \to \infty$ we get the desired estimate. The inequality (13) follows from the trivial pointwise bound

$$\sup_j 2^j |P_j f| \leq \left(\sum_j |P_j f|^2 \right)^{1/2}$$

and the Littlewood-Paley inequality.) Now this suggest one to consider an initial splitting

$$f = \sum_j P_j f \chi_{\{2^j|P_jf|>C\sum_{k<j} 2^k|P_kf|\}} + \sum_j P_j f \chi_{\{2^j|P_jf|\leq C\sum_{k<j} 2^k|P_kf|\}} \tag{14}$$

and approximate the two sums separately. To approximate the first, one would try invoking the above general construction, taking something like $f_j = P_j f \chi_{\{2^j|P_jf|>C\sum_{k<j} 2^k|P_kf|\}}$ and $G_j = |P_j f| \chi_{\{2^j|P_jf|>C\sum_{k<j} 2^k|P_kf|\}}$, hoping to end up with an estimate of form (11) in place of (10); this does not

work since both f_j and G_j are not even continuous, and cannot be differentiated in the framework of (2), (3), (4), (5), (6), (7), (8) and (9). Nevertheless, one can try to smooth this out a bit, and give it another go. This would basically work, except that one still does not gain when one differentiate in the good directions.

1.2. Special directions: Introducing the controlling functions ω_j.

To gain when one differentiates in the good directions, one needs to introduce some auxillary controlling functions ω_j that controls $|P_j f|$ in the sense that

$$|P_j f| \leq \omega_j \leq \|P_j f\|_{L^\infty},$$

and has small derivatives in the good directions in the sense that

$$|\partial_i \omega_j| \leq 2^{j-\sigma}\omega_j \quad \text{for } i = 2, \dots, n, \quad \text{and} \quad |\partial_1 \omega_j| \leq 2^j \omega_j.$$

where $\sigma >> 0$ is a large integer to be chosen. The price to pay then is that we only have

$$\left\| \sup_j 2^j \omega_j \right\|_{L^n} \leq C 2^{\sigma(n-1)/n} \|\nabla f\|_{L^n}$$

where the right hand side gets big as σ gets big; c.f. (13). The crucial thing here is that the power of 2^σ on the right hand side, namely $(n-1)/n$, is strictly less than 1. These will be used to define the G_j's when we want to approximate the first sum in (14)[2]. There will then be the second sum in (14) that needs to be approximated, again using the scheme given by equations (2), (3), (4), (5), (6), (7), (8) and (9), but this time it is easier since these are terms where the high frequencies are dominated by the low frequencies (because of the support of the relevant characteristic functions), and when one differentiate low frequencies one gains. The relevant estimates will be made using Littlewood-Paley theory. For that reason, below we first turn to some Littlewood-Paley theory we shall use, and then describe the construction of these auxillary controlling functions ω_j that allows one to pick up good derivatives in all but one directions.

2. Preliminaries on Littlewood-Paley theory

We recall here a few well-known facts about Littlewood-Paley theory and vector-valued singular integrals. Let $1 < p < \infty$.

(1) For $\Phi \in L^p$,

$$\|\Phi\|_{L^p} \simeq \left\| \left(\sum_j |P_j \Phi|^2 \right)^{\frac{1}{2}} \right\|_{L^p}.$$

This is because $\Phi \mapsto \{P_j \Phi\}_{j \in \mathbb{Z}}$ is a vector-valued singular integral taking values in $l^2(\mathbb{Z})$.

[2]To comply with the original notations of Bourgain-Brezis, below we shall give these G_j's for the first sum a different name, and call them U_j instead.

(2) Also from the theory of vector-valued singular integrals, we have, for any sequence of functions Φ_j,

$$\left\| \left(\sum_j |\nabla P_j \Phi_j|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \simeq \left\| \left(\sum_j |2^j P_j \Phi_j|^2 \right)^{\frac{1}{2}} \right\|_{L^p}.$$

This is because the vector-valued multipliers $\xi \mapsto \{2^{-j}|\xi|\chi(2^{-j}\xi)\}_{j\in\mathbb{Z}}$ and $\xi \mapsto \{2^j|\xi|^{-1}\chi(2^{-j}\xi)\}_{j\in\mathbb{Z}}$ both behave as if they were homogeneous of degree 0. Here χ is a smooth cut-off function that is 1 the annulus $\{1/2 \leq |\xi| \leq 2\}$ and is supported in a slightly larger annulus.

(3) For any sequence Φ_j, we have

$$\left\| \left(\sum_j |P_j \Phi_j|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \leq C \left\| \left(\sum_j |\Phi_j|^2 \right)^{\frac{1}{2}} \right\|_{L^p},$$

again by the theory of vector-valued singular integrals.

(4) For any sequence Φ_j, we have

$$\left\| \left(\sum_j |M \Phi_j|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \leq C \left\| \left(\sum_j |\Phi_j|^2 \right)^{\frac{1}{2}} \right\|_{L^p}$$

where M is the standard maximal function operator

$$M\Phi(x) = \sup_{r>0} \frac{1}{r^n} \int_{|y|<r} \Phi(x+y)dy.$$

This is Fefferman-Stein's inequality for the vector-valued maximal function.

(5) Derivatives (and indeed any multiplier operators) commute with the Littlewood-Paley projections P_j, i.e.

$$\partial_i(P_j f) = P_j(\partial_i f)$$

for all i and j.

We shall also use some special kernels on $\mathbb{R}^n$. Let K_j^0 be a kernel on $\mathbb{R}$ such that $\widehat{K_j^0}$ is piecewise linear, equals 1 in $[-2^{-j-1}, 2^{j+1}]$, vanishes outside $[-2^{j+2}, 2^{j+2}]$ and is linear in between; in other words, let

$$K_j^0(x) = (1 + e^{2^{j+2}\pi i x} + e^{-2^{j+2}\pi i x}) F_j^0(x)$$

where F_j^0 is the Fejer kernels on $\mathbb{R}$, satisfying

$$\widehat{F_j^0}(\xi) = \left(1 - \frac{|\xi|}{2^{j+1}}\right)_+.$$

Note that F_j^0 has an explicit expression

$$F_j^0(x) = \frac{1}{2^{j+1}} \left(\frac{\sin 2^{j+1}\pi x}{\pi x} \right)^2$$

that shows that it is non-negative, and hence $\|F_j^0\|_{L^1} = \widehat{F_j^0}(0) = 1$. Note also that $|K_j^0(x)| \leq 3F_j^0(x)$. Next define on $\mathbb{R}^n$

$$K_j(x) = K_j^0(x_1)K_j^0(x_2)\dots K_j^0(x_n) \tag{15}$$

and

$$F_j(x) = F_j^0(x_1)F_j^0(x_2)\dots F_j^0(x_n). \tag{16}$$

Then

$$F_j(x) \geq 0, \quad \|F_j\|_{L^1} = 1 \quad \text{and} \quad |K_j(x)| \leq 3^n F_j(x).$$

We shall often need the fact that for $f \in \dot{W}^{1,n}$, we have

$$|P_j f| \leq 3^n |P_j f| * F_j.$$

This is because

$$P_j f = P_j f * K_j,$$

which in turn follows from the fact that $\widehat{K_j} \equiv 1$ on the support of $\widehat{P_j f}$.

3. Properties of ω_j

Let σ be a large integer. Given $f \in \dot{W}^{1,n}$, we shall introduce an auxillary controling function ω_j that basically plays the role of $P_j f$, except that it has better derivatives in all but one direction. More precisely, the ω_j we define will satisfy the following properties:

$$|P_j f| \leq \omega_j \leq \|P_j f\|_{L^\infty}; \tag{17}$$

$$|\partial_i \omega_j| \leq 2^{j-\sigma}\omega_j \quad \text{for } i = 2, \dots, n, \quad \text{and} \quad |\partial_1 \omega_j| \leq 2^j \omega_j. \tag{18}$$

We shall also need the following crucial property:

$$\left\| \sup_j 2^j \omega_j \right\|_{L^n} \leq C 2^{\sigma(n-1)/n} \|\nabla f\|_{L^n}. \tag{19}$$

The key here is that the power of 2^σ on the right hand side is strictly less than 1. We point out again that the analgous property for $P_j f$ is very easy to prove; see (13). We lose powers of 2^σ here because we have good derivatives in $(n-1)$ directions; in fact for each good direction one loses a factor $2^{\sigma/n}$. We shall also need the facts that

$$\left\| \left(\sum_j (2^j \omega_j)^2 \right)^{1/2} \right\|_{L^n} \leq C 2^{\sigma(2n-1)} \|\nabla f\|_{L^n} \tag{20}$$

and that ω_j is 'locally constant' on the natural scale:

$$\omega_j(x+y) \leq C\omega_j(x) \qquad \text{if } |y_1| \leq 2^{-j} \text{ and } |y'| \leq 2^{-(j-\sigma)}, \tag{21}$$

where C is a constant ($C = e^n$ will do).

The construction of ω_j is as follows. Given $f \in \dot{W}^{1,n}$, define

$$\omega_j(x) = \sup_{y \in \mathbb{R}^n} |P_j f(x-y)| e^{-2^j |y_1| - 2^{j-\sigma}|y'|} \tag{22}$$

where we wrote $y \in \mathbb{R}^n$ as $y = (y_1, y')$, with $y_1 \in \mathbb{R}$ and $y' \in \mathbb{R}^{n-1}$. This is like taking a convolution, except that the integral is replaced by a sup norm. The main advantage of this over an honest convolution is that then (17) becomes obvious (just take $y = 0$ in the supremum for the first inequality), and this is a quality that is absent if we had taken convolutions. This is important because if we want to bound the L^∞ norm of a function that one constructs via Lemma 4, then it is essential to have pointwise estimates (c.f. (4)). We shall think of ω_j as some smoothed out version of $|P_j f|$, and this serves as a useful guide of seeing what estimates are reasonable for ω_j.

On the other hand, one disadvantage of this over an honest convolution is that ω_j is no longer smooth; it is only Lipschitz. Nonetheless, rewriting

$$\omega_j(x) = \sup_{y \in \mathbb{R}^n} |P_j f(y)| e^{-2^j |x_1 - y_1| - 2^{j-\sigma}|x'-y'|} \tag{23}$$

by a change of variable, and differentiating under the supremum, one sees that (18) holds a.e. Hence we gain when we differentiate in the good directions.

Note that (21) also follows from the alternative expression of ω_j in (23).

Now to prove (19), first we observe the following:

$$\omega_j(x) \leq C \sup_{r \in \mathbb{Z}^n} |P_j f| * t_j(x - 2^{-j} r) e^{-|r_1| - 2^{-\sigma}|r'|} \tag{24}$$

where

$$t_j(x) := 2^j \min\{1, (2^j|x|)^{-2}\}^n \geq F_j(x).$$

In other words, it is possible to discretize the supremum defining ω_j. This is because if $|y| < 2^{-j}$, then

$$|P_j f(x+y)| \leq 3^n |P_j f| * F_j(x-y) \leq 3^n |P_j f| * t_j(x-y) \leq C |P_j f| * t_j(x),$$

the last inequality following from the fact that

$$t_j(x+y) \leq C t_j(x)$$

uniformly in j, x and y if $|y| < 2^{-j}$. It follows that for any $y \in \mathbb{R}^n$, if we take $r \in \mathbb{Z}^n$ such that $|y - 2^{-j} r| < 2^{-j}$, then

$$\begin{aligned} |P_j f|(x-y) e^{-2^j|y_1| - 2^{j-\sigma}|y'|} &\leq C |P_j f| * t_j(x - 2^{-j} r) e^{-2^j|y_1| - 2^{j-\sigma}|y'|} \\ &\leq C |P_j f| * t_j(x - 2^{-j} r) e^{-|r_1| - 2^{-\sigma}|r'|} \end{aligned}$$

and hence the desired discrete estimate for ω_j.

We can now prove the estimate for $\| \sup_j 2^j \omega_j \|_{L^n}$. Observe that

$$\sup_j 2^j \omega_j(x) \leq \sup_j \sup_{r \in \mathbb{Z}^n} 2^j |P_j f| * t_j(x - 2^{-j} r) e^{-|r_1| - 2^{-\sigma}|r'|}$$

so

$$\begin{aligned}&\int \left(\sup_j 2^j \omega_j(x)\right)^n dx\\ \le &\int \sup_j \sup_{r\in\mathbb{Z}^n} \left(2^j |P_j f| * t_j(x-2^{-j}r) e^{-|r_1|-2^{-\sigma}|r'|}\right)^n dx\\ \le &\sum_j \sum_{r\in\mathbb{Z}^n} \int \left(2^j |P_j f| * t_j(x-2^{-j}r) e^{-|r_1|-2^{-\sigma}|r'|}\right)^n dx.\end{aligned}$$

It is crucial here that we have discretized the sup to a discrete one; only so one can estimate the integral of a sup by the integral of a sum. It is also important that we are replacing the sup by the sum only after we put the power into the expression being maximized, because this gives a smaller sum. Now

$$\int \left(|P_j f| * t_j(x-2^{-j}r)\right)^n dx \le \int |P_j f|^n \left(\int t_j\right)^n \le \int |P_j f|^n$$

and

$$\sum_{r\in\mathbb{Z}^n} e^{-n|r_1|-n2^{-\sigma}|r'|} = C2^{\sigma(n-1)}.$$

Hence

$$\begin{aligned}\left\|\sup_j 2^j \omega_j\right\|_{L^n} &\le C2^{\sigma(n-1)/n} \left(\int \sum_j (2^j |P_j f|)^n\right)^{1/n}\\ &\le C2^{\sigma(n-1)/n} \left(\int \left(\sum_j (2^j |P_j f|)^2\right)^{n/2}\right)^{1/n}\\ &\le C2^{\sigma(n-1)/n} \|\nabla f\|_{L^n},\end{aligned}$$

the second-to-last inequality holding because $n \ge 2$.

Finally, from (24) again,

$$\omega_j(x) \le C \sum_{r\in\mathbb{Z}^n} |P_j f| * t_j(x-2^{-j}r) e^{-|r_1|-2^{-\sigma}|r'|}.$$

Hence taking square function in j and L^n norm in space,

$$\begin{aligned}&\left\|\left(\sum_j (2^j \omega_j)^2\right)^{\frac{1}{2}}\right\|_{L^n}\\ \le &C \sum_{r\in\mathbb{Z}^n} e^{-|r_1|-2^{-\sigma}|r'|} \left\|\left(\sum_j (2^j |P_j f| * t_j(x-2^{-j}r))^2\right)^{\frac{1}{2}}\right\|_{L^n}\end{aligned}$$

But the kernel $t_j(\cdot - 2^{-j}r)$ has a radial decreasing majorant of integral $|r|^n$, and dominating by the maximal function, we get that the above is bounded by

$$\sum_{r\in\mathbb{Z}^n} e^{-|r_1|-2^{-\sigma}|r'|}|r|^n \left\| \left(\sum_j (2^j M(P_j f)(x))^2 \right)^{\frac{1}{2}} \right\|_{L^n}$$

$$\leq C2^{\sigma n} \int_{y\in\mathbb{R}^n} e^{-|y_1|-2^{-\sigma}|y'|}(|y_1| + 2^{-\sigma}|y'|)^n dy \left\| \left(\sum_j |2^j P_j f(x)|^2 \right)^{\frac{1}{2}} \right\|_{L^n}$$

$$\leq C2^{\sigma(2n-1)} \|\nabla f\|_{L^n} .$$

This proves (20).

4. Attempt 1: Leibniz rule

We now describe two model constructions that illustrates some of the techniques we shall use. In the first one, we take $f_j = P_j f$ and approximate $f = \sum_j f_j$ by F constructed in (3) where $G_j := \omega_j$. We shall assume that $\|\nabla f\|_{L^n}$ is sufficiently small as in the statement of Lemma 2, so that $\|P_j f\|_{L^\infty} \leq 1$ for all j by Bernstein's inequality, from which it follows that both $|f_j|$ and $G_j = \omega_j$ are bounded by 1 in L^∞ by (17).

In that case, F is automatically bounded by 1, since (4) is satisfied. To estimate $\|\partial_i(f - F)\|_{L^n}$, where $i = 2, \dots, n$, we use (9), noting that in our case

$$|\partial_i f_j| \leq C2^j M(P_j f) \quad \text{and} \quad |\partial_i G_j| \leq 2^{j-\sigma}\omega_j.$$

(The first inequality follows because $\partial_i f_j = \partial_i(f_j * K_j) = 2^j f_j * (\partial_i K)_j$ where K_j is the reproducing kernel introduced above, the second inequality is (18).) It follows that both of them are bounded by $C2^j\|\nabla f\|_{L^n}$, and thus

$$\begin{aligned} |\partial_i(f - F)| &\leq \sum_j 2^{j-\sigma}\omega_j + \sum_j \omega_j \sum_{j'<j} C2^{j'}\|\nabla f\|_{L^n} \\ &= 2^{-\sigma}\sum_j 2^j\omega_j + C\sum_j 2^j\omega_j\|\nabla f\|_{L^n}. \end{aligned}$$

Note the small factor $2^{-\sigma}$ one gains in the first term, and the extra $\|\nabla f\|_{L^n}$ in the second term which will contribute to the quadratic nature of the desired estimate. Now one has trouble estimating $\sum_j 2^j\omega_j$ in L^n, because even the smaller sum $\sum_j 2^j|P_j f|$ cannot be estimated in L^n. Nevertheless, the kind of splitting as in (14) (or more precisely, a smoothed out version of that) will allow us to replace any L^n norm of $\sum_j 2^j\omega_j$ by

$$\left\| \sum_j 2^j\omega_j\chi_{\{2^j\omega_j > C\sum_{k<j} 2^k\omega_k\}} \right\|_{L^n},$$

which could then be bounded by

$$(C+1)\|\sup_j 2^j \omega_j\|_{L^n} \lesssim 2^{\sigma(n-1)/n}\|\nabla f\|_{L^n}$$

using (19) since pointwisely $\sum_j 2^j \omega_j \chi_{\{2^j\omega_j > C\sum_{k<j} 2^k \omega_k\}} \leq (C+1)\sup_j 2^j \omega_j$ as in the derivation of (12). If we are allowed to make such a cheat here, we would then have $\|\partial_i(f-F)\|_{L^n} \leq C 2^{-\sigma/n}\|\nabla f\|_{L^n} + C\|\nabla f\|_{L^n}^2$ (note the count of the powers of 2^σ here), which would be the form of inequality we would want to prove in Lemma 2, since the power of 2^σ we get here is negative, and σ can be taken to be big.

5. Attempt 2: Littlewood-Paley theory

Let's take another naive attempt, in which we make estimates using Littlewood-Paley theory. We still take $f = \sum_j f_j$ where $f_j = P_j f$, and to approximate this we let F be defined by (3), where $G_j := 3^n |P_j f| * F_j$ and F_j are the Fejer kernels introduced above. This time we do not expect to gain in the good directions, since there is no distinction between different directions; nevertheless it is intructive to see how the frequency localization in f_j and G_j (note the convolution with the Fejer kernel in G_j) will help one make estimates using Littlewood-Paley theory.

First, we still assume that $\|\nabla f\|_{L^n}$ is sufficiently small, so that $\|G_j\|_\infty \leq 1$ still holds by Bernstein's inequality. Now observe that $f_j = P_j f = P_j f * K_j$, from which it follows that $|f_j| \leq |P_j f| * (3^n F_j) = G_j$. As a result, (4) holds, and thus $\|F\|_{L^\infty} \leq 1$.

Now we bound $\|\partial_i(f-F)\|_{L^n}$ using Littlewood-Paley theory. Recall from (5) and (6) that $f - F = \sum_j G_j H_j$, where $H_j = \sum_{j'<j} f_{j'} \prod_{j'<j''<j}(1-G_{j''})$. Since both f_j and G_j are supported in frequency in a ball of radius 2^j, H_j is also supported on a ball of frequency $\sim 2^j$; in fact

$$P_k(G_j H_j) = 0 \quad \text{whenever } k > j+2.$$

(Note how the frequency support of the non-negative kernel F_j comes into play here.) As a result,

$$\begin{aligned}
\|\nabla(f-F)\|_{L^n} &\simeq \left\| \left(\sum_k |\nabla P_k (f-F)|^2 \right)^{\frac{1}{2}} \right\|_{L^n} \\
&= \left\| \left(\sum_k \left| \sum_{s\geq -2} \nabla P_k (G_{k+s} H_{k+s}) \right|^2 \right)^{\frac{1}{2}} \right\|_{L^n} \\
&\leq \sum_{s\geq -2} \left\| \left(\sum_k |\nabla P_k (G_{k+s} H_{k+s})|^2 \right)^{\frac{1}{2}} \right\|_{L^n}
\end{aligned}$$

There are now two ways to proceed: the first one is to differentiate $G_{k+s}H_{k+s}$ and make the estimate

$$\left\| \left(\sum_k |\nabla P_k(G_{k+s}H_{k+s})|^2 \right)^{\frac{1}{2}} \right\|_{L^n} \leq C \left\| \left(\sum_k |\nabla (G_{k+s}H_{k+s})|^2 \right)^{\frac{1}{2}} \right\|_{L^n};$$

the other is to differentiate the Littlewood-Paley projections and make the estimate

$$\left\| \left(\sum_k |\nabla P_k(G_{k+s}H_{k+s})|^2 \right)^{\frac{1}{2}} \right\|_{L^n} \leq C \left\| \left(\sum_k \left| 2^k P_k(G_{k+s}H_{k+s}) \right|^2 \right)^{\frac{1}{2}} \right\|_{L^n}.$$

The first approach gives us in general a factor of 2^{k+s} that goes with $G_k H_k$, while the second approach gives us the better factor of 2^k. Hence we adopt the second approach, and arrive at

$$\begin{aligned} \|\nabla(f-F)\|_{L^n} &\leq \sum_{s\geq -2} \left\| \left(\sum_k \left| 2^k P_k(G_{k+s}H_{k+s}) \right|^2 \right)^{\frac{1}{2}} \right\|_{L^n} \\ &\leq \sum_{s\geq -2} 2^{-s} \left\| \left(\sum_k \left| 2^{k+s} G_{k+s}H_{k+s} \right|^2 \right)^{\frac{1}{2}} \right\|_{L^n} \\ &= \left\| \left(\sum_k \left| 2^k G_k H_k \right|^2 \right)^{\frac{1}{2}} \right\|_{L^n}. \end{aligned}$$

Now $\|H_k\|_{L^\infty} \leq 1$ by (7), and $|G_k| = 3^n |P_k f| * F_k \leq 3^n M(P_k f)$. Hence

$$\begin{aligned} \|\nabla(f-F)\|_{L^n} &\leq 3^n \left\| \left(\sum_k \left(2^k M(P_k f) \right)^2 \right)^{\frac{1}{2}} \right\|_{L^n} \\ &\leq C \left\| \left(\sum_k \left(2^k |P_k f| \right)^2 \right)^{\frac{1}{2}} \right\|_{L^n} \\ &\leq C \|\nabla f\|_{L^n} \end{aligned}$$

as desired.

In reality, to gain a small factor like δ in estimates like this, we can only sum over large values of s, say $s \geq R$ where R is another very big positive integer. In that case one gains powers of 2^{-R}. One will then need to figure out some other way in which the sum over small s can be dealt with. To do that one need to replace G_j by something whose derivatives are small; in fact we need something like $|\nabla G_j|$ to be of the order 2^{j-R}. Approximations of this kind will be used to deal with the second sum in (a smoothed out version of) (14), where intuitively speaking the high frequencies are dominated by the sum of lower frequencies. In effect Littlewood-Paley theory will allow one to shift the derivatives on the high frequency components to low frequency ones, thereby gaining the desired factors of 2^{-R} where R is as above. The additional complication, on the other hand, is that we will need to take G_j to be an infinite sum (with the terms corresponding to the frequencies lower than 2^j), and one need to gain an additional convergence factor to account for the convergence of this additional summation when making estimates.

6. The actual proof

We have finally most of the ingredients to prove Lemma 2. Below we begin afresh; the notations in the previous sections should not be confused with the ones defined below.

Proof of Lemma 2. Let $f \in \dot{W}^{1,n}(\mathbb{R}^n)$ be such that $\|\nabla f\|_{L^n} \leq c_n$. If c_n is a sufficiently small dimensional constant, then by Bernstein's inequality, we have $\|P_j f\|_{L^\infty} \leq 3^{-n}$ for all j. Fix such a c_n from now on. Suppose in addition that $P_j f$ is identically equal to zero unless j belongs to an arithmetic progression of length R, where R is a very large integer to be determined. This assumption can be easily removed at the end of the argument, since a general function f is the sum of R such functions.

Now let $\sigma = R/4n$, and define ω_j by (22), so that (17), (18), (19), (20) and (21) are all satisfied. We split f into a sum,

$$f = g + h,$$

where

$$g = \sum_j g_j, \quad h = \sum_j h_j,$$

$$g_j = (P_j f \cdot \chi_{\{2^j \omega_j \leq \sum_{k<j} 2^k \omega_k\}}) * K_j,$$

and

$$h_j = (P_j f \cdot \chi_{\{2^j \omega_j > \sum_{k<j} 2^k \omega_k\}}) * K_j.$$

Here K_j are the reproducing kernels defined by (15) and χ denotes the characteristic function of a set. The extra convolutions with K_j does not affect the fact that $g_j + h_j = P_j f$, because K_j are reproducing. But the extra convolutions smoothes out the product of $P_j f$ with the characteristic functions, and at the same time localizes such in frequency. This will turn out be very handy for us, as was hinted in the discussions in the previous two sections.

Note that ω_j, g_j, h_j are all identically zero unless j belongs to the special arithmetic progression of common difference R that we had at the beginning.

We first approximate $h = \sum_j h_j$, using the paradigm we introduced in Section 1. More precisely, we will find functions U_j such that

$$|h_j| \leq U_j \leq 1 \quad \text{pointwisely for all } j, \tag{25}$$

and approximate h by

$$\tilde{h} := \sum_j h_j \prod_{j'<j} (1 - U_{j'}). \tag{26}$$

Note that such an $\tilde{h}$ must be in L^∞; in fact

$$\|\tilde{h}\|_{L^\infty} \leq 1 \tag{27}$$

by Lemma 4 and (25). We shall prove that

$$\|\partial_i (h - \tilde{h})\|_{L^n} \leq C 2^{-\sigma/n} \|\nabla f\|_{L^n} + C 2^{\sigma(n-1)/n} \|\nabla f\|_{L^n}^2 \quad \text{if } i = 2, \dots, n, \tag{28}$$

and

$$\|\nabla (h - \tilde{h})\|_{L^n} \leq C 2^{\sigma(n-1)/n} \|\nabla f\|_{L^n} \tag{29}$$

upon a suitable choice of U_j. These estimates will be established using Leibniz rule, in a similar spirit as what we did in Section 4. The key here is that the coefficient of the linear factor in $\|\nabla f\|_{L^n}$ in (28) is small when R (and hence σ) is big.

First observe that

$$|h_j| \leq 3^n (\omega_j \chi_{\{2^j \omega_j > \sum_{k<j} 2^k \omega_k\}}) * F_j,$$

and F_j are the Fejer kernels defined by (16). Hence one is tempted to take U_j as the right hand side above and run the paradigm we introduced in Section 1. However, one would then need to estimate $\|\sum_j \partial_i U_j\|_{L^n}$, and for that one needs to have some control on the support of U_j (or its constituents). To do that, we introduce smooth cut-off functions ψ_j on $\mathbb{R}$ such that

$$0 \leq \psi_j \leq 1, \quad \psi_j = 0 \text{ outside } [-2^{-j}, 2^{-j}], \quad \psi_j(0) = 1, \quad |\psi_j'| \leq 2^j.$$

Then define a second auxillary function

$$u_j(x) = \sup_{y \in \mathbb{R}^n} \left(\omega_j \chi_{\{2^j \omega_j > \sum_{k<j} 2^k \omega_k\}} \right) (x - y) \psi_j(y_1) \psi_{j-\sigma}(y_2) \dots \psi_{j-\sigma}(y_n),$$

smoothing out $\omega_j \chi_{\{2^j \omega_j > \sum_{k<j} 2^k \omega_k\}}$. Note the similarity with the construction of ω_j. The advantage of doing that is that one now has control on the support of the derivative of u_j: in fact since ψ_j has compact support, $u(x)$ depends only on the values of ω_j near x. Now by (21),

$$u_j(x) \leq e^n \omega_j(x).$$

Indeed a more precise estimate is possible: if

$$2^j \omega_j(x) \leq e^{-2n} \sum_{k<j} 2^k \omega_k(x)$$

then for all y with $|y_1| < 2^{-j}$ and $|y_2|, \dots, |y_n| < 2^{-(j-\sigma)}$, we have

$$\begin{aligned} 2^j \omega_j(x+y) &\leq e^n 2^j \omega_j(x) \\ &\leq e^n e^{-2n} \sum_{k<j} 2^k \omega_k(x) \\ &\leq e^n e^{-2n} \sum_{k<j} 2^k \omega_k(x+y) e^n \\ &\leq \sum_{k<j} 2^k \omega_k(x+y) \end{aligned}$$

and hence $u_j(x) = 0$. It follows that

$$u_j(x) \leq e^n (\omega_j \chi_{\{2^j \omega_j > e^{-2n} \sum_{k<j} 2^k \omega_k\}})(x)$$

improving our previous estimate. Similarly, the derivatives can be estimated: for $i \neq 1$,

$$\begin{aligned} &|\partial_i u_j(x)| \\ &\leq \sup_y (\omega_j \chi_{\{2^j \omega_j > \sum_{k<j} 2^k \omega_k\}})(y) |\partial_i \psi_j(x_1 - y_1) \psi_{j-\sigma}(x_2 - y_2) \dots \psi_{j-\sigma}(x_n - y_n)| \\ &\leq C 2^{j-\sigma} \sup_{|y_1| < 2^{-j}, |y_2|, \dots, |y_n| < 2^{-(j-\sigma)}} (\omega_j \chi_{\{2^j \omega_j > \sum_{k<j} 2^k \omega_k\}})(x-y) \\ &\leq C 2^{j-\sigma} e^n (\omega_j \chi_{\{2^j \omega_j > e^{-2n} \sum_{k<j} 2^k \omega_k\}})(x), \end{aligned}$$

and

$$|\nabla u_j(x)| \leq C 2^j e^n (\omega_j \chi_{\{2^j \omega_j > e^{-2n} \sum_{k<j} 2^k \omega_k\}})(x).$$

Notice how we obtained control on the support of these derivatives.

Now observe that

$$|h_j(x)| \leq 3^n (\omega_j \chi_{\{2^j \omega_j > \sum_{k<j} 2^k \omega_k\}}) * F_j(x) \leq 3^n u_j * F_j(x)$$

and that

$$\|3^n u_j * F_j(x)\|_{L^\infty} \leq 3^n \|u_j\|_{L^\infty} \leq 3^n \|\omega_j\|_{L^\infty} \leq 3^n \|P_j f\|_{L^n} \leq 1.$$

Hence we define

$$U_j = 3^n u_j * F_j$$

and this completes the definition of $\tilde{h}$ by (26). Note (25) and hence (27) is satisfied. Now by the same paradigm that leads to the proof of (9), we get

$$|\partial_i (h - \tilde{h})| \leq \sum_j |\partial_i U_j| + \sum_j U_j \sum_{j' < j} \left(|\nabla h_{j'}| + |\nabla U_{j'}| \right).$$

But

$$|\nabla h_j| + |\nabla U_j| \leq C 2^j \|\nabla f\|_{L^n},$$

since

$$|\nabla h_j| \leq C 2^j M(P_j f) \leq C 2^j \|P_j f\|_{L^\infty},$$
$$|\nabla U_j| \leq C 2^j M(u_j) \leq C 2^j \|u_j\|_{L^\infty} \leq C 2^j \|\omega_j\|_{L^\infty} \leq C 2^j \|P_j f\|_{L^\infty},$$

and $\|P_j f\|_{L^\infty} \leq C\|\nabla f\|_{L^n}$ by Bernstein. Also, for $i \neq 1$,

$$|\partial_i U_j| \leq C|\partial_i u_j| * F_j \leq C 2^{j-\sigma}(\omega_j \chi_{\{2^j\omega_j > e^{-2n}\sum_{k<j} 2^k \omega_k\}}) * F_j.$$

Hence for $i \neq 1$,

$$\begin{aligned} |\partial_i (h - \tilde{h})| \leq & C 2^{-\sigma} \sum_j 2^j (\omega_j \chi_{\{2^j\omega_j > e^{-2n}\sum_{k<j} 2^k \omega_k\}}) * F_j \\ & + C\|\nabla f\|_{L^n} \sum_j 2^j (\omega_j \chi_{\{2^j\omega_j > e^{-2n}\sum_{k<j} 2^k \omega_k\}}) * F_j. \end{aligned}$$

We need to estimate

$$\left\| \sum_j 2^j (\omega_j \chi_{2^j\omega_j > e^{-2n}\sum_{k<j} 2^k \omega_k}) * F_j \right\|_{L^n}. \tag{30}$$

Here we need a lemma that says the frequency localization by convolution against F_j is harmless here[3]:

Lemma 5. *If Φ_j is a sequence of non-negative functions, then for $1 \leq p \leq \infty$,*

$$\left\| \sum_j \Phi_j * F_j \right\|_{L^p} \leq \left\| \sum_j \Phi_j \right\|_{L^p}.$$

Proof of Lemma 5. This is because when $\Phi_j \geq 0$,

$$\left| \sum_j \Phi_j * F_j \right| \leq \sum_j M(\Phi_j) = M\left(\sum_j \Phi_j \right)$$

and the maximal function is bounded on L^p if $1 < p \leq \infty$. By duality we can extend the estimate to $p = 1$. □

Hence (30) is bounded by

$$C \left\| \sum_j 2^j \omega_j \chi_{\{2^j\omega_j > e^{-2n}\sum_{k<j} 2^k \omega_k\}} \right\|_{L^n}. \tag{31}$$

Now pointwisely,

$$\sum_j 2^j \omega_j \chi_{\{2^j\omega_j > e^{-2n}\sum_{k<j} 2^k \omega_k\}} \leq C \sup_j 2^j \omega_j.$$

This can be proved in the same way that (12) is proved. Hence (31) is bounded by

$$C \left\| \sup_j 2^j \omega_j \right\|_{L^n} \leq C 2^{\sigma(n-1)/n} \|\nabla f\|_{L^n}$$

[3] We state it also for $p = 1$ for interest only; we only need to use it for $p = n$, for which the argument is easy.

by (19). Putting these together, we see that for $i \neq 1$, (28) holds. Note how we squeezed a small factor $2^{-\sigma/n}$ in front of the linear term in $\|\nabla f\|_{L^n}$ in this estimate. In general, if we differentiate in the bad (i.e. ∂_1) direction, the above arguments give

$$\|\nabla(h-\tilde{h})\|_{L^n} \leq C2^{\sigma(n-1)/n}\|\nabla f\|_{L^n} + C2^{\sigma(n-1)/n}\|\nabla f\|_{L^n}^2$$

which implies (29) since $\|\nabla f\|_{L^n}$ was assumed to be bounded by a dimensional constant. This completes our approximation for h.

Next we first approximate $g = \sum_j g_j$, again using the paradigm we introduced in Section 1. More precisely, we will find functions G_j such that

$$|g_j| \leq G_j \leq 1 \quad \text{pointwisely for all } j, \tag{32}$$

and approximate g by

$$\tilde{g} := \sum_j g_j \prod_{j'<j} (1 - G_{j'}). \tag{33}$$

Note that such an $\tilde{g}$ must be in L^∞; in fact

$$\|\tilde{g}\|_{L^\infty} \leq 1 \tag{34}$$

by Lemma 4 and (32). We shall prove that

$$\|\nabla(g-\tilde{g})\|_{L^n} \leq CR2^{-R}2^{\sigma(2n-1)}\|\nabla f\|_{L^n} + CR2^{\sigma(2n-1)}\|\nabla f\|_{L^n}^2. \tag{35}$$

upon a suitable choice of G_j. These estimates will be established using Littlewood-Paley theory, in a similar spirit as what we did in Section 5. Note we do not need to distinguish between the good and bad derivatives; all of them will be controlled in the same way. Note also that the coefficient of the linear factor in $\|\nabla f\|_{L^n}$ in (35) is small when R is big, since $\sigma = R/4n$ and thus $R2^{-R}2^{\sigma(2n-1)} \leq R2^{-R/2}$.

First there is a pointwise domination of g_j, given by

$$\begin{aligned} |g_j| &\leq (\omega_j \chi_{\{2^j\omega_j \leq \sum_{k<j} 2^k \omega_k\}}) * 3^n F_j \\ &\leq 3^n \sum_{k<j} 2^{k-j} \omega_k * F_j. \end{aligned}$$

Remember g_j and ω_j are both identically zero unless j is in a certain arithmetic progression of common difference R. Hence we could have also written

$$|g_j| \leq 3^n \sum_{t \geq R} 2^{-t} \omega_{j-t} * F_j$$

and we define G_j to be the right hand side of the above inequality. Note that

$$\|G_j\|_{L^\infty} \leq 3^n \sum_{t\geq R} 2^{-t} \|\omega_{j-t} * F_j\|_{L^\infty} \leq 3^n \|P_j f\|_{L^\infty} \leq 1.$$

Now we estimate $\|\nabla(g-\tilde{g})\|_{L^n}$: note that

$$g - \tilde{g} = \sum_j G_j H_j$$

where

$$H_j = \sum_{j'<j} g_{j'} \prod_{j'<j''<j} (1 - G_{j''}).$$

For future reference, we remark that

$$|H_j| \le 1$$

which follows from Lemma 4 and (32),

$$|\nabla G_j| \le 3^n \sum_{t \ge R} 2^{-t} |\nabla \omega_{j-t}| * F_j \le C \sum_{t>0} 2^{-t} 2^{j-t} \omega_{j-t} * F_j,$$

and

$$\begin{aligned} |\nabla H_k| &\le \sum_{l>0} (|\nabla g_{k-l}| + |\nabla G_{k-l}|) \\ &\le C \sum_{l>0} 2^{k-l} M \omega_{k-l} + C \sum_{l>0} \sum_{t>0} 2^{-t} 2^{k-l-t} M \omega_{k-l-t} \\ &\le C \sum_{l>0} 2^{k-l} M \omega_{k-l}. \end{aligned} \tag{36}$$

Since G_j and H_j are both compactly supported in frequency, $P_k(G_j H_j) = 0$ if $k > j + 2$. Hence

$$\begin{aligned} \|\nabla(g - \tilde{g})\|_{L^n} &\simeq \left\| \left(\sum_k \left| \nabla P_k \left(\sum_j G_j H_j \right) \right|^2 \right)^{\frac{1}{2}} \right\|_{L^n} \\ &\le \sum_{s \ge -2} \left\| \left(\sum_k |\nabla P_k (G_{k+s} H_{k+s})|^2 \right)^{\frac{1}{2}} \right\|_{L^n} \end{aligned}$$

There are now two ways to proceed: we can differentiate $G_{k+s} H_{k+s}$, or we can differentiate the Littlewood-Paley projections. But G_{k+s} is a sum of components whose derivatives get smaller and smaller: indeed

$$G_{k+s} = 3^n \sum_{t \ge R} 2^{-t} \omega_{k+s-t} * F_{k+s}$$

and if one differentiate ω_{k+s-t} in the sum, one gets a factor 2^{k+s-t} which is better than the factor 2^k that one gets from differentiating the Littlewood-Paley projections. Hence it is natural to split G_j into two parts and deal with them differently.

Let now s be fixed, and consider

$$\left\| \left(\sum_k |\nabla P_k (G_{k+s} H_{k+s})|^2 \right)^{\frac{1}{2}} \right\|_{L^n}. \tag{37}$$

Let $\bar{s}$ be a non-negative function of s to be determined, and let

$$G_j^{(1)} = 3^n \sum_{R \leq t < \bar{s}} 2^{-t} \omega_{j-t} * F_j,$$

$$G_j^{(2)} = 3^n \sum_{t \geq \max\{\bar{s}, R\}} 2^{-t} \omega_{j-t} * F_j.$$

Then (37) is bounded by

$$\left\| \left(\sum_k \left| \nabla P_k (G_{k+s}^{(1)} H_{k+s}) \right|^2 \right)^{\frac{1}{2}} \right\|_{L^n} + \left\| \left(\sum_k \left| \nabla P_k (G_{k+s}^{(2)} H_{k+s}) \right|^2 \right)^{\frac{1}{2}} \right\|_{L^n} \tag{38}$$

According to the heuristics above, we estimate the first term by

$$C \left\| \left(\sum_k \left| 2^k P_k (G_{k+s}^{(1)} H_{k+s}) \right|^2 \right)^{\frac{1}{2}} \right\|_{L^n} \leq C 2^{-s} \left\| \left(\sum_k \left| 2^k G_k^{(1)} H_k \right|^2 \right)^{\frac{1}{2}} \right\|_{L^n}$$

and from

$$|G_k^{(1)} H_k| \leq |G_k^{(1)}| \leq 3^n \sum_{R \leq t < \bar{s}} 2^{-t} \omega_{k-t} * F_k \leq 3^n \sum_{R \leq t < \bar{s}} 2^{-t} M \omega_{k-t},$$

we conclude that the first term of (38) is bounded by

$$C 2^{-s} \sum_{R \leq t < \bar{s}} \left\| \left(\sum_k (2^{k-t} M \omega_{k-t})^2 \right)^{\frac{1}{2}} \right\|_{L^n} \leq C 2^{-s} \bar{s} \left\| \left(\sum_k (2^k \omega_k)^2 \right)^{\frac{1}{2}} \right\|_{L^n},$$

which is then bounded by

$$C 2^{-s} \bar{s} 2^{\sigma(2n-1)} \|\nabla f\|_{L^n} \tag{39}$$

by (20).

To estimate the second term in (38) involving $G_j^{(2)}$, we use

$$\left\| \left(\sum_k \left| \nabla P_k (G_{k+s}^{(2)} H_{k+s}) \right|^2 \right)^{\frac{1}{2}} \right\|_{L^n} \leq C \left\| \left(\sum_k \left| \nabla (G_k^{(2)} H_k) \right|^2 \right)^{\frac{1}{2}} \right\|_{L^n}.$$

Now

$$|\nabla (G_k^{(2)} H_k)| \leq |\nabla G_k^{(2)}| + |G_k^{(2)}| |\nabla H_k|,$$

and

$$\left\|\left(\sum_k \left|\nabla G_k^{(2)}\right|^2\right)^{\frac{1}{2}}\right\|_{L^n} \leq \sum_{t>\max\{\bar{s},R\}} 2^{-t}\left\|\left(\sum_k |(\nabla\omega_{k-t}) * F_k|^2\right)^{\frac{1}{2}}\right\|_{L^n}$$

$$\leq C\sum_{t>\max\{\bar{s},R\}} 2^{-t}\left\|\left(\sum_k \left(2^{k-t}\omega_{k-t} * F_k\right)^2\right)^{\frac{1}{2}}\right\|_{L^n}$$

$$\leq C\sum_{t>\max\{\bar{s},R\}} 2^{-t}\left\|\left(\sum_k \left(2^{k-t}M\omega_{k-t}\right)^2\right)^{\frac{1}{2}}\right\|_{L^n}$$

$$\leq C2^{-\max\{\bar{s},R\}}\left\|\left(\sum_k (2^k\omega_k)^2\right)^{\frac{1}{2}}\right\|_{L^n}$$

$$\leq C2^{-\max\{\bar{s},R\}}2^{\sigma(2n-1)}\|\nabla f\|_{L^n} \tag{40}$$

Also, by (36),

$$\left\|\left(\sum_k \left|G_k^{(2)}\nabla H_k\right|^2\right)^{\frac{1}{2}}\right\|_{L^n} \leq \sum_{l\geq 0}\left\|\left(\sum_k (2^{k-l}M\omega_{k-l}G_k^{(2)})^2\right)^{\frac{1}{2}}\right\|_{L^n}$$

$$\leq \sum_{l\geq 0}\sum_{t>\bar{s}} 2^{-t-l}\left\|\left(\sum_k (2^kM\omega_{k-l}M\omega_{k-t})^2\right)^{\frac{1}{2}}\right\|_{L^n}.$$

Now we split this sum into two parts, one where $t > l$, and another where $t \leq l$. In the first sum we estimate $M\omega_{k-t}$ by $C\|\nabla f\|_{L^n}$, and in the second sum we estimate $M\omega_{k-l}$ by $C\|\nabla f\|_{L^n}$. Then the two sums are bounded by

$$C\|\nabla f\|_{L^n}\sum_{l\geq 0}\sum_{t>\max\{l,\bar{s}\}} 2^{-t}\left\|\left(\sum_k (2^{k-l}M\omega_{k-l})^2\right)^{\frac{1}{2}}\right\|_{L^n}$$

$$+ C\|\nabla f\|_{L^n}\sum_{t>\bar{s}}\sum_{l\geq t} 2^{-l}\left\|\left(\sum_k (2^{k-t}M\omega_{k-t})^2\right)^{\frac{1}{2}}\right\|_{L^n}$$

$$\leq C(1+\bar{s})2^{-\bar{s}}\|\nabla f\|_{L^n}\left\|\left(\sum_k (2^k\omega_k)^2\right)^{\frac{1}{2}}\right\|_{L^n}$$

$$\leq C(1+\bar{s})2^{-\bar{s}}2^{\sigma(2n-1)}\|\nabla f\|_{L^n}^2. \tag{41}$$

Putting the estimates (39), (40) and (41) together, we get

$$\|\nabla(g-\tilde{g})\|_{L^n}$$
$$\leq C\sum_{s\geq -2}\left((2^{-s}\bar{s} + 2^{-\max\{\bar{s},R\}})2^{\sigma(2n-1)}\|\nabla f\|_{L^n} + (1+\bar{s})2^{-\bar{s}}2^{\sigma(2n-1)}\|\nabla f\|_{L^n}^2\right).$$

Recall $\bar{s}$ was a non-negative function of s to be determined. If we now pick $\bar{s} = 0$ if $s \leq R$ and $\bar{s} = s$ if $s > R$, then summing the above we get (35).

Altogether, if we now set

$$F = \tilde{g} + \tilde{h}$$

then

$$\|F\|_{L^\infty} \leq \|\tilde{g}\|_{L^\infty} + \|\tilde{h}\|_{L^\infty} \leq 2$$

and we have now for $i \neq 1$,

$$\begin{aligned}
\|\partial_i(f - F)\|_{L^n} &\leq \|\nabla(g - \tilde{g})\|_{L^n} + \|\partial_i(h - \tilde{h})\|_{L^n} \\
&\leq CR2^{-R}2^{\sigma(2n-1)}\|\nabla f\|_{L^n} + CR2^{\sigma(2n-1)}\|\nabla f\|_{L^n}^2 \\
&\quad + C2^{-\sigma/n}\|\nabla f\|_{L^n} + C2^{\sigma(n-1)/n}\|\nabla f\|_{L^n}^2 \\
&\leq CR2^{-R/4n^2}\|\nabla f\|_{L^n} + CR2^{R/2}\|\nabla f\|_{L^n}^2.
\end{aligned}$$

(The last inequality follows because $\sigma = R/4n$.) Also,

$$\begin{aligned}
\|\nabla(f - F)\|_{L^n} &\leq CR2^{-R}2^{\sigma(2n-1)}\|\nabla f\|_{L^n} + CR2^{\sigma(2n-1)}\|\nabla f\|_{L^n}^2 \\
&\quad + C2^{\sigma(n-1)/n}\|\nabla f\|_{L^n} \\
&\leq CR2^{R/2}\|\nabla f\|_{L^n}.
\end{aligned}$$

These are true whenever $\|\nabla f\|_{L^n} \leq c_n$, and $P_j f$ vanishes identically except for j in an arithmetic progression of common difference R. Now given a general f with $\|\nabla f\|_{L^n} \leq c_n$, it can be written as the sum of R functions with the previous property. Hence what we have proved implies that given any general f with $\|\nabla f\|_{L^n} \leq c_n$, there exists a function $F \in \dot{W}^{1,n} \cap L^\infty$ satisfying

$$\|F\|_{L^\infty} \leq 2R,$$

$$\|\partial_i(f - F)\|_{L^n} \leq CR2^{-R/4n^2}\|\nabla f\|_{L^n} + CR2^{R/2}\|\nabla f\|_{L^n}^2 \quad \text{for } i = 2, \dots, n$$

and

$$\|\nabla(f - F)\|_{L^n} \leq CR2^{R/2}\|\nabla f\|_{L^n}.$$

(We just multiply each corresponding bound by R.) Since this is true for any large R, by picking R big enough, we complete the proof of Lemma 2. □

References

[1] Jean Bourgain and Haïm Brezis, *New estimates for elliptic equations and Hodge type systems*, J. Eur. Math. Soc. (JEMS) **9** (2007), no. 2, 277–315.