Hierarchical Construction of Bounded Solutions in Critical Regularity Spaces

EITAN TADMOR
University of Maryland

Whenever you can settle a question by explicit construction, be not satisfied with purely existential arguments.
Hermann Weyl [43, p. 326]

Abstract

We construct uniformly bounded solutions for the equations $\operatorname{div} U = f$ and $\operatorname{curl} U = \mathbf{f}$ in the critical cases $f \in L^d_\#(\mathbb{T}^d, \mathbb{R})$ and $\mathbf{f} \in L^3_\#(\mathbb{T}^3, \mathbb{R}^3)$, respectively. Criticality in this context manifests itself by the lack of a linear solution operator mapping $L^d_\# \mapsto L^\infty(\mathbb{T}^d)$. Thus, the intriguing aspect here is that although the problems are linear, construction of their solutions is not.

Our constructions are special cases of a general framework for solving linear equations of the form $\mathcal{L}U = f$, where $\mathcal{L}$ 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 $\mathcal{L}^*$. 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

$$\mathbf{u}_{j+1} = \arg\min_{\mathbf{u}} \left\{ \|\mathbf{u}\|_{\mathbb{B}} + \lambda_{j+1} \|r_j - \mathcal{L}\mathbf{u}\|^p_{L^p(\Omega)} \right\},$$

where the residuals $r_j := f - \mathcal{L}(\sum_{k=1}^j \mathbf{u}_k)$ are resolved in terms of a dyadic sequence of scales $\lambda_{j+1} := \lambda_1 2^j$ with large enough $\lambda_1 \gtrsim \|f\|^{1-p}_{L^p}$. The nonlinear aspect of this construction is a counterpart of the fact that one cannot linearly solve $\mathcal{L}U = f$ in critical regularity spaces. © 2016 Wiley Periodicals, Inc.

1 Introduction and Statement of 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 of the problem

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

The classical elliptic solution of the first half of (1.1), $U = \nabla\Delta^{-1} f$, lies in $W^{1,d}_\#(\mathbb{T}^d)$, which may fail to satisfy the uniform bound sought in the second

Communications on Pure and Applied Mathematics, Vol. LXIX, 1087–1109 (2016)

half of (1.1). Thus, the question is whether (1.1) admits a solution that gains uniform boundedness, $\|U\|_{L^\infty} \lesssim \|f\|_{L^d}$, at the expense of giving up on the irrotationality condition $\operatorname{curl} U = 0$. This question was addressed by Bourgain and Brezis [7, prop. 1]. They proved that (1.1) admits uniformly bounded solutions for all $f \in L^d_\#(\mathbb{T}^d)$, with the intricate aspect that a solution operator mapping $L^d_\# \mapsto L^\infty$ *must be nonlinear*; in particular, therefore, the uniform boundedness of an irrotational elliptic solution *must* fail. The existence of such uniformly bounded solutions was proved in [7] by using a straightforward duality argument based on the closed range theorem.

The purpose of this paper is to present an alternative approach for the existence of such solutions. Our approach is *constructive*: the solution U is constructed as a *hierarchical* sum, $U = \sum_{j=1}^\infty \mathbf{u}_j$, where the $\{\mathbf{u}_j\}$'s are computed recursively as appropriate minimizers,

$$\mathbf{u}_{j+1} = \arg\min_{\mathbf{u}} \left\{ \|\mathbf{u}\|_{L^\infty} + \lambda_1 2^j \left\| f - \operatorname{div}\left(\sum_{k=1}^{j} \mathbf{u}_k + \mathbf{u} \right) \right\|^d_{L^d} \right\}, \quad j = 0, 1, \dots,$$

and λ_1 is a sufficiently large parameter specified below. As an example, we refer to [37] for the computation of a uniformly bounded hierarchical solution of the equation $\operatorname{div} U = \Delta G$ with $G := x_1 |\ln r|^{1/3} \zeta(r) \in L^2_\#(\mathbb{T}^2)$ where $\zeta(\cdot)$ is a radial cutoff away from the origin [7, sec. 3, remark 7]. The elliptic solution, $U = \nabla G$, has a fractional logarithmic growth at the origin, whereas the computation confirms that the hierarchical solution $U = \sum \mathbf{u}_j$ remains uniformly bounded, $\|U\|_{L^\infty} \lesssim \|\Delta G\|_{L^2} < \infty$.

The above construction is in fact a special case of our main result that applies to general linear problems of the form

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

Here, $\mathcal{L} : \mathbb{B} \mapsto L^p_\#(\Omega)$ 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,

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

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

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

The closed range theorem combined with the open mapping principle tells us that if the a priori duality estimate assumed in (1.3) holds, then equation (1.2) admits a solution, $\|U\|_{\mathbb{B}} \lesssim \gamma \|f\|_{L^p}$ with a constant $\gamma = \gamma(\beta, p, d)$. Our main result explains the existence of such U's by *explicit construction*.

THEOREM 1.1. *Fix $1 < p < \infty$ and assume that the a priori estimate* (1.3) *holds. There exists $\gamma < \infty$ (depending on p and linearly on β) such that for any given*

$f \in L^p_\#(\Omega)$, the equation $\mathcal{L}U = f$ admits a hierarchical solution of the form $U = \sum_{j=1}^{\infty} \mathbf{u}_j \in \mathbb{B}$,

$$\|U\|_{\mathbb{B}} \leq \gamma \|f\|_{L^p}, \quad \gamma < \infty. \tag{1.4}$$

Here, the $\{\mathbf{u}_j\}$ are constructed recursively as minimizers of

$$\mathbf{u}_{j+1} = \arg\min_{\mathbf{u}} \left\{ \|\mathbf{u}\|_{\mathbb{B}} + \lambda_{j+1} \left\| f - \mathcal{L}\left(\sum_{k=1}^{j} \mathbf{u}_k + \mathbf{u} \right) \right\|^p_{L^p} \right\}, \quad j = 0, 1, \ldots, \tag{1.5}$$

where $\{\lambda_j\}_{j \geq 1}$ is a dyadic sequence, $\lambda_{j+1} := \lambda_1 2^j$ with $\lambda_1 \geq \beta \|f\|^{1-p}_{L^p}$.

Remark 1.2 (Exponential Convergence). The description of U as the sum $U = \sum \mathbf{u}_j$ provides a multiscale *hierarchical decomposition* of a solution for a (1.2) for rapidly increasing sequence of scales, $\lambda_{j+1} = \lambda_1 \zeta^j$ with any $\zeta > 1$. The role of the $\{\lambda_j\}$ as the different scales associated with the $\mathbf{u}_j$ is reflected through the exponential decay bound (consult (4.23) below)

$$\|\mathbf{u}_j\|_{\mathbb{B}} \lesssim \frac{\lambda_{j+1}}{\lambda_j^{p'}} \sim \|f\|_{L^p_\#} \zeta^{-\frac{j}{p-1}}, \quad \lambda_{j+1} = \lambda_1 \zeta^j \sim \frac{\zeta^j}{\|f\|^{p-1}_{L^p_\#}}, \quad 1 < p < \infty.$$

For simplicity, we limit our discussion to the dyadic case $\zeta = 2$.

Remark 1.3 (On the A Priori Duality Estimate (1.3)). The a priori estimate (1.3) is exactly what is needed for the hierarchical construction $\sum \mathbf{u}_j$ to converge. It should be emphasized, however, that the construction does *not* require knowledge of the precise value of the constant β appearing in estimate (1.3). Indeed, the parameter β enters through the initial scale, λ_1, which is to be chosen large enough,

$$\lambda_1 \geq \beta \|f\|^{1-p}_{L^p},$$

so that by Lemma A.3, it dictates a nontrivial first hierarchical step,

$$\mathbf{u}_1 = \arg\min_{\mathbf{u}} \left\{ \|\mathbf{u}\|_{\mathbb{B}} + \lambda_1 \|f - \mathcal{L}\mathbf{u}\|^p_{L^p} \right\}.$$

What happens if the initial scale λ_1 is underestimated relative to an unknown value of β? Then, as noted in Lemma A.2 below, the variational statement (1.5) will yield zero hierarchical terms, $\mathbf{u}_j \equiv 0$ for an increasing sequence of scales $\lambda_1 2^j$, $j = 1, 2, \ldots$, until reaching the critical scale such that $\lambda_1 2^{j_0} \gtrsim \beta \|f\|^{1-p}_{L^p}$, which will dictate the first nontrivial step of the hierarchical decomposition, $U = \sum_{j=j_0} \mathbf{u}_j$. In this sense, the construction of a hierarchical solution $U = \sum \mathbf{u}_j$ is *independent* of the precise value of β in (1.3): the latter is only needed to guarantee that the hierarchical construction will indeed pick up the first nontrivial minimizer after finitely many steps j_0.

Remark 1.4 (The Limiting Cases $p = 1, \infty$). The L^p-valued hierarchical constructions in Theorem 1.1 can be extended to a more general setup of operators valued in Orlicz spaces (outlined in Remark 4.8 below). The limiting cases, however, are excluded; for example, there exist no $\dot{W}^{1,p}$ solutions of $\operatorname{div} U = f$ for general $f \in L^p$ with $p = 1, \infty$ [7, sec. 2], [14]. The iterative aspect of the hierarchical construction is reminiscent of the Artola and Tartar construction of $L^p(\mathbb{R})$-functions for the end case $p = 1$ as a limiting case for interpolation of $W^{1,1}(\mathbb{R}^2)$-traces [39, sec. II], [17].

L^2-based hierarchical decompositions were introduced by us in the context of image processing [35, 36] and motivated the present construction of solutions in the more general setup of the closed range theorem. We demonstrate such hierarchical constructions of the solution to two important examples of critical regularity studied by Bourgain and Brezis [7, 8]. These are the constructions of uniformly bounded solutions to $\operatorname{div} U = f \in L^d_\#(\mathbb{T}^d)$, discussed in Section 2, and the construction of uniformly bounded solutions to $\operatorname{curl} U = \mathbf{f} \in L^3_\#(\mathbb{T}^3, \mathbb{R}^3)$, discussed in Section 3. The critical regularity in these cases manifests itself in terms of the lack of right inverses for $\mathcal{L}$ bounded on the corresponding critical L^p spaces, or equivalently, $\operatorname{Ker} \mathcal{L}$ that cannot be complemented in L^∞ [7, sec. 3], [3, sec. 5.3].

The main novelty of theorem 1.1 is using these hierarchical decompositions for explicit construction of solutions for general equations governed by operators with a closed range in $L^p_\#$, $1 < p < \infty$. The proof of the special case $p = 2$ is given in Section 4.1: 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 L^p-spaces in Section 4. Finally, the characterization of minimizers, such as those encountered in (1.5), is summarized in the Appendix on $\vee$-minimizers.

2 Bounded Solutions of $\operatorname{div} U = f \in L^p_\#(\Omega, \mathbb{R})$

Let $\mathcal{P}$ denote the averaging projection $\mathcal{P}g := \bar{g}$, where $\bar{g}$ is the average value of g. Given $f \in L^p_\#(\mathbb{T}^d) := \{g \in L^p(\mathbb{T}^d) \mid \bar{g} = 0\}$, then according to Theorem 1.1, we can construct hierarchical solutions of

$$\operatorname{div} U = f, \quad f \in L^p_\#(\mathbb{T}^d), \quad 1 < p < \infty, \tag{2.1}$$

in an appropriate Banach space, $U \in \mathbb{B}$, provided the corresponding a priori 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

$$\|g - \bar{g}\|_{L^{p'}} \le \beta \|\nabla g\|_{\mathbb{B}^*} \quad \forall g \in L^{p'}(\mathbb{T}^d). \tag{2.2}$$

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

Case 1. Solution of $\operatorname{div} U = f \in L^p_\#$ with $U \in \dot{W}^{1,p}$.

Since

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

we can construct hierarchical solutions of (2.1) 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 (2.1), $\nabla \Delta^{-1} f \in \dot{W}^{1,p}(\mathbb{T}^d, \mathbb{R}^d)$.

Case 2. Solution of $\operatorname{div} U = f \in L^p_\#$ with $U \in L^{p^*}$.

By the Sobolev inequality, for all $g \in L^{p'}(\mathbb{T}^d)$ there holds

$$\|g-\overline{g}\|_{L^{p'}(\mathbb{T}^d)} \le \beta \|\nabla g\|_{L^{(p^*)'}(\mathbb{T}^d,\mathbb{R}^d)}, \quad \frac{1}{p^*} = \frac{1}{p} - \frac{1}{d}, \quad d \le p < \infty; \tag{2.3}$$

the case $p = d$ corresponds to the isoperimetric Gagliardo-Nirenberg inequality [13, 15] $\|g - \overline{g}\|_{L^{d'}(\mathbb{T}^d)} \le \beta \|g\|_{BV(\mathbb{T}^d)}$. We distinguish between two cases.

(i) The case $d < p < \infty$: the equation $\operatorname{div} U = f \in L^p_\#(\mathbb{T}^d)$ 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 $\operatorname{div} U = f \in L^d_\#(\mathbb{T}^d)$ has a solution $U \in L^\infty(\mathbb{T}^d, \mathbb{R}^d)$. This is the prototype example discussed in the beginning of the introduction. According to the intriguing observation of Bourgain and Brezis [7, prop. 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. The computation of such L^∞-solutions using hierarchical iterations in the two-dimensional critical case was carried out in [37].

Remark 2.1 (Homogeneity). We rewrite the hierarchical iteration (1.5) with $\lambda_1 = C\|f\|^{1-p}_{L^p}$ in the form

$$[\mathbf{u}_{j+1}, r_{j+1}] = \operatorname*{arg\,min}_{\mathcal{L}\mathbf{u}+r=r_j} \left\{ \|\mathbf{u}\|_{\mathbb{B}} + C2^j \frac{\|r\|^p_{L^p}}{\|f\|^{p-1}_{L^p}} \right\},$$

$$r_j := \begin{cases} f, & j = 0, \\ f - \mathcal{L}\big(\sum_{k=1}^j \mathbf{u}_k\big), & j \ge 1. \end{cases}$$

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 [36, remark 1.1], the hierarchical solution is homogeneous of degree 1: if $U = U_f$ specifies the (nonlinear) dependence of the hierarchical solution on f, then $U_{\{\alpha f\}} = \alpha U_f$.

Case 3. Solution of $\operatorname{div} U = f \in L^d_\#$ with $U \in L^\infty \cap \dot{W}^{1,d}$.

A central question raised and answered in [7] is whether (2.1) has a solution that 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 (2.2), which now reads

$$\|g - \overline{g}\|_{L^{d'}(\mathbb{T}^d)} \le \beta \|\nabla g\|_{L^1 + \dot{W}^{-1,d'}(\mathbb{T}^d,\mathbb{R}^d)} \quad \forall g \in L^{d'}(\mathbb{T}^d). \tag{2.4}$$

This key estimate was proved in [7]. Thus, Theorem 1.1 converts (2.4) into a constructive proof of the following:

COROLLARY 2.2. *The equation* $\operatorname{div} U = f \in L^d_{\#}(\mathbb{T}^d)$ *admits a solution* $U \in L^\infty \cap \dot W^{1,d}(\mathbb{T}^d, \mathbb{R}^d)$*, which is given by the hierarchical decomposition* $U = \sum_{j=1} \mathbf{u}_j$*. This is constructed by the refinement step*

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

with sufficiently large $\lambda_1 \geq \beta \|f\|^{1-d}_{L^d}$.

Remark 2.3. We comment here on the key role of the duality estimate (2.4). The case $d = 2$ was proved by a direct method outlined in [7, sec. 4]; alternative two-dimensional proofs were given by Maz′ya [26] and Mironescu [28]. For $d > 2$, however, the derivation of (2.4) was proved in Bourgain and Brezis [7, theorem 1] as a byproduct of their *construction* of $L^\infty \cap \dot W^{1,d}$ solutions for $\operatorname{div} U = f$(!). The construction, based on an intricate Littlewood-Paley decomposition, is rather involved [7, sec. 6], and to our knowledge, a simpler, *direct* derivation of (2.4) is still open. Thus, Corollary 2.2—which still depends on the construction of Bourgain and Brezis to justify (2.4)—offers a simpler alternative for the construction of such $L^\infty \cap \dot W^{1,d}$-bounded solutions in terms of the minimizers,

$$\vee_{\operatorname{div}}(r, \lambda) := \inf_{\mathbf{u}} \left\{ \|\mathbf{u}\|_{L^\infty \cap \dot W^{1,d}} + \lambda \|r - \operatorname{div} \mathbf{u}\|^d_{L^d} \right\}.$$

Computations of the related L^∞-based minimizers were carried out in [18,23], and it would be desirable to develop efficient algorithms to compute the corresponding minimizers of $\vee_{\operatorname{div}}(r, \lambda; L^\infty \cap \dot W^{1,d})$. Spectral approximation of such minimizers was discussed in [25].

Since the proof of the dual estimate (2.4) in $d > 2$ dimensions is indirect, a specific value of β is not known. As noted in remark 1.3, however, the hierarchical construction can proceed without a priori knowledge of the precise value of β: if one sets $\lambda_1 = \|f\|^{1-d}_{L^d}$, and this initial scale underestimates a correct value of, say, $\beta > 1$, then it will take at most $j_0 \sim \log(\beta)$ steps before picking up nontrivial terms in the hierarchical decomposition, $U = \sum_{j=j_0} \mathbf{u}_j$.

Case 4. Solution of $\operatorname{div} U = f \in L^d_{\#}(\Omega)$ with $U \in L^\infty \cap \dot W^{1,d}_0(\Omega)$.

The constructions of bounded solutions for (2.1) extend to the case of Lipschitz domains, $\Omega \subset \mathbb{R}^d$; see [7, sec. 7.2]. For future reference we state the following:

COROLLARY 2.4. *Given* $f \in L^d_{\#}(\Omega) := \{g \in L^d(\Omega) \mid \int_\Omega g(x)dx = 0\}$*, the equation* $\operatorname{div} U = f$ *admits a solution* $U \in L^\infty \cap \dot W^{1,d}_0(\Omega, \mathbb{R}^d)$ *such that*

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

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

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

with sufficiently large $\lambda_1 \gtrsim \beta \|f\|_{L^d(\Omega)}^{1-d}$.

3 Bounded Solution of curl $U = \mathbf{f} \in L^3_\#(\mathbb{T}^3, \mathbb{R}^3)$

Let $L^3_\#(\mathbb{T}^3, \mathbb{R}^3)$ denote the L^3-subspace of divergence-free 3-vectors with zero mean. We seek solutions of

$$\operatorname{curl} U = \mathbf{f}, \quad \mathbf{f} \in L^3_\#(\mathbb{T}^3, \mathbb{R}^3), \tag{3.1}$$

in an appropriate Banach space $U \in \mathbb{B}$. We appeal to the framework of hierarchical solutions in Theorem 1.1, where $\mathcal{P} : L^3(\mathbb{T}^3, \mathbb{R}^3) \mapsto L^3(\mathbb{T}^3, \mathbb{R}^3)$ is the irrotational portion of Hodge decomposition with a dual, $\mathcal{P}^*\mathbf{g} := \nabla \Delta^{-1} \operatorname{div} \mathbf{g} - \overline{\mathbf{g}}$. According to Theorem 1.1, we can construct hierarchical solutions $U \in \mathbb{B}$ of (3.1) provided (1.3) holds,

$$\|\mathbf{g} - \mathcal{P}^*\mathbf{g}\|_{L^{3/2}} \leq \beta \|\operatorname{curl} \mathbf{g}\|_{\mathbb{B}^*}, \quad \mathbf{g} \in L^{3/2}(\mathbb{T}^3, \mathbb{R}^3). \tag{3.2}$$

Since $\|\mathbf{g} - \mathcal{P}^*\mathbf{g}\|_{L^{3/2}} \lesssim \|\operatorname{curl} \mathbf{g}\|_{\dot{W}^{-1,3/2}}$, we can construct hierarchical solutions of (3.1) in $\dot{W}^{1,3}$. This has the same integrability as the divergence-free solution of (3.1), $(-\Delta)^{-1} \operatorname{curl} \mathbf{f}$. A more intricate question is whether (3.1) 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 [8, cor. 8′], which in turn imply the key a priori estimate: for all $\mathbf{g} \in L^{3/2}(\mathbb{T}^3, \mathbb{R}^3)$ such that $\operatorname{div} \mathbf{g} = \overline{\mathbf{g}} = 0$ there holds,

$$\|\mathbf{g} - \mathcal{P}^*\mathbf{g}\|_{L^{3/2}(\mathbb{T}^3, \mathbb{R}^3)} \leq \beta \|\operatorname{curl} \mathbf{g}\|_{L^1 + \dot{W}^{-1,3/2}(\mathbb{T}^3, \mathbb{R}^3)}. \tag{3.3}$$

Granted (3.3), Theorem 1.1 offers a simpler alternative to the construction in [8] based on the following hierarchical decomposition:

COROLLARY 3.1. *The equation* curl $U = \mathbf{f} \in L^3_\#(\mathbb{T}^3, \mathbb{R}^3)$ *admits a solution* $U \in L^\infty \cap \dot{W}^{1,3}(\mathbb{T}^3, \mathbb{R}^3)$,

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

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

$$\mathbf{u}_{j+1} = \underset{\mathbf{u}}{\arg\min} \left\{ \|\mathbf{u}\|_{L^\infty \cap \dot{W}^{1,3}} + \lambda_1 2^j \left\| \mathbf{f} - \operatorname{curl}\left(\sum_{k=1}^{j} \mathbf{u}_k + \mathbf{u} \right) \right\|_{L^3(\mathbb{T}^3, \mathbb{R}^3)}^3 \right\},$$
$$j = 0, 1, \ldots,$$

with sufficiently large $\lambda_1 \geq \beta \|\mathbf{f}\|_{L^3(\mathbb{T}^3, \mathbb{R}^3)}^{-2}$.

4 Construction of Hierarchical Solutions for $\mathcal{L}U = f \in L^p_\#(\Omega)$

4.1 A Prototype 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

$$\operatorname{div} U = f, \quad f \in L^2_\#(\mathbb{T}^2) := \left\{ g \in L^2(\mathbb{T}^2) \,\Big|\, \int_{\mathbb{T}^2} g(x)dx = 0 \right\}. \tag{4.1}$$

Our starting point for the construction of a uniformly bounded solution of (4.1) is a decomposition of f,

$$f = \operatorname{div} \mathbf{u}_1 + r_1, \quad f \in L^2_\#(\mathbb{T}^2), \tag{4.2a}$$

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

$$[\mathbf{u}_1, r_1] = \underset{\operatorname{div}\mathbf{u}+r=f}{\arg\min} \left\{ \|\mathbf{u}\|_{L^\infty} + \lambda_1 \|r\|^2_{L^2} \,\big|\, \mathbf{u} \in C^0(\mathbb{T}^2),\ r \in L^2_\#(\mathbb{T}^2) \right\}. \tag{4.2b}$$

Here λ_1 is a fixed parameter at our disposal: if we choose λ_1 large enough, $\lambda_1 > 1/(2\|f\|_{BV})$, then according to Lemma A.3 below (with $\varphi(r) = 2r$ and $\mathcal{L}^* = -\nabla$), (4.2b) admits a minimizer $[\mathbf{u}_1, r_1]$ satisfying

$$\|\nabla r_1\|_{\mathcal{M}} = \|r_1\|_{BV} = \frac{1}{2\lambda_1},$$

where the space of Radon measures $\mathcal{M}$ arises here as the dual of C^0. To proceed we invoke the isoperimetric Gagliardo-Nirenberg inequality, which states that there exists $\beta > 0$ (any $\beta \geq 1/\sqrt{4\pi}$ will do) such that for all bounded variation g's with zero mean,

$$\|g\|_{L^2} \leq \beta \|g\|_{BV}, \quad \int_{\mathbb{T}^2} g(x)dx = 0. \tag{4.3}$$

Since f has a zero mean, so does the residual r_1 and (4.3) yields

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

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 (4.2) and use it to decompose r_1 with scale $\lambda = \lambda_2 > \lambda_1 = 1/(2\|r_1)\|_{BV}$. This yields

$$r_1 = \operatorname{div} \mathbf{u}_2 + r_2,$$
$$[\mathbf{u}_2, r_2] = \underset{\operatorname{div}\mathbf{u}+r=r_1}{\arg\min} \left\{ \|\mathbf{u}\|_{L^\infty} + \lambda_2 \|r\|^2_{L^2} \,\big|\, \mathbf{u} \in C^0(\mathbb{T}^2),\ r \in L^2_\#(\mathbb{T}^2) \right\}.$$

Combining this with (4.2a) we obtain $f = \operatorname{div} U_2 + r_2$, where $U_2 := \mathbf{u}_1 + \mathbf{u}_2$ is viewed as an improved *approximate solution* of (4.1) in the sense that it has a smaller residual r_2,

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

when compared with the previous residual $\beta\|r_1\|_{BV} = \beta/(2\lambda_1)$. This process can be repeated: if $r_j \in L^2_\#(\mathbb{T}^2)$ is the residual at step j, then we decompose it

$$r_j = \operatorname{div} \mathbf{u}_{j+1} + r_{j+1}, \tag{4.4a}$$

where $[\mathbf{u}_{j+1}, r_{j+1}]$ is a minimizer over all pairs $(\mathbf{u}, r) \in (C^0(\mathbb{T}^2), L^2_\#(\mathbb{T}^2))$,

$$[\mathbf{u}_{j+1}, r_{j+1}] = \underset{\operatorname{div}\mathbf{u}+r=r_j}{\arg\min} \left\{ \|\mathbf{u}\|_{L^\infty} + \lambda_{j+1}\|r\|^2_{L^2} \right\}, \quad j = 0, 1, \ldots. \tag{4.4b}$$

For $j = 0$, the decomposition (4.4) is interpreted as (4.2) by setting $r_0 := f$. Note that the recursive decomposition (4.4a) 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 (4.3) $r_{j+1} \in L^2_\#(\mathbb{T}^2)$. The iterative process depends on a sequence of increasing scales, $\lambda_1 < \lambda_2 < \cdots < \lambda_{j+1}$, which are yet to be determined.

The telescoping sum of the first k steps in (4.4a) yields an improved approximate solution $U_k := \sum_{j=1}^k \mathbf{u}_j$:

$$f = \operatorname{div} U_k + r_k, \quad \|r_k\|_{L^2} \leq \beta\|r_k\|_{BV} = \frac{\beta}{2\lambda_k} \downarrow 0, \quad k = 1, 2, \ldots. \tag{4.5}$$

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 (4.4b) with the trivial pair $[\mathbf{u} \equiv 0, r_j]$ implies, in view of (4.5),

$$\begin{aligned} \|\mathbf{u}_{j+1}\|_{L^\infty} + \lambda_{j+1}\|r_{j+1}\|^2_{L^2} &\leq \lambda_{j+1}\|r_j\|^2_{L^2} \\ &\leq \begin{cases} \lambda_1\|f\|^2_{L^2}, & j = 0, \\ \frac{\beta^2\lambda_{j+1}}{4\lambda_j^2}, & j = 1, 2, \ldots. \end{cases} \end{aligned} \tag{4.6}$$

We conclude by choosing a sufficiently rapidly increasing λ_j 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 4.1. *Fix β such that* (4.3) *holds. Then, for any given $f \in L^2_\#(\mathbb{T}^2)$, there exists a uniformly bounded solution of* (4.1),

$$\operatorname{div} U = f, \quad \|U\|_{L^\infty} \leq 2\beta\|f\|_{L^2}. \tag{4.7}$$

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

$$(4.8)\qquad \begin{aligned} [\mathbf{u}_{j+1}, r_{j+1}] &= \underset{\operatorname{div}\mathbf{u}+r=r_j}{\arg\min} \left\{ \|\mathbf{u}\|_{L^\infty} + \lambda_1 2^j \|r\|_{L^2}^2 \right\}, \\ r_0 &:= f, \quad \lambda_1 = \frac{\beta}{\|f\|_{L^2}}. \end{aligned}$$

PROOF. With $\lambda_j = \lambda_1 2^{j-1}$ we have $\|U_k - U_\ell\|_{L^\infty} \lesssim \sum \lambda_{j+1}\lambda_j^{-2} \lesssim 2^{-\ell}$, $k > \ell \gg 1$. Let U be the limit of the Cauchy sequence $\{U_k\}$; then $\|U_j - U\|_{L^\infty} + \|\operatorname{div} U_j - f\|_{L^2} \lesssim 2^{-j} \to 0$, and since div has a closed graph on its domain $D := \{\mathbf{u} \in L^\infty : \operatorname{div}\mathbf{u} \in L^2(\mathbb{T}^2)\}$, it follows that $\operatorname{div} U = f$. By (4.6) we have

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

Here $\lambda_1 > 1/(2\|f\|_{BV})$ is a free parameter at our disposal: we choose $\lambda_1 := \beta/\|f\|_{L^2}$, which by (4.3) is admissible, $\lambda_1 = \beta/(\|f\|_{L^2}) > 1/(2\|f\|_{BV})$, and (4.7) follows. □

Remark 4.2 (Energy Decomposition). A telescoping summation of the left inequality of (4.6) yields

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

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

$$(4.9)\qquad \sum_{j=1}^{\infty} \frac{1}{2^{j-1}} \|\mathbf{u}_j\|_{L^\infty} \le \frac{\beta}{2} \|f\|_{L^2}.$$

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

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

A telescoping sum of the last equality yields the following:

Corollary 4.3. *Let $U = \sum_{j=1}^{\infty} \mathbf{u}_j \in L^\infty$ be a hierarchical solution of* $\operatorname{div} U = f$, $f \in L^2_\#(\mathbb{T}^2)$. *Then*

$$(4.10)\qquad \frac{1}{\lambda_1} \sum_{j=1}^{\infty} \frac{1}{2^{j-1}} \|\mathbf{u}_j\|_{L^\infty} + \sum_{j=1}^{\infty} \|\operatorname{div}\mathbf{u}_j\|_{L^2_\#(\mathbb{T}^2)}^2 = \|f\|_{L^2_\#(\mathbb{T}^2)}^2, \quad \lambda_1 = \frac{\beta}{2\|f\|_{L^2}}.$$

We mention two examples related to the two-dimensional setup of Theorem 4.1.

Oscillations and Image Processing

As noted earlier, there exists no linear construction of solutions of (4.1) for *general* $f \in L^2$. Yet, for the "slightly smaller" Lorenz space $L^{2,1}$, we have

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

(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 [7, prop. 2]). Thus, the nonlinear aspect of constructing hierarchical solutions for (4.1) becomes essential for highly oscillatory functions such that $f \in L^2 \setminus L^{2,1}$ (and in particular, $f \notin \mathrm{BV}(\mathbb{T}^2)$). Such f's are encountered in image processing in the form of noise, texture, and blurry images [10, 27]. Hierarchical decompositions in this context of images were introduced by us in [35] and were found to be effective tools in image denoising, image deblurring, and image registration, [5,10,21,31,32,36,38], including graph-based signals [19, 20]. Here, we are given a noisy and possibly blurry observed image, $f = \mathcal{L}U + r \in L^2(\mathbb{R}^2)$, and the purpose is to recover a faithful description of the underlying "clean" image, $U \sim$ "$\mathcal{L}^{-1}$"f, by denoising r and deblurring $\mathcal{L}$. The inverse "$\mathcal{L}^{-1}$"f should be properly interpreted, say, in the smaller space $\mathrm{BV}(\mathbb{R}^2) \subset L^2(\mathbb{R}^2)$, which is known to be well-adapted to represent edges. The resulting inverse problem can solved by the corresponding variational problem of [11, 12, 34],

$$[\mathbf{u}, r] = \underset{\mathcal{L}\mathbf{u}+r=f}{\arg\min} \left\{ \|\mathbf{u}\|_{BV} + \lambda \|r\|^2_{L^2(\mathbb{R}^2)} \right\}, \tag{4.11}$$

which is a special case of *Tikhonov regularization* [29, 30, 40]. The (BV, L^2)-hierarchical decomposition corresponding to (4.11) reads [35, 36]

$$\begin{gathered} f \simeq \mathcal{L}U_m, \quad U_m = \sum_{j=1}^{m} \mathbf{u}_j, \\ [\mathbf{u}_{j+1}, r_{j+1}] = \underset{\mathcal{L}\mathbf{u}+r=r_j}{\arg\min} \left\{ \|\mathbf{u}\|_{BV} + \lambda_1 2^j \|r\|^2_{L^2} \right\}. \end{gathered} \tag{4.12}$$

The oscillatory nature of noise and texture in images was addressed by Y. Meyer [27], who advocated replacing L^2 with the larger space of "images" $G := \{f \mid \operatorname{div}\mathbf{u} = f,\ \mathbf{u} \in L^{\infty}\}$. The equation $\operatorname{div}\mathbf{u} = f$ arises here with *one-signed* measure f's, and its L^{∞} solutions were characterized in [27, sec. 1.14], [33]: the space G_+ coincides with the Morrey space $M^2_+(\Omega)$:

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

For one-signed measures, $M^2_+(\Omega)$ coincides with Besov space $\dot{B}^{-1,\infty}_\infty$. The corresponding Meyer's energy functional then reads

$$[\mathbf{u}, r] = \underset{\mathcal{L}\mathbf{u}+r=f}{\arg\min} \left\{ \|\mathbf{u}\|_{BV(\Omega)} + \lambda \|r\|_{\dot{B}^{-1,\infty}_\infty(\Omega)} \right\};$$

numerical simulations with this model are found in [42].

$L^1(\mathbb{T}^2)$-Bounds and $\dot{H}^{-1}(\mathbb{T}^2)$-Compactness

Here is a simple application of Theorem 4.1. 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 4.1, there exist bounded $U_{ij} \in L^\infty(\mathbb{T}^2)$ such that

$$\begin{cases} \xi_1 \widehat{g}(\xi) = \xi_1 \widehat{U}_{11}(\xi) + \xi_2 \widehat{U}_{12}(\xi), \\ \xi_2 \widehat{g}(\xi) = \xi_1 \widehat{U}_{21}(\xi) + \xi_2 \widehat{U}_{22}(\xi), \end{cases} \qquad \|U_{ij}\|_{L^\infty} \lesssim \|g\|_{\dot{H}^1(\mathbb{T}^2)}.$$

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

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

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

$$g = U_1 + R_1 R_2 U_2, \quad \|U_1\|_{L^\infty} + \|U_2\|_{L^\infty} \lesssim \|g\|_{\dot{H}^1(\mathbb{T}^2)} \quad \text{for all } g \in \dot{H}^1(\mathbb{T}^2).$$

Here, U_1, U_2 are given by a linear combination of the U_{ij}'s in their Cartesian and their rotated coordinates. The last representation shows that although an $L^1(\mathbb{T}^2)$-bound of f does not imply $f \in \dot{H}^{-1}(\mathbb{T}^2)$, f does belong to $\dot{H}^{-1}$ if f *and* its repeated Riesz transform, $R_1 R_2 f$, are L^1-bounded.

COROLLARY 4.4. *The following bound holds*:

$$\|f\|_{\dot{H}^{-1}(\mathbb{T}^2)} \lesssim \|f\|_{L^1(\mathbb{T}^2)} + \|R_1 R_2 f\|_{L^1(\mathbb{T}^2)}. \tag{4.13}$$

As an example, consider a family of divergence-free two-vector fields, $\mathbf{u}^\epsilon(t,\cdot) \in L^2(\mathbb{T}^2, \mathbb{R}^2)$, 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 [16]. It was shown in [24] that $\{\mathbf{u}^\epsilon\}$ converges to such a weak solution if the vorticity, $\omega^\epsilon(t\cdot) := \partial_1 u_2^\epsilon(t,\cdot) - \partial_2 u_1^\epsilon(t,\cdot)$, is compactly embedded in $H^{-1}(\mathbb{T}^2)$. By Corollary 4.4, H^{-1}-compactness holds if $\{R_1 R_2 \omega^\epsilon(t,\cdot)\} \hookrightarrow L^1(\mathbb{T}^2)$; consult [41].

4.2 Hierarchical Solutions for $\mathcal{L}U = f \in L^p_\#(\Omega)$: Approximate Solutions

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

$$\mathcal{L}U = f, \quad f \in L^p_\#(\Omega), \quad 1 < p < \infty. \tag{4.14}$$

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

The linear operator $\mathcal{L}$ is assumed to be densely defined on $\mathbb{B}$ with a closed graph in $L^p_\# := L^p \cap \mathrm{Ker}(\mathcal{P})$ for an appropriate $\mathcal{P} : L^p \mapsto L^p$. We let $\mathcal{L}^* : L^{p'} \mapsto \mathbb{B}^*$ denote the formal dual of $\mathcal{L}$, acting on $L^{p'}$ with the natural pairing (effectively, $\mathcal{L}^*$ is acting on $L^{p'}_\# := L^{p'} \cap \mathrm{Ker}(\mathcal{P})$, since $R(\mathcal{P}^*)$ is in the null of $\mathcal{L}^*$)

$$\langle \mathcal{L}^* g, \mathbf{u} \rangle = (g, \mathcal{L}\mathbf{u}), \quad g \in D(\mathcal{L}^*),\ \mathbf{u} \in \mathbb{B},$$

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

$$\|\mathcal{L}^* g\|_{\mathbb{B}^*} := \sup_{\mathbf{u}\neq \mathbf{0}} \frac{\langle \mathcal{L}^* g, \mathbf{u}\rangle}{\|\mathbf{u}\|_{\mathbb{B}}}, \quad g \in D(\mathcal{L}^*).$$

We begin by constructing an *approximate solution* of (4.14), $U_\lambda : \mathcal{L}U_\lambda \approx f$, such that the residual $r_\lambda := f - \mathcal{L}U_\lambda$ 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,

$$\vee(f,\lambda) := \inf_{\mathcal{L}\mathbf{u}+r=f} \left\{ \|\mathbf{u}\|_{\mathbb{B}} + \lambda \|r\|^p_{L^p} : \mathbf{u} \in \mathbb{B},\ r \in L^p_\# \right\}. \tag{4.15}$$

In Theorem A.1 below, we show that if λ is chosen sufficiently large,

$$\lambda > \frac{1}{\|\mathcal{L}^*\varphi(f)\|_{\mathbb{B}^*}}, \quad \varphi(f) := p\,\mathrm{sgn}(f)|f|^{p-1}, \tag{4.16}$$

then the functional $\vee(f,\lambda)$ in (4.15) admits a minimizer, $\mathbf{u} = \mathbf{u}_\lambda$, such that the size of the residual, $r_\lambda := f - \mathcal{L}\mathbf{u}_\lambda$, is dictated by the dual statement

$$\|\mathcal{L}^*\varphi(r_\lambda)\|_{\mathbb{B}^*} = \frac{1}{\lambda}. \tag{4.17}$$

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

$$f = \mathcal{L}\mathbf{u}_1 + r_1, \quad [\mathbf{u}_1, r_1] = \underset{\mathcal{L}\mathbf{u}+r=f}{\arg\min}\ \vee(f,\lambda_1)$$

Borrowing the terminology from image processing, we note that the corresponding residual r_1 contains "small" features that 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 $r_1 \in L^p_\#$ so that we can repeat the $\vee$-decomposition of r_1, this time at the refined scale λ_2:

$$r_1 = \mathcal{L}\mathbf{u}_2 + r_2, \quad [\mathbf{u}_2, r_2] = \underset{\mathcal{L}\mathbf{u}+r=r_1}{\arg\min}\ \vee(r_1,\lambda_2).$$

Combining the last two steps, 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 $\mathcal{L}U_2 \approx f$.

Features below scale λ_2 remain unresolved in U_2, but the process can be continued by successive applications of the refinement step,

$$r_j = \mathcal{L}\mathbf{u}_{j+1} + r_{j+1}, \quad [\mathbf{u}_{j+1}, r_{j+1}] := \underset{\mathcal{L}\mathbf{u}+r=r_j}{\arg\min} \vee(r_j, \lambda_{j+1}), \quad j = 1, 2, \ldots. \tag{4.18}$$

To enable this process we require the residuals r_j to remain in $L^p_\#$. In view of the dual bound (4.17), we therefore make the following assumption:

Assumption (A Closure Bound). There exists a constant $\eta = \eta(p, d) < \infty$ such that the following a priori estimate holds:

$$\|g\|^p_{L^p} \le \eta \|\mathcal{L}^*\varphi(g)\|^{p'}_{\mathbb{B}^*}, \quad \varphi(g) = p\,\mathrm{sgn}(g)|g|^{p-1}. \tag{4.19}$$

We postpone the discussion of this bound to Theorem 4.5 below and continue with the generic hierarchical step where $[\mathbf{u}_{j+1}, r_{j+1}]$ is constructed as a minimizing pair of $\vee(r_j, \lambda_{j+1})$: since this minimizer is characterized by satisfying $\|\mathcal{L}^*\varphi(r_{j+1})\|_{\mathbb{B}^*} = 1/\lambda_{j+1}$, the closure bound (4.19) implies that $r_{j+1} \in L^p$; moreover, since r_j and $R(\mathcal{L})$ are in $\mathrm{Ker}(\mathcal{P})$,

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

and we conclude that $r_{j+1} \in L^p_\#$. In this manner, the iteration step $[\mathbf{u}_j, r_j] \mapsto [\mathbf{u}_{j+1}, r_{j+1}]$, is well-defined on $\mathbb{B} \times L^p_\#$. After k such steps we have

$$\begin{aligned} f &= \mathcal{L}\mathbf{u}_1 + r_1 \\ &= \mathcal{L}\mathbf{u}_1 + \mathcal{L}\mathbf{u}_2 + r_2 \\ &\;\vdots \\ &= \mathcal{L}\mathbf{u}_1 + \mathcal{L}\mathbf{u}_2 + \cdots + \mathcal{L}\mathbf{u}_k + r_k. \end{aligned}$$

We end up with a multiscale *hierarchical representation* of an approximate solution of (4.14) $U_k := \sum_{j=1}^k \mathbf{u}_j \in \mathbb{B}$ such that $\mathcal{L}U_k \simeq f$. Here the approximate equality $\simeq$ is interpreted as the convergence of the residuals,

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

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

THEOREM 4.5 (Approximate Solutions). *Consider $\mathcal{L} : \mathbb{B} \mapsto L^p_\#(\Omega)$ and assume its dual is injective so that* (1.3) *holds,*

$$\|g - \mathcal{P}^* g\|_{L^{p'}} \le \beta \|\mathcal{L}^* g\|_{\mathbb{B}^*} \quad \forall g \in L^{p'}(\Omega),$$

for some $\mathcal{P} : L^p \mapsto L^p$ whose range is "compatible" with the range of $\mathcal{L}$. Then, the equation $\mathcal{L}U = f \in L^p_\#(\Omega)$ admits an approximate solution $U_k \in \mathbb{B}$ *such that*

$\mathcal{L}U_k \simeq f$ in the sense that the residuals $r_k := f - \mathcal{L}U_k$ satisfy

$$\|\mathcal{L}^*\varphi(r_k)\|_{\mathbb{B}^*} = \frac{1}{\lambda_k}, \quad r_k := f - \mathcal{L}U_k. \tag{4.20}$$

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,

$$[\mathbf{u}_{j+1}, r_{j+1}] = \underset{\mathcal{L}\mathbf{u}+r=r_j}{\arg\min} \left\{\|\mathbf{u}\|_{\mathbb{B}} + \lambda_{j+1}\|r\|^p_{L^p}\right\}, \quad r_0 = f.$$

PROOF. We verify that the a priori duality estimate assumed in (1.3) implies the closure bound sought in (4.19). Fix $g \in L^p_\#(\Omega)$. Then $\varphi(g) := p\,\mathrm{sgn}(g)|g|^{p-1} \in L^{p'}(\Omega)$, and since $g \in \mathrm{Ker}(\mathcal{P})$ we find

$$\begin{aligned} p\int_\Omega |g|^p\,dx = \int_\Omega g\varphi(g)dx &= \int_\Omega g(\varphi(g) - \mathcal{P}^*\varphi(g))dx \\ &\le \|g\|_{L^p}\|\varphi(g) - \mathcal{P}^*\varphi(g)\|_{L^{p'}}. \end{aligned}$$

The a priori dual estimate assumed in (1.3) then yields

$$\begin{aligned} p\|g\|^p_{L^p} &\le \|g\|_{L^p}\|\varphi(g) - \mathcal{P}^*\varphi(g)\|_{L^{p'}} \\ &\le \beta\|g\|_{L^p}\|\mathcal{L}^*\varphi(g)\|_{\mathbb{B}^*} \quad \forall g \in L^p_\#(\Omega), \end{aligned}$$

and the closure bound (4.19) follows, with $\eta := (\beta/p)^{p'}$,

$$\|g\|^{p-1}_{L^p} \le \frac{\beta}{p}\|\mathcal{L}^*\varphi(g)\|_{\mathbb{B}^*}. \tag{4.21}$$

This allows us to proceed with the hierarchical iterations (4.18),

$$\begin{aligned} [\mathbf{u}_j, r_j] \in \mathbb{B} \times L^p_\# &\mapsto \mathbf{u}_{j+1}, r_{j+1}] \\ &:= \underset{\mathcal{L}\mathbf{u}+r=r_j}{\arg\min} \vee(r_j, \lambda_{j+1}) \in \mathbb{B} \times L^p_\#, \quad j = 1, 2, \dots, \end{aligned}$$

starting with $[\mathbf{u}_0, r_0] = [0, f]$. A telescoping summation of (4.18) yields an approximate solution $U_k = \sum_{j=1}^k \mathbf{u}_j$ such that its residual $r_k = f - \mathcal{L}U_k$ satisfies (4.20). □

Remark 4.6 (On the Closure Bound). As an example for the closure bound (4.21) for $\mathcal{L}$'s with an injective dual, consider the critical case of $\mathcal{L} = \mathrm{div} : L^\infty \mapsto L^d(\mathbb{T}^d)$, and let $\mathcal{P}$ denote the zero averaging projection $\mathcal{P}g = g - \overline{g}$. The corresponding dual estimate (1.3) reads

$$\|g - \overline{g}\|_{L^{d'}} \lesssim \|\mathcal{L}^*g\|_{BV}.$$

This is the isoperimetric Gagliardo-Nirenberg inequality, and it implies, along the lines of Theorem 4.5, the following closure bound corresponding to (4.19):

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

Equivalently, we can rewrite this inequality in terms of $\varphi(g) = p\,\mathrm{sgn}(g)|g|^{p-1}$ as $\|\varphi(g)\|_{L^{d'}} \lesssim \|\varphi(g)\|_{BV(\mathbb{T}^d)}$. The observant reader will notice that the latter is a slight variant of the Gagliardo-Nirenberg inequality, since for $d > 2$, $\varphi(g)$ need not have zero average; only g does.

4.3 From Approximate to Exact Solutions

We now turn to showing that the approximate solutions, $U_k = \sum_{j=1}^{k} \mathbf{u}_j$, converge to a limit $U = \sum_{j=1}^{\infty} \mathbf{u}_j$, which is an exact solution sought for (4.14), uniformly bounded in $\mathbb{B}$.

We start by comparing the minimizer $[\mathbf{u}_{j+1}, r_{j+1}]$ of $\vee(r_j, \lambda_{j+1})$ in (4.18) with the trivial pair $[\mathbf{u} \equiv 0, r_j]$, which yields the key refinement estimate

$$\|r_j\|_{L^p}^p \geq \frac{1}{\lambda_{j+1}} \|\mathbf{u}_{j+1}\|_{\mathbb{B}} + \|r_{j+1}\|_{L^p}^p, \quad j = 0, 1, \ldots. \tag{4.22}$$

In particular, the closure bound (4.19) followed by (4.20) implies

$$\|\mathbf{u}_{j+1}\|_{\mathbb{B}} \leq \lambda_{j+1}\|r_j\|_{L^p}^p \begin{cases} = \lambda_1 \|f\|_{L^p}^p, & j = 0, \\ \leq \lambda_{j+1}\eta \|\mathcal{L}^*\varphi(r_j)\|_{\mathbb{B}^*}^{1/p'} \leq \dfrac{\lambda_{j+1}\eta}{\lambda_j^{p'}}, & j \geq 1, \end{cases} \tag{4.23}$$

where $\{\lambda_j\}$ is an increasing sequence of scales at our disposal. Setting $\lambda_j = \lambda_1 2^{j-1}$, we conclude that the approximate solutions, $U_k = \sum_1^k \mathbf{u}_j$, form a Cauchy sequence:

$$\|U_k - U_\ell\|_{\mathbb{B}} \lesssim \sum_{j=\ell+1}^{k} 2^{j(1-p')} \ll 1, \quad k > \ell \gg 1,$$

which has a limit, $U = \sum_{j=1}^{\infty} \mathbf{u}_j$, such that $\|\mathcal{L}U_j - f\|_{L^p}^p \to 0$. Since $\mathcal{L}$ has a closed graph in L^p, $\mathcal{L}U = f$. It remains to show that the limit U has a finite $\mathbb{B}$-norm, which brings us to the following proof:

PROOF OF THEOREM 1.1. Using (4.23) with $\eta = (\beta/p)^{p'}$ yields

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

Set $\lambda_1 := \beta \|f\|_{L^p}^{1-p}$. Such a choice of λ_1 satisfies the admissibility requirement (4.16). Indeed, according to (4.21), $\|g\|_{L^p}^{p-1} \leq \frac{\beta}{p} \|\mathcal{L}^*\varphi(g)\|_{\mathbb{B}^*}$; hence

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

and the uniform bound (1.4) follows:

$$\|U\|_{\mathbb{B}} \leq \gamma \|f\|_{L^p_{\#}}, \quad \gamma = \beta\left(1 + \left(\frac{2}{p}\right)^{p'} \frac{1}{2^{p'-1}-1}\right). \tag{4.24}$$ □

Remark 4.7. We summarize the two main aspects in the hierarchical construction.

(i) The existence of minimizers $\{\mathbf{u}_j\}$ of $\vee(r_{j-1}, \lambda_j)$ follows from basic principles in uniformly convex Banach spaces. We use here the mere existence of such minimizers instead of standard duality-based existence arguments in the closed range theorem, e.g., [46, VII.5], [44, I.A.13-14], [9, theorem 2.20]. We note in passing that existence of minimizers and duality principles in uniformly convex Banach spaces can be deduced from each other [22].

(ii) The exponential decay of these minimizers and hence the uniform bound of their sum, $\|U\|_{\mathbb{B}} \leq \sum \|\mathbf{u}_j\|_{\mathbb{B}} \lesssim \|f\|_{L^p_{\#}}$, follow from the key a priori dual estimate (1.3) used in the refinement step (4.23).

Remark 4.8 (Extension to Orlicz Spaces). The hierarchical construction extends to equations valued in more general Orlicz spaces,

$$\begin{aligned} &\mathcal{L}U = f \in L^{\Phi}_{\#} := L^{\Phi} \cap \operatorname{Ker}(\mathcal{P}^*), \\ &L^{\Phi} = \left\{ f : [f]_{\Phi} := \int_{\Omega} \Phi(|f|)dx < \infty \right\}, \end{aligned} \tag{4.25}$$

for a proper N-function Φ satisfying the Δ_2 condition [2, sec. 8], [6, sec. 4.8].

Assume that the following a priori closure bound holds: there exists an increasing function $\eta : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that

$$[g]_{\Phi} \leq \eta(\|\mathcal{L}^*\varphi(g)\|_{\mathbb{B}}), \quad \int_{s=0}^{1} \frac{\eta(s)}{s^2}\, ds < \infty.$$

Then the problem (4.25) admits the bounded hierarchical solution $U = \sum \mathbf{u}_j$ such that $\|U\|_{\mathbb{B}} \lesssim [f]_{\Phi}$. The closure bound enters through the initial scale $\lambda_1 \gtrsim 1/\eta^{-1}([f]_{\Phi})$. The L^p setup corresponds to $\Phi(t) = t^p$ and $\eta(s) \sim s^{p'}$.

Remark 4.9 (Sharp Bounds). The bound (4.24) with $p = 2$ shows that if $\mathcal{L}^*$ is injective so that (1.3) holds with constant β, then $\mathcal{L}U = f \in L^2$ admits a solution $\|U\|_{\mathbb{B}} \leq \gamma \|f\|_{L^2}$, with twice the bound $\gamma = 2\beta$ (in agreement with the L^2-case in Theorem 4.1). Using a rapidly growing scale $\lambda_{j+1} = \lambda_1 \zeta^j$ with $\zeta \gg 1$ yields a tighter bound γ. A sharp form of the $\mathbb{B}$-bound (4.24) for general $1 < p < \infty$,

$$\|U\|_{\mathbb{B}} \leq \gamma \|f\|_{L^p_{\#}} \quad \text{for } any\ \gamma > \beta, \tag{4.26$_\gamma$}$$

can be argued by invoking the Hahn-Banach theorem. To this end, we reproduce here a slight generalization of [7, prop. 1]. Normalize $\|f\|_{L^p} = 1$ and consider the two nonempty convex sets: the ball

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

and $C := \{U \in \mathbb{B} : \mathcal{L}U = f\}$. The claim is that $B_{\gamma_\epsilon} \cap C \neq \varnothing$ and the desired estimate $(4.26)_\gamma$, $\gamma = \gamma_\epsilon$, then follows with arbitrarily small ϵ. If not, $B_{\gamma_\epsilon} \cap C = \varnothing$, and by Hahn-Banach there exists a nontrivial $g^* \in L^{p'}$ such that for some $\alpha \in \mathbb{R}_+$

$$\langle g^*, \mathbf{u}\rangle \leq \alpha \quad \forall \mathbf{u} \in B_{\gamma_\epsilon} \tag{4.27a}$$

and

$$\langle g^*, U\rangle \geq \alpha \quad \forall U \in C. \tag{4.27b}$$

If $V \in \mathrm{Ker}(\mathcal{L})$ then by applying (4.27b) with $U \mapsto U \pm \delta V \in C$ we obtain $\pm\delta\langle g^*, V\rangle \geq 0$ or $\langle g^*, V\rangle = 0$; that is, $g^* \in \mathrm{Ker}(\mathcal{L})^\perp = \mathrm{R}(\mathcal{L}^*)$ is of the form $g^* = \mathcal{L}^* g$ for some $g \in D(\mathcal{L}^*) \subset L^{p'}$.

Now, by (4.27a)

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

and the a priori estimate assumed in (1.3) implies

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

But this leads to a contradiction: pick $U \in C$ (which we recall is not empty); then (4.27b) implies

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

Appendix: On ∨-minimizers

To study the hierarchical expansions (4.18), we characterize the minimizers of the ∨-functionals (4.15),

$$\begin{aligned} [\mathbf{u}, r] &:= \underset{\mathcal{L}\mathbf{u}+r=f}{\arg\min} \vee(f,\lambda), \\ \vee(f,\lambda) &:= \inf_{\mathcal{L}\mathbf{u}+r=f} \left\{\|\mathbf{u}\|_{\mathbb{B}} + \lambda\|r\|^p_{L^p} : \mathbf{u} \in \mathbb{B}\right\}. \end{aligned} \tag{A.1}$$

Here $\mathcal{L} : \mathbb{B} \mapsto L^p_\#(\Omega)$ is densely defined into a subspace of $L^p(\Omega)$ over a Lipschitz domain $\Omega \subset \mathbb{R}^d$. The characterization summarized below extends related results that can be found in [27, theorem 4], [4, chap. 1], [36, theorem 2.3].

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

$$\langle \mathcal{L}^* g, \mathbf{u}\rangle \leq \|\mathbf{u}\|_{\mathbb{B}}\|\mathcal{L}^* g\|_{\mathbb{B}^*}, \quad g \in D(\mathcal{L}^*),\ \mathbf{u} \in \mathbb{B} \tag{A.2}$$

(and the convention that $\|\mathcal{L}^* g\|_{\mathbb{B}^*} = \infty$ if g lies outside (the closure of) $D(\mathcal{L}^*)$). We say that $\mathbf{u}$ and $\mathcal{L}^* g$ are an *extremal pair* if equality holds above. The theorem below characterizes $[\mathbf{u}, r]$ as a minimizer of the ∨-functional if and only if $\mathbf{u}$ and $\mathcal{L}^*\varphi(r)$ form an extremal pair.

THEOREM A.1. *Let* $\mathcal{L}: \mathbb{B} \to L^p_{\#}(\Omega)$ *be a linear operator with dual* $\mathcal{L}^*$*, and let* $\vee(f, \lambda)$ *denote the associated functional* (4.15).

(i) *The variational problem* (A.1) *admits a minimizer* $\mathbf{u}$. *Moreover, if* $\|\cdot\|_{\mathbb{B}}$ *is* strictly convex*, then the minimizer* $\mathbf{u}$ *is unique.*

(ii) $\mathbf{u} \in \mathbb{B}$ *is a minimizer of* (A.1) *if and only if the residual* $r := f - \mathcal{L}\mathbf{u}$ *satisfies*

$$\text{(A.3)} \qquad \begin{aligned} \langle \mathcal{L}^*\varphi(r), \mathbf{u}\rangle &= \|\mathbf{u}\|_{\mathbb{B}} \cdot \|\mathcal{L}^*\varphi(r)\|_{\mathbb{B}^*} = \frac{\|\mathbf{u}\|_{\mathbb{B}}}{\lambda}, \\ \varphi(r) &:= p\,\mathrm{sgn}(r)|r|^{p-1} \in L^{p'}. \end{aligned}$$

PROOF.

(i) The existence of a minimizer for the $\vee$-functional follows from standard arguments which we omit, consult [1, 27]. We address the issue of uniqueness. Assume $\mathbf{u}_1$ and $\mathbf{u}_2$ are minimizers with the corresponding residuals $r_1 = f - \mathcal{L}\mathbf{u}_1$ and $r_2 = f - \mathcal{L}\mathbf{u}_2$

$$\|\mathbf{u}_i\|_{\mathbb{B}} + \lambda\|r_i\|^p_{L^p} = v_{\min}, \quad i = 1, 2.$$

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

$$\begin{aligned} v_{\min} &\le \|\mathbf{u}_\theta\|_{\mathbb{B}} + \lambda\|r_\theta\|^p_{L^p} \\ &\le \theta\|\mathbf{u}_2\|_{\mathbb{B}} + (1-\theta)\|\mathbf{u}_1\|_{\mathbb{B}} + \theta\lambda\|r_2\|^p_{L^p} + (1-\theta)\lambda\|r_1\|^p_{L^p} = v_{\min}. \end{aligned}$$

Consequently, $\|r_\theta\|^p_{L^p} = \theta\|r_2\|^p_{L^p} + (1-\theta)\|r_1\|^p_{L^p}$ and hence $r_1 = r_2$. In particular, $\|r_1\|^p_{L^p} = \|r_2\|^p_{L^p}$ 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 strict convexity must be the trivial segment, i.e., $\mathbf{u}_2 = \mathbf{u}_1$. We note in passing that strict convexity is in fact *necessary* for uniqueness, e.g., the counterexample of lack of uniqueness over the ℓ^∞-unit ball [27, p. 40].

(ii) If $\mathbf{u}$ is a minimizer of (A.1), then for any $\mathbf{v}$ in the domain of $\mathcal{L}$ we have

$$\text{(A.4)} \qquad \begin{aligned} \|\mathbf{u}\|_{\mathbb{B}} + \lambda\|f - \mathcal{L}\mathbf{u}\|^p_{L^p} &= \vee(\mathbf{u}, \lambda) \\ &\le \vee(\mathbf{u} + \epsilon\mathbf{v}, \lambda) \\ &= \|\mathbf{u} + \epsilon\mathbf{v}\|_{\mathbb{B}} + \lambda\|f - \mathcal{L}(\mathbf{u} + \epsilon\mathbf{v})\|^p_{L^p} \\ &\le \|\mathbf{u}\|_{\mathbb{B}} + |\epsilon| \cdot \|\mathbf{v}\|_{\mathbb{B}} + \lambda\|f - \mathcal{L}\mathbf{u}\|^p_{L^p} \\ &\quad - p\lambda\epsilon\big(\mathrm{sgn}(f - \mathcal{L}\mathbf{u})|f - \mathcal{L}\mathbf{u}|^{p-1}, \mathcal{L}\mathbf{v}\big) + o(\epsilon). \end{aligned}$$

Since the domain of $\mathcal{L}$ is assumed to be densely defined in $\mathbb{B}$, it follows that for all $\mathbf{v} \in \mathbb{B}$,

$$|\langle \mathcal{L}^*\varphi(r), \mathbf{v}\rangle| \le \frac{1}{\lambda}\|\mathbf{v}\|_{\mathbb{B}} + o(1), \quad \varphi(r) = p\,\mathrm{sgn}(r)|r|^{p-1},\ r := f - \mathcal{L}\mathbf{u},$$

and by letting $\epsilon \to 0$

$$\|\mathcal{L}^*\varphi(r)\|_{\mathbb{B}^*} \le \frac{1}{\lambda}. \tag{A.5}$$

To verify the reverse inequality, we set $\mathbf{v} = \pm\mathbf{u}$ and $0 < \epsilon < 1$ in (A.4), yielding

$$\|\mathbf{u}\|_{\mathbb{B}} + \lambda\|f - \mathcal{L}\mathbf{u}\|^p_{L^p} \le (1 \pm \epsilon)\|\mathbf{u}\|_{\mathbb{B}} + \lambda\|f - \mathcal{L}\mathbf{u} \mp \epsilon\mathcal{L}\mathbf{u}\|^p_{L^p},$$

and hence $\pm\epsilon\|\mathbf{u}\|_{\mathbb{B}} \mp \lambda\epsilon(\varphi(f - \mathcal{L}\mathbf{u}), \mathcal{L}\mathbf{u}) + o(\epsilon) \ge 0$. Dividing by ϵ and letting $\epsilon \downarrow 0_+$, we obtain $\|\mathbf{u}\|_{\mathbb{B}} = \lambda\langle\mathcal{L}^*\varphi(r), \mathbf{u}\rangle$ and (A.3) follows:

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

Conversely, we show that if (A.3) holds, then $\mathbf{u}$ is a minimizer. The convexity of L^p yields

$$\begin{aligned}\|f - \mathcal{L}(\mathbf{u} + \mathbf{v})\|^p_{L^p} &= \|r - \mathcal{L}\mathbf{v}\|^p_{L^p} \\ &\ge \|r\|^p_{L^p(\Omega)} - p\big(\mathrm{sgn}(r)|r|^{p-1}, \mathcal{L}(\mathbf{u} + \mathbf{v})\big) \\ &\quad + p\big(\mathrm{sgn}(r)|r|^{p-1}, \mathcal{L}\mathbf{u}\big) \\ &= \|f - \mathcal{L}\mathbf{u}\|^p_{L^p} - \underbrace{\langle\mathcal{L}^*\varphi(r), (\mathbf{u} + \mathbf{v})\rangle}_{\#1} + \underbrace{\langle\mathcal{L}^*\varphi(r), \mathbf{u}\rangle}_{\#2}.\end{aligned}$$

By the equalities assumed in (A.3), $\|\mathcal{L}^*\varphi(r)\|_{\mathbb{B}^*} = 1/\lambda \rightsquigarrow -\lambda(\#1) \ge -\|\mathbf{u} + \mathbf{v}\|_{\mathbb{B}}$, and moreover, $\lambda(\#2) = \|\mathbf{u}\|_{\mathbb{B}}$. We conclude that for any $\mathbf{v} \in \mathbb{B}$,

$$\begin{aligned}\vee(\mathbf{u} + \mathbf{v}, \lambda) &= \|\mathbf{u} + \mathbf{v}\|_{\mathbb{B}} + \lambda\|f - \mathcal{L}(\mathbf{u} + \mathbf{v})\|^p_{L^p} \\ &\ge \|\mathbf{u} + \mathbf{v}\|_{\mathbb{B}} + \lambda\|f - \mathcal{L}\mathbf{u}\|^p_{L^p} - p\lambda(\#1) + p\lambda(\#2) \\ &\ge \|\mathbf{u}\|_{\mathbb{B}} + \lambda\|f - \mathcal{L}\mathbf{u}\|^p_{L^p} = \vee(\mathbf{u}, \lambda).\end{aligned}$$

Thus, $\mathbf{u}$ is a minimizer of (A.1). □

The next two assertions are a refinement of Theorem A.1, depending on the size of $\|\mathcal{L}^*\varphi(f)\|_{\mathbb{B}^*}$.

LEMMA A.2 (The Case $\|\mathcal{L}^*\varphi(f)\|_{\mathbb{B}^*} \le 1/\lambda$). *Let $\mathcal{L} : \mathbb{B} \to L^p_\#(\Omega)$ with adjoint $\mathcal{L}^*$ and let $\vee$ denote the associated functional* (4.15). *Then $\lambda\|\mathcal{L}^*\varphi(f)\|_{\mathbb{B}^*} \le 1$ if and only if $\mathbf{u} \equiv 0$ is a minimizer of* (A.1).

PROOF. Assume $\|\mathcal{L}^*\varphi(f)\|_{\mathbb{B}^*} \le 1/\lambda$. Then by convexity of L^p

$$\begin{aligned}&\|\mathbf{u}\|_{\mathbb{B}} + \lambda\|f - \mathcal{L}\mathbf{u}\|^p_{L^p} \\ &\quad \ge \|\mathbf{u}\|_{\mathbb{B}} + \lambda\int_\Omega |f|^p\,dx - \lambda\int_\Omega (\varphi(f), \mathcal{L}\mathbf{u})dx \\ &\quad \ge \|\mathbf{u}\|_{\mathbb{B}} + \lambda\int_\Omega |f|^p\,dx - \lambda\|\mathcal{L}^*\varphi(f)\|_{\mathbb{B}^*}\|\mathbf{u}\|_{\mathbb{B}} \ge \lambda\|f\|^p_{L^p},\end{aligned}$$

which tells us that $\mathbf{u} \equiv 0$ is a minimizer of (4.15). Conversely, if $\mathbf{u} \equiv 0$ is a minimizer of (A.1), then $\epsilon\|\mathbf{u}\|_{\mathbb{B}} + \lambda\|f - \epsilon\mathcal{L}\mathbf{u}\|^p_{L^p} \geq \lambda\|f\|^p_{L^p}$ for all $\mathbf{u} \in \mathbb{B}$. It follows that

$$\epsilon\|\mathbf{u}\|_{\mathbb{B}} - \epsilon\lambda \int_{\Omega} (\varphi(f), \mathcal{L}\mathbf{u})dx + o(\epsilon) \geq 0.$$

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

LEMMA A.3 (The Case $\|\mathcal{L}^*\varphi(f)\|_{\mathbb{B}^*} > 1/\lambda$). *Let* $\mathcal{L} : \mathbb{B} \to L^p_\#(\Omega)$ *with adjoint* $\mathcal{L}^*$*, and let* $\vee$ *denote the associated functional* (4.15)*. If* $1 < \lambda\|\mathcal{L}^*\varphi(f)\|_{\mathbb{B}^*} \leq \infty$*, then* $\mathbf{u}$ *is a minimizer of* (A.1) *if and only if* $\mathcal{L}\mathbf{u}$ *and* $\varphi(r)$ *are an extremal pair and*

$$\|\mathcal{L}^*\varphi(r)\|_{\mathbb{B}^*} = \frac{1}{\lambda}, \quad \langle\mathbf{u}, \mathcal{L}^*\varphi(r)\rangle = \frac{\|\mathbf{u}\|_{\mathbb{B}}}{\lambda}. \tag{A.6}$$

PROOF. Since $\|\mathcal{L}^*\varphi(f)\|_{\mathbb{B}^*} > 1/\lambda$, we have $\|\mathbf{u}\|_{\mathbb{B}} \neq 0$ and can now divide the equality on the right of (A.3) by $\|\mathbf{u}\|_{\mathbb{B}} \neq 0$ and (A.6) follows. □

Acknowledgments. I am indebted to Haim Brezis for discussions on the works [7, 8], to François Golse, Giuseppe Savaré, and Terence Tao, who, respectively, brought to my attention references [22], [17, 39], and [3], and to Przemyslaw Wojtaszczyk for his comments on an earlier version of this paper [45]. Research was supported by National Science Foundation grants DMS10-08397 and RNMS11-07444 (KI-Net) and Office of Naval Research Grant #N00014-1210318.

Bibliography

[1] Acar, R.; Vogel, C. R. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Problems* **10** (1994), no. 6, 1217–1229.

[2] Adams, R. A.; Fournier, J. J. F. *Sobolev spaces.* Second edition. Pure and Applied Mathematics (Amsterdam), 140. Elsevier/Academic Press, Amsterdam, 2003.

[3] Ajiev, S. On Chebyshev centres, retractions, metric projections and homogeneous inverses for Besov, Lizorkin-Triebel, Sobolev and other Banach spaces. *East J. Approx.* **15** (2009), no. 3, 375–428.

[4] Andreu-Vaillo, F.; Caselles, V.; Mazón, J. M. *Parabolic quasilinear equations minimizing linear growth functionals.* Progress in Mathematics, 223. Birkhäuser, Basel, 2004. doi:10.1007/978-3-0348-7928-6

[5] Athavale, P.; Xu, R.; Radau, P.; Nachman, A.; Wright, G. Multiscale TV flow with applications to fast denoising and registration. *Proc. SPIE* **8669** (2013), 86692K–86692K-7. doi:10.1117/12.2007190

[6] Bennett, C.; Sharpley, R. *Interpolation of operators.* Pure and Applied Mathematics, 129. Academic Press, Boston, 1988.

[7] Bourgain, J.; Brezis, H. On the equation div $Y = f$ and application to control of phases. *J. Amer. Math. Soc.* **16** (2003), no. 2, 393–426 (electronic). doi:10.1090/S0894-0347-02-00411-3

[8] Bourgain, J.; Brezis, H. New estimates for elliptic equations and Hodge type systems. *J. Eur. Math. Soc. (JEMS)* **9** (2007), no. 2, 277–315. doi:10.4171/JEMS/80

[9] Brezis, H. *Functional analysis, Sobolev spaces and partial differential equations.* Universitext. Springer, New York, 2011.

[10] Buades, A.; Coll, B.; Morel, J. M. A review of image denoising algorithms, with a new one. *Multiscale Model. Simul.* **4** (2005), no. 2, 490–530. doi:10.1137/040616024

[11] Chambolle, A.; Lions, P.-L. Image restoration by constrained total variation minimization and variants. *Proc. SPIE* **2567** (1995), 50–59. doi:10.1117/12.218480

[12] Chambolle, A.; Lions, P.-L. Image recovery via total variation minimization and related problems. *Numer. Math.* **76** (1997), no. 2, 167–188. doi:10.1007/s002110050258

[13] Cordero-Erausquin, D.; Nazaret, B.; Villani, C. A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities. *Adv. Math.* **182** (2004), no. 2, 307–332. doi:10.1016/S0001-8708(03)00080-X

[14] Dacorogna, B.; Fusco, N.; Tartar, L. On the solvability of the equation div $u = f$ in L^1 and in C^0. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.* **14** (2003), no. 3, 239–245 (2004).

[15] Del Pino, M.; Dolbeault, J. Best constants for Gagliardo-Nirenberg inequalities and applications to nonlinear diffusions. *J. Math. Pures Appl. (9)* **81** (2002), no. 9, 847–875. doi:10.1016/S0021-7824(02)01266-7

[16] DiPerna, R. J.; Majda, A. J. Concentrations in regularizations for 2-D incompressible flow. *Comm. Pure Appl. Math.* **40** (1987), no. 3, 301–345. doi:10.1002/cpa.3160400304

[17] Gaglirado, E. Caratterizzazioni delle tracce sulla frontiera relative ad alcuni classi di funzioni in n variabili. *Rend. Scm. Mat. Univ. Padova* **27** (1957), 284–305.

[18] Garnett, J. B.; Le, T. M.; Meyer, Y.; Vese, L. A. Image decompositions using bounded variation and generalized homogeneous Besov spaces. *Appl. Comput. Harmon. Anal.* **23** (2007), no. 1, 25–56. doi:10.1016/j.acha.2007.01.005

[19] Hidane, M.; Lézoray, O.; Elmoataz, A. Nonlinear multilayered representation of graph-signals. *J. Math. Imaging Vision* **45** (2013), no. 2, 114–137. doi:10.1007/s10851-012-0348-9

[20] Hidane, M.; Lézoray, O.; Ta, V.-T.; Elmoataz, A. Nonlocal multiscale hierarchical decomposition on graphs. *Computer Vision – ECCV 2010*, 638–650. Lecture Notes in Computer Science, 6314. Springer, Berlin-Heidelberg, 2010. doi:10.1007/978-3-642-15561-1_46

[21] Hintermüller, M.; Rincon-Camacho, M. M. Expected absolute value estimators for a spatially adapted regularization parameter choice rule in L1-TV-based image restoration. *Inverse Problems* **26** (2010), no. 8, 085005, 30 pp. doi:10.1088/0266-5611/26/8/085005

[22] James, R. C. Orthogonality and linear functionals in normed linear spaces. *Trans. Amer. Math. Soc.* **61** (1947), 265–292. doi:10.2307/1990220

[23] Le, T. M.; Vese, L. A. Image decomposition using total variation and div(BMO). *Multiscale Model. Simul.* **4** (2005), no. 2, 390–423. doi:10.1137/040610052

[24] Lopes Filho, M. C.; Nussenzveig Lopes, H. J.; Tadmor, E. Approximate solutions of the incompressible Euler equations with no concentrations. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **17** (2000), no. 3, 371–412. doi:10.1016/S0294-1449(00)00113-X

[25] Maday, Y. L^∞-stable approximation of a solution to div$(\mathbf{Y}) = f$ for $f \in L^2$ in two dimensions. *J. Sci. Comput.* **28** (2006), no. 2-3, 451–458. doi:10.1007/s10915-006-9073-0

[26] Maz′ya, V. Bourgain-Brezis type inequality with explicit constants. *Interpolation theory and applications*, 247–252. Contemporary Mathematics, 445. American Mathematical Society, Providence, R.I., 2007. doi:10.1090/conm/445/08605

[27] Meyer, Y. *Oscillating patterns in image processing and nonlinear evolution equations.* University Lecture Series, 22. American Mathematical Society, Providence, R.I., 2001.

[28] Mironescu, P. On some inequalities of Bourgain, Brezis, Maz′ya, and Shaposhnikova related to L^1 vector fields. *C. R. Math. Acad. Sci. Paris* **348** (2010), no. 9-10, 513–515. doi:10.1016/j.crma.2010.03.019

[29] Morozov, V. A. *Methods for solving incorrectly posed problems.* Springer, New York, 1984. doi:10.1007/978-1-4612-5280-1

[30] Morozov, V. A. *Regularization methods for ill-posed problems.* CRC Press, Boca Raton, Fla., 1993.

[31] Paquin, D.; Levy, D.; Schreibmann, E.; Xing, L. Multiscale image registration. *Math. Biosci. Eng.* **3** (2006), no. 2, 389–418. doi:10.3934/mbe.2006.3.389

[32] Paquin, D.; Levy, D.; Xing, L. Hybrid multiscale landmark and deformable image registration. *Math. Biosci. Eng.* **4** (2007), no. 4, 711–737. doi:10.3934/mbe.2007.4.711

[33] Phuc, N. C.; Torres, M. Characterizations of the existence and removable singularities of divergence-measure vector fields. *Indiana Univ. Math. J.* **57** (2008), no. 4, 1573–1597. doi:10.1512/iumj.2008.57.3312

[34] Rudin, L.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Physica D* **60** (1992), no. 1-4, 259–268. doi:10.1016/0167-2789(92)90242-F

[35] Tadmor, E.; Nezzar, S.; Vese, L. A multiscale image representation using hierarchical (BV, L^2) decompositions. *Multiscale Model. Simul.* **2** (2004), no. 4, 554–579 (electronic). doi:10.1137/030600448

[36] Tadmor, E.; Nezzar, S.; Vese, L. Multiscale hierarchical decomposition of images with applications to deblurring, denoising and segmentation. *Commun. Math. Sci.* **6** (2008), no. 2, 281–307.

[37] Tadmor, E.; Tan, C. Hierarchical construction of bounded solutions of div $U = F$ in critical regularity spaces. *Nonlinear partial differential equations*, 255–269. Abel Symposia, 7. Springer, Berlin-Heidelberg, 2012. doi:10.1007/978-3-642-25361-4_14

[38] Tang, L.; He, C. Multiscale texture extraction with hierarchical (BV, G_p, L^2) decomposition. *J. Math. Imaging Vision* **45** (2013), no. 2, 148–163. doi:10.1007/s10851-012-0351-1

[39] Tartar, L. Remarks on some interpolation spaces. *Boundary value problems for partial differential equations and applications*, 229–252. RMA: Research Notes in Applied Mathematics, 29. Masson, Paris, 1993.

[40] Tikhonov, A. N.; Arsenin, V. Y. *Solutions of ill-posed problems.* Scripta Series in Mathematics. V. H. Winston, Washington, D.C.: Wiley, New York–Toronto–London, 1977.

[41] Vecchi, I. Gagliardo-Nirenberg imbeddings and DiPerna-Majda sequences. *Proc. Roy. Soc. London Ser. A* **439** (1992), no. 1906, 337–340. doi:10.1098/rspa.1992.0153

[42] Vese, L. A.; Osher, S. J. Image denoising and decomposition with total variation minimization and oscillatory functions. *J. Math. Imaging Vision* **20** (2004), no. 1-2, 7–18. doi:10.1023/B:JMIV.0000011316.54027.6a

[43] Weyl, H. The Princeton University Bicentennial Conference on the Problems of Mathematics. *A century of mathematics in America. Part II*, 309–333. History of Mathematics, 2. American Mathematical Society, Providence, R.I., 1989.

[44] Wojtaszczyk, P. *Banach spaces for analysts.* Cambridge Studies in Advanced Mathematics, 25. Cambridge University Press, Cambridge, 1991. doi:10.1017/CBO9780511608735

[45] Wojtaszczyk, P. Private communication, 2010.

[46] Yosida, K. *Functional analysis.* Sixth edition. Grundlehren der mathematischen Wissenschaften, 123. Springer, Berlin–New York, 1980.

EITAN TADMOR
Department of Mathematics
Center of Scientific Computation
and Mathematical Modeling
and
Institute for Physical Science
and Technology
University of Maryland
College Park, MD 20742
E-mail: tadmor@cscamm.umd.edu

Received July 2014.
Revised October 2014.