HIERARCHICAL SOLUTIONS FOR LINEAR EQUATIONS: A CONSTRUCTIVE PROOF OF THE CLOSED RANGE THEOREM

EITAN TADMOR

ABSTRACT. We construct uniformly bounded solutions for the equations $div\,U = f$ and $curl\,U = \mathbf{f}$ in the critical cases $f \in L^d_\#(\mathbb{T}^d, \mathbb{R})$ and $\mathbf{f} \in L^3_\#(\mathbb{R}^3, \mathbb{R}^3)$. Bourgain & Brezis, [BB03, BB07], have shown that there exists no *linear* construction for such solutions. Our constructions are special cases of a general framework for solving linear equations of the form $\mathscr{L}U = f$, where $\mathscr{L}$ is a linear operator densely defined in Banach space $\mathbb{B}$ with a closed range in a (proper subspace) of Lebesgue space $L^p_\#(\Omega)$, and with an injective dual $\mathscr{L}^*$. The solutions are realized in terms of a multiscale *hierarchical representation*, $U = \sum_{j=1}^\infty \mathbf{u}_j$, interesting for its own sake. Here, the $\mathbf{u}_j$'s are constructed *recursively* as minimizers of the iterative refinement step, $\mathbf{u}_{j+1} = arg\,min_{\mathbf{u}} \left\{ \|\mathbf{u}\|_{\mathbb{B}} + \lambda_{j+1} \|r_j - \mathscr{L}\mathbf{u}\|^p_{L^p(\Omega)} \right\}$, where $r_j := f - \mathscr{L}(\sum_{k=1}^j \mathbf{u}_k)$, are resolved in terms of an exponentially increasing sequence of scales $\lambda_{j+1} := \lambda_1 \zeta^j$. The resulting hierarchical decompositions, $U = \sum_{j=1}^\infty \mathbf{u}_j$, are nonlinear.

CONTENTS

1. INTRODUCTION AND STATEMENT OF THE MAIN RESULTS

We begin with a prototype example for the class of equations alluded to in the title of the paper. Let $L^d_\#(\mathbb{T}^d)$ denote the Lebesgue space of periodic functions with zero mean over the d-dimensional torus $\mathbb{T}^d$. Given $f \in L^d_\#(\mathbb{T}^d)$, we seek a uniformly bounded solution, $U \in L^\infty(\mathbb{T}^d, \mathbb{R}^d)$, of the problem

$$div\,U = f, \qquad U \in L^\infty(\mathbb{T}^d, \mathbb{R}^d). \tag{1.1a}$$

July 18, 2010

1991 *Mathematics Subject Classification.* 35C10, 35F05, 46A30, 49M27, 68U10.

Key words and phrases. Closed range theorem, hierarchical solutions, duality, Sobolev estimates, divergence, curl, multiscale expansion, image processing, critical regularity.

Research was supported by grants from National Science Foundation. I thank Haim Brezis for helpful discussions.

A classical solution is given by $U = \nabla\Delta^{-1} f$, which in addition to (1.1a), satisfies the irotationality constraint,

$$curl\, U = 0. \tag{1.1b}$$

Clearly, this solution lies in $W^1(L^d_\#)$. But since $W^1(L^d_\#)$ is *not contained* in L^∞, the solution may — and in fact does fail to satisfy the uniform bound sought for the solution of (1.1a), [BB03, §3, Remark 7]. Thus, the question is whether (1.1a) admits a uniformly bounded solution by giving up on the additional constraint of irotationality (1.1b). The existence of such solutions was proved by Bourgain and Brezis, [BB03, Proposition 1] using a straightforward but non-constructive duality argument based on the closed range theorem. We present here another duality-based approach for the existence of such solutions. Our approach is *constructive*: the solution U is constructed as the sum, $U = \sum \mathbf{u}_j$, where the $\{\mathbf{u}_j\}$'s are computed *recursively* as appropriate minimizers,

$$\mathbf{u}_{j+1} = \underset{\mathbf{u}}{arg\,min}\Big\{\|\mathbf{u}\|_{L^\infty} + \lambda_1 2^j \big\|f - div\big(\sum_{k=1}^{j}\mathbf{u}_k\big) - div\,\mathbf{u}\big\|^d_{L^d}\Big\}, \quad j = 0,1,\ldots,$$

and λ_1 is a sufficiently large parameter specified below.

This construction is in fact a special case of our main result which applies to general linear problems of the form

$$\mathscr{L}U = f, \quad f \in L^p_\#(\Omega), \quad \Omega \subset \mathbb{R}^d,\ 1 < p < \infty. \tag{1.2}$$

Here, $\mathscr{L} : \mathbb{B} \mapsto L^p_\#(\Omega)$ is a linear operator densely defined on a Banach space $\mathbb{B}$ with a closed range in $L^p_\#(\Omega)$. The subscript $\{\cdot\}_\#$ indicates an appropriate subspace of L^p,

$$L^p_\#(\Omega) = L^p(\Omega) \cap \mathrm{Ker}(\mathscr{P}),$$

where $\mathscr{P} : L^p \mapsto L^p$ is a linear operator whose null is "compatible" with the range of $\mathscr{L}$ so that the dual of $\mathscr{L}$ is injective, namely, there exists $\beta > 0$ such that

$$\|g - \mathscr{P}^* g\|_{L^{p'}} \le \beta \|\mathscr{L}^* g\|_{\mathbb{B}^*}, \qquad \forall g \in L^{p'}(\Omega). \tag{1.3}$$

The closed range theorem combined with the open mapping principle tell us that equation (1.2) has a bounded solution, $U \in \mathbb{B}$. Our main result is a constructive proof for the existence of such solutions in this setup.

Theorem 1.1. *Assume that the apriori estimate* (1.3) *holds. Then, for any given $f \in L^p_\#(\Omega), 1 < p < \infty$, equation* (1.2) *admits a solution of the form $U = \sum_{j=1}^\infty \mathbf{u}_j \in \mathbb{B}$,*

$$\|U\|_{\mathbb{B}} \lesssim \beta \|f\|_{L^p}, \tag{1.4}$$

where the $\{\mathbf{u}_j\}$'s are constructed recursively as minimizers of

$$\mathbf{u}_{j+1} = \underset{\mathbf{u}}{arg\,min}\big\{\|\mathbf{u}\|_{\mathbb{B}} + \lambda_{j+1}\big\|r_j - \mathscr{L}\mathbf{u}\big\|^p_{L^p}\big\}, \quad r_j := f - \mathscr{L}\big(\sum_{k=1}^{j}\mathbf{u}_k\big), \ \ j = 0,1,\ldots. \tag{1.5}$$

Here, $\{\lambda_j\}_{j\ge1}$ can be taken as any exponentially increasing sequence, $\lambda_{j+1} := \lambda_1 \zeta^j,\ j = 0,1,\ldots$ with a sufficiency large $\lambda_1 \gtrsim \dfrac{\beta}{\|f\|^{p-1}_{L^p}}$.

Remark 1.1. The description of U as the sum $U = \sum \mathbf{u}_j$ provides a multiscale *hierarchical decomposition* of a solution for (1.2). The role $\{\lambda_j\}$'s as the different scales associated with the $\mathbf{u}_j$'s, is realized in terms of the "energy bound"

$$\sum_{j=1}^{\infty} \frac{1}{\zeta^j} \|\mathbf{u}_j\|_{\mathbb{B}} \lesssim \beta \|f\|_{L^p}, \quad f \in L^p_{\#}(\Omega), \qquad \lambda_{j+1} = \lambda_1 \zeta^j. \tag{1.6}$$

The particular choice of dyadic scales, $\zeta = 2^{p-1}$, for example, yields the specific $\mathbb{B}$-bound in (1.4),

$$\|U\|_{\mathbb{B}} \leq \frac{\beta 2^p}{p} \|f\|_{L^p}, \qquad U = \sum_{j=1}^{\infty} \mathbf{u}_j, \quad \lambda_{j+1} = \frac{\beta 2^{(p-1)j}}{\|f\|_{L^p}^{p-1}}, \ j = 0, 1, \ldots. \tag{1.7}$$

Remark 1.2. We emphasize that the hierarchical construction $U = \sum \mathbf{u}_j$ does *not* require apriori knowledge of the constant β appearing in the duality estimate (1.3). The parameter β enters through the initial scale, λ_1, which is chosen sufficiently large so that,

$$\lambda_1 \gtrsim \frac{\beta}{\|f\|_{L^p}^{p-1}} > \frac{1}{\|\mathscr{L}^*(sgn(f)|f|^{p-1})\|_{\mathbb{B}^*}}.$$

By lemma 5.3, it dictates a non-trivial first hierarchical step, $\mathbf{u}_1 = \underset{\mathbf{u}}{arg\,min}\left\{\|\mathbf{u}\|_{\mathbb{B}} + \lambda_1 \left\|f - \mathscr{L}\mathbf{u}\right\|_{L^p}^p\right\}$. If the initial scale λ_1 is underestimated, however, then as already noted in [TNV08, Remark 2.1] (consult lemma 5.2 below), the hierarchical expansion will yield zero hierarchical terms, $\mathbf{u}_j \equiv 0, j = 1, 2, \ldots$, until reaching the critical scale $\lambda_1 2^{j_0} \gtrsim \beta \|f\|_{L^p}^{1-p}$ which will dictate the initial non-zero step of the hierarchical decomposition, $U = \sum_{j=j_0} \mathbf{u}_j$. In this sense, the construction of hierarchical solution, $U = \sum \mathbf{u}_j$ is *independent* of the apriori estimate (1.3): the latter is only needed to guarantee that there exists a β so that the hierarchical construction will indeed pick up at some finite scale λ_{j_0}.

The main novelty of theorem 1.1 is a constructive proof of the closed range theorem for operators with a closed range in $L^p_{\#}$, $1 < p < \infty$. The proof of the special case $p = 2$ is given in section 2: here, we trace precise bounds and clarify their role in the hierarchical construction. The L^2-case serves as the prototype case for the general setup of hierarchical constructions in Orlicz spaces, outlined in section 3. The general L^p-case is then deduced in section 4. Finally, the characterization of minimizers, such as those encountered in (1.5), is summarized in section 5.

L^2-based hierarchical decompositions were introduced by us in the context of image processing, [TNV04], and we realize that their construction in the more general context of the closed range theorem could be useful in many applications. We demonstrate the hierarchical constructions in two important examples recently studied by Bourgain & Brezis, [BB03, BB07].

1.1. **Bounded solutions of** $div\,U = f \in L^p_{\#}$. Let $\mathscr{P}$ denote the averaging projection, $\mathscr{P}g := \overline{g}$ where $\overline{g}$ is the average value of g. Given $f \in L^p_{\#}(\mathbb{T}^d) := \left\{g \in L^p(\mathbb{T}^d) \mid \overline{g} = 0\right\}$, then according to theorem 1.1, we can construct hierarchical solutions of

$$div\,U = f, \qquad f \in L^p_{\#}(\mathbb{T}^d), \quad 1 < p < \infty, \tag{1.8}$$

in an appropriate Banach space, $U \in \mathbb{B}$, provided the corresponding apriori estimate (1.3) holds, namely, there exists a constant $\beta > 0$ (which may vary of course, depending on p, d and $\mathbb{B}$), such that

$$\|g - \overline{g}\|_{L^{p'}} \leq \beta \|\nabla g\|_{\mathbb{B}^*}, \qquad \forall g \in L^{p'}(\mathbb{T}^d). \tag{1.9}$$

We specify four cases of such relevant $\mathbb{B}$'s.

#1. **Solutions of (1.8) in $\dot{W}^{1,p}$.** Since

$$\|g-\overline{g}\|_{L^{p'}(\mathbb{T}^d)} \leq \|\nabla g\|_{\dot{W}^{-1,p'}(\mathbb{T}^d,\mathbb{R}^d)}, \quad \forall g \in L^{p'}(\mathbb{T}^d),$$

we can construct hierarchical solutions of (1.8) in $\mathbb{B} = \dot{W}^{1,p}(\mathbb{T}^d,\mathbb{R}^d), 1<p<\infty$. This is the same integrability space of the irrotational solution of (1.8), $\nabla\Delta^{-1}f \in \dot{W}^{1,p}(\mathbb{T}^d,\mathbb{R}^d)$.

Remark 1.3. The L^p-valued hierarchical constructions in theorem 1.1 is derived from a more general result of operators valued in Orlicz spaces, L^Φ, associated with proper N-function Φ. This is carried out in our main theorem 3.2 below, under the assumption that (3.13) holds. The end cases $p=1,\infty$ are excluded from theorem 1.1 since the corresponding L^p-spaces are not associated with proper N-functions and the closure assumption (3.13) fails. Observe that the end cases $p=1,\infty$ do *not* admit $\dot{W}^{1,p}$ solutions for $div\, U = f \in L^p$, [BB03, Section 2], [DFT05]. Instead, theorem 3.2 suggests that the corresponding end cases apply to critical $L(\log L)$ and $\exp(L)$ spaces, but we shall not pursue this here.

#2. **Solutions of (1.8) in L^{p^*}.** By Sobolev inequality

$$\|g-\overline{g}\|_{L^{p'}(\mathbb{T}^d)} \leq \beta\|\nabla g\|_{L^{(p^*)'}(\mathbb{T}^d,\mathbb{R}^d)}, \qquad \frac{1}{p^*}=\frac{1}{p}-\frac{1}{d}, \qquad d \leq p < \infty, \quad \forall g \in L^{p'}(\mathbb{T}^d), \tag{1.10}$$

where the case $p=d$ corresponds to the is the Sobolev-Nirenberg inequality, $\|g-\overline{g}\|_{L^{d'}(\mathbb{T}^d)} \leq \beta\|g\|_{BV(\mathbb{T}^d)}$. We distinguish between two cases.

(i) The case $d<p<\infty$: the equation $div\, U = f \in L^p_\#(\mathbb{T}^d)$ has a solution $U \in L^{p^*}(\mathbb{T}^d,\mathbb{R}^d)$. This is the same integrability space of the irrotational solution $\nabla\Delta^{-1}f \in W^{1,p}(\mathbb{T}^d,\mathbb{R}^d) \subset L^{p^*}(\mathbb{T}^d,\mathbb{R}^d)$.

(ii) The case $d=p$: the equation $div\, U = f \in L^d_\#(\mathbb{T}^d)$ has a solution $U \in L^\infty(\mathbb{T}^d,\mathbb{R}^d)$. This is the the prototype example discussed in the beginning of the introduction. According to the intriguing observation of Bourgain & Brezis, [BB03, Proposition 2], there exists no bounded right inverse $K: L^d_\# \mapsto L^\infty$ for the operator div, and therefore, there exists no *linear* construction of solutions $f \mapsto U$ (in particular, $\nabla\Delta^{-1}f$ cannot be uniformly bounded). Theorem 1.1 provides a *nonlinear* hierarchical construction of such solutions.

Remark 1.4. We rewrite the hierarchical iteration (1.5) with $\lambda_1 = C\|f\|_{L^p}^{1-p}$ in the form

$$[\mathbf{u}_{j+1},r_{j+1}] = \underset{\mathscr{L}\mathbf{u}+r=r_j}{arg\,min}\left\{\|\mathbf{u}\|_{\mathbb{B}} + C\zeta^j\frac{\|r\|^p_{L^p}}{\|f\|^{p-1}_{L^p}}\right\}, \quad r_j := \begin{cases} f, & j=0 \\ f-\mathscr{L}\big(\sum_{k=1}^j \mathbf{u}_k\big), & j=1,2\ldots. \end{cases}$$

Observe that if $[\mathbf{u}_1,r_1]$ is the first minimizer associated with $r_0=f$, then the corresponding first minimizer associated with αf is $[\alpha\mathbf{u}_1,\alpha r_1]$, and recursively, the next hierarchical components are $[\alpha\mathbf{u}_j,\alpha r_j]$. Thus, as noted in [TNV08, remark 1.1], the hierarchical decomposition is homogeneous of degree one: if we let $\mathbf{u}_j \equiv \mathbf{u}_j[f]$ specify the nonlinear f-dependence of the hierarchical componenets on f, then $\mathbf{u}_j[\alpha f] = \alpha\mathbf{u}_j[f]$.

#3. **Solutions of (1.8) in $L^\infty \cap \dot{W}^{1,d}$.** A central question raised and answered in [BB03] is whether (1.8) has a solution which captures the *joint* regularity, $U \in \mathbb{B} = L^\infty \cap \dot{W}^{1,d}(\mathbb{T}^d,\mathbb{R}^d)$. To this end, one needs to verify the duality estimate (1.9), which now reads

$$\|g-\overline{g}\|_{L^{d'}(\mathbb{T}^d)} \leq \beta\|\nabla g\|_{L^1+\dot{W}^{-1,d'}(\mathbb{T}^d,\mathbb{R}^d)}, \qquad \forall g \in L^{d'}(\mathbb{T}^d). \tag{1.11}$$

This key estimate was proved in [BB03]. Thus, theorem 1.1 converts (1.11) into a constructive proof of:

Corollary 1.2. *The equation* $div\,U = f \in L^d_\#(\mathbb{T}^d)$ *admits a solution* $U \in L^\infty \cap \dot{W}^{1,d}(\mathbb{T}^d, \mathbb{R}^d)$*, given by the hierarchical decomposition* $U = \sum_{j=1} \mathbf{u}_j$*, which is constructed by the refinement step,*

$$\mathbf{u}_{j+1} = \underset{\mathbf{u}}{arg\,min} \left\{ \|\mathbf{u}\|_{L^\infty \cap \dot{W}^{1,d}} + \lambda_1 \zeta^j \big\| f - div\big(\sum_{k=1}^{j} \mathbf{u}_k\big) - div\,\mathbf{u} \big\|^d_{L^d} \right\}, \quad j = 0, 1, 2 \ldots,$$

with $\zeta > 1$ *and* $\lambda_1 \gtrsim \beta \|f\|^{1-d}_{L^d}$.

Remark 1.5. We note the key role of the duality estimate (1.11). In $d = 2$ it was proved by Bourgain & Brezis using a direct method outlined in [BB03, Section 4]. The two-dimensional estimate was revisited by direct proofs of Maz'ya [Ma07] and Mironescu [Mi10]. For $d > 2$, however, (1.11) is in fact a byproduct of Bourgain & Brezis *construction* of $L^\infty \cap \dot{W}^{1,d}$ solutions for $div\,U = f$, [BB03, theorem 1]: their construction, which is based on Littlewood-Paley decomposition, is rather involved. Corollary 1.2 offers a simpler construction of such solutions which could be implemented in actual computations based on the construction of minimizers for, $\vee_{div}(r, \lambda; L^\infty, L^d_\#) := \inf_{\mathbf{u}} \big\{ \|\mathbf{u}\|_{L^\infty} + \lambda \|r - div\,\mathbf{u}\|^d_{L^d} \big\}$, [GLMV07, LV05]. It would be desirable to develop efficient algorithms to compute minimizers of $\vee_{div}(r, \lambda; L^\infty \cap \dot{W}^{1,d}, L^p_\#)$. Spectral approximation of such minimizers was discussed in [Ma06].

Since the proof of the dual estimate (1.11) in $d > 2$ dimensions is indirect, a specific value of β is not known. As noted earlier, however, the hierarchical construction can proceed without a precise knowledge of β: if one sets $\lambda_1 = \|f\|^{1-d}_{L^d}$ and this initial scale underestimates a correct value of $\beta > 1$, then it will take at most $j_0 \sim \log(\beta)$ steps before picking-up non-trivial terms in the hierarchical decomposition, $U = \sum_{j=j_0} \mathbf{u}_j$.

#4. **Solutions of (1.8) in** $L^\infty \cap \dot{W}^{1,d}_0(\Omega)$. The above constructions of bounded solutions for (1.8) extends to the case of Lipschitz domains, $\Omega \subset \mathbb{R}^d$, see [BB03, section 7.2]. For future reference we state the following.

Corollary 1.3. *Given* $f \in L^d_\#(\Omega) := \big\{ f \in L^p(\Omega) \mid \int_\Omega f(x)dx = 0 \big\}$*, then the equation* $div\,U = f$ *admits a solution* $U \in L^\infty \cap \dot{W}^{1,d}_0(\Omega, \mathbb{R}^d)$*, such that*

$$\|U\|_{L^\infty \cap \dot{W}^{1,d}(\Omega)} \leq \beta \|f\|_{L^d(\Omega)}.$$

It is given by the hierarchical decomposition, $U = \sum_{j=1} \mathbf{u}_j$*, which is constructed by the refinement step,*

$$\mathbf{u}_{j+1} = \underset{\mathbf{u}:\ \mathbf{u}_{|\partial\Omega}=0}{arg\,min} \left\{ \|\mathbf{u}\|_{L^\infty \cap \dot{W}^{1,d}(\Omega)} + \lambda_1 \zeta^j \big\| f - div\big(\sum_{k=1}^{j} \mathbf{u}_k\big) - div\,\mathbf{u} \big\|^d_{L^d(\Omega)} \right\}, \ j = 0, 1, 2 \ldots,$$

with $\zeta > 1$ *and* $\lambda_1 \gtrsim \beta \|f\|^{1-d}_{L^d(\Omega)}$.

The bound on the hierarchical solution in (1.7) is not sharp: the existence of solution with an optimal bound, $\|U\|_{\mathbb{B}} \leq \beta' \|f\|_{L^p}$, $\forall \beta' > \beta$, can be derived using Hahn-Banach theorem, which in turn can be used to obtain optimal (least) upper-bounds in the Sobolev inequality (1.10). This is carried out in section 4 below. As an example, we have the following sharp version of the Gagliardo-Nirenberg inequality.

Corollary 1.4. *Let Ω be a Lipschitz domain in $\mathbb{R}^d$. The least upper-bound in the Gagliardo-Nirenberg inequality, $\|g-\overline{g}\|_{L^{d'}(\Omega)} \leq \beta\|g\|_{BV(\Omega)}$, is given by*

$$\beta \geq \frac{|\Omega|^{1/d'}}{|\partial\Omega|}.$$

Note that by the isoperimetric inequality, we conclude a uniform least upper-bound, $\beta = 1/(d\omega_d^{1/d})$, where ω_d is the volume of the unit ball in $\mathbb{R}^d$, [Zi89, §2.7], [DPD02, CNV04].

1.2. **Bounded solution of** $curl\, U = \mathbf{f} \in L^3_\#(\mathbb{R}^3,\mathbb{R}^3)$**.** Let $L^3_\#(\mathbb{R}^3,\mathbb{R}^3)$ denote the L^3-subspace of divergence-free 3-vectors with zero mean. We seek solutions of

$$curl\, U = \mathbf{f}, \qquad \mathbf{f} \in L^3_\#(\mathbb{R}^3,\mathbb{R}^3), \tag{1.12}$$

in an appropriate Banach space, $U \in \mathbb{B}$. We appeal to the framework of hierarchical solutions in theorem 1.1, where $\mathscr{P} : L^3(\mathbb{R}^3,\mathbb{R}^3) \mapsto L^3(\mathbb{R}^3,\mathbb{R}^3)$ is the irrotational portion of Hodge decomposition with a dual, $\mathscr{P}^*\mathbf{g} := \nabla\Delta^{-1} div\,\mathbf{g} - \overline{\mathbf{g}}$. According to theorem 1.1, we can construct hierarchical solutions, $U \in \mathbb{B}$ of (1.12), provided (1.3) holds

$$\|\mathbf{g} - \mathscr{P}^*\mathbf{g}\|_{L^{3/2}} \leq \beta\|curl\,\mathbf{g}\|_{\mathbb{B}^*}, \qquad \mathbf{g} \in L^{3/2}(\mathbb{R}^3,\mathbb{R}^3). \tag{1.13}$$

Since $\|\mathbf{g} - \mathscr{P}^*\mathbf{g}\|_{L^{3/2}} \lesssim \|curl\,\mathbf{g}\|_{\dot{W}^{-1,3/2}}$, we can construct hierarchical solutions of (1.12) in $\dot{W}^{1,3}$. This has the same integrability as the divergence-free solution of (1.12), $(-\Delta)^{-1} curl\,\mathbf{f}$. A more intricate question is whether (1.12) admits a uniformly bounded solution, since such a solution *cannot* be constructed by a linear procedure. These solutions were constructed by Bourgain and Brezis in [BB07, Corollary 8'], which in turn imply the key estimate,

$$\|\mathbf{g} - \mathscr{P}^*\mathbf{g}\|_{L^{3/2}(\mathbb{R}^3,\mathbb{R}^3)} \leq \beta\|curl\,\mathbf{g}\|_{L^1+\dot{W}^{-1,3/2}(\mathbb{R}^3,\mathbb{R}^3)}, \qquad \forall \mathbf{g} \in L^{3/2}(\mathbb{R}^3,\mathbb{R}^3) : \ div\,\mathbf{g} = \overline{\mathbf{g}} = 0. \tag{1.14}$$

Granted (1.14), theorem 1.1 offers a simpler alternative to the construction in [BB07] based on the following hierarchical decomposition.

Corollary 1.5. *The equation $curl\, U = \mathbf{f} \in L^3_\#(\mathbb{R}^3,\mathbb{R}^3)$, admits a solution $U \in L^\infty \cap \dot{W}^{1,3}(\mathbb{R}^3,\mathbb{R}^3)$,*

$$\|U\|_{L^\infty\cap\dot{W}^{1,3}(\mathbb{R}^3,\mathbb{R}^3)} \leq \beta\|\mathbf{f}\|_{L^3(\mathbb{R}^3,\mathbb{R}^3)},$$

which can be constructed by the (nonlinear) hierarchical expansion, $U = \sum \mathbf{u}_j$,

$$\mathbf{u}_{j+1} = \underset{\mathbf{u}}{arg\,min}\left\{ \|\mathbf{u}\|_{L^\infty\cap\dot{W}^{1,3}} + \lambda_1\zeta^j \Big\|\mathbf{f} - curl\,\Big(\sum_{k=1}^{j}\mathbf{u}_k\Big) - curl\,\mathbf{u}\Big\|^3_{L^3(\mathbb{R}^3,\mathbb{R}^3)} \right\}, \quad j = 0,1,\ldots,$$

with $\zeta > 1$ and $\lambda_1 \gtrsim \beta/\|\mathbf{f}\|^2_{L^3(\mathbb{R}^3,\mathbb{R}^3)}$.

2. AN EXAMPLE: HIERARCHICAL SOLUTION OF $div\, U = f \in L^2_\#(\mathbb{T}^2)$

We begin our discussion on hierarchical constrictions with a two-dimensional prototype example of

$$div\, U = f, \qquad f \in L^2_\#(\mathbb{T}^2). \tag{2.1}$$

Our starting point for the construction of a uniformly bounded solution of (2.1), $U \in L^\infty(\mathbb{T}^2,\mathbb{R}^2)$, is a decomposition of f,

$$f = div\,\mathbf{u}_1 + r_1, \qquad f \in L^2_\#(\mathbb{T}^2) := \Big\{ g \in L^2(\mathbb{T}^2) \ \Big| \ \int_{\mathbb{T}^2} g(x)dx = 0 \Big\}, \tag{2.2a}$$

where $[\mathbf{u}_1, r_1]$ is a minimizing pair of the functional,

$$[\mathbf{u}_1, r_1] = \underset{div\,\mathbf{u}+r=f}{arg\,min} \left\{ \|\mathbf{u}\|_{L^\infty} + \lambda_1 \|r\|^2_{L^2} \right\}. \tag{2.2b}$$

Here λ_1 is a fixed parameter at our disposal: if we choose λ_1 large enough, $\lambda_1 > \dfrac{1}{2\|f\|_{BV}}$, then according to lemma 5.3 below, (2.2b) admits a minimizer, $[\mathbf{u}_1, r_1]$, satisfying,

$$\|r_1\|_{BV} = \frac{1}{2\lambda_1}.$$

To proceed we invoke the Gagliardo-Nirenberg isoperimetric inequality, which states that there exists $\beta > 0$ (any $\beta \geq 1/\sqrt{4\pi}$ will do, [Zi89, §2.7]), such that for all bounded variation g's with zero mean,

$$\|g\|_{L^2} \leq \beta \|g\|_{BV}, \qquad \int_{\mathbb{T}^2} g(x)dx = 0 \tag{2.3}$$

Since f has a zero mean so does the residual r_1 and (2.3) yields

$$\|r_1\|_{L^2} \leq \beta \|r_1\|_{BV} = \frac{\beta}{2\lambda_1}. \tag{2.4}$$

We conclude that the residual $r_1 \in L^2_\#(\mathbb{T}^2)$, and we can therefore implement the same variational decomposition of f in (2.2), and use it to decompose r_1 with scale $\lambda = \lambda_2 > \lambda_1 = \dfrac{1}{2\|r_1\|_{BV}}$. This yields

$$r_1 = div\,\mathbf{u}_2 + r_2, \qquad [\mathbf{u}_2, r_2] = \underset{div\,\mathbf{u}+r=r_1}{arg\,min} \left\{ \|\mathbf{u}\|_{L^\infty} + \lambda_2 \|r\|^2_{L^2} \right\}. \tag{2.5}$$

Combining (2.5) with (2.2a) we obtain $f = div\,U_2 + r_2$, where $U_2 := \mathbf{u}_1 + \mathbf{u}_2$ is viewed as an improved *approximate solution* of (2.1) in the sense that it has a smaller residual r_2,

$$\|r_2\|_{L^2} \leq \beta \|r_2\|_{BV} = \frac{\beta}{2\lambda_2}.$$

This process can be repeated: if $r_j \in L^2_\#(\mathbb{T}^2)$ is the residual at step j, then we decompose it

$$r_j = div\,\mathbf{u}_{j+1} + r_{j+1}, \tag{2.6a}$$

where $[\mathbf{u}_{j+1}, r_{j+1}]$ is a minimizing pair of

$$[\mathbf{u}_{j+1}, r_{j+1}] = \underset{div\,\mathbf{u}+r=r_j}{arg\,min} \left\{ \|\mathbf{u}\|_{L^\infty} + \lambda_{j+1} \|r\|^2_{L^2} \right\}, \qquad j = 0, 1, \ldots. \tag{2.6b}$$

For $j = 0$, the decomposition (2.6) is interpreted as (2.2a) by setting $r_0 := f$. Note that the recursive decomposition (2.6a) depends on the invariance of $r_j \in L^2_\#(\mathbb{T}^2)$: if r_j has a zero mean then so does r_{j+1}, and by (2.3) $r_{j+1} \in L^2_\#(\mathbb{T}^2)$. The iterative process depends on a sequence of increasing scales, $\lambda_1 < \lambda_2 < \ldots \lambda_{j+1}$, which are yet to be determined.

The telescoping sum of the first k steps in (2.6a) yields an improved approximate solution, $U_k := \sum_{j=1}^k \mathbf{u}_j$:

$$f = div\,U_k + r_k, \qquad \|r_k\|_{L^2} \leq \beta \|r_k\|_{BV} = \frac{\beta}{2\lambda_k} \downarrow 0, \qquad k = 1, 2, \ldots. \tag{2.7}$$

The key question is whether the U_k's remain uniformly bounded, and it is here that we use the freedom in choosing the scaling parameters λ_k: comparing the minimizing pair $[\mathbf{u}_{j+1}, r_{j+1}]$ of (2.6b) with the trivial pair $[\mathbf{u}\equiv 0, r_j]$ implies, in view of (2.7),

$$\|\mathbf{u}_{j+1}\|_{L^\infty} + \lambda_{j+1}\|r_{j+1}\|^2_{L^2} \le \lambda_{j+1}\|r_j\|^2_{L^2} \le \begin{cases} \lambda_1\|f\|^2_{L^2}, & j=0, \\[2mm] \dfrac{\beta^2\lambda_{j+1}}{4\lambda_j^2}, & j=1,2,\ldots. \end{cases} \tag{2.8}$$

We conclude that by choosing a sufficiently fast increasing λ_j's such that $\sum_j \lambda_{j+1}\lambda_j^{-2} < \infty$, then the approximate solutions $U_k = \sum_1^k \mathbf{u}_j$ form a Cauchy sequence in L^∞ whose limit, $U = \sum_1^\infty \mathbf{u}_j$, satisfies the following.

Theorem 2.1. *Fix β such that* (2.3) *holds. Then, for any given $f \in L^2_\#(\mathbb{T}^2)$, there exists a uniformly bounded solution of* (2.1),

$$div\, U = f, \qquad \|U\|_{L^\infty} \le 2\beta\|f\|_{L^2}. \tag{2.9}$$

The solution U is given by $U = \sum_{j=1}^\infty \mathbf{u}_j$, where the $\{\mathbf{u}_j\}$'s are constructed recursively as minimizers of

$$[\mathbf{u}_{j+1}, r_{j+1}] = \underset{div\,\mathbf{u}+r=r_j}{arg\,min} \left\{\|\mathbf{u}\|_{L^\infty} + \lambda_1 2^j\|r\|^2_{L^2}\right\}, \qquad r_0 := f, \quad \lambda_1 = \frac{\beta}{\|f\|_{L^2}}. \tag{2.10}$$

Proof. Set $\lambda_j = \lambda_1 2^{j-1}, j=1,2,\ldots$, then, $\|U_k - U_\ell\|_{L^\infty} \lesssim 2^{-k},\ k > \ell \gg 1$. Let U be the limit of the Cauchy sequence $\{U_k\}$ then $\|U_j - U\|_{L^\infty} + \|div\,U_j - f\|_{L^2} \lesssim 2^{-j} \to 0$, and since div has a closed graph on its domain $D := \{\mathbf{u}\in L^\infty :\ div\,\mathbf{u} \in L^2(\Omega)\}$, it follows that $div\,U = f$. By (2.8) we have

$$\|U\|_{L^\infty} \le \sum_{j=1}^\infty \|\mathbf{u}_j\|_{L^\infty} \le \lambda_1\|f\|^2_{L^2} + \frac{\beta^2}{4\lambda_1}\sum_{j=2}^\infty \frac{1}{2^{j-3}} = \lambda_1\|f\|^2_{L^2} + \frac{\beta^2}{\lambda_1}.$$

Here $\lambda_1 > \dfrac{1}{2\|f\|_{BV}}$ is a free parameter at our disposal: we choose $\lambda_1 := \beta/\|f\|_{L^2}$ which by (2.3) is admissible, $\lambda_1 = \dfrac{\beta}{\|f\|_{L^2}} > \dfrac{1}{2\|f\|_{BV}}$, and (2.9) follows. $\square$

Remark 2.1. [Energy decomposition] A telescoping summation of the left inequality of (2.8) yields

$$\sum_{j=1}^\infty \frac{1}{\lambda_j}\|\mathbf{u}_j\|_{L^\infty} \le \|f\|^2_{L^2};$$

setting $\lambda_j = \beta 2^{j-1}/\|f\|_{L^2}$, we conclude the "energy bound"

$$\sum_{j=1}^\infty \frac{1}{2^{j-1}}\|\mathbf{u}_j\|_{L^\infty} \le \beta\|f\|_{L^2}. \tag{2.11}$$

In fact, a precise energy *equality* can be formulated in this case, using the characterization of the minimizing pair (consult theorem 5.1 below), $2(r_{j+1}, div\,\mathbf{u}_{j+1}) = \|\mathbf{u}_{j+1}\|_{L^\infty}/\lambda_{j+1}$: by squaring the refinement step $r_j = r_{j+1} + \mathscr{L}\mathbf{u}_{j+1}$ we find

$$\|r_j\|^2_{L^2} - \|r_{j+1}\|^2_{L^2} = 2(r_{j+1}, div\,\mathbf{u}_{j+1}) + \|div\,\mathbf{u}_{j+1}\|^2_{L^2} = \frac{1}{\lambda_{j+1}}\|\mathbf{u}_{j+1}\|_{L^\infty} + \|div\,\mathbf{u}_{j+1}\|^2_{L^2}.$$

A telescoping sum of the last equality yields

Corollary 2.2. *Let $U=\sum_{j=1}^{\infty}\mathbf{u}_j \in L^\infty$ be a hierarchical solution of $div\, U = f,\ f\in L^2_\#(\mathscr{L}^2)$. Then*

$$\frac{1}{\lambda_1}\sum_{j=1}^{\infty}\frac{1}{2^{j-1}}\|\mathbf{u}_j\|_{L^\infty}+\sum_{j=1}^{\infty}\|div\,\mathbf{u}_j\|^2_{L^2_\#(\mathbb{T}^2)}=\|f\|^2_{L^2_\#(\mathbb{T}^2)} \tag{2.12}$$

2.1. **Oscillations and image processing.** Bourgain and Brezis [BB03, Proposition 2] have shown that there exists no linear construction of solutions of (2.1) for *general* $f\in L^2$. Yet, for the 'slightly smaller' Lorenz space, $L^{2,1}$, we have

$$\nabla\Delta^{-1}f\in L^\infty, \qquad f\in L^{2,1}_\#(\mathbb{T}^2).$$

(we note in passing that $L^{2,1}$ is a limiting case for the linearity of $f\mapsto U$ to survive the $L^{2,\infty}$-based nonlinearity result argued in the proof of [BB03, propositoin 2]). Thus, the nonlinear aspect of constructing hierarchical solutions for (2.1) becomes essential for highly oscillatory functions such that $f\in L^2\backslash L^{2,1}$ (and in particular, $f\notin BV(\mathbb{T}^2)$). Such f's are encountered in image processing in the form of noise, texture, and blurry images, [Me02, BCM05]. Hierarchical decompositions in this context of images were introduced by us in [TNV04] and were found to be effective tools in image denoising, image deblurring and image registration, [BCM05, LPSX06, PL07, TNV08]. Here, we are given a noisy and possibly blurry observed image, $f=\mathscr{L}U+r\in L^2(\mathbb{R}^2)$, and the purpose is to recover a faithful description of the underlying 'clean' image, $U\sim$ "$\mathscr{L}^{-1}$"f, by de-noising r and de-blurring $\mathscr{L}$. The inverse "$\mathscr{L}^{-1}$"f should be properly interpreted, say, in the smaller space $BV(\mathbb{R}^2)\subset L^2(\mathbb{R}^2)$ which is known to be well-adapted to represent edges. The resulting inverse problem can solved by corresponding variational problem of [ROF92, CL95, CL97]

$$[\mathbf{u},r]=\underset{\mathscr{L}\mathbf{u}+r=f}{arg\,min}\left\{\|\mathbf{u}\|_{BV}+\lambda\|r\|^2_{L^2(\mathbb{R}^2)}\right\}, \tag{2.13}$$

which is a special case of *Tikhonov-regularization*, [TA77, Mo84, Mo93]. The (BV,L^2)-hierarchical decomposition corresponding to (2.13) reads, [TNV04, TNV08],

$$f\cong\mathscr{L}U_m, \quad U_m=\sum_{j=1}^{m}\mathbf{u}_j, \qquad [\mathbf{u}_{j+1},r_{j+1}]=\underset{\mathscr{L}\mathbf{u}+r=r_j}{arg\,min}\left\{\|\mathbf{u}\|_{BV}+\lambda_1 2^j\|r\|^2_{L^2}\right\}. \tag{2.14}$$

The oscillatory nature of noise and texture in images was addressed by Y. Meyer who advocated, [Me02], to replace L^2 with the larger space of "images" $G:=\{f\ |div\,\mathbf{u}=f,\mathbf{u}\in L^\infty\}$. The equation $div\,\mathbf{u}=f$ arises here with *one-signed* measure f's, and its L^∞ solutions were characterized in [Me02, §1.14],[PT08]: the space G_+ coincides with Morrey space $M^2_+(\Omega)$:

$$M^2(\Omega)=\left\{\mu\in\mathscr{M}\ \Big|\int_{B_r}d\mu\lesssim r, \quad \forall B_r\subset\Omega\right\}.$$

For one-signed measure, $M^2_+(\Omega)$ coincides with Besov space $\dot{B}^{-1,\infty}_\infty$. The corresponding Meyer's energy functional then reads, $[\mathbf{u},r]=\underset{\mathscr{L}\mathbf{u}+r=f}{arg\,min}\left\{\|\mathbf{u}\|_{BV(\Omega)}+\lambda\|r\|_{\dot{B}^{-1,\infty}_\infty(\Omega)}\right\}$; numerical simulations with this model are found in [VO04].

2.2. $L^1(\mathbb{T}^2)$**-bounds and** $\dot{H}^{-1}(\mathbb{T}^2)$**-compactness.** Here is a simple application of theorem 3.2. Let $f\in\dot{H}^{-1}(\mathbb{T}^2)$ be given. For arbitrary $g\in\dot{H}^1(\mathbb{T}^2)$ we have $\xi_j\widehat{g}(\xi)\in L^2_\#(\mathbb{T}^2)$, and by theorem 3.2, there exist bounded $U_{ij}\in L^\infty(\mathbb{T}^2)$, such that

$$\begin{aligned}\xi_1\widehat{g}(\xi) &= \xi_1\widehat{U}_{11}(\xi)+\xi_2\widehat{U}_{12}(\xi),\\ \xi_2\widehat{g}(\xi) &= \xi_1\widehat{U}_{21}(\xi)+\xi_2\widehat{U}_{22}(\xi).\end{aligned}$$

Thus, expressed in terms of the Riesz transforms, $\widehat{R_j\psi}(\xi) := \widehat{\psi}(\xi)\xi_j/|\xi|$, we have

$$g = \frac{1}{2}(U_{11}+U_{22}) + \frac{1}{2}\left(R_1^2 - R_2^2\right)(U_{11}-U_{22}) + R_1R_2(U_{12}+U_{21});$$

Since $R_1^2 - R_2^2$ and R_1R_2 agree up to rotation, we conclude that: every $g \in \dot{H}^1(\mathbb{T}^2)$ can be written as the sum

$$g \in \dot{H}^1(\mathbb{T}^2): \qquad g = U_1 + R_1R_2U_2, \quad U_1, U_2 \in L^\infty(\mathbb{T}^2).$$

The last representation shows that although an $L^1(\mathbb{T}^2)$-bound of f does not imply $f \in \dot{H}^{-1}(\mathbb{T}^2)$, then f does belong to $\dot{H}^{-1}$ if f *and* its repeated Riesz transform, R_1R_2f, are L^1-bounded.

Corollary 2.3. *The following bound holds*

$$\|f\|_{\dot{H}^{-1}(\mathbb{T}^2)} \lesssim \|f\|_{L^1(\mathbb{T}^2)} + \|R_1R_2f\|_{L^1(\mathbb{T}^2)}. \tag{2.15}$$

As an example, consider a family of divergence-free 2-vector fields, $\mathbf{u}^\varepsilon(t,\cdot) \in L^2(\mathbb{T}^2,\mathbb{R}^2)$, which are approximate solutions of two-dimensional incompressible Euler's equations. One is interested in their convergence to a proper weak solution, with no concentration effects, [DM87]. It was shown in [LNT00] that $\{\mathbf{u}^\varepsilon\}$ converges to such a weak solution if the vorticity, $\omega^\varepsilon(t\cdot) := \partial_1 u_2^\varepsilon(t,\cdot) - \partial_2 u_1^\varepsilon(t,\cdot)$, is compactly embedded in $H^{-1}(\mathbb{T}^2)$. By corollary 2.3, H^{-1}-compactness holds if $\{R_1R_2\omega^\varepsilon(t,\cdot)\} \hookrightarrow L^1(\mathbb{T}^2)$; consult [Ve92].

3. The general case: construction of hierarchical solutions for $\mathscr{L}U = f \in L^\Phi_\#$

We turn our attention to the construction of hierarchical solutions for equations of the general form

$$\mathscr{L}U = f, \qquad f \in L^\Phi_\#(\Omega). \tag{3.1}$$

A solution U is sought in a Banach space $\mathbb{B} := \{U : \|U\|_{\mathbb{B}} < \infty\}$. The general framework, involving two linear operators, $\mathscr{L}$ and $\mathscr{P}$, is outlined below.

The linear operator $\mathscr{L}$ is densely defined on $\mathbb{B}$ with a closed range in $L^\Phi_\# := L^\Phi \cap \mathrm{Ker}(\mathscr{P})$ with appropriate $\mathscr{P} : L^\Phi \mapsto L^\Phi$. Here, $L^\Phi = L^\Phi(\Omega)$ is an Orlicz function space associated with a C^1 N-function, $\Phi : \mathbb{R}_+ \to \mathbb{R}_+$, satisfying the Δ_2-condition, $\Phi(2\xi) < 2\Phi(\xi)$, [AF03, §8],[BS88, §4.8],

$$\mathscr{L} : \mathbb{B} \mapsto L^\Phi_\# := L^\Phi \cap \mathrm{Ker}(\mathscr{P}), \qquad L^\Phi = \left\{f \,:\, [f]_\Phi := \int_\Omega \Phi(|f|)dx < \infty\right\}. \tag{3.2}$$

The Orlicz space L^Ψ, the dual of L^Φ, is associated with the complementary N-function $\Psi(s) = \sup_t(st - \Phi(t))$. Finally, we let $\mathscr{L}^* : L^\Psi \mapsto \mathbb{B}^*$ denote the formal dual of $\mathscr{L}$, acting on L^Ψ with the natural pairing (effectively, $\mathscr{L}^*$ is acting on $L^\Psi_\# := L^\Psi \cap \mathrm{Ker}(\mathscr{P})$, since $R(\mathscr{P}^*)$ is in the null of $\mathscr{L}^*$)

$$\langle \mathscr{L}^* g, \mathbf{u}\rangle = (g, \mathscr{L}\mathbf{u}), \qquad g \in L^\Psi,\ \mathbf{u} \in \mathbb{B},$$

and let $\|\cdot\|_{\mathbb{B}^*}$ denote the dual norm

$$\|\mathscr{L}^* g\|_{\mathbb{B}^*} := \sup_{\mathbf{u}\neq 0} \frac{\langle \mathscr{L}^* g, \mathbf{u}\rangle}{\|\mathbf{u}\|_{\mathbb{B}}}, \qquad g \in L^\Psi.$$

3.1. **Approximate solutions.** We begin by constructing an *approximate solution* of (3.1), $U_\lambda : \mathscr{L}U_\lambda \approx f$, such that the residual $r_\lambda := f - \mathscr{L}U_\lambda$ is driven to be small by a proper choice of a *scaling parameter* λ at our disposal. The approximate solution is obtained in terms of minimizers of the variational problem,

$$\vee_{\mathscr{L}}(f,\lambda;\mathbb{B},\Phi) := \inf_{\mathscr{L}\mathbf{u}+r=f} \Big\{ \|\mathbf{u}\|_{\mathbb{B}} + \lambda[r]_\Phi \,:\, \mathbf{u}\in\mathbb{B},\ r\in L^\Phi_\# \Big\}. \tag{3.3}$$

In theorem 5.1 below we show if λ is chosen sufficiently large,

$$\lambda > \frac{1}{\|\mathscr{L}^*\varphi(f)\|_{\mathbb{B}^*}}, \qquad \varphi(f) := \operatorname{sgn}(f)\Phi'(|f|), \tag{3.4}$$

then the $\vee_{\mathscr{L}}(f,\lambda)$-functional in (3.3) admits a minimizer, $\mathbf{u}=\mathbf{u}_\lambda$, such that the size of the residual, $r_\lambda := f - \mathscr{L}\mathbf{u}_\lambda$, is dictated by the dual statement

$$\|\mathscr{L}^*\varphi(r_\lambda)\|_{\mathbb{B}^*} = \frac{1}{\lambda}. \tag{3.5}$$

Fix the scale $\lambda = \lambda_1 > 1/\|\mathscr{L}^*\varphi(f)\|_{\mathbb{B}^*}$. We construct an approximate solution, $\mathscr{L}U_1 \approx f$, $U_1 := \mathbf{u}_1$, where $\mathbf{u}_1$ is a minimizer of $\vee_{\mathscr{L}}(f,\lambda_1)$,

$$f = \mathscr{L}\mathbf{u}_1 + r_1, \qquad [\mathbf{u}_1, r_1] = \underset{\mathscr{L}\mathbf{u}+r=f}{\arg\min} \vee_{\mathscr{L}}(f,\lambda_1;\mathbb{B},\Phi) \tag{3.6}$$

Borrowing the terminology from image processing we note that the corresponding residual r_1 contains 'small' features which were left out of $\mathbf{u}_1$. Of course, whatever is interpreted as 'small' features at a given λ_1-scale, may contain significant features when viewed under a refined scale, say $\lambda_2 > \lambda_1$. To this end we *assume* that the residual $r_1 \in L^\Phi_\#$ so that we can repeat the $\vee_{\mathscr{L}}$-decomposition of r_1, this time at the refined scale λ_2:

$$r_1 = \mathscr{L}\mathbf{u}_2 + r_2, \qquad [\mathbf{u}_2, r_2] = \underset{\mathscr{L}\mathbf{u}+r=r_1}{\arg\min} \vee_{\mathscr{L}}(r_1,\lambda_2;\mathbb{B},\Phi). \tag{3.7}$$

Combining (3.6) with (3.7) we arrive at a better two-scale representation of U given by $U_2 := \mathbf{u}_1 + \mathbf{u}_2$, as an improved approximate solution of $\mathscr{L}U_2 \approx f$. Features below scale λ_2 remain unresolved in U_2, but the process can be continued by successive application of the refinement step,

$$r_j = \mathscr{L}\mathbf{u}_{j+1} + r_{j+1}, \quad [\mathbf{u}_{j+1}, r_{j+1}] := \underset{\mathscr{L}\mathbf{u}+r=r_j}{\arg\min} \vee_{\mathscr{L}}(r_j,\lambda_{j+1};\mathbb{B},\Phi), \quad j=1,2,\ldots. \tag{3.8}$$

To enable this process we require that the residuals r_j remain in $L^\Phi_\#$; thus, in view of the dual bound (3.5) we make the following *closure assumption*.

Assumption 3.1. *The following apriori bound holds:*

$$\|\mathscr{L}^*\varphi(g)\|_{\mathbb{B}^*} < \infty \longrightarrow [g]_\Phi = \int_\Omega \Phi(|g(x)|)dx < \infty, \quad \varphi(g) = sgn(g)\Phi'(|g|).$$

Given a residual term $r_j \in L^\Phi_\#$, we consider the generic hierarchical step where $[\mathbf{u}_{j+1}, r_{j+1}]$ is constructed as a minimizing pair of $\vee_{\mathscr{L}}(r_j,\lambda_{j+1})$. Since $\|\mathscr{L}^*\varphi(r_{j+1})\|_{\mathbb{B}^*} = 1/\lambda_{j+1}$ then by assumption 3.1, $[r_{j+1}]_\Phi$ is finite; moreover, since r_j and $R(\mathscr{L})$ are in $\operatorname{Ker}(\mathscr{P})$ then,

$$r_{j+1} = r_j - \mathscr{L}\mathbf{u}_{j+1} \in \operatorname{Ker}(\mathscr{P}),$$

and we conclude that $r_{j+1} \in L^\Phi_\#$. In this manner, the iteration step

$$[\mathbf{u}_j, r_j] \mapsto [\mathbf{u}_{j+1}, r_{j+1}],$$

is well-defined on $\mathbb{B}\times L^{\Phi}_{\#}$. After k such steps we have,

$$
\begin{aligned}
f &= \mathscr{L}\mathbf{u}_1 + r_1 = \\
&= \mathscr{L}\mathbf{u}_1 + \mathscr{L}\mathbf{u}_2 + r_2 = \\
&= \ldots\ldots \qquad\qquad = \\
&= \mathscr{L}\mathbf{u}_1 + \mathscr{L}\mathbf{u}_2 + \cdots + \mathscr{L}\mathbf{u}_k + r_k.
\end{aligned} \tag{3.9}
$$

We end up with a multiscale, *hierarchical representation* of an approximate solution of (3.1) $U_k := \sum_{j=1}^{k}\mathbf{u}_j \in \mathbb{B}$ such that $\mathscr{L}U_k \cong f$. Here, the approximate equality $\cong$ is interpreted as the convergence of the residuals,

$$
\|\mathscr{L}^*\varphi(r_k)\|_{\mathbb{B}^*} = \frac{1}{\lambda_k} \to 0, \qquad r_k := f - \mathscr{L}U_k,
$$

dictated by the sequence of scales, $\lambda_1 < \lambda_2 < \ldots < \lambda_k$, which is at our disposal. We summarize in the following theorem.

Theorem 3.1. *[Approximate solutions] Consider $\mathscr{L}:\mathbb{B}\mapsto L^{\Phi}_{\#}(\Omega)$ and assume that the closure assumption* (3.1) *holds. Then the equation $\mathscr{L}U = f, f \in L^{\Phi}_{\#}(\Omega)$ admits an* approximate solution*, $U_k \in \mathbb{B}:\ \mathscr{L}U_k \cong f$, such that the residual, $r_k := f - \mathscr{L}U_k$, satisfy*

$$
\|\mathscr{L}^*\varphi(r_k)\|_{\mathbb{B}^*} = \frac{1}{\lambda_k}, \qquad\qquad r_k := f - \mathscr{L}U_k. \tag{3.10}
$$

The approximate solution admits the hierarchical expansion, $U_k = \sum_{j=1}^{k}\mathbf{u}_j$, where the $\mathbf{u}_j$'s are constructed as minimizers,

$$
[\mathbf{u}_{j+1}, r_{j+1}] = \mathop{arg\,min}_{\mathscr{L}\mathbf{u}+r=r_j}\left\{\|\mathbf{u}\|_{\mathbb{B}} + \lambda_{j+1}[r]_\Phi\right\}, \qquad r_0 = f
$$

and satisfy

$$
\sum_{j=1}^{k}\frac{1}{\lambda_j}\|\mathbf{u}_j\|_{\mathbb{B}} < [f]_\Phi. \tag{3.11}
$$

Proof. Compare the minimizer $[\mathbf{u}_{j+1}, r_{j+1}]$ of $\vee_{\mathscr{L}}(r_j, \lambda_{j+1})$ with the trivial pair $[\mathbf{u}\equiv 0, r_j]$ which yields the key refinment estiamte

$$
[r_j]_\Phi \geq \frac{1}{\lambda_{j+1}}\|\mathbf{u}_{j+1}\|_{\mathbb{B}} + [r_{j+1}]_\Phi, \qquad j = 0, 1, \ldots. \tag{3.12}
$$

Observe that the last inequality holds for $j = 0$ with $r_0 = f$. A telescoping summation of (3.12) yields (3.11):

$$
\sum_{j=1}^{k}\frac{1}{\lambda_j}\|\mathbf{u}_j\|_{\mathbb{B}} + [r_k]_\Phi \leq [r_0]_\Phi = [f]_\Phi.
$$

□

3.2. **From approximate to exact solutions.** According to (3.12), $\|U_k\|_{\mathbb{B}} \leq \lambda_k[f]_\Phi$. Do the approximate solutions U_k remain uniformly bounded in $\mathbb{B}$? to this end, we need to quantify the closure assumption 3.1.

Assumption 3.2. *There exists an increasing function $\eta:\mathbb{R}_+\mapsto\mathbb{R}_+$ satisfying*

$$
\int_{s=0}^{1}\frac{\eta(s)}{s^2}ds < \infty, \tag{3.13a}
$$

such that the following apriori bound holds:

$$[g]_\Phi \leq \eta\left(\|\mathscr{L}^*\varphi(g)\|_{\mathbb{B}^*}\right), \qquad \forall g \in L^\Phi_\#, \quad \varphi(g) = sgn(g)\Phi'(|g|) \in L^\Psi. \tag{3.13b}$$

Using (3.12) we find

$$\|\mathbf{u}_{j+1}\|_{\mathbb{B}} \leq \lambda_{j+1}[r_j]_\Phi \leq \begin{cases} \lambda_1[f]_\Phi, & j = 0, \\ \lambda_{j+1}\eta\left(\dfrac{1}{\lambda_j}\right), & j = 1,2,\ldots, \end{cases} \tag{3.14}$$

where $\{\lambda_j\}$ is an increasing sequence of scales at our disposal. Setting $\lambda_j = \lambda_1\zeta^{j-1}$, we conclude that the approximate solutions, $U_k = \sum_1^k \mathbf{u}_j$ is a Cauchy sequence whose limit, $U = \sum_1^\infty \mathbf{u}_j$, satisfies the following.

Theorem 3.2. *Assume that* (3.13) *holds. Then, given $f \in L^\Phi_\#(\Omega)$, there exists a solution of* (3.1)*, $\mathscr{L}U = f$. The solution U is given by the hierarchical decomposition,*

$$U = \sum_{j=1}^\infty \mathbf{u}_j, \tag{3.15a}$$

where the $\{\mathbf{u}_j\}$'s are constructed recursively as minimizers of

$$[\mathbf{u}_{j+1}, r_{j+1}] = \underset{\mathscr{L}\mathbf{u}+r=r_j}{arg\,min}\left\{\|\mathbf{u}\|_{\mathbb{B}} + \lambda_1\zeta^j[r]_\Phi \ : \mathbf{u} \in \mathbb{B},\ r \in L^\Phi_\#\right\}, \qquad r_0 := f, \ \ \zeta > 1, \tag{3.15b}$$

and the following $\mathbb{B}$-bound holds

$$\|U\|_{\mathbb{B}} \lesssim \lambda_1[f]_\Phi + \int_0^{1/\zeta\lambda_1} \frac{\eta(s)}{s^2}ds < \infty, \quad \lambda_1 \sim \frac{1}{\eta^{-1}([f]_\Phi)}. \tag{3.16}$$

Proof. For simplicity, we consider the dyadic case, $\zeta = 2$. The upper-bound (3.14) implies that $\{U_k = \sum_{j=1}^k \mathbf{u}_j\}$ is a Cauchy sequence:

$$\begin{aligned} \|U_k - U_\ell\|_{\mathbb{B}} &\leq \sum_{j=\ell+1}^k \lambda_1 2^j \eta\left(\frac{1}{\lambda_1 2^{j-1}}\right) \leq 4\lambda_1 \sum_{j=2^\ell}^{2^{k-1}} \eta\left(\frac{1}{\lambda_1 j}\right) \\ &\lesssim 4\int_{t=\lambda_1 2^\ell}^{\lambda_1 2^{k-1}} \eta\left(\frac{1}{t}\right)dt = 4\int_{1/\lambda_k}^{1/\lambda_{\ell+1}} \frac{\eta(s)}{s^2}ds, \end{aligned}$$

which is sufficiently small for $k > \ell \gg 1$ large enough. Hence U_j has a limit, $U = \sum_{j=1}^\infty \mathbf{u}_j$, such that $[\mathscr{L}U_j - f]_\Phi \to 0$, and since $\mathscr{L}$ has a closed range, $\mathscr{L}U = f$. Moreover, the limit U satisfies

$$\|U\|_{\mathbb{B}} \leq \sum_{j=1}^\infty \|\mathbf{u}_j\|_{\mathbb{B}} \leq \lambda_1[f]_\Phi + \sum_{j=2}^\infty \lambda_1 2^j \eta\left(\frac{1}{\lambda_1 2^{j-1}}\right) \leq \lambda_1[f]_\Phi + 4\int_0^{1/2\lambda_1} \frac{\eta(s)}{s^2}ds.$$

Here λ_1 is a free parameter at our disposal subject to (3.4): we choose $\lambda_1 := C/\eta^{-1}([f]_\Phi)$ with a sufficiently large constant C so that (3.4) holds,

$$\lambda_1 := \frac{C}{\eta^{-1}([f]_\Phi)} > \frac{1}{\|\mathscr{L}^*\varphi(f)\|_{\mathbb{B}^*}},$$

and the bound (3.16) follows. $\square$

3.3. **Energy decomposition.** We consider Orlicz spaces, L^Φ with a *strictly* convex N- function

$$\Phi'' \geq \kappa > 0. \tag{3.17}$$

Starting with the basic refinement step $r_j = r_{j+1} + \mathscr{L}\mathbf{u}_{j+1}$ we then find

$$[r_j]_\Phi = \int_\Omega \Phi\left(|r_{j+1} + \mathscr{L}\mathbf{u}_{j+1}|\right) dx \geq \int_\Omega \left(\Phi(|r_{j+1}|) + (\varphi(r_{j+1}), \mathscr{L}\mathbf{u}_{j+1}) + \frac{\kappa}{2}\|\mathscr{L}\mathbf{u}_{j+1}\|^2\right) dx.$$

By corollary 5.3, since $[\mathbf{u}_{j+1}, r_{j+1}]$ is a minimizer of $\vee_{\mathscr{L}}(r_j, \lambda_{j+1})$ it is an extremal pair, namely,

$$(\varphi(r_{j+1}), \mathscr{L}\mathbf{u}_{j+1}) = \langle \mathscr{L}^*\varphi(r_{j+1}), \mathbf{u}_{j+1}\rangle = \frac{1}{\lambda_{j+1}}\|\mathbf{u}_{j+1}\|_{\mathbb{B}}, \tag{3.18}$$

and we end up with the following improved version of the key estimate (3.12),

$$[r_j]_\Phi - [r_{j+1}]_\Phi \geq \frac{1}{\lambda_{j+1}}\|\mathbf{u}_{j+1}\|_{\mathbb{B}} + \frac{\kappa}{2}\|\mathscr{L}\mathbf{u}_{j+1}\|^2_{L^2(\Omega)}. \tag{3.19}$$

A telescoping sum of this improvement yields the following.

Corollary 3.3. *Let $U = \sum_{j=1}^\infty \mathbf{u}_j \in \mathbb{B}$ be a hierarchical solution of $\mathscr{L}U = f$, $f \in L^\Phi_\#$ outlined in theorem 3.2, subject to a strictly convex Φ, (3.17). Then*

$$\frac{1}{\lambda_1}\sum_{j=1}^\infty \frac{1}{2^{j-1}}\|\mathbf{u}_j\|_{\mathbb{B}} + \frac{\kappa}{2}\sum_{j=1}^\infty \|\mathscr{L}\mathbf{u}_j\|^2_{L^2(\Omega)} \leq \int_\Omega \Phi(|f|)dx. \tag{3.20}$$

The last energy bound turns into a precise energy equality in quadratic case, when $L^\Phi_\# = L^2_\#$: arguing along the lines of corollary 2.2 we have

$$\frac{1}{\lambda_1}\sum_{j=1}^\infty \frac{1}{2^{j-1}}\|\mathbf{u}_j\|_{\mathbb{B}} + \sum_{j=1}^\infty \|\mathscr{L}\mathbf{u}_j\|^2_{L^2_\#(\Omega)} = \|f\|^2_{L^2_\#(\Omega)} \tag{3.21}$$

4. HIERARCHICAL SOLUTIONS FOR $\mathscr{L}U = f \in L^p_\#(\Omega)$

We focus our attention on the case where L^Φ is the Lebesgue space $L^p_\#(\Omega)$, seeking solutions of

$$\mathscr{L}U = f, \qquad f \in L^p_\#(\Omega), \quad 1 < p < \infty, \tag{4.1}$$

in an appropriate Banach space $U \in \mathbb{B}$. According to theorem 3.2, equation (4.1) admits a solution $U \in \mathbb{B}$, provided assumption 3.2 holds, namely, there exists an appropriate $\eta : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that,

$$\|g\|^p_{L^p} \lesssim \eta\left(\|\mathscr{L}^*\varphi(g)\|_{\mathbb{B}^*}\right), \qquad \forall g \in L^p_\#, \quad \varphi(g) := p \cdot \mathrm{sgn}(g)|g|^{p-1}, \ \int_0^1 \frac{\eta(s)}{s^2}ds < \infty. \tag{4.2}$$

The following lemma provides a useful simplification of the closure assumption (4.2) for the L^p-case.

Lemma 4.1. *Let $\mathscr{P} : L^\Phi \mapsto L^\Phi$ with dual $\mathscr{P}^*$. Assume that there exists $\beta > 0$ such that the following apriori estimate holds,*

$$\|g - \mathscr{P}^*g\|_{L^{p'}} \leq \beta\|\mathscr{L}^*g\|_{\mathbb{B}^*}, \qquad \forall g \in L^{p'}(\Omega), \quad 1 < p < \infty. \tag{4.3}$$

Then assumption 3.2 is fulfilled with $L^p_\# = L^p \cap Ker(\mathscr{P})$ and $\eta_p(s) \sim s^{p'}$, $1 < p < \infty$.

Proof. Fix $g \in L^p_\#(\Omega)$. Then $\varphi(g) = p \cdot \mathrm{sgn}(g)|g|^{p-1} \in L^{p'}(\Omega)$ and since $g \in \mathrm{Ker}(\mathscr{P})$ we find

$$\int_\Omega |g|^p dx = \frac{1}{p}\int_\Omega g\varphi(g)dx = \frac{1}{p}\int_\Omega g\big(\varphi(g) - \mathscr{P}^*\varphi(g)\big)dx \le \frac{1}{p}\|g\|_{L^p}\|\varphi(g) - \mathscr{P}^*\varphi(g)\|_{L^{p'}}.$$

The apriori estimate assumed in (4.3) yields

$$\|g\|^{p-1}_{L^p} < \frac{\beta}{p}\|\mathscr{L}^*\varphi(g)\|_{\mathbb{B}^*}, \qquad \forall g \in L^p_\#(\Omega), \tag{4.4}$$

and the closure estimate (4.2) follows with $\eta_p(s) \sim s^{p'}$, $1 < p < \infty$. □

Remark 4.1. Observe that the proof of lemma 4.1 — and as a consequence, the proof of the main theorem 1.1 derived below, are limited to $1 < p < \infty$. The end cases $p = 1, \infty$ are excluded since (3.13a) fails to hold for $\eta(s) \sim s^{p'}$. This, in turn, is related to the fact that the scaling functions Φ of L^p are not proper N-functions for $p = 1, \infty$.

Remark 4.2. As an example with $\mathscr{L} = div$ and $\mathscr{P}$ denoting the zero averaging projection $\mathscr{P}g = g - \overline{g}$, lemma 4.1 implies a generalized Sobolev-Nirenberg inequality

$$\|g\|^{d-1}_{L^d(\mathbb{T}^d)} \lesssim \|\mathrm{sgn}(g)|g|^{d-1}\|_{BV(\mathbb{T}^d)}, \quad \forall g \in L^d_\#(\mathbb{T}^d).$$

Proof of theorem 1.1. Granted (4.3), then by lemma 4.1, theorem 3.2 applies with $\eta(s) = \eta_p(s) = s^{p'}$, and (3.16) implies

$$\|U\|_{\mathbb{B}} \lesssim \lambda_1\|f\|^p_{L^p} + \int_0^{1/\zeta\lambda_1} s^{p'-2}ds \lesssim \lambda_1\|f\|^p_{L^p} + \frac{1}{(\lambda_1\zeta)^{1/(p-1)}}.$$

Choose $\lambda_1 := C\|f\|^{-(p-1)}_{L^p}$: according to (4.4), such a choice of λ_1 with a sufficiently large C satisfies the admissibility requirement (3.4),

$$\lambda_1 = \frac{C}{\|f\|^{p-1}_{L^p}} > \frac{1}{\|\mathscr{L}^*\varphi(f)\|_{\mathbb{B}^*}},$$

and the uniform bound (1.4) follows, $\|U\|_{\mathbb{B}} \lesssim \|f\|_{L^p}$. □

Since the above proof was deduced from the general main theorem 3.2, we did not keep track of the precise dependence on the constant β and it remains to verify the more precise bound asserted in (1.7), namely, by specifying $\lambda_{j+1} = \lambda_1\zeta^j$ with $\zeta = 2^{p-1}$ we have

$$\|U\|_{\mathbb{B}} \le \frac{\beta 2^p}{p}\|f\|_{L^p}, \qquad \lambda_{j+1} = \lambda_1 2^{(p-1)j}, \ \lambda_1 = \frac{\beta}{\|f\|^p_{L^p}}. \tag{4.5}$$

Proof of remark 1.1. We argue along the lines of theorem 2.1 which proves (4.5) in the case $p = 2$. The extension for $p > 2$ is based on (4.4) and (3.10), which yield the basic iterative estimate corresponding to (2.8),

$$\|\mathbf{u}_{j+1}\|_{\mathbb{B}} + \lambda_{j+1}\|r_{j+1}\|^p_{L^p} \le \lambda_{j+1}\|r_j\|^p_{L^p}, \qquad \|r_j\|^{p-1}_{L^p} \le \frac{\beta}{p}\|\mathscr{L}^*\varphi(r_j)\|_{\mathbb{B}^*} = \frac{\beta}{p\lambda_j}, \ j = 1, 2, \ldots.$$

Setting $\lambda_{j+1} = \lambda_1 2^{(p-1)j}$ implies,

$$\|\mathbf{u}_{j+1}\|_{\mathbb{B}} \le \lambda_{j+1}\left(\frac{\beta}{\lambda_j p}\right)^{\frac{p}{p-1}} \le \frac{\lambda_1^{-\frac{1}{p-1}}}{2^{p(j-1)}}\left(\frac{\beta}{p}\right)^{\frac{p}{p-1}}, \quad j = 1, 2, \ldots,$$

and we conclude

$$\begin{aligned}\|U\|_{\mathbb{B}} &\leq \|\mathbf{u}_1\|_{\mathbb{B}} + \sum_{j=2}^{\infty}\|\mathbf{u}_j\|_{\mathbb{B}} \\ &\leq \lambda_1\|f\|^p_{L^p} + \lambda_1^{-\frac{1}{p-1}}\left(\frac{\beta}{p}\right)^{\frac{p}{p-1}}\sum_{j=2}^{\infty}\frac{1}{2^{j-p-1}} = \lambda_1^{\frac{1}{p-1}}\|f\|^p_{L^p} + \lambda_1^{-\frac{1}{p-1}}2^p\left(\frac{\beta}{p}\right)^{\frac{p}{p-1}}.\end{aligned}$$

Finally, the choice of $\lambda_1 = \beta/\|f\|^{p-1}_{L^p}$ yields (4.5). $\square$

The bound in (4.5) is not sharp: if $\mathscr{L}U = f \in L^p$ admits a solution such that $\|U\|_{\mathbb{B}} \leq \beta\|f\|_{L^p}$, then the apriori estimate (4.3) follows by a straightforward duality argument. A sharp form of the inverse implication, (4.3) $\mapsto \|U\|_{\mathbb{B}} \leq \beta\|f\|_{L^p}$, follows from an argument based on Hahn-Banach theorem (outlined below). It would be desirable to obtain such a solution by hierarchical constructions. As a consequence, one can derive sharp version for the Sobolev inequalities, which we demonstrate in the context of the Gagliardo-Nirenberg inequality (2.3)

$$\|f\|_{L^{d'}(\Omega)} \leq \beta\|f\|_{BV(\Omega)}, \quad f \in L^p_{\#}(\Omega). \tag{4.6}$$

What is the best (least) possible β for (4.6) to hold? we turn to the

Proof of corollary 1.4. Let U denote the solution with the corresponding sharp bound

$$div\,U = f, \qquad \|U\|_{L^\infty} \leq \beta\|f\|_{L^d(\Omega)}.$$

It follows that

$$\int_\Omega f dx = \int_{\partial\Omega} U\cdot\mathbf{n}dS \leq \|U\|_{L^\infty}\cdot|\partial\Omega| \leq \beta\|f\|_{L^d(\Omega)}|\partial\Omega|.$$

We now fix the extremal $f = 1_\Omega(x)$ such that $\int_\Omega f dx = \|f\|_{L^d(\Omega)}|\partial\Omega|^{1/d'}$; the last inequality then yields

$$\beta \geq \frac{|\Omega|^{1/d'}}{|\partial\Omega|}.$$

By the isoperimetric inequality we conclude

$$\sup_{f\in L^d_{\#}}\frac{\|f\|_{L^{d'}(\Omega)}}{\|f\|_{BV(\Omega)}} = \sup_U\frac{\|U\|_{L^\infty(\Omega)}}{\|div\,U\|_{L^d(\Omega)}} \leq \frac{|\Omega|^{1/d'}}{|\partial\Omega|} \leq \frac{1}{d\omega_d^{1/d}}, \tag{4.7}$$

where ω_d denotes the volume of the unit ball in $\mathbb{R}^d$. $\square$

Remark 4.3. That the upper bound on the right of (4.7) is sharp can be verified by the extremal $f = 1_B(x)$ over balls, [DPD02, CNV04],

$$\sup_{f\in L^{d'}_{\#}}\frac{\|f\|_{L^{d'}}}{\|f\|_{BV}} = \frac{1}{d\omega_d^{1/d}}.$$

The proof of corollary 1.4 suggests that $f = 1_\Omega(x)$ is the extremal equalizer to obtain an Ω-dependent bound, such that inequality on the left of (4.7) becomes sharp,

$$\sup_{f\in L^{d'}_{\#}}\frac{\|f\|_{L^{d'}(\Omega)}}{\|f\|_{BV(\Omega)}} = \frac{|\Omega|^{1/d'}}{|\partial\Omega|}.$$

Proof of the sharp bound (4.8). This is essentially due to Bourgain & Brezis. We reproduce here (a slight generalization of) their non-constructive proof in [BB03, proposition 1]: if (4.3) is holds then $\mathscr{L}U = f$ admits a solution with the following *sharp* bound

$$\|U\|_{\mathbb{B}} \leq \beta' \|f\|_{L^p}, \qquad \beta' > \beta. \tag{4.8}$$

Normalize $\|f\|_{L^p} = 1$ and consider the two non-empty convex sets: the ball

$$B_{\beta_\varepsilon} := \{\mathbf{u} \in \mathbb{B} : \ \|\mathbf{u}\|_{\mathbb{B}} < \beta_\varepsilon\}, \quad \beta_\varepsilon := (1+\varepsilon)\beta > \beta$$

and $C := \{U \in \mathbb{B} : \ \mathscr{L}U = f\}$. The claim is that $B_{\beta_\varepsilon} \cap C \neq \emptyset$ and the desired estimate (4.8) then follows with sufficiently small $\varepsilon \ll 1$. If not, $B_{\beta_\varepsilon} \cap C = \emptyset$, and by Hahn-Banach there exists a non-trivial $g^* \in L^{p'}$ such that for some $\alpha \in \mathbb{R}_+$

$$\langle g^*, \mathbf{u} \rangle \leq \alpha, \quad \forall \mathbf{u} \in B_{\beta_\varepsilon} \tag{4.9a}$$

and

$$\langle g^*, U \rangle \geq \alpha, \quad \forall U \in C. \tag{4.9b}$$

If $V \in \mathrm{Ker}(\mathscr{L})$ then application of (4.9b) with $U \mapsto U \pm \delta V \in C$ yields $\pm\delta\langle g^*, V\rangle \geq 0$, or $\langle g^*, V\rangle = 0$; that is, $g^* \in \mathrm{Ker}(\mathscr{L})^\perp = \mathrm{R}(\mathscr{L}^*)$ is of the form $g^* = \mathscr{L}^* g$ for some $g \in D(\mathscr{L}^*) \subset L^{p'}$. Now, by (4.9a)

$$\|g^*\|_{\mathbb{B}^*} = \sup_{\|\mathbf{u}\|_{\mathbb{B}} = \beta_{\varepsilon/2}} \frac{\langle g^*, \mathbf{u}\rangle}{\beta_{\varepsilon/2}} \leq \frac{\alpha}{\beta_{\varepsilon/2}},$$

and the apriori estimate assumed in (4.3) implies

$$\|g\|_{L^{p'}} \leq \beta \|\mathscr{L}^* g\|_{\mathbb{B}^*} = \beta \|g^*\|_{\mathbb{B}^*} \leq \frac{\alpha}{1+\varepsilon/2}.$$

But this leads to a contradiction: pick $U \in C$ (which we recall is not empty) then (4.9b) implies,

$$\alpha \leq \langle g^*, U\rangle = \langle \mathscr{L}^* g, U\rangle = \langle g, f\rangle \leq \|g\|_{L^{p'}} \|f\|_{L^p} \leq \frac{\alpha}{1+\varepsilon/2}. \qquad \square$$

5. AN APPENDIX ON $\vee_{\mathscr{L}}$-MINIMIZERS

To study the hierarchical expansions (3.9), we characterize the minimizers of the $\vee_{\mathscr{L}}$-functionals (3.3)

$$[\mathbf{u}, r] := \mathop{arg\,min}_{\mathscr{L}\mathbf{u}+r=f} \vee_{\mathscr{L}}(f, \lambda; \mathbb{B}, \Phi), \qquad \vee_{\mathscr{L}}(f, \lambda; \mathbb{B}, \Phi) := \inf_{\mathscr{L}\mathbf{u}+r=f} \Big\{ \|\mathbf{u}\|_{\mathbb{B}} + \lambda [r]_\Phi \, : \, \mathbf{u} \in \mathbb{B} \Big\}. \tag{5.1}$$

Here $\mathscr{L} : \mathbb{B} \mapsto L^\Phi_\#(\Omega)$ is densely defined into a subspace of Orlicz space over a Lipschitz domain $\Omega \subset \mathbb{R}^d$. We shall often abbreviate $\vee_{\mathscr{L}}(f, \lambda)$ for $\vee_{\mathscr{L}}(f, \lambda; \mathbb{B}, \Phi)$. The characterization summarized below extends related results which can be found in [Me02, Theorem 4], [ACM04, Chapter1], [TNV08, Theorem 2.3].

Recall that $\|\cdot\|_{\mathbb{B}^*}$ denotes the *dual* norm, $\|\mathscr{L}^* g\|_{\mathbb{B}^*} = \langle \mathscr{L}^* g, \mathbf{u}\rangle / \|\mathbf{u}\|_{\mathbb{B}}$, so that the usual duality bound holds

$$\langle \mathscr{L}^* g, \mathbf{u}\rangle \leq \|\mathbf{u}\|_{\mathbb{B}} \|\mathscr{L}^* g\|_{\mathbb{B}^*}, \qquad g \in L^\Psi, \ \mathbf{u} \in \mathbb{B}. \tag{5.2}$$

We say that $\mathbf{u}$ and $\mathscr{L}^* g$ is an *extremal pair* if equality holds above. The theorem below characterizes $[\mathbf{u}, r]$ as a minimizer of the $\vee_{\mathscr{L}}$-functional if and only if $\mathbf{u}$ and $\mathscr{L}^* \varphi(r)$ form an extremal pair.

Theorem 5.1. *Let $\mathscr{L}:\mathbb{B}\to L^{\Phi}_{\#}(\Omega)$ be a linear operator with dual $\mathscr{L}^*$ and let $\vee_{\mathscr{L}}(f,\lambda;\mathbb{B},\Phi)$ denote the associated functional* (3.3)*.*

(i) The variational problem (5.1) *admits a minimizer* $\mathbf{u}$*. Moreover, if $\|\cdot\|_{\mathbb{B}}$ is* uniformly convex*, then a minimizer $\mathbf{u}$ with $\|\mathbf{u}\|_{\mathbb{B}}\neq 0$ is unique.*

(ii) $\mathbf{u}$ is a minimizer of (5.1) *if and only if the residual $r:=f-\mathscr{L}\mathbf{u}$ satisfies*

$$\langle \mathscr{L}^*\varphi(r),\mathbf{u}\rangle = \|\mathbf{u}\|_{\mathbb{B}}\cdot\|\mathscr{L}^*\varphi(r)\|_{\mathbb{B}^*}=\frac{\|\mathbf{u}\|_{\mathbb{B}}}{\lambda},\qquad \mathbf{u}\in L^{\Phi}_{\#},\ \ \varphi(r):=sgn(r)\Phi'(|r|)\in L^{\Psi}. \tag{5.3}$$

Proof. (i) The existence of a minimizer for the $\vee_{\mathscr{L}}$-functional follows from standard arguments which we omit, consult [AV94, Me02]. We address the issue of uniqueness. Assume $\mathbf{u}_1$ and $\mathbf{u}_2$ are minimizers with the corresponding residuals $r_1=f-\mathscr{L}\mathbf{u}_1$ an $r_2=f-\mathscr{L}\mathbf{u}_2$

$$\|\mathbf{u}_i\|_{\mathbb{B}}+\lambda[r_i]_{\Phi}=v_{min},\quad i=1,2$$

We then end up with the one-parameter family of minimizers, $\mathbf{u}_\theta:=\mathbf{u}_1+\theta(\mathbf{u}_2-\mathbf{u}_1),\ \theta\in[0,1]$,

$$v_{min}\le\|\mathbf{u}_\theta\|_{\mathbb{B}}+\lambda[r_\theta]_{\Phi}\le\theta\|\mathbf{u}_2\|_{\mathbb{B}}+(1-\theta)\|\mathbf{u}_1\|_{\mathbb{B}}+\theta\lambda[r_2]_{\Phi}+(1-\theta)\lambda[r_1]_{\Phi}=v_{min}.$$

Consequently, $[r_\theta]_\Phi=\theta[r_2]_\Phi+(1-\theta)[r_1]_\Phi$ and hence $r_1=r_2$. In particular, $[r_1]_\Phi=[r_2]_\Phi$ implies that the two minimizers satisfy $\|\mathbf{u}_1\|_{\mathbb{B}}=\|\mathbf{u}_2\|_{\mathbb{B}}$ and we conclude that the ball $\|\mathbf{u}\|_{\mathbb{B}}=\|\mathbf{u}_1\|_{\mathbb{B}}\neq 0$ contains the segment $\{\mathbf{u}_\theta,\ \theta\in[0,1]\}$, which by uniform convexity, must be the trivial segment, i.e., $\mathbf{u}_2=\mathbf{u}_1$.

(ii) If $\mathbf{u}$ is a minimizer of (5.1) then for any $\mathbf{v}\in\mathbb{B}$ we have

$$\begin{aligned}\vee_{\mathscr{L}}(\mathbf{u},\lambda)&=\|\mathbf{u}\|_{\mathbb{B}}+\lambda[f-\mathscr{L}\mathbf{u}]_{\Phi}\\ &\le\vee_{\mathscr{L}}(\mathbf{u}+\varepsilon\mathbf{v},\lambda)=\|\mathbf{u}+\varepsilon\mathbf{v}\|_{\mathbb{B}}+\lambda[f-\mathscr{L}(\mathbf{u}+\varepsilon\mathbf{v})]_{\Phi}\\ &\le\|\mathbf{u}\|_{\mathbb{B}}+|\varepsilon|\cdot\|\mathbf{v}\|_{\mathbb{B}}+\lambda[f-\mathscr{L}\mathbf{u}]_{\Phi}-\lambda\varepsilon\Big(\mathrm{sgn}(f-\mathscr{L}\mathbf{u})\Phi'(|f-\mathscr{L}\mathbf{u}|),\mathscr{L}\mathbf{v}\Big)+\mathscr{O}(\varepsilon^2).\end{aligned}\tag{5.4}$$

It follows that for all $\mathbf{v}\in\mathbb{B}$,

$$\Big|\Big\langle\mathscr{L}^*\varphi(r),\mathbf{v}\Big\rangle\Big|\le\frac{1}{\lambda}\|\mathbf{v}\|_{\mathbb{B}}+\mathscr{O}(|\varepsilon|),\qquad\varphi(r)=\mathrm{sgn}(r)\Phi'(|r|),\ r:=f-\mathscr{L}\mathbf{u},$$

and by letting $\varepsilon\to 0$

$$\|\mathscr{L}^*\varphi(r)\|_{\mathbb{B}^*}\le\frac{1}{\lambda}. \tag{5.5}$$

To verify the reverse inequality, we set $\mathbf{v}=\pm\mathbf{u}$ and $0<\varepsilon<1$ in (5.4), yielding

$$\|\mathbf{u}\|_{\mathbb{B}}+\lambda[f-\mathscr{L}\mathbf{u}]_{\Phi}\le(1\pm\varepsilon)\|\mathbf{u}\|_{\mathbb{B}}+\lambda[f-\mathscr{L}\mathbf{u}\mp\varepsilon\mathscr{L}\mathbf{u}]_{\Phi},$$

and hence $\pm\varepsilon\|\mathbf{u}\|_{\mathbb{B}}\mp\lambda\varepsilon\big(\varphi(f-\mathscr{L}\mathbf{u}),\mathscr{L}\mathbf{u}\big)+\mathscr{O}(\varepsilon^2)\ge 0$. Dividing by ε and letting $\varepsilon\downarrow 0_+$, we obtain $\|\mathbf{u}\|_{\mathbb{B}}=\lambda\langle\mathscr{L}^*\varphi(r),\mathbf{u}\rangle$ and (5.3) follows:

$$\frac{1}{\lambda}\|\mathbf{u}\|_{\mathbb{B}}=\langle\mathscr{L}^*\varphi(r),\mathbf{u}\rangle\le\|\mathscr{L}^*\varphi(r)\|_{\mathbb{B}^*}\|\mathbf{u}\|_{\mathbb{B}}\le\frac{1}{\lambda}\|\mathbf{u}\|_{\mathbb{B}}.$$

Conversely, we show that if (5.3) holds then $\mathbf{u}$ is a minimizer. The convexity of Φ yields

$$\begin{aligned}&[f-\mathscr{L}(\mathbf{u}+\mathbf{v})]_{\Phi}=[r-\mathscr{L}\mathbf{v}]_{\Phi}=\\ &\quad\ge\int_\Omega\Phi(|r|)dx-\big(\mathrm{sgn}(r)\Phi'(|r|),\mathscr{L}(\mathbf{u}+\mathbf{v})\big)+\big(sgn(r)\Phi'(|r|),\mathscr{L}\mathbf{u}\big)\\ &\quad=[f-\mathscr{L}\mathbf{u}]_{\Phi}-\overbrace{\big\langle\mathscr{L}^*\varphi(r),(\mathbf{u}+\mathbf{v})\big\rangle}^{\#1}+\overbrace{\big\langle\mathscr{L}^*\varphi(r),\mathbf{u}\big\rangle}^{\#2}.\end{aligned}$$

By (5.3) $\|\mathscr{L}^*\varphi(r)\|_{\mathbb{B}^*} = 1/\lambda$, which implies

$$-\lambda(\#1) \geq -\|\mathbf{u}+\mathbf{v}\|_{\mathbb{B}} \quad \text{and} \quad \lambda(\#2) = \|\mathbf{u}\|_{\mathbb{B}}.$$

We conclude that for any $\mathbf{v} \in \mathbb{B}$,

$$\begin{aligned} \vee_{\mathscr{L}}(\mathbf{u}+\mathbf{v},\lambda) = \|\mathbf{u}+\mathbf{v}\|_{\mathbb{B}} + \lambda[f-\mathscr{L}(\mathbf{u}+\mathbf{v})]_{\Phi} &\geq \|\mathbf{u}+\mathbf{v}\|_{\mathbb{B}} + \lambda[f-\mathscr{L}\mathbf{u}]_{\Phi} - \lambda(\#1) + \lambda(\#2) \\ &\geq \|\mathbf{u}\|_{\mathbb{B}} + \lambda[f-\mathscr{L}\mathbf{u}]_{\Phi} = \vee_{\mathscr{L}}(\mathbf{u},\lambda). \end{aligned}$$

Thus, $\mathbf{u}$ is a minimizer of (5.1). □

Remark 5.1. A lack of uniqueness is demonstrated in an example of [Me02, pp. 40], using the ℓ^∞-unit ball, which in turn lacks strict convexity. Thus, strict convexity is necessary and sufficient for uniqueness.

The next two assertions are a refinement of Theorem 5.1, depending on the size of $\|\mathscr{L}^*\varphi(f)\|_{\mathbb{B}^*}$.

Lemma 5.2. *[The case $\|\mathscr{L}^*\varphi(f)\|_{\mathbb{B}^*} \leq 1/\lambda$]. Let $\mathscr{L} : \mathbb{B} \to L^\Phi_\#(\Omega)$ with adjoint $\mathscr{L}^*$ and let $\vee_{\mathscr{L}}$ denote the associated functional* (3.3)*. Then $\lambda\|\mathscr{L}^*\varphi(f)\|_{\mathbb{B}^*} \leq 1$ if and only if $\mathbf{u} \equiv 0$ is a minimizer of* (5.1)*.*

Proof. Assume $\|\mathscr{L}^*\varphi(f)\|_{\mathbb{B}^*} \leq 1/\lambda$. Then by the convexity of Φ we have

$$\begin{aligned} \|\mathbf{u}\|_{\mathbb{B}} + \lambda[f-\mathscr{L}\mathbf{u}]_{\Phi} &\geq \|\mathbf{u}\|_{\mathbb{B}} + \lambda\int_\Omega \Phi(|f|)dx - \lambda\int_\Omega \big(\varphi(f),\mathscr{L}\mathbf{u}\big)dx \\ &\geq \|\mathbf{u}\|_{\mathbb{B}} + \lambda\int_\Omega \Phi(|f|)dx - \lambda\|\mathscr{L}^*\varphi(f)\|_{\mathbb{B}^*}\|\mathbf{u}\|_{\mathbb{B}} \geq \lambda[f]_{\Phi} \end{aligned}$$

which tells us that $\mathbf{u} \equiv 0$ is a minimizer of (3.3). Conversely, if $\mathbf{u} \equiv 0$ is a minimizer of (5.1), then $\varepsilon\|\mathbf{u}\|_{\mathbb{B}} + \lambda[f-\varepsilon\mathscr{L}\mathbf{u}]_{\Phi} \geq [f]_{\Phi}$ for all $\mathbf{u} \in \mathbb{B}$. It follows that

$$\varepsilon\|\mathbf{u}\|_{\mathbb{B}} - \varepsilon\lambda\int_\Omega (\varphi(f),\mathscr{L}\mathbf{u})dx + \mathscr{O}(\varepsilon^2) \geq 0.$$

Letting $\varepsilon \downarrow 0$ we conclude $\lambda\langle\mathscr{L}^*\varphi(f),\mathbf{u}\rangle \leq \|\mathbf{u}\|_{\mathbb{B}}$, hence $\|\mathscr{L}^*\varphi(f)\|_{\mathbb{B}^*} \leq 1/\lambda$. □

Lemma 5.3. *[The case $\|\mathscr{L}^*\varphi(f)\|_{\mathbb{B}^*} > 1/\lambda$]. Let $\mathscr{L} : \mathbb{B} \to L^\Phi_\#(\Omega)$ with adjoint $\mathscr{L}^*$ and let $\vee_{\mathscr{L}}$ denote the associated functional* (3.3)*. If $1 < \lambda\|\mathscr{L}^*\varphi(f)\|_{\mathbb{B}^*} < \infty$, then $\mathbf{u}$ is a minimizer of* (5.1) *if and only if $\mathscr{L}\mathbf{u}$ and $\varphi(r)$ is an extremal pair and*

$$\|\mathscr{L}^*\varphi(r)\|_{\mathbb{B}^*} = \frac{1}{\lambda}, \qquad \langle\mathbf{u},\mathscr{L}^*\varphi(r)\rangle = \frac{\|\mathbf{u}\|_{\mathbb{B}}}{\lambda}. \tag{5.6}$$

Proof. Since $\|\mathscr{L}^*\varphi(f)\|_{\mathbb{B}^*} > 1/\lambda$ then $\|\mathbf{u}\|_{\mathbb{B}} \neq 0$ and we can now divide the equality on the right of (5.3) by $\|\mathbf{u}\|_{\mathbb{B}} \neq 0$ and (5.6) follows. □

REFERENCES

[AV94] R. ACAR AND C. R. VOGEL, *Analysis of Bounded Variation Penalty Methods of Ill-Posed Problems*, Inverse Problems, 10 (1994), pp. 1217-1229.

[AF03] R. ADAMS AND J. FOURNIER, Sobolev Spaces, 2nd ed., Academic Press, 2003.

[BS88] C. BENNET AND R. SHARPLEY, Interpolation of Operators, Academic Press, 1988.

[ACM04] F. ANDREU-VAILLO, V. CASELLES, AND J.M. MAZÓN, Parabolic Quasilinear Equations Minimizing Linear Growth Functionals, Birkhäuser Verlag, 2004.

[AV97] G. AUBERT AND L. VESE, *A Variational Method in Image Recovery*, SIAM J. Numer. Anal., 34 (1997), pp. 1948-1979.

[BCM05] A. BAUDES, B. COLL AND J.-M. MOREL, *A review of image denoising algorithms, with a new one*, Multiscale Model. Simul. 4 (2005), no. 2, 490–530.

[BB03] J. BOURGAIN AND H. BREZIS, *On the equation* $\mathrm{div} Y = f$ *and application to control of phases*, J. Amer. Math. Soc. 16 (2003), no. 2, 393–426.

[BB07] J. BOURGAIN AND H. BREZIS, *New estimates for elliptic equations and Hodge type systems* J. Eur. Math. Soc. (JEMS) 9 (2007), no. 2, 277–315.

[CL95] A. CHAMBOLLE AND P.L. LIONS, *Image restoration by constrained total variation minimization and variants*, SPIE Electronic Imaging Proceedings, Vol. 2567, 1995.

[CL97] A. CHAMBOLLE AND P. L. LIONS, *Image recovery via total variation minimization and related problems*, Numer. Math., 76 (1997), pp. 167-188.

[CNV04] D. CORDERO-ERAUQUIN, B. NAZARET AND C. VILLANI, *A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities*, Adv. Math. 182 (2004), no. 2, 307–332.

[DFT05] B. DACOROGNA, N. FUSCO AND L. TARTAR, *On the solvability of the equation* $div\, u = f$ *in* L^1 *and in* C^0, Rend. Mat. Acc. Lincei. s. 9 v. 14 (2003), pp. 239-245.

[DPD02] M. DEL PINO AND J. DOLBEAULT, *Best constant for Galiardo-Nirenberg inequalities and applications to nonlinear diffusions*, J. Math. Pures Appl. 81 (2002), 847-875.

[DT84] F. DEMENGEL AND R. TEMAM, *Convex functions of a measure and applications*, Indiana Univ. Mathematics Journal 33(5): 673-709, 1984.

[DM87] R. DIPERNA AND A. MAJDA, *Concentrations in regularizations for 2D incompressible flow*, Comm. Pure and Appl. Math., 40 (1987), pp. 301-345.

[GLMV07] JOHN B. GARNETT, TRIET M. LE, YVES MEYER AND LUMINITA A. VESE, *Image decompositions using bounded variation and generalized homogeneous Besov spaces*, Appl. Comput. Harmon. Anal. 23 (2007), 25-56.

[LV05] TRIET M. LE AND LUMINITA A. VESE, *Image Decomposition Using Total Variation and div(BMO)*, Multiscale Modeling and Simulation, 4(2) (2005), 390-423.

[LPSX06] D. LEVY, D. PAQUIN, E. SCHREIBMANN, AND L. XING, *Multiscale Image Registration, Mathematical Biosciences and Engineering*, 3, 2006, pp. 389-418.

[LNT00] M. LOPES FILHO, H. J. NUSSENZVEIG AND E. TADMOR, *Approximate solutios of the incompressible Euler equations with no concentrations*, Annales De L'institut Henri Poincare (c) Non Linear Analysis, 17 (2000) pp. 371-412.

[Ma06] Y. MADAY, L^∞*-stable approximation of a solution to* $\mathrm{div}(\mathbf{Y}) = f$ *for* $f \in L^2$ *in two dimensions*, J. Sci. Comput. 28(2-3) (2006), pp. 451-458.

[Ma07] V. MAZ'YA, *Bourgain-Brezis type inequality with explicit constants*, Interpolation theory and applications, 247–252, Contemp. Math., 445, AMS Providence, RI, 2007.

[Me02] Y. MEYER, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, University Lecture Series Volume 22, AMS 2002.

[Mi10] P. MIRONESCU, *On some inequalities of Bourgain, Brezis, Maz'ya, and Shaposhnikova related to* L^1 *vector fields*, C. R. Math., Math. Anal., in press.

[Mo84] V.A. MOROZOV, Methods for Solving Incorrectly Posed Problems, Springer Verlag, Heidelberg, 1984.

[Mo93] V.A. MOROZOV, Regularization Methods for Ill-Posed Problems, Florida, CRC Press, 1993.

[PL07] D. PAQUIN, D. LEVY AND L. XING, *Hybrid Multiscale Landmark and Deformable Registration*, Mathematical Biosciences and Engineering, 4 (2007), pp. 711-737.

[PT08] N.C. PHUC AND M. TORRES, Characterizations of the existence and removable singularities of divergence-measure vector fields, , Indiana Univ. Math. J. 57(4), 1573-1598, 2008.

[ROF92] L. RUDIN, S. OSHER AND E. FATEMI, *Nonlinear Total Variation Based Noise Removal Algorithms*, Physica D, 60, pp. 259-268, 1992.

[TNV04] E. TADMOR, S. NEZZAR AND L. VESE, *A Multiscale Image Representation Using Hierarchical (BV,L2) Decomposition*, Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal, Volume 2, Number 4, pp. 554-579, 2004.

[TNV08] E. TADMOR, S. NEZZAR AND L. VESE, *hierarchical decomposition of images with applications to deblurring, denoising and segmentation*, Communications in Math. Sciences 6(2) (2008) 281-307.

[TA77] A. N. TIKHONOV AND V. A. ARSENIN, Solution of Ill-posed Problems, Winston & Sons, Washington, 1977.

[Ve92] I. VECCHI, *Gagliardo-Nirenberg imbeddings and DiPerna-Majda sequences*, Proc. Roy. Soc. London Ser. A 439 (1992), no. 1906, pp. 337-340.

[VO04] L. VESE AND S. OSHER, Image denoising and decomposition with total variation minimization and oscillatory functions. Special issue on mathematics and image analysis. J. Math. Imaging Vision 20(1-2) (2004), pp. 7-18.

[Zi89] W. ZIEMER, Weakly Differentiable Functions, Graduate Texts in Math., Springer, vol. 120, 1989.

DEPARTMENT OF MATHEMATICS, CENTER OF SCIENTIFIC COMPUTATION AND MATHEMATICAL MODELING (CSCAMM) AND INSTITUTE FOR PHYSICAL SCIENCE AND TECHNOLOGY (IPST), UNIVERSITY OF MARYLAND, MD 20742.

E-mail address: `tadmor@cscamm.umd.edu`