# REGULARITY VIA DUALITY IN CALCULUS OF VARIATIONS AND DEGENERATE ELLIPTIC PDES

FILIPPO SANTAMBROGIO

Abstract. A technique based on duality to obtain $H^1$ or other Sobolev regularity results for solutions of convex variational problems is presented. This technique, first developed in order to study the regularity of the pressure in the variational formulation of the Incompressible Euler equation, has been recently re-employed in Mean Field Games. Here, it is shown how to apply it to classical problems in relation with degenerate elliptic PDEs of $p$-Laplace type. This allows to recover many classical results via a different point of view, and to have inspiration for new ones. The applications include, among others, variational models for traffic congestion and more general minimization problems under divergence constraints, but the most interesting results are obtained in dynamical problems such as Mean Field Games with density constraints or density penalizations.

**Keywords:** $p$-Laplacian; traffic congestion; convex duality; fractional Sobolev spaces.

## 1. Introduction

This paper presents a technique to obtain Sobolev regularity results for solutions of convex variational problems (or of the corresponding Euler-Lagrange equations) using their optimality and, more precisely, using duality techniques. The main object of the paper are pairs of optimization problems in duality, such as

$$\min\left\{\int H(v)\ :\ \nabla\cdot v=f\right\}\qquad\text{and}\qquad\min\left\{\int fu+\int H^*(\nabla u)\right\}.$$

These two problems have been written as two minimization problems (while usually the dual of a min is a max), just for simplicity, and the sum of the two minimal values is 0. The Euler-Lagrange equation of the second problem is

$$\nabla\cdot(\nabla H^*(\nabla u))=f,$$

and the optimal $v$ in the first problem is given by $v=\nabla H^*(\nabla u)$ (when $H^*\in C^1$, of course).

The method that we will analyze strongly recalls the Nirenberg's method of incremental ratios, where $H^1$ estimates on some quantities $V=G(\nabla u)$ are obtained by proving $L^2$ estimates on the increment $V_\delta-V$ (where $V_\delta:=V(\cdot+\delta)$). The standard procedure is to test both the equation satisfied by $u$ and that satisfied by $u_\delta$ against $u_\delta-u$. A crucial tool is to dispose of a function $G:\mathbb{R}^d\to\mathbb{R}^d$ such that the following inequality is satisfied for all $w_0,w_1\in\mathbb{R}^d$:

$$(\nabla H^*(w_0)-\nabla H^*(w_1))\cdot(w_0-w_1)\geq c|G(w_0)-G(w_1)|^2. \tag{1.1}$$

In these notes the idea is similar, but not based on the equation. We will "test" the optimality of $u$ in the second problem using the fact that $v=\nabla H^*(\nabla u)$ must be optimal in the first one, the fact that the sum of the two minimal values is 0, and using a translation $u_\delta$ in the second problem, estimating how much $u_\delta$ is not optimal. The key observation is that, if

$$h \mapsto g(\delta) := \int f u_\delta + \int H^*(\nabla u_\delta)$$

is smooth, then, by optimality of $u$, we must necessarily have $g(\delta) - g(0) \leq C|\delta|^2$. This means that the smoothness of $g$ automatically gives information on how much $u_\delta$ is not optimal. The structural assumption on $H$ will now based on a different inequality than (1.1), and more precisely on

$$H(v) + H^*(w) \geq v \cdot w + c|F(v) - G(w)|^2. \tag{1.2}$$

Note that (1.1) is a quantified improvement of the monotonicity of $\nabla H$, while (1.2) is a quantified improvement of the generalized Young inequality $H(v) + H^*(w) \geq v \cdot w$ (the two inequalities are in some sense one the integrated version of the other, which explains why they look often similar, at least in homogeneous cases). However, writing (1.2) does not require differentiability of $H^*$.

Most of the computations will be similar to what is usually done in elliptic PDEs by using (1.1) and incremental ratios, but the point of view is slightly different. In particular, we do not really need to write and use the PDE, which could allow to deal with more degenerate or singular problems where it would not be clear if (or in which sense) the solutions solves a PDE.

This is indeed the framework where this technique arose, and more precisely in time-dependent problems coming from fluid mechanics. In [14], Brenier introduced a convex relaxed variational formulation to deal with Arnold's interpretation of the incompressible Euler equation for inviscid fluids as a geodesic equation in the group of measure-preserving diffeomorphisms. An important question in such a model was the regularity of the pressure field $p$ that appeared as a Lagrange multiplier and solved a dual problem. By some very refined estimates, using comparison between smooth and almost-optimal pressures in the dual problem, and optimal incompressible evolutions in the primal, Brenier managed to prove a bound on $\int_0^1 \int_\Omega |\nabla p|$, thus obtaining BV (in space) regularity for $p$ (an exponent 2 is lost in some estimates, which explains why we do not finally obtain $H^1$). One of the key ingredient of the estimate was the quantified almost-optimality of translations; another one was an inequality of the form (1.2) applied to a quadratic term appearing in the kinetic energy. The result was later improved by Ambrosio and Figalli, [2] who obtained $L^2$ in time estimates on the BV norm, but the stategy was essentially the same (the result confirmed indeed a first conjecture formally evoked by Brenier in [14]). Using $p \in L^2([0,1]; BV(\Omega))$, which implies in particular that $p$ is a function (and not only a measure, or a more singular distribution), the same authors also managed in [3] to give a rigorous meaning to the fact that almost every particle moves following the equation $x''(t) = \nabla p(t, x(t))$, which was unattainable before.

We can dare to say that the fact that the techniques developed in [14, 2] were actually a very general approach to the question of regularity was not clear at that time, and indeed they have not been re-employed for long. They re-apparead in the framework of Mean Field Games (MFG), a theory introduced by Lasri and Lions where the goal is to find equilibrium configurations in the movement of a continuum of players, whose choices are influenced by the density of the other players. These problems admit a fluid mechanics formulation which is variational in some interesting cases, and amounts to a convex minimization whose dual plays an important role in the theory. When the model prescribes a density constraint (see [17]) the situation is somewhat similar to that of incompressible fluid mechanics and indeed requires similar techniques. The success of the duality method for the regularity in the case of MFG is due to the reason we evoked above. Indeed, the PDEs which describe the optimal configuration, which is also an equilibrium in a suitable sense, are a combination of a Hamilton-Jacobi solved in an a.e. sense and a transport equation solved in distributional sense, and are difficult to exploit; on the contrary, the minimization problem and its dual are well-known. Moreover, it was important, in order to give a rigorous meaning to the equilibrium condition, to improve the regularity of the pressure $p$ and for this the same kind of results as in [3] were necessary: this means that a very small improvement in the regularity was seeked, and that the one provided by this method was enough (with no need for a bootstrap argument proving higher differentiability, for instance).

Later, it has been investigated whether duality methods can give interesting results in simpler Mean Field Games, where for instance density constraints are replaced by density penalizations (see [26]). Finally, these notes present the applications of this method to more classical variational problems and PDEs, in particular the ($p$-)Laplace equation, or other very degenerate PDEs arising from traffic models (as in [5, 10, 11, 18]).

Before entering in details about the structure and the content of these notes, it is useful to compare the notions we use to more classical notion about Bregman distances. First, let us recall the definition of Bregman distance associated to a (smooth and strictly convex, for simplicity) convex function $H$: we define

$$d_H(x,y) := H(x) - (H(y) + \nabla H(y) \cdot (x-y)).$$

This value (asymmetric in $x$ and $y$) is strictly positive unless $x = y$, and comparable to $|x-y|^2$ if the Hessian of $H$ is bounded from below and above (just by using a simple Taylor expansion). It measures a sort of distance between $x$ and $y$, more adapted than other quantities in problems involving $H$. It is useful to observe that we have

$$H(v) + H^*(w) - v \cdot w = d_H(v, \nabla H^*(w)) = d_{H^*}(\nabla H(v), w),$$

which means that (1.2) is an assumption on lower bounds of the Bregman distance $d_H$. Moreover, it is interesting to compare the strategy and the content of these notes with a recent paper by Burger, [16], which, among other results, proves and uses a statement that roughly sounds as "consider a convex minimization problem involving a parameter $f$ and a convex cost $H$; then, given two different values of the parameter, say $f_0$ and $f_1$, we compare the two corresponding optimal solutions $x_0$ and $x_1$, and we can estimante $d_H(x_0, x_1)$ in terms of some distances between $f_0$ and $f_1$". If we consider the minimization problem $\min \left\{ \int fu + \int H^*(\nabla u) \right\}$ on a torus, the only parameter is the function $f$ and it is clear that translating $f$ into $f_\delta$, has the effect of translating in the same way the solution $u$ into $u_\delta$. An estimate of $||u_\delta - u||$ in terms of $||f_\delta - f||$ (for suitable norms or distances) exactly means deducing the regularity of $u$ from that of $f$!

Moreover, there is another advantage in the procedure presented in [16] (which, on the contrary, does not explicitly present applications to regularity issues in variational problems): boundary conditions could be part of the parameter, and hence could be modified. Indeed, in the study presented in these notes one should always use $u_\delta$ as a competitor in the same minimization problem as $u$, which means that cut-off functions have to be used to avoid modifying the boundary data. Thus, results are local when a boundary is actually present. In time-dependent problems such as in MFG, the initial datum $\rho_0$ is prescribed, and most results do not extend to $t = 0$ (even extending them up to $t = T$ is not trivial, but it is done in some particular cases in [26]); on the contrary, one could think that, translating in space the initial datum $\rho_0$ and obtaining a translation of the original solution, could give, under some assumptions on $\rho_0$ and using the technique of [16], interesting regularity results up to $t = 0$.

The present paper wants to introduce the reader to this technique for regularity based on convex duality. It is based on some short lecture notes for master students, and keeps some features of pedagogical notes rathen than a research article. Indeed, most of the results it presents are not new, and only the point of view is slightly different to what usually considered. After this introduction, Section 2 sets the basis of convex duality in the formulation we need, and provides the corresponding proofs. Proofs are detailed, as they follow a path which is not always considered as standard, and is based on a technique developed for instance in [8]. The duality result is presented for both Neumann and Dirichlet boundary conditions. Then, Section 3 is the core of the paper, as it introduces the main strategy, details the role of the inequality (1.2), provides examples of functions $H$ which satisfy it, and gives application to several classical results in degenerate PDEs. In particular, at the end of Section 3 two more advanced results are presented: they are matter of recent studies, where non-trivial estimates would be necessary to handle cut-off and lower order terms; on the contrary, global estimates on the torus can be easily obtained in some lines. As all the results of this section, they are not specifically related to duality methods and could be obtained in other ways; however, we believe that some of them are easier to "see" via duality methods. In particular this is the case for new estimate about the $p$-Laplacian, which was truly inspired by the duality approach: we can say that without this approach it would have not been easy to remark that such an estimate was possible, even if later [13] has presented full details with local estimates via a different method. Section 4 presents two more complicated variants of the theory, in order to handle bounded domains (and prove interior regularity) or variable coefficients: the goal is to convince the reader that this theory can be adapted to these more difficult frameworks. Finally, Section 5 comes back to the original motivation, i.e. time-dependent problems, and gives a brief and informal presentation of the role of duality in the regularity for MFG and incompressible fluid mechanics.

## 2. Convex duality

Given a Banach space $X$ and a function $H : X \to \mathbb{R} \cup \{+\infty\}$, let us recall the definition of $H^*$, Legendre transform of $H$, defined on the dual space $X'$

$$H^*(w) = \sup_{v \in X} \langle w, v\rangle - H(v).$$

We will use reflexive spaces for simplicity; in this case it is clear that $H^{**}$ is also defined on $X$. We recall the important result stating that, if $H : X \to \mathbb{R}$ is convex and l.s.c., then $H^{**} = H$. For the main facts about conjugate convex functions, see for instance [15, 19].

Finally, we also recall the formula for the Legendre transform of $H(v) = \frac{1}{q}|v|^q$, defined on $\mathbb{R}^d$, where we get $H^*(w) = \frac{1}{p}|w|^p$, where $p$ and $q$ are conjugate exponents, i.e $q = p' = p/(p-1)$, characterized by $\frac{1}{p} + \frac{1}{q} = 1$.

Now, consider a bounded connected domain $\Omega$ which could be either a subset of $\mathbb{R}^d$ or the torus $\mathbb{T}^d$. Let us also consider a function $H : \Omega \times \mathbb{R}^d \to \mathbb{R}$ which is convex in the second variable

$$\text{(Hyp1)} \qquad v \mapsto H(x, v) \text{ is convex for every } x$$

and satisfying the following uniform bounds:

$$\text{(Hyp2)} \quad \frac{c_0}{q}|v|^q - h_0(x) \le H(x, v) \le \frac{c_1}{q}|v|^q + h_1(x),$$

where $h_0, h_1$ are $L^1$ functions on $\Omega$, $c_0, c_1 > 0$ are given finite constants, and $q \in (1, +\infty)$ is a given exponent. For functions of this form, when we write $H^*(x, w)$ we mean the Legendre transform in the second variable, i.e. $H^*(x, w) = \sup_v w \cdot v - H(x, v)$. It is useful to see that Assumption (Hyp2), which gives a growth of order $q$ on $H$, implies the same, but of order $p = q' = p/(p-1)$, as far as $H^*$ is intended.

In these notes, we will consider calculus of variations problem of the form

$$\min \left\{ \int_\Omega H(x, v(x)) dx \; : \; \nabla \cdot v = f \right\}. \tag{2.1}$$

Before giving rigorous results, we want to show how to build the dual problem of (2.1), with an informal derivation. This can be done in the following way: the constraint $\nabla \cdot v = f$ can be written, in weak form, as $-\int v \cdot \nabla u = \int fu$ for every $u$ (let us be sloppy about the regularity of the test functions now). This means that we can rewrite the above problem in the min-max form

$$\min \left\{ \int H(x, v) + \sup_u - \int fu - \int v \cdot \nabla u \right\}.$$

Indeed, the last sup is 0 if the constraint is satisfied and $+\infty$ if not. Now, we have a min-max problem and the dual problem can be obtained just by inverting inf and sup. In this case we get

$$\sup \left\{ - \int fu + \inf_v \int H(x, v) - \int v \cdot \nabla u \right\}.$$

Since $\inf_v \int H(x, v) - \int v \cdot \nabla u = -\sup_v \int \nabla u \cdot v - \int H(x, v) = \int H^*(x, \nabla u)$, the problem becomes

$$\sup \left\{ - \int fu - \int H^*(x, \nabla u) \right\}. \tag{2.2}$$

In the following, we will see precise statements about the duality between the two problems, and also provide a variant for the case of Dirichlet conditions. The duality proof, based on the above convex analysis tools, is essentially inspired to the method used in [8] . Other proofs are obviously possible, using for instance Flenchel-Rockafellar's duality result (see Chapter 1 in [15]).

Before going on with duality, we want to provide a statement which guarantees that problems such as (2.1) have a finite value (i.e., under which condition on $f$ there exists at least an admissible $v$ giving a finite value in (2.1)). The reader may note how this proof itself is derived from a sort of duality between (2.1) and (2.2).

We consider the space $W^{1,p}(\Omega)$, with its dual $(W^{1,p})'$, and the space $(W^{1,p})'_\diamond(\Omega) \subset (W^{1,p})'(\Omega)$ composed by those $f$ such that $\langle f, 1 \rangle = 0$ (i.e. those $f$ with zero mean). $(W^{1,p})'_\diamond(\Omega)$ will be endowed with the same dual norm as $(W^{1,p})'(\Omega)$. In case of no ambiguity, we will omit the domain $\Omega$ and just write $(W^{1,p})'_\diamond$ and $(W^{1,p})'$.

Note that for every $v \in L^q(\Omega; \mathbb{R}^d)$, the distribution $\nabla \cdot v$, defined through

$$\langle \nabla \cdot v, \phi \rangle := - \int_\Omega v \cdot \nabla \phi$$

naturally belongs to $(W^{1,p})'_\diamond(\Omega)$. This will be by the way the definition that we will use of the divergence operator (in weak form), and it includes a natural Neumann boundary condition on $\partial\Omega$. However, consider that we will often use $\Omega$ to be the torus, which gets rid of many boundary issues.

**Lemma 2.1.** *Given $f \in (W^{1,p})'_\diamond(\Omega)$ there exists $v \in L^q(\Omega;\mathbb{R}^d)$ such that $f = \nabla \cdot v$ and $||v||_{L^q} \leq C||f||_{(W^{1,p})'}$, where $C$ is a universal constant (depending on $\Omega, p$ and $d$, but not on $f$).*

*Proof.* Consider the classical minimization problem

$$\min \left\{ \frac{1}{p} \int_\Omega |\nabla \phi|^p dx + \langle f, \phi \rangle \ : \ \phi \in W^{1,p}(\Omega) \right\}.$$

It is easy to prove that this problem admits a solution, as the minimization can be restricted to the set of functions in $W^{1,p}$ with zero mean (and apply a Poincaré-Wirtinger inequality to prove a bound on minimizing sequences). This solution $\phi$ satisfies

$$-\int_\Omega (\nabla \phi)^{p-1} \cdot \nabla \psi = \langle f, \psi \rangle$$

for all $\psi \in W^{1,p}(\Omega)$ (pay attention to the notation: for every vector $v$ we denote by $w^\alpha$ the vector with modulus equal to $|w|^\alpha$, and same direction as $w$, i.e. $w^\alpha := |w|^{\alpha-1}w$). This exactly means $\nabla \cdot v = f$ for $v = (\nabla\phi)^{p-1}$. Moreover, testing against $\phi$, we get

$$\begin{aligned} ||v||^q_{L^q} = \int_\Omega |v|^q = \int_\Omega |\nabla \phi|^p = -\langle f, \phi \rangle \leq ||f||_{(W^{1,p})'} ||\phi||_{W^{1,p}} \\ \leq C||f||_{(W^{1,p})'} ||\nabla \phi||_{L^p} = C||f||_{(W^{1,p})'} ||v||^{q-1}_{L^q}, \end{aligned}$$

which gives the desired bound on $||v||_{L^q}$. □

## 2.1. **Neumann boundary conditions.**

We will prove the following duality result.

**Theorem 2.2.** *Suppose that $\Omega$ is smooth enough and that $H$ satisfies* Hyp1 *and* Hyp2*. Then, for any $f \in (W^{1,p})'_\diamond(\Omega)$, we have*

$$\begin{aligned} \min \left\{ \int_\Omega H(x, v(x)) dx \ : \ v \in L^q(\Omega;\mathbb{R}^d), \nabla \cdot v = f \right\} \\ = \max \left\{ -\int_\Omega H^*(x, \nabla u(x)) dx - \langle f, u \rangle \ : \ u \in W^{1,p}(\Omega) \right\} \end{aligned}$$

*Proof.* We will define a function $\mathcal{F} : (W^{1,p})' \to \mathbb{R}$ in the following way

$$\mathcal{F}(h) := \min \left\{ \int_\Omega H(x, v(x)) dx \ : \ v \in L^q(\Omega;\mathbb{R}^d), \nabla \cdot v = f + h \right\}.$$

Note that if $h \notin (W^{1,p})'_\diamond \subset (W^{1,p})'$, then $\mathcal{F}(h) = +\infty$, as there is no $v \in L^q$ with $\nabla \cdot v = f + h$. On the other hand, if $h \in (W^{1,p})'_\diamond$, then $\mathcal{F}(h)$ is well-defined and real-valued since $\int_\Omega H(x, v(x))dx$ is comparable to the $L^q$ norm, and we can use Lemma 2.1.

We now compute $\mathcal{F}^* : W^{1,p} \to \mathbb{R}$:

$$\begin{aligned}\mathcal{F}^*(u) &= \sup_h \langle h, u\rangle - \mathcal{F}(h)\\ &= \sup_{h,v\,:\,\nabla\cdot v = f+h} \langle h, u\rangle - \int_\Omega H(x, v(x))dx\\ &= \sup_{h,v\,:\,\nabla\cdot v = f+h} \langle h+f, u\rangle - \langle f, u\rangle - \int_\Omega H(x, v(x))dx\\ &= \sup_v -\langle f, u\rangle - \int_\Omega H(x, v(x))dx - \int (v\cdot\nabla u)dx\\ &= -\langle f, u\rangle + \int_\Omega H^*(x, -\nabla u(x))dx.\end{aligned}$$

Now we want to use the fact that $\mathcal{F}^{**}(0) = \sup -\mathcal{F}^*$. Note that $\sup -\mathcal{F}^* = +\infty$ if $f \notin (W^{1,p})'_\diamond$, as it is possible to add an arbitrary constant to $u$, without changing the gradient term, and letting the term $-\langle f, u\rangle$ tend to $-\infty$. On the other hand, if $f \in (W^{1,p})'_\diamond$, then in the above optimization problem we can impose that $u$ is of zero mean.

By taking the sup on $-u$ instead of $u$ we also have

$$\mathcal{F}^{**}(0) = \sup_u \; -\langle f, u\rangle - \int_\Omega H^*(x, \nabla u(x))dx = -\inf_u \langle f, u\rangle + \int_\Omega H^*(x, \nabla u(x))dx.$$

Finally, if we prove that $\mathcal{F}$ is convex and l.s.c., then we also have $\mathcal{F}^{**}(0) = \mathcal{F}(0)$, which proves the claim.

Convexity of $\mathcal{F}$ is easy. We just need to take $h_0, h_1 \in (W^{1,p})'_\diamond(\Omega)$ and set $h_t := (1-t)h_0 + th_1$. Let $v_0, v_1$ be optimal in the definition of $\mathcal{F}(h_0)$ and $\mathcal{F}(h_1)$, i.e. $\int H(x, v_i(x))dx = \mathcal{F}(h_i)$ and $\nabla\cdot v_i = f + h_i$. Let $v_t := (1-t)v_0 + tv_1$. Of course we have $\nabla\cdot v_t = f + h_t$ and, by convexity of $H(x,\cdot)$ we have

$$\mathcal{F}(h_t) \le \int H(x, v_t(x))dx \le (1-t)\int H(x, v_0(x))dx + t\int H(x, v_1(x))dx = (1-t)\mathcal{F}(h_0) + t\mathcal{F}(h_1),$$

and the convexity is proven.

For the semicontinuity, we take a sequence $h_n \to h$ in $(W^{1,p})'$. We can suppose that $\mathcal{F}(h_n) \le C < +\infty$ otherwise there is nothing to prove. In particular, $h_n \in (W^{1,p})'_\diamond(\Omega)$. Take the corresponding optimal vector fields $v_n \in L^q$, i.e. $\int H(x, v_n(x))dx = \mathcal{F}(h_n)$. We can extract a subsequence such that $\lim_k \mathcal{F}(h_{n_k}) = \liminf_n \mathcal{F}(h_n)$. Moreover, from the bound on $H$ we can see that the $L^q$ norm of $v_n$ is bounded in terms of the values of $\mathcal{F}(h_n)$, which are themselves bounded by assumption. Hence, up to an extra subsequence extraction, we can assume $v_{n_k} \rightharpoonup v$. Obviously we have $\nabla\cdot v = f+h$ and, by semicontinuity of the integral functional $v \mapsto \int H(x,v)dx$, we get

$$\mathcal{F}(h) \le \int H(x, v(x))dx \le \liminf_k \int H(x, v_{n_k}(x))dx = \lim_k \mathcal{F}(p_{n_k}) = \liminf_n \mathcal{F}(h_n),$$

which gives the desired result. □

The duality result that we proved will be used in the rest of these notes written in the following form

$$\min\{A(v)\} + \min\{B(u)\} = 0, \tag{2.3}$$

where $A$ is defined on $L^q(\Omega;\mathbb{R}^d)$ and $B$ on $W^{1,p}(\Omega)$ via

$$A(v) := \begin{cases} \int_\Omega H(x,v(x))dx & \text{if } \nabla\cdot v = f, \\ +\infty & \text{otherwise,} \end{cases}$$

and

$$B(u) = \int_\Omega H^*(x,\nabla u(x))dx + \langle f,u\rangle.$$

2.2. **Dirichlet boundary conditions.** We also want to provide a variant of Theorem 2.2 in the case where the values of $u$ are prescribed on $\partial\Omega$. In this case, besides the space $W^{1,p}$ and its dual $(W^{1,p})'$, we also need to consider the space $X(\partial\Omega)$ defined as those elements $\pi$ of $(W^{1,p})'$ such that $\langle \pi,u\rangle = 0$ for all $u\in W^{1,p}_0(\Omega)$ (in practice, these are the elements of $(W^{1,p})'$ which are concentrated on the boundary $\partial\Omega$).

We first note the following fact: for every $f\in (W^{1,p})'$ there exists $\pi\in X(\partial\Omega)$ such that

$$f+\pi\in (W^{1,p})'_\diamond,\ ||\pi||_{(W^{1,p})'}\le C||f||_{(W^{1,p})'}. \tag{2.4}$$

This can be done explicitly whenever $\Omega$ is smooth enough, by taking

$$\langle \pi,\phi\rangle := -\frac{\int_{\partial\Omega}\phi d\mathcal{H}^{d-1}}{\mathcal{H}^{d-1}(\partial\Omega)}\langle f,1\rangle$$

(this require the use of the trace operator, and $\mathcal{H}^{d-1}(\partial\Omega)<+\infty$). Otherwise, we can do it by using the Hahn-Banach Theorem (see for instance the first chapter in [15]) in the following way: there exists an element $\pi\in (W^{1,p})'$ with the following properties $\langle\pi,\phi\rangle = 0$ for every $\phi\in W^{1,p}_0$, $\langle \pi,1\rangle = -\langle f,1\rangle$ and $||\pi||_{(W^{1,p})'}\le |\langle f,1\rangle|/||1||_{W^{1,p}}$ (not that we only use $1\notin W^{1,p}_0$, which is true for every domain $\Omega$).

This, together with Lemma 2.1, guarantees finiteness of the minimum in the left hand-side of the following statement.

**Theorem 2.3.** *Suppose that $\Omega$ is smooth enough and that $H$ satisfies* Hyp1 *and* Hyp2*. Then, for any $f\in (W^{1,p})'(\Omega)$ and $\bar u\in W^{1,p}(\Omega)$, we have*

$$\begin{aligned} \min\left\{\int_\Omega H(x,v(x))dx + \langle\pi,\bar u\rangle \ :\ v\in L^q(\Omega;\mathbb{R}^d),\pi\in X(\partial\Omega),\nabla\cdot v = f+\pi\right\} \\ = \max\left\{-\int_\Omega H^*(x,\nabla u(x))dx - \langle f,u\rangle \ :\ u\in W^{1,p}(\Omega), u-\bar u\in W^{1,p}_0(\Omega)\right\} \end{aligned} \tag{2.5}$$

*Proof.* The proof will be very similar to that of Theorem 2.2. We define

$$\mathcal{F}(h) := \min\left\{\int_\Omega H(x,v(x))dx + \langle\pi,\bar u\rangle \ :\ v\in L^q(\Omega;\mathbb{R}^d),\pi\in X(\partial\Omega),\nabla\cdot v = f+h+\pi\right\}.$$

We now compute $\mathcal{F}^* : W^{1,p} \to \mathbb{R}$:

$$\begin{aligned}
\mathcal{F}^*(u) &= \sup_h \langle h, u\rangle - \mathcal{F}(h) \\
&= \sup_{h,v,\pi\,:\,\nabla\cdot v = f+h+\pi} \langle h, u\rangle - \int_\Omega H(x, v(x))dx - \langle \pi, \bar{u}\rangle \\
&= \sup_{h,v,\pi\,:\,\nabla\cdot v = f+h} \langle h+f+\pi, u\rangle - \langle f, u\rangle - \int_\Omega H(x, v(x))dx - \langle \pi, u+\bar{u}\rangle \\
&= \sup_{v,\pi} -\langle f, u\rangle - \int_\Omega H(x, v(x))dx - \int (v\cdot\nabla u)dx - \langle \pi, u+\bar{u}\rangle \\
&= \sup_{\pi} -\langle f, u\rangle + \int_\Omega H^*(x, -\nabla u(x))dx - \langle \pi, u+\bar{u}\rangle \\
&= \begin{cases} -\langle f, u\rangle + \int_\Omega H^*(x, -\nabla u(x))dx & \text{if } u+\bar{u} \in W^{1,p}_0(\Omega), \\ +\infty & \text{if not.} \end{cases}
\end{aligned}$$

Again, we will conclude by using $\mathcal{F}^{**}(0) = \sup -\mathcal{F}^*$ and taking the sup on $-u$ instead of $u$.

We need to prove that $\mathcal{F}$ is convex and l.s.c.

Convexity of $\mathcal{F}$ follows the same scheme as in Theorem 2.2. Take $h_0, h_1 \in (W^{1,p})'(\Omega)$ and define $h_t := (1-t)h_0 + th_1$. Let $(v_0, \pi_0)$ and $(v_1, \pi_1)$ be optimal in the definition of $\mathcal{F}(h_0)$ and $\mathcal{F}(h_1)$ and use $v_t := (1-t)v_0 + tv_1, \pi_t := (1-t)\pi_0 + t\pi_1$.

For the semicontinuity, we take a sequence $h_n \to h$ in $(W^{1,p})'$, with the corresponding optimal $(v_n, \pi_n)$. We may suppose that $h_n$ is bounded in $(W^{1,p})'$ and that $\mathcal{F}(h_n)$ is also bounded.

We observe that we have

$$\langle \pi_n, \bar{u}\rangle = -\langle f + h_n, \bar{u}\rangle - \int v_n \cdot \nabla\bar{u}\, dx \geq -C - C||v_n||_{L^q}$$

(where we used the bound on $||h_n||_{(W^{1,p})'}$) and

$$\frac{c_0}{p}||v_n||^q_{L^q} - C \leq \int H(x, v_n(x))dx.$$

This allows to give a bound on $||v_n||_{L^q}$ in terms of $\mathcal{F}(h_n)$. Once we have a bound on $v_n$, the bound on $\pi_n$ comes from the constraint $\nabla\cdot v_n = f + h_n + \pi_n$.

The proof can be completed as in Theorem 2.2. □

**Remark 2.1.** *In the case $\bar{u} = 0$ (i.e. with homogeneous Dirichlet conditions), Problem (2.5) becomes easier, as $\pi$ does not appear in the functional: in this case we just need to minimize $\int H(x, v(x))dx$ among vector fields with $\nabla\cdot v = f$ inside $\Omega$ (i.e. $\int v\cdot\nabla\phi = -\int \phi f$ for every $\phi \in W^{1,p}_0(\Omega)$).*

## 3. Regularity via duality

In this section we will use the relation (2.3) to produce Sobolev regularity results for solutions of the minimization problems $\min A$ or $\min B$.

We will start by describing the general strategy. We consider a function $H : \mathbb{R}^d \to \mathbb{R}$, and we suppose that an inequality of the following form is true

$$\text{(Hyp3)} \quad H(v) + H^*(w) \geq v\cdot w + c|F(v) - G(w)|^2$$

for some given functions $F, G : \mathbb{R}^d \to \mathbb{R}^d$. This is an improvement of the Young inequality $H(v) + H^*(w) \geq v \cdot w$ (which is just a consequence of the definition of $H^*$). Of course this is always true taking $F = G = 0$, but the interesting cases are the ones where $F$ and $G$ are non-trivial.

To simplify the computations, we will suppose that $\Omega$ is the flat $d$-dimensional torus $\mathbb{T}^d$ (and we will omit the indication of the domain). We start from the following observations, that we collect in a lemma. For the sake of the notations, we call $\hat{v}$ and $\hat{u}$ the minimizers (or some minimizers, in case there is no uniqueness) of $A$ and $B$, respectively, and we denote by $\hat{u}_\delta$ the function $\hat{u}_\delta(x) := \hat{u}(x+\delta)$. We define a function $g : \mathbb{R}^d \to \mathbb{R}$ given by[1]

$$g(\delta) := \int f(x)\hat{u}(x+\delta)dx - \int f(x)\hat{u}(x)dx.$$

**Lemma 3.1.** *Suppose $H$ satisfies* Hyp1, 2, 3 *and let $\hat{v}$ and $\hat{u}$ be optimal. Then*

(1) *$F(\hat{v}) = G(\nabla\hat{u})$.*
(2) *$c\int |G(\nabla\hat{u}_\delta) - G(\nabla\hat{u})|^2 dx \leq g(\delta)$.*
(3) *If $g(\delta) = O(|\delta|^2)$, then $G(\nabla\hat{u}) \in H^1$.*
(4) *If $g$ is $C^{1,1}$, then $g(\delta) = O(|\delta|^2)$ and $G(\nabla\hat{u}) \in H^1$.*
(5) *If $f \in W^{1,q}(\Omega)$, then $g \in C^{1,1}$ and hence $G(\nabla\hat{u}) \in H^1$*

*Proof.* First, we compute for arbitrary $v$ and $u$ admissible in the primal and dual problems (i.e. we need $\nabla \cdot v = f$), the sum $A(v) + B(u)$:

$$A(v)+B(u) = \int (H(v)+H^*(\nabla u)+fu)dx = \int (H(v)+H^*(\nabla u)-v\cdot\nabla u)dx \geq c\int |F(v)-G(\nabla u)|^2 dx.$$

If we take $v = \hat{v}$ and $u = \hat{u}$, then $A(v) = \min A$, $B(u) = \min B$ and $A(v) + B(u) = 0$. Hence, we deduce $F(\hat{v}) = G(\nabla\hat{u})$, i.e. the Part (1) in the statement.

Now, let us fix $v = \hat{v}$ but $u = \hat{u}_\delta$. We obtain

$$c\int |G(\nabla\hat{u}) - G(\nabla\hat{u}_\delta)|^2 dx = c\int |F(\hat{v}) - G(\nabla\hat{u}_\delta)|^2 dx \leq A(\hat{v}) + B(\hat{u}_\delta) = B(\hat{u}_\delta) - B(\hat{u}).$$

In computing $B(\hat{u}_\delta) - B(\hat{u})$, we see that the terms $\int H^*(\nabla\hat{u}_\delta)$ and $\int H^*(\nabla\hat{u})$ are equal, as one can see from an easy change-of-variable $x \mapsto x + \delta$. Hence,

$$B(\hat{u}_\delta) - B(\hat{u}) = \int f\hat{u}_\delta - \int f\hat{u} = g(\delta),$$

which gives part (2).

Part (3) of the statement is an easy consequence of classical characterization of Sobolev spaces. Part (4) comes from the optimality of $\hat{u}$, which means that $g(0) = 0$ and $g(\delta) \geq 0$ for all $\delta$. This implies, as soon as $g \in C^{1,1}$, $\nabla g(0) = 0$ and $g(\delta) = O(|\delta|^2)$.

For Part (5), we first differentiate $g(\delta)$, thus getting

$$\nabla g(\delta) = \int f(x)\nabla\hat{u}(x+\delta)dx.$$

Note that this provides the first-order optimalty condition for $\hat{u}$: using $\delta = 0$, we get

$$0 = \nabla g(0) = \int f(x)\nabla\hat{u}(x)dx. \tag{3.1}$$

[1]From now on, $f \in (W^{1,p})'_\diamond$ will actually be function, typically in $L^q$ or even in $W^{1,q}$, and we will write terms of the form $\langle f, u\rangle$ in the form of integrals.

If we want to differentiate once more, we use the regularity assumption on $f$: we write

$$\int f(x)\nabla\hat{u}(x+\delta)dx = \int f(x-\delta)\nabla\hat{u}(x)dx$$

and then

$$D^2 g(\delta) = -\int \nabla f(x-\delta)\otimes\nabla\hat{u}(x)dx.$$

Note that $\hat{u}$ naturally belongs to $W^{1,p}$, hence the integral above. Morover, we obtain $|D^2 g| \leq ||\nabla f||_{L^q}||\nabla\hat{u}||_{L^p}$, and $g \in C^{1,1}$. □

Unfortunately, the last assumption ($f \in W^{1,q}$) is quite restrictive, but we want to provide a case where it is reasonable to use it. Before this, let us find interesting cases of functions $H$ and $H^*$ for which we can provide non-trivial functions $F$ and $G$.

3.1. **Pointwise vector inequalities.** The first interesting case is the quadratic case. Take $H(v) = \frac{1}{2}|v|^2$ with $H^*(w) = \frac{1}{2}|w|^2$. In this case we have easily

$$H(v) + H^*(w) = \frac{1}{2}|v|^2 + \frac{1}{2}|w|^2 = v\cdot w + \frac{1}{2}|v-w|^2.$$

Hence, one can take $F(v) = v$ and $G(w) = w$.

Then, we pass to another interesting case, the case of other powers. Take $H(v) = \frac{1}{q}|v|^q$ with $H^*(w) = \frac{1}{p}|w|^p$. We claim that in this case we can take $F(v) = v^{q/2}$ and $G(w) = w^{p/2}$ (remember the notation for powers of vectors).

**Lemma 3.2.** *For any $v, w \in \mathbb{R}^d$ we have*

$$\frac{1}{q}|v|^q + \frac{1}{p}|w|^p \geq v\cdot w + \frac{1}{2\max\{p,q\}}|v^{q/2} - w^{p/2}|^2.$$

*Proof.* First we write $a = v^{q/2}$ and $b = w^{p/2}$ and we express the inequality in terms of $a, b$. Hence we try to prove $\frac{1}{q}|a|^2 + \frac{1}{p}|b|^2 \geq a^{2/q}\cdot b^{2/p} + \frac{1}{2\max\{p,q\}}|a-b|^2$. In this way the inequality is homogeneous, as it is of order 2 in all its terms (remember $1/p + 1/q = 1$). Then we notice that we can also write the expression in terms of $|a|, |b|$ and $\cos\theta$, where $\theta$ is the angle between $a$ and $b$ (which is the same as the one between $v = a^{2/q}$ and $w = b^{2/p}$). Hence, we want to prove

$$\frac{1}{q}|a|^2 + \frac{1}{q}|b|^2 \geq \cos\theta\left(|a|^{2/q}|b|^{2/p} - \frac{1}{\max\{p,q\}}|a||b|\right) + \frac{1}{2\max\{p,q\}}(|a|^2+|b|^2).$$

Since this depends linearly in $\cos\theta$, it is enough to prove the inequality in the two limit cases $\cos\theta = \pm 1$. For simplicity, due to the simmetry in $p$ and $q$ of the claim, we suppose $p \geq 2 \geq q$. We start from the case $\cos\theta = 1$, i.e. $b = ta$, with $t \geq 0$ (the case $a = 0$ is trivial). In this case the l.h.s. of the inequality becomes

$$|a|^2(\frac{1}{q}+\frac{1}{p}t^2) = |a|^2(\frac{1}{q}+\frac{1}{p}(1+(t-1))^2) = |a|^2(1+\frac{2}{p}(t-1)+\frac{1}{p}(t-1)^2) \geq |a|^2(t^{2/p}+\frac{1}{p}(t-1)^2),$$

where we used the concavity of $t \mapsto t^{2/p}$, which provides $1 + \frac{2}{p}(t-1) \geq t^{2/p}$. This inequality is even stronger than the one we wanted to prove, as we get a factor $1/p$ instead of $1/(2p)$ in the r.h.s..

The factor $1/(2p)$ appears in the case $\cos\theta = -1$, i.e. $b = -ta$, $t \geq 0$ (we do not claim that this coefficient is optimal, anyway). In this case we start from the r.h.s.

$$|a|^2(\frac{1}{2p}(1+t)^2 - t^{2/p}) \leq |a|^2\frac{1}{2p}(1+t)^2 \leq |a|^2\frac{2}{2p}(1+t^2) \leq |a|^2(\frac{1}{q}+\frac{1}{p}t^2),$$

which gives the claim. □

**Remark 3.1.** *The above inequality replaces, in this duality-based approach, the usual vector inequality that PDE methods require to handle equations involving $\Delta_p$, i.e.*

$$(w_0^{p-1} - w_1^{p-1})\cdot(w_0 - w_1) \geq c|w_0^{p/2} - w_1^{p/2}|^2, \tag{3.2}$$

*which is an improved version of the monotonicity of the gradient of $w \mapsto \frac{1}{p}|w|^p$. Note that the proof of Lemma 3.2 is quite short, if compared to some classical proofs of* (3.2) *(see for instance* [25]*). However, alternative proofs for* (3.2) *are also possible, as in the appendix of* [12].

We finish with a general consideration, the case where $H$ is uniformly convex, in the sense that $D^2H$ is uniformly bounded from below (which also implies $H^* \in C^1$). In this case we have

**Lemma 3.3.** *If $D^2H \geq \lambda I$ for $\lambda > 0$, we have*

$$H(v) + H^*(w) \geq v\cdot w + \frac{\lambda}{2}|v - \nabla H^*(w)|^2.$$

*Proof.* Just consider

$$H^*(w) = \max_v \; v\cdot w - H(v) = \nabla H^*(w)\cdot w - H(\nabla H^*(w)), \tag{3.3}$$

where we used the fact that the optimal $v$ is characterized by $w = \nabla H(v)$, which can be turned into $\nabla H^*(w) = v$. Then, we use

$$H(v) \geq H(v_0) + \nabla H(v_0)\cdot(v - v_0) + \frac{\lambda}{2}|v - v_0|^2, \tag{3.4}$$

which is just a consequence of second-order Taylor expansion, for any $v_0$. Choosing $v_0 = \nabla H^*(w)$, using $\nabla H(v_0) = \nabla H(\nabla H^*(w)) = w$, and summing up (3.3) and (3.4) we get

$$H(v)+H^*(w) \geq \nabla H^*(w)\cdot w - H(\nabla H^*(w)) + H(\nabla H^*(w)) + w\cdot(v - \nabla H^*(w)) + \frac{\lambda}{2}|v - \nabla H^*(w)|^2,$$

which is the claim. □

Lemma 3.3 shows that Hyp(3) is typical of uniformly convex functions, while Lemma 3.2 shows that some variants exist form more degenerate convex functions.

3.2. **Very degenerate PDEs.** Consider for instance the case $H(v) = |v| + \frac{1}{q}|v|^q$. In this case, we can use $F(v) = v^{q/2}$ and $G(w) = (w-1)_+^{p/2}$ (again, we use this weird notation: the vector $(w-1)_+^{p/2}$ is the vector with norm equal to $(|w|-1)_+^{p/2}$ and same direction as $w$, i.e. $G(w) = (|w|-1)_+^{p/2}w/|w|)$. Indeed, we have

$$H^*(w) = \sup_v \; v\cdot w - |v| - \frac{1}{q}|v|^q = \frac{1}{p}(|w|-1)_+^p$$

and

$$H(v) + H^*(w) = |v| + \frac{1}{q}|v|^q + \frac{1}{p}(|w|-1)_+^p \geq |v| + v\cdot(w-1)_+ + c|v^{q/2} - (w-1)_+^{p/2}|^2.$$

We only need to prove $|v| + v \cdot (w-1)_+ \geq v \cdot w$. This can be done by writing

$$|v| + v \cdot (w-1)_+ = |v|(1 + (|w|-1)_+ \cos\theta).$$

If $|w| \geq 1$ then we go on with

$$|v|(1 + (|w|-1)_+ \cos\theta) \geq |v|\cos\theta(1 + (|w|-1)_+) = |v|\cos\theta|w| = v \cdot w.$$

If $|w| \leq 1$ then we simply use

$$|v|(1 + (|w|-1)_+ \cos\theta) = |v| \geq v \cdot w.$$

As a consequence we get the following result[2]:

**Proposition 3.4.** *Let $H$ be given by $H(v) = |v| + \frac{1}{q}|v|^q$ and $H^*(w) = \frac{1}{p}(|w|-1)_+^p$. Suppose that $\Omega$ is the flat torus and $f \in W^{1,q}(\Omega)$. Let $\hat{v}$ is a solution of $\min A$ and $\hat{u}$ a solution of $\min B$ (equivalently, suppose that $\hat{u}$ solves $\nabla \cdot ((\nabla\hat{u} - 1)_+^{p-1}) = f$). Then $\hat{v}^{q/2} = (\nabla\hat{u} - 1)_+^{p/2} \in H^1$.*

This result is the same proven in [11], where it was approached with PDE methods. The equation $\nabla \cdot ((\nabla u - 1)_+^{p-1}) = f$, which can be written,

$$\nabla \cdot \left( (|\nabla u| - 1)_+^{p-1} \frac{\nabla u}{|\nabla u|} \right) = f,$$

is very degenerate in the sense that the coefficient $\frac{(|\nabla u|-1)_+^{p-1}}{|\nabla u|}$ vanishes on the whole set where $|\nabla u| \leq 1$.

This equation and these minimization problems arise in traffic congestion (see [5, 18, 11]) and the choice of the function $H$ is very natural: we need a superlinear function of the form $H(v) = |v|h(|v|)$, with $h \geq 1$). This automatically implies the degeneracy of $H^*$.

### 3.3. The Laplacian case: $\Delta u = f$.

The case of the Poisson equation $\Delta u = f$, corresponding to the minimization of $\int \frac{1}{2}|\nabla u|^2 + fu$, and hence to $H(v) = \frac{1}{2}|v|^2$ and $H^*(w) = \frac{1}{2}|w|^2$, deserves special attention. It is possible to treat this case by the same techniques as in the degenerate case above, but the result is disappointing. Indeed, from these techniques we just obtain $f \in H^1 \Rightarrow \nabla u \in H^1$, while it is well-known that $f \in L^2$ should be enough for the same result. However, with some more attention it is also possible to treat the $L^2$ case.

**Proposition 3.5.** *Suppose that $\Omega$ is the flat torus and $\Delta u = f \in L^2(\Omega)$. Then $\nabla u \in H^1$.*

*Proof.* We use the variational framework we presented before, with $H(v) = \frac{1}{2}|v|^2$. We have

$$\frac{1}{2}||\nabla u_\delta - \nabla u||_{L^2}^2 \leq g(\delta). \tag{3.5}$$

Now, set $\omega_t := \sup\{||\nabla u_\delta - \nabla u||_{L^2} \ : \ |\delta| \leq t\}$. From (3.5) we have

$$\omega_t^2 \leq \sup_{\delta:|\delta|\leq t} 2g(\delta) \leq 2t \sup_{\delta:|\delta|\leq t} |\nabla g(\delta)|.$$

From $\nabla g(\delta) = \nabla g(\delta) - \nabla g(0) = \int f(\nabla u_\delta - \nabla u)$ we deduce $|\nabla g(\delta)| \leq ||f||_{L^2}||\nabla u_\delta - \nabla u||_{L^2} \leq ||f||_{L^2}\omega_t$, hence $\omega_t^2 \leq 2t||f||_{L^2}\omega_t$, which implies $\omega_t \leq 2t||f||_{L^2}$ and hence $\nabla u \in H^1$. □

[2] For the rest of Section 3 we will stop using $\hat{v}$ and $\hat{u}$ to denote the optimal solutions of the primal and dual problems, as we do not need to explicitly compare them to arbitrary competitors $v$ and $u$.

It is important to note that, in the framework of duality-based regularity, the main tool to prove $\nabla u \in H^1$ would be the $C^{1,1}$ behavior of $\delta \mapsto g(\delta)$; for this to be true, the natural assumption would be $f \in W^{1,q}$ (when $H^*$ has growth of exponent $p$, which implies $u \in W^{1,p}$). This means that the result of proposition 3.5 is in some sense more difficult, as it requires a step further, if one wants $\nabla u \in H^1$ using only $f \in L^2$. Compare this to standard PDE methods: the typical strategy requires integrating by parts a term such as $-\int \nabla f \cdot \nabla u$, getting $\int f \Delta u$, which is almost automatical, and then estimating via Cauchy-Schwartz inequalities and absorbing the second order terms in the left hand side. We can infer a peculiarity of the duality-based method presented in these notes: if on the one hand it gives satisfactory results even for very degenerate cost functions (here, $H^*$) once the data (here, $f$) are smooth enough, on the other hand improving the results for non-smooth data is harder even in the case where $H^*$ is purely quadratic. This means that in this last case ($f$ non smooth and $H^*$ quadratic) one needs an extra non-trivial step; yet, once we learn how to do it, we can replicate it in non-quadratic cases. This is what will be done in the next sections of the paper, and can give unexpected results (which are definitely provable without duality, but not easy to guess a priori).

3.4. **The $p-$Laplacian case: $\Delta_p u = f$.** If we look at the case $H(v) = \frac{1}{q}|v|^q$, we have $H^*(w) = \frac{1}{p}|w|^p$ and the solutions of $\Delta_p u = f$ (where $\Delta_p u := \nabla \cdot ((\nabla u)^{p-1})$) are the minimizers of $\int \frac{1}{p}|\nabla u|^p + fu$. Classical references on the $p$-Laplacian regularity question are, for instance, [7, 25, 21, 24].

From the consideration of the previous sections we easily obtain the following.

**Proposition 3.6.** *Suppose that $\Omega$ is the flat torus and $\Delta_p u = f \in W^{1,q}(\Omega)$. Then $(\nabla u)^{p/2} \in H^1$.*

This result is quite classical (in the case $f = 0$ it dates back to Uhlenbeck, [30], for the general case one can look at [25] and for other generalizations to [24]). Yet, it is not very satisfactory, since if we set $p = q = 2$ we get the result $\Delta u \in H^1 \Rightarrow \nabla u \in H^1$ which, as we said, is very disappointing.

This is why we also look at the following other classical result. We recall, before stating it, the clasical notion of fractional Sobolev spaces (see, for instance, [1]), with a simplified definition which works in the case of the flat torus.

**Definition 3.1.** *When $\Omega = \mathbb{T}^d$, fix $R > 0$, $1 < p < +\infty$ and $0 < s < 1$; then, the space $W^{s,p}(\Omega)$ is defined as*

$$W^{s,p}(\Omega) = \left\{ u \in L^p(\Omega) \,:\, [u]^p_{s,p} := \int_{B(0,R)} \frac{||u_\delta - u||^p_{L^p}}{|\delta|^{d+sp}} d\delta < +\infty \right\}$$

*and its norm is given by $||u||_{L^p} + [u]_{s,p}$. The space $H^s$ is defined as $W^{s,2}$.*

Note that an inequality of the form $||u_\delta - u||_{L^p} \leq C|\delta|^s$ implies $u \in W^{r,p}$ for every $r < s$.

We can now give a duality-based proof of another classical result (see [28]).

**Proposition 3.7.** *Suppose that $\Omega$ is the flat torus and $\Delta_p u = f \in L^q(\Omega)$, with $p > 2$. Then $||(\nabla u_\delta)^{p/2} - (\nabla u)^{p/2}||_{L^2} \leq C|\delta|^{q/2}$, which implies in particular $(\nabla u)^{p/2} \in H^s$ for $s < q/2 < 1$.*

*Proof.* We use the same strategy as in Proposition 3.5. For simplicity, we set $V := (\nabla u)^{p/2}$. As in Proposition 3.5, we set $\omega_t := \sup_{\delta:|\delta| \leq t} ||V_\delta - V||_{L^2}$. We have $||V_\delta - V||^2_{L^2} \leq Cg(\delta)$, which implies

$$\omega_t^2 \le Ct \sup_{\delta:|\delta|\le t} |\nabla g(\delta) - \nabla g(0)| \le Ct||f||_{L^q} \sup_{\delta:|\delta|\le t} ||\nabla u_\delta - \nabla u||_{L^p}.$$

From the $\alpha$-Hölder behavior of the vector map $w \mapsto w^\alpha$ in $\mathbb{R}^d$ (see Lemma 3.8 below), with $\alpha = 2/p < 1$, we deduce, using $\nabla u = V^\alpha$,

$$||\nabla u_\delta - \nabla u||_{L^p}^p = \int |\nabla u_\delta - \nabla u|^p dx \le C \int |V_\delta - V|^2 dx = C||V_\delta - V||_{L^2}^2.$$

Hence, we have

$$\omega_t^2 \le Ct||f||_{L^q} \omega_t^{2/p},$$

which implies

$$\omega_t^{2/q} \le Ct||f||_{L^q},$$

i.e. the claim. □

**Lemma 3.8.** *For $0 < \alpha < 1$, the map $w \mapsto w^\alpha$ is $\alpha$-Hölder continuous in $\mathbb{R}^d$.*

*Proof.* Let $a, b \in \mathbb{R}^d$. We write

$$|a^\alpha - b^\alpha| = \left| |a|^\alpha \frac{a}{|a|} - |a|^\alpha \frac{b}{|b|} + |a|^\alpha \frac{b}{|b|} - |b|^\alpha \frac{b}{|b|} \right| \le |a|^\alpha \left| \frac{a}{|a|} - \frac{b}{|b|} \right| + ||a|^\alpha - |b|^\alpha|.$$

For the second term in the r.h.s., we use the $\alpha$-Hölder behaviour of $t \mapsto t^\alpha$ in $\mathbb{R}_+$ and get

$$||a|^\alpha - |b|^\alpha| \le ||a| - |b||^\alpha \le |a - b|^\alpha.$$

For the first term in the r.h.s., we use the inequality

$$\left| \frac{a}{|a|} - \frac{b}{|b|} \right| = \left| \frac{a}{|a|} - \frac{b}{|a|} + \frac{b}{|a|} - \frac{b}{|b|} \right| \le \frac{|a-b|}{|a|} + |b| \frac{||b| - |a||}{|a||b|} \le 2 \frac{|a-b|}{|a|}$$

and get

$$|a|^\alpha \left| \frac{a}{|a|} - \frac{b}{|b|} \right| \le 2|a|^{\alpha-1} |a - b|.$$

If we choose $a$ to be such $|a| \ge |b|$ (which is possible w.l.o.g.), we have $2|a| \ge |a - b|$ and hence $2^{\alpha-1}|a|^{\alpha-1} \le |a - b|^{\alpha-1}$, i.e. $2|a|^{\alpha-1}|a - b| \le 2^{2-\alpha}|a - b|^\alpha$.

Summing up, we have

$$|a^\alpha - b^\alpha| \le (2^{2-\alpha} + 1)|a - b|^\alpha. \quad \square$$

**Remark 3.2.** *The constant $2^{2-\alpha} + 1$ that we found in the proof of Lemma 3.8 is not optimal: in [9] the same result is proven with constant $2^{1-\alpha}$, which is indeed optimal.*

**Remark 3.3.** *Note that the result of Proposition 3.7 is also classical, and quite sharp. Indeed, one can informally consider the following example. Take $u(x) \approx |x|^r$ as $x \approx 0$ (and then multiply times a cut-off function out of $0$). In this case we have*

$$\nabla u(x) \approx |x|^{r-1}, \quad (\nabla u(x))^{p-1} \approx |x|^{(r-1)(p-1)}, \quad f(x) := \Delta_p u(x) \approx |x|^{(r-1)(p-1)-1}.$$

*Hence, $f \in L^q$ if and only if $((r - 1)(p - 1) - 1)q > -d$, i.e. $(r - 1)p - q > -d$. On the other hand, the fractional Sobolev regularity can be observed by considering that "differentiating $s$ times" means subtracting $s$ from the exponent, hence*

$$(\nabla u(x))^{p/2} \approx |x|^{p(r-1)/2} \Rightarrow (\nabla u)^{p/2} \in H^s \Leftrightarrow |x|^{p(r-1)/2-s} \in L^2 \Leftrightarrow p(r - 1) - 2s > -d.$$

*If we want this last condition to be true for arbitrary $s < q/2$, then it amounts to $p(r-1)-q > -d$, which is the same condition as above.*

3.5. **Sharp $H^1$ regularity of $(\nabla u)^{p/2}$ in the $p$-Laplace equation $\Delta_p u = f$.** We already saw the very classical result of Proposition 3.6 which states that the solution of $\Delta_p u = f$ satisfies $(\nabla u)^{p/2} \in H^1$ as soon as $f \in W^{1,q}$. In the same spirit of the observation we did in Section 3.3, we can say that this result is far from being optimal, as it is not coherent with what we know in the limit case $p = 2$, where the sharp assumption to prove $\nabla u \in H^1$ is $f \in L^2$, and not $f \in H^1$.

Indeed, in [13] a new result is presented: the same $H^1$ regularity is true, under a fractional regularity assumption on $f$, which is supposed to be $W^{s,q}$, for $s > (p-2)/p$ (with counterexamples, of the same type of those of our Remark 3.3, for $s < (p-2)/p$). Up to the fact that the limit case $s = (p-2)/p$ is not studied in [13], this assumption is coherent with the case $p = 2$, as in this case we get exactly the classical $L^2$ assumption on $f$. The result in [13] is local, and obtained via more classical methods from elliptic PDE techniques. Yet, it can also be obtained, at least in the global case on $\Omega = \mathbb{T}^d$, via duality methods. Actually, the proof via duality method that we present here was the first to be found, but was not easy to adapt to local regularity. The considerations that we did at the end of Section 3.3 should clarify to the reader in which sense the duality method allows more easily to guess the valdity of this result.

**Proposition 3.9.** *Suppose that $\Omega$ is the flat torus and $\Delta_p u = f \in W^{s,q}(\Omega)$ for $s > (p-2)/p$ and $p > 2$. Then $(\nabla u)^{p/2} \in H^1$.*

*Proof.* Set $V := (\nabla u)^{p/2}$. Using the same strategy that the reader knows well now we can say

$$||V_\delta - V||^2_{L^2} \leq C \left( \int f u_\delta - \int f u \right).$$

We use

$$\int f u_\delta - \int f u = \int_0^1 dt \int f \nabla u(x + t\delta) \cdot \delta \, dx,$$

which can also be written as

$$\begin{aligned} \int f u_\delta - \int f u &= \int_0^1 dt \int f \left( \nabla u(x + t\delta) - \nabla u(x) \right) \cdot \delta \, dx \\ &= \int_0^1 dt \int \left( f(x) - f(x - t\delta) \right) \nabla u(x) \cdot \delta \, dx, \end{aligned}$$

thanks to $\int f \nabla u = 0$ (which is the first-order optimality condition of $u$, see (3.1)), and then

$$\int f u_\delta - \int f u = \int_0^1 dt \int_0^1 d\tau \int \nabla f(x - t\tau\delta) \cdot \delta \, \nabla u(x) \cdot \delta \, dx.$$

This allows to estimate, for arbitrary $0 < r < 1$:

$$\int f u_\delta - \int f u \leq |\delta|^2 ||\nabla u||_{W^{r,p}} ||\nabla f(\cdot - t\tau\delta)||_{(W^{r,p})'}.$$

In the estimate above we used an easy translation invariance to replace $||\nabla f(\cdot - t\tau\delta)||_{(W^{r,p})'}$ with a constant term $||\nabla f||_{(W^{r,p})'}$.

Using Lemma 3.10 below, with $g(z) = z^{2/p}$ (which is $C^{0,\alpha}$ for $\alpha = 2/p$) and $u = g \circ V$, we get

$$||\nabla u||_{W^{r,p}} \leq C ||V||^{2/p}_{H^1}$$

provided $r < 2/p$. This provides

$$\frac{||V_\delta - V||^2_{L^2}}{|\delta|^2} \leq C||G||^{2/p}_{H^1}||\nabla f||_{(W^{r,p})'}.$$

The supremum over $\delta \in \mathbb{R}^d \setminus \{0\}$ in the left hand side gives exactly $||V||^2_{H^1}$, a term which also appears on the right-hand side but with a smaller exponent. Hence the above estimate provides a uniform bound on $||V||_{H^1}$ (note that, in order to give a rigorous proof, one should first approximate so as to guarantee that $V$ is actually $H^1$ and provide an a priori estimate on its norm, which is possible by approximating $f$ with $W^{1,q}$ functions, for instance) in terms of $||\nabla f||_{(W^{r,p})'}$ We then use

$$||\nabla f||_{(W^{r,p})'} \leq C||f||_{W^{1-r,q}},$$

an estimate which sounds natural if we think of $(W^{r,p}) \approx W^{-r,q}$ (see the Appendix of [13] for a precise proof via interpolation arguments; alternative proofs can be obtained via Fourier transform in the case $p = 2$, or by solving fractional Laplace equations and using regularity estimates for the corresponding solutions) and prove the claim as soon as $f \in W^{s,q}$ for an exponent $s$ of the form $s = 1 - r$, $r < 2/p$, which is our assumption. □

**Lemma 3.10.** *Consider a map $g \in C^{0,\alpha}(\mathbb{R}^d; \mathbb{R}^d)$ and a Sobolev function $V \in W^{1,p_0}(\Omega; \mathbb{R}^d)$. Then $g \circ V \in W^{r,p_1}$ provided $\alpha p_1 = p_0$ and $r < \alpha$; moreover we have*

$$[g \circ V]_{r,p_1} \leq C||\nabla V||^\alpha_{L^{p_0}},$$

*for a constant $C$ depending on $d, s, \alpha$ and on the Hölder constant of $g$.*

*Proof.* We recall the formula

$$[u]^p_{r,p} := \int_{B(0,R)} \frac{||u_\delta - u||^p_{L^p}}{|\delta|^{d+rp}} d\delta.$$

We use

$$|(g \circ V)_\delta - g \circ V| \leq C|V_\delta - V|^\alpha,$$

which implies

$$||(g \circ V)_\delta - g \circ V||^{p_1}_{L^{p_1}} \leq C||V_\delta - V||^{\alpha p_1}_{L^{\alpha p_1}}.$$

Using $V \in W^{1,p_0}$ we also have

$$||V_\delta - V||_{L^{p_0}} \leq C|\delta|||\nabla V||_{L^{p_0}},$$

which provides, when $\alpha p_1 = p_0$,

$$[g \circ V]^{p_1}_{r,p_1} \leq C||\nabla V||^{p_0}_{L^{p_0}} \int_{B(0,R)} \frac{|\delta|^{p_0}}{|\delta|^{d+rp}} d\delta. \tag{3.6}$$

The result is proven once we note that the integral

$$\int_{B(0,R)} \frac{|\delta|^{p_0}}{|\delta|^{d+rp_1}} d\delta$$

converges as soon as $p_0 = \alpha p_1 > rp_1$ (i.e. $r < \alpha$). Taking the $p_1$-th root on both sides of (3.6) we get the desired estimate. □

3.6. **Anisotropic orthotropic equations.** Consider the case where $H^*(w) = \sum_i H_i^*(w_i)$, for some convex functions of one variable $H_i$. Note that this is the case for $H^*(w) = \frac{1}{2}|w|^2$ but not for $H^*(w) = \frac{1}{p}|w|^p$; on the contrary, $H^*(w) = \sum_i \frac{1}{p}|w_i|^p$ is of this form. PDEs and variational problems involving these functions arise, for instance, in traffic congestion, when taking limits of problems on networks: the corresponding equations are not isotropic and "remember" the shape of the network (and this precise structure is obtained in the case of a Manhattan metric, i.e. of a cartesian lattice network; see [4, 10, 20] for details about these problems). Their difficulty is due to the fact that degeneracies sum up: if $\frac{1}{p}|w|^p$ is only degenerate at $w = 0$, the Hessian of $H^*(w) = \sum_i \frac{1}{p}|w_i|^p$ is singular at every point $w$ where at least one component $w_i$ vanishes. This especially creates difficulty when studying Lipschitz bounds for the solution$u$, as the degeneracy set becomes unbounded. For Sobolev regularity this is less important, but difficulties arise when, for instance, one considers $H^*(w) = \sum_i \frac{1}{p_i}|w_i|^{p_i}$, for different exponents $p_i$.

We will consider the following framework:

$$H^*(w) = \sum_i h_i^*(w_i), \qquad \text{with each } h_i \text{ satifsying (Hyp2) for an exponent } q_i.$$

We will avoid dependence on $x$ for simplicity and suppose that for each $i$ there are two functions $f_i, g_i : \mathbb{R} \to \mathbb{R}$ such that

$$h_i(t) + h_i^*(s) \geq ts + c|f_i(t) - g_i(s)|^2.$$

Then we have the following

**Proposition 3.11.** *Let $u$ be a solution of the PDE*

$$\nabla \cdot (\nabla H^*(\nabla u)) = f$$

*on the flat torus $\Omega = \mathbb{T}^d$, and suppose that $f$ is a Sobolev function such that, for each $i$, we have $\partial_i f \in L^{q_i}$, with $q_i = p_i'$. Let $V : \Omega \to \mathbb{R}^d$ be the function defined via $V_i = g_i(\partial_i u)$.*

*Then $V \in H^1(\Omega)$.*

*Proof.* From the definitions of $V$, $H^*$ and $g_i$, applying the standard strategy of this section and the computations of Proposition 3.9, we easily get

$$||V_\delta - V||_{L^2}^2 \leq C\left(\int f\hat{u}_\delta - \int fu\right) = C\int_0^1 dt \int_0^1 d\tau \int \nabla f(x - t\tau\delta) \cdot \delta \, \nabla\hat{u}(x) \cdot \delta \, dx. \tag{3.7}$$

In order to prove $V \in H^1$ it is enough to prove $||V_\delta - V||_{L^2}^2 \leq C|\delta|^2$ for those vectors $h$ parallel to one of the coordinate axes. When selecting one such a vector, say $h = \lambda e_i$ with $|\lambda| = |\delta|$, then we can estimate the right hand side in (3.7) using the integrabilities of $\partial_i u$ and $\partial_i f$ and obtain

$$||V_\delta - V||_{L^2}^2 \leq C|\delta|^2 ||\nabla f(\cdot - t\tau\delta) \cdot e_i||_{L^{p_i'}} ||\nabla\hat{u} \cdot e_i||_{L^{p_i}},$$

which allows to conclude. $\square$

**Remark 3.4.** *In the anistropic orthotropic p-Laplacian case, i.e. $H^*(w) = \sum_i \frac{1}{p_i}|w_i|^{p_i}$ (corresponding to $H(v) = \sum_i \frac{1}{q_i}|w_i|^{q_i}$ for $q_i = p_i'$), the above Proposition easily allows to prove*

$$(\partial_i u)^{p_i/2} \in H^1(\Omega) \quad \textit{for every } i.$$

As we already said in the introduction, it is important to stress that the above result does not require any assumptions on the anisotropy, i.e. it works for arbitrary exponents $p_i$, even very far from each other. This is due to the global character of this estimate, which is much easier to get that local estimates (which are the object of [12]), and not on the approach by dualiity, which is just a quick way of presenting the result.

## 4. Non–spatially-homogeneous cases: the effects of the boundary or of space-dependance

The attentive reader has for sure observed that the analysis of the previous sections is strongly simplified by the fact that we have

$$\int H^*(\nabla \hat{u}_\delta)dx = \int H^*(\nabla \hat{u})dx, \tag{4.1}$$

which is a consequence of an easy change of variables $x \mapsto x + \delta$. Not only this is due to the absence of explicit dependance on $x$ (the equation is autonomous), but also to the fact that there are no boundary issues. Indeed, even when there is no dependence on $x$, we cannot say that all the points are the same for a PDE, or for the solution of a PDE, when some of the points are closer to the boundary than others. In practice, the use of a translation $\hat{u}_\delta$ raises some difficulty if we don't have the equality (4.1), which could be due to the presence of the boundary, or to space heterogeneity.

In this section we will see how to adapt the analysis we presented so far to handle these difficulties. First we will analyze the case of a bounded domain, with boundary, where to study local regularity, and then a case where some coefficients depend on $x$.

4.1. **Variant – Local regularity.** We provide here a result concerning local Sobolev regularity. As the result is local, boundary conditions should not be very important. Yet, we need anyway to fix a variational problem and consider its dual, which requires to choose the boundary conditions we use. We will use Dirichlet boundary conditions, but the analysis can be adapted to the case of Neumann conditions. This is, by the way, one of the reasons why we introduced duality for the Dirichlet case.

We will only provide the following result, in the easiest case $p = 2$.

**Theorem 4.1.** *Let $H, H^*, F$ and $G$ satisfy* Hyp1, 2, 3 *with $p = 2$. Suppose $f \in H^1$. Suppose also $H^* \in C^{1,1}$ and $G \in C^{0,1}$. Suppose $\nabla \cdot (\nabla H^*(\nabla \hat{u})) = f$ in $\Omega$. Then, $G(\nabla \hat{u}) \in H^1_{loc}(\Omega)$.*

*Proof.* The condition $\nabla \cdot \nabla(H^*(\nabla \hat{u})) = f$ is equivalent to the fact that $\hat{u}$ is solution of

$$\min \left\{ \int_\Omega H^*(\nabla u)dx + \int_\Omega fu\,dx \ : \ u \in W^{1,p}(\Omega), u - \bar{u} \in W^{1,p}_0(\Omega) \right\},$$

for $\bar{u} = \hat{u}$ (i.e. $\hat{u}$ minimizes under its own optimality conditions). We will also use the dual problem presented in Theorem 2.3. We set $A(v, \pi) := \int H(v) + \langle \pi, \bar{u} \rangle$ with the constraint $\nabla \cdot v = f + \pi$. As usual, we sum $A(v, \pi) + B(u)$ and we get $A(v, \pi) + B(u) = \int (H(v) + H^*(\nabla u) - v \cdot \nabla u)dx \geq c \int |F(v) - G(\nabla u)|^2 dx$.

The strategy is the same: use the optimal $v$ and $\pi$ together with a translation of $u$. Yet, in order not to have boundary problems, we need to use a cut-off function $\eta \in C^\infty_c(\Omega)$ and define

$$\hat{u}_\delta(x) = \hat{u}(x + \delta\eta(x))$$

(note that this does not change the boundary value $\bar{u}$). In this case it is no longer true that $\tilde{g}(\delta) := \int H^*(\nabla \hat{u}_\delta)dx$ equals $\int H^*(\nabla \hat{u})dx$. If this term is not constant as a function of $\delta$, then we need to prove that it is a $C^{1,1}$. To do this, and to avoid differentiating $\nabla\hat{u}$, we use a change-of-variable. Set $y = x + \delta\eta(x)$. We have $\nabla(\hat{u}_\delta)(x) = (\nabla\hat{u})(y)(I + \delta \otimes \nabla\eta(x))$, hence

$$\tilde{g}(\delta) = \int_\Omega H^*(\nabla\hat{u}_\delta)dx = \int_\Omega H^*\big(\nabla\hat{u}(y) + (\nabla\hat{u}(y)\cdot\delta)\nabla\eta(x)\big)\frac{1}{1+\delta\cdot\nabla\eta(x)}dy,$$

where $x = X(\delta, y)$ is a function of $\delta$ and $y$ obtained by inverting $x \mapsto x + \delta\eta(x)$; we also used $\det(I + \delta\otimes\nabla\eta(x)) = 1 + \delta\cdot\nabla\eta(x)$. The function $X$ is $C^\infty$ by the implicit function theorem, and all the other ingredient of the above integral are at least $C^{1,1}$ in $\delta$. This proves that $\tilde{g}$ is $C^{1,1}$. The regularity of the term $g(\delta) = \int f\hat{u}_\delta$ should also be considered. Differentiating once we get $\nabla g(\delta) = \int f(x)\nabla\hat{u}(x + \delta\eta(x))\eta(x)dx$. To differentiate once more, we use the same change-of-variable, thus getting

$$\nabla g(\delta) = \int f(X(\delta,y))\nabla\hat{u}(y)\eta(X(\delta,y))\frac{1}{1+h\cdot\nabla\eta(x)}dy.$$

From $y = X(\delta,y) + \delta\eta(X(\delta,y))$ we get a formula for $D_\delta X(\delta,y)$, i.e.

$$0 = D_\delta X(\delta,y) + \eta(X(\delta,y))I + \delta\otimes\nabla\eta(\eta(X(\delta,y))D_\delta X(\delta,y).$$

This allows to differentiate once more the function $g$ and proves $g \in C^{1,1}$.

Finally, we come back to the duality estimate. What we can easily get is

$$c||G(\nabla(\hat{u}_\delta)) - G(\nabla\hat{u})||^2_{L^2} \le g(\delta) + \tilde{g}(\delta) = O(|\delta|^2).$$

The problem is that $G(\nabla(\hat{u}_\delta))$ is not the translation of $G(\nabla\hat{u})$! Yet, it is almost true. Indeed, if we put the subscript $\delta$ every time that we compose with $x + \delta\eta(x)$, we have

$$\nabla(\hat{u}_\delta) = (\nabla\hat{u})_\delta + \delta\cdot(\nabla\hat{u})_\delta\eta.$$

Since $G$ is supposed to be Lipschitz continuous, then

$$|G(\nabla(\hat{u}_\delta)) - G((\nabla\hat{u})_\delta)| \le C|\delta||\nabla\hat{u}|_\delta\eta.$$

Hence, we have

$$||G((\nabla\hat{u})_\delta) - G(\nabla\hat{u})||_{L^2} \le ||G(\nabla(\hat{u}_\delta)) - G(\nabla\hat{u})||_{L^2} + C|\delta|||\nabla\hat{u}||_{L^2},$$

which is enough to show that this increment is of order $|\delta|$, since $\hat{u} \in H^1$ (this depends on the fact that $H^*$ is quadratic). Hence, as in Lemma 3.1 (4), we get $G(\nabla\hat{u}) \in H^1$. □

The reader can see that the above method for local regularity required some structure on the function $H$ and $H^*$: it could be adapted under suitable growth condition, but cannot easily handle the case studied in [12] (and in our Section 3.6), where the functional has anisotropic growth.

4.2. **Variant – Dependence on** $x$**.** The duality theory has been presented in the case where $H$ and $H^*$ could also depend on $x$, while for the moment regularity results have only be presented under the assumption that they not. In this section, we will see how to handle the following particular case, corresponding to the minimization problem

$$\min\left\{\frac{1}{p}\int_\Omega a(x)|\nabla u(x)|^p dx + \int_\Omega f(x)u(x)dx \ : \ u \in W^{1,p}(\Omega)\right\}. \tag{4.2}$$

We will use $\Omega = \mathbb{T}^d$ to avoid cumulating difficulties (boundary issues and dependence on $x$) and we will omit the indication of the domain. Note that the PDE corresponding to the above minimization problem is $\nabla\cdot(a(x)(\nabla u)^{p-1}) = f$.

First, we need to compute the transform of $w \mapsto H^*(w) := \frac{a}{p}|w|^p$. Set $b = a^{1/(p-1)}$. It is easy to obtain $H(v) = \frac{1}{bq}|v|^q$. Also, we can check (just by scaling the inequality of Lemma 3.2), that we have

$$\frac{1}{bq}|v|^q + \frac{b^{p-1}}{p}|w|^p \geq v \cdot w + b^{p-1}\left|w^{p/2} - \frac{v^{q/2}}{b^{p/2}}\right|^2.$$

In particular, if we suppose that $a(x)$ is bounded from below by a positive constant, and we set $H^*(x, w) = \frac{a(x)}{p}|w|^p$ then we get

$$H(x, v) + H^*(x, w) \geq v \cdot w + c|F(x, v) - G(w)|^2$$

where $G(w) = w^{p/2}$.

We can now prove the following theorem.

**Theorem 4.2.** *Suppose $f \in W^{1,q}$ and $a \in \mathrm{Lip}, a \geq a_0$, and let $\hat{u}$ be the minimizer of (4.2). Then $V := (\nabla\hat{u})^{p/2} \in H^1$.*

*Proof.* Our usual computations show that

$$c||V_\delta - V||^2_{L^2} \leq g(\delta) + \tilde{g}(\delta),$$

where $g(\delta) = \int f\hat{u}_\delta - \int f\hat{u}$ and $\tilde{g}(\delta) = \int \frac{a(x)}{p}|\nabla\hat{u}_\delta|^p - \int \frac{a(x)}{p}|\nabla\hat{u}|^p$. With our assumptions, $g \in C^{1,1}$. As for $\tilde{g}(\delta)$, we write

$$\int \frac{a(x)}{p}|\nabla\hat{u}_\delta|^p = \int \frac{a(x-\delta)}{p}|\nabla\hat{u}|^p$$

and hence

$$\nabla\tilde{g}(\delta) = \int \frac{\nabla a(x-\delta)}{p}|\nabla\hat{u}|^p = \int \frac{\nabla a(x)}{p}|\nabla\hat{u}_\delta|^p.$$

Hence,

$$|\nabla\tilde{g}(\delta) - \nabla\tilde{g}(0)| \leq \int \frac{|\nabla a(x)|}{p}\,||\nabla\hat{u}_\delta|^p - |\nabla\hat{u}|^p| \leq C\int ||G_\delta|^2 - |G|^2| \leq C||V_\delta - V||_{L^2}||G_\delta + G||_{L^2}.$$

Here we used the $L^\infty$ bound on $|\nabla a|$. Then, from the lower bound on $a$, we also know $G \in L^2$, hence we get $|\nabla\tilde{g}(\delta) - \nabla\tilde{g}(0)| \leq C||V_\delta - V||_{L^2}$.

Now, we define as usual $\omega_t := \sup_{\delta:|\delta|\leq t}||V_\delta - V||_{L^2}$ and we get

$$\begin{aligned}\omega_t^2 \leq C\sup_{\delta:|\delta|\leq t} g(\delta) + \tilde{g}(\delta) &\leq Ct\sup_{\delta:|\delta|\leq t}|\nabla g(\delta) + \nabla\tilde{g}(\delta)| \\ &= Ct\sup_{\delta:|\delta|\leq t}|\nabla g(\delta) - \nabla g(0) + \nabla\tilde{g}(\delta) - \nabla\tilde{g}(0)| \leq Ct^2 + Ct\omega_t,\end{aligned}$$

which allows to deduce $\omega_t \leq Ct$ and hence $V \in H^1$. □

We also provide the following theorem, which is also interesting for $p = 2$.

**Theorem 4.3.** *Suppose $p \geq 2$, $f \in L^q$ and $a \in \mathrm{Lip}, a \geq a_0$, and let $\hat{u}$ be the minimizer of (4.2). Then $V := (\nabla\hat{u})^{p/2}$ satisfies $||V_\delta - V||_{L^2} \leq C|\delta|^{q/2}$. In particular, $V \in H^1$ when $p = 2$ and $V \in H^s$ for all $s < q/2$ when $p > 2$.*

*Proof.* The only difference with the previous case is that we cannot say that $g$ is $C^{1,1}$ but we should stick to the computation of $\nabla g$. We use as usual

$$|\nabla g(\delta) - \nabla g(0)| \leq ||f||_{L^q} ||\nabla \hat{u}_\delta - \nabla \hat{u}||_{L^p}.$$

As we are forced to let the norm $||\nabla \hat{u}_\delta - \nabla \hat{u}||_{L^p}$ appear, we will use it also in $\tilde{g}$. Indeed, we can observe that we can estimate

$$|\nabla \tilde{g}(\delta) - \nabla \tilde{g}(0)| \leq \int \frac{|\nabla a(x)|}{p} \, ||\nabla \hat{u}_\delta|^p - |\nabla \hat{u}|^p| \leq C \int (|\nabla \hat{u}_\delta|^{p-1} + |\nabla \hat{u}|^{p-1}) |\nabla \hat{u}_\delta - \nabla \hat{u}|$$
$$\leq C ||(\nabla \hat{u})^{p-1}||_{L^q} ||\nabla \hat{u}_\delta - \nabla \hat{u}||_{L^p}.$$

We then use $||(\nabla \hat{u})^{p-1}||_{L^q} = ||\nabla \hat{u}||_{L^p}^{p-1}$ and conclude

$$|\nabla \tilde{g}(\delta) - \nabla \tilde{g}(0)| \leq C ||\nabla \hat{u}_\delta - \nabla \hat{u}||_{L^p}.$$

This gives, defining $\omega_t$ as usual,

$$\omega_t \leq Ct \sup_{\delta : |\delta| \leq t} |\nabla g(\delta) - \nabla g(0) + \nabla \tilde{g}(\delta) - \nabla \tilde{g}(0)| \leq Ct \sup_{\delta : |\delta| \leq t} ||\nabla \hat{u}_\delta - \nabla \hat{u}||_{L^p}$$

and hence

$$\omega_t^2 \leq Ct \omega_t^{2/p}$$

as in Proposition 3.7. □

## 5. Time-dependent problems

The technique that we saw in the previous sections to prove Sobolev regularity provides in general classical results, through a slightly different point of view than the usual PDE-based tools. Yet, it has the advantage that it requires only the optimality, with no need to write a PDE, and could be useful in some very degenerate or non-smooth cases. As we underlined in the introduction, the first use (to the best of our knowledge) of duality-based methods to prove regularity was in [14], in the study of variational models for the incompressible Euler Equation. This has been later adapted in [17] to density-constrained Mean Field Games.

In this last section we only want to give an idea of where these estimates could be really useful, concentrating, without entering into details, on the case of an easier MFG. This takes the following form

Consider the following minimization problem

$$\min \left\{ \mathcal{A}(\rho, v) := \int_0^T \int_\Omega \left( \frac{1}{2} \rho_t |v_t|^2 + H(\rho_t) \right) + \int_\Omega \Psi \rho_T \right\}$$

among pairs $(\rho, v)$ such that $\partial_t \rho + \nabla \cdot (\rho v) = 0$, with given $\rho_0$, where $H$ is a given convex function.

Note that this problem is convex in the variables $(\rho, E := \rho v)$ (while it is not convex in $(\rho, v)$) and it recalls the Benamou-Brenier formulation for optimal transport ([6]). Moreover, in these variables, it exactly corresponds to a problem with constraints on the divergence (indeed, $\partial_t \rho + \nabla \cdot E$ is the space-time divergence of $(\rho, E)$).

As all convex minimization problem, $\min \mathcal{A}$ admits a dual problem, formally obtained by interchanging inf and sup in

$$\min_{\rho, v} \left\{ \mathcal{A}(\rho, v) + \sup_\phi \int_0^T \int_\Omega (\rho \partial_t \phi + \nabla \phi \cdot \rho v) + \int_\Omega \phi_0 \rho_0 - \int_\Omega \phi_T \rho_T \right\}.$$

We get

$$\sup\left\{-\mathcal{B}(\phi,p):=\int_\Omega \phi_0\rho_0-\int_0^T\!\!\int_\Omega H^*(p_+)\ :\ \phi_T\le\Psi,\ -\partial_t\phi+\frac{1}{2}|\nabla\phi|^2=p\right\},$$

where $H^*$ is the Legendre transform of $H$ (the positive part is due to the constraint $\rho\ge 0$). Note that the problem could be written in terms of $\phi$ only (as $p$ depends on $\phi$), but in this way there is more simmetry with the primal problem, as in both case we have two variables ($(\rho,v)$ or $(\phi,p)$), with a PDE constraint.

Now, we can do our usual computation taking arbitrary $(\rho,v)$ and $(\phi,p)$ admissible in the primal and dual problem. Compute

$$\mathcal{A}(\rho,v)+\mathcal{B}(\phi,p)=\int_\Omega(\Psi-\phi_T)\rho_T+\int_0^T\!\!\int_\Omega(H(\rho)+H^*(p_+)-p\rho)+\frac{1}{2}\int_0^T\!\!\int_\Omega\rho|v+\nabla\phi|^2. \tag{5.1}$$

Notice $(H(\rho)+H^*(p_+)-p\rho)\ge\frac{\lambda}{2}|\rho-(H')^{-1}(p_+)|^2$ where $\lambda=\inf H''$ (with equality if and only if $\rho=(H')^{-1}(p_+)$ and $\rho p=\rho p_+$). Suppose $\lambda>0$.

Supposing for simplicity $\Omega=\mathbb{T}^d$ to be the flat torus, using

$$\mathcal{A}(\rho,v)+\mathcal{B}(\phi,p)\ge c\int_0^T\!\!\int_\Omega|\rho-(H')^{-1}(p_+)|^2$$

we can deduce, with the same technique as in the rest of the paper, $\rho\in H^1$ (we can get both regularity in space and local in time, see [26] for an improvement up to $t=T$). By the way, using the last term in (5.1), we can also get $\iint\rho|D^2\phi|^2<\infty$.

It is important to observe that this is just formal, as the computation in (5.1) would require smooth functions $(\phi,p)$ (indeed, $(\rho,v)$ solves the equation in distributional sense, and maximizers $(\phi,p)$ could not be regular enough; in some particular MFG it is not even evident that they do exist), but this can be fixed using pairs $(\phi_\varepsilon,p_\varepsilon)$ which are not optimal but optimal up to an error $\varepsilon>0$. This is a very powerful feature of the duality method (see [26] or directly [14]).

The above computation is important as it gives regularity for $\rho$, but this implies regularity for $p_+=H'(\rho)$, and $p=p_+$ on $\{\rho>0\}$ (i.e. on the set where there is actually some mass). The function $p$ appears in the Hamilton-Jacobi equation $-\partial_t\phi+\frac{1}{2}|\nabla\phi|^2=p$. Indeed, the solution $(\rho,v)$ represents the motion of a population $\rho$, where each individual follows the velocity $v=-\nabla\phi$, and $\phi$ is the value function of the control problem

$$\min\left\{\int_0^T\left(\frac{|x'(t)|^2}{2}+p(t,x(t))\right)dt+\Psi(x(T))\right\}. \tag{5.2}$$

This explains the name mean-field games: we look for a global motion configuration, where each individual chooses his trajectory by optimizing a criterion where $p$ (and hence $\rho$) appears, i.e. where the criterion depends, through a sort of mean-field effect, on the choice of the others. The mathematical difficulty is that we need to integrate $p$ over a trajectoriy, i.e. a neglible set, which requires a little bit of regularity.

The situation is even more complicated when we try to study the case where the density penalization $H(\rho)$ is replaced by the constraint $\rho\le 1$. If we look at the variational problem

$$\min\left\{\int_0^T\!\!\int_\Omega\frac{1}{2}\rho_t|v_t|^2+\int_\Omega\Psi\rho_T\ :\ \rho\le 1\right\}$$

we can compute the dual

$$\sup\left\{\int_\Omega \phi_0\rho_0 - \int_0^T\!\!\int_\Omega p_+ \ :\ \phi_T \leq \Psi,\ -\partial_t\phi + \frac{1}{2}|\nabla\phi|^2 = p\right\}.$$

Note that here $p$ is a pressure arising from the incompressibility constraint $\rho \leq 1$ (at the optimum we will have $p = 0$ on $\{\rho < 1\}$ and $p \geq 0$ on $\{\rho = 1\}$), but finally acts as a price, in the sense that every agent pays in (5.2) a cost $p(x)$ when passying through a saturated point $x$ with $\rho(x) = 1$. Again, in order to give a meaning to (5.2) we need a bit of regularity. The situation is much trickier, since a priori $p$, which only appears with linear growth in the dual problem, is only supposed to be a measure. The same kind of duality arguments as above, with some loss of exponents due to the linear behavior, allow [17] to get

$$p \in L^2_{loc}((0,T); BV(\Omega)),$$

which is the same result as the one obtained in [2] for the incompressible Euler equation.

To finish this very informal section, we also try to give an idea of how the incompressible Euler equation can fit this framework. The easiest way is to look at the Eulerian-Lagrangian framework of [14] and [3], which is a multi-phasic formulation: consider a family of densities $\rho_\alpha$, each following a continuity equation $\partial_t\rho_\alpha + \nabla\cdot(\rho_\alpha v_\alpha) = 0$, with given $\rho_\alpha$ at times $t = 0$ and $t = 1$. Each $\rho_\alpha$ represents the density of a particle (or of a phase) with label $\alpha$. Then impose the global incompressibility constraint $\int \rho_\alpha(t,x)\,d\alpha = 1$ for every $(t,x)$, and minimize the total kinetic energy:

$$\min\left\{\int d\alpha \int_0^T dt \int_\Omega \frac{1}{2}\rho_\alpha(t,x)|v_\alpha(t,x)|^2 dx \ :\ \begin{array}{c}\int \rho_\alpha\, d\alpha = 1 \text{ for every } (t,x)\\ \partial_t\rho_\alpha + \nabla\cdot(\rho_\alpha v_\alpha) = 0 \text{ for every } \alpha\end{array}\right\}.$$

This provides a similar structure, where the penalization $+H(\rho)$ or the constraint $\rho \leq 1$ are replaced with a multiphasic constaint on $\int \rho_\alpha\, d\alpha$, and allows to perform a similar analysis which was, as we said, the starting point for the method presented in these notes.

**Acknowledgments.** The author wants to thank L. Brasco for very fruitful discussions and interesting remarks on a first version of the manuscript, and S. Dweik for help in revision and in particular for pointing out a simplification of some proofs. The author also acknowledges the support of the ANR project ISOTACE (ANR-12-MONU-0013), of the iCODE project "Strategic Crowds", funded by IDEX Paris-Saclay and of the project "Variational and PDE methods in Mean Field Games" (VarPDEMFG), funded by the Programme Gaspard Monge pour l'Optimisation of EDF and of the Hadamard Foundation.

## References

[1] R. A. Adams, *Sobolev spaces*, Pure and Applied Mathematics, 65, Academic Press, New York-London, 1975
[2] L. Ambrosio and A. Figalli, On the regularity of the pressure field of Brenier's weak solutions to incompressible Euler equations, *Calc. Var. PDE*, 31 (2008) No. 4, 497-509.
[3] L. Ambrosio and A. Figalli, Geodesics in the space of measure-preserving maps and plans, *Arch. Rational Mech. Anal.*, 194 (2009), 421-462.
[4] J.-B. Baillon and G. Carlier, From discrete to continuous Wardrop equilibria, *Net. Heterogenous Media*, 7 (2012), 219–241.
[5] M. Beckmann, A continuous model of transportation, *Econometrica* 20 (1952), 643-660.
[6] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.*, 84, 375–393, 2000.

[7] B. Bojarski and T. Iwaniec, *p-harmonic equation and quasiregular mappings*, Partial Differential Equations (Warsaw 1984), Banach Center Publications, 19 (1987).

[8] G. Bouchitté and G. Buttazzo, Characterization of optimal shapes and masses through Monge-Kantorovich equation *J. Eur. Math. Soc.* 3 (2), 139–168, 2001.

[9] P. Bousquet and L. Brasco, $C^1$ regularity of orthotropic $p$-harmonic functions in the plane, preprint, 2016, available at `http://cvgmt.sns.it/paper/3113/`

[10] L. Brasco and G. Carlier On certain anisotropic elliptic equations arising in congested optimal transport: local gradient bounds, *Adv. Calc. Var.,* 7 (2014), 379–407.

[11] L. Brasco, G. Carlier and F. Santambrogio, Congested traffic dynamics, weak flows and very degenerate elliptic equations, *J. Math. Pures and Appl.*, 93, No 6, 2010, 652–671.

[12] L. Brasco, C. Leone, G. Pisante and A. Verde Sobolev and Lipschitz regularity for local minimizers of widely degenerate anisotropic functionals, *Nonlinear Anal.*, available online, in press.

[13] L. Brasco and F. Santambrogio A sharp estimate à la Calderón-Zygmund for the $p$-Laplacian, *Comm. Cont. Math.*, to appear.

[14] Y. Brenier, Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations, *Comm. Pure Appl. Math.*, 52 (1999) 4, 411-452.

[15] H. Brezis, *Functional analysis. Theory and applications* (French), Masson, Paris, 1983.

[16] M. Burger Bregman Distances in Inverse Problems and Partial Differential Equation, *Advances in Mathematical Modeling, Optimization and Optimal Control* Volume 109 of the series Springer Optimization and Its Applications (2016), 3-33.

[17] P. Cardaliaguet, A. R. Mészáros and F. Santambrogio First order Mean Field Games with density constraints: pressure equals price. *SIAM J. Contr. Opt.*, Vol. 54, No. 5, pp. 2672-2709, 2016.

[18] G. Carlier, C. Jimenez and F. Santambrogio, Optimal transportation with traffic congestion and Wardrop equilibria, *SIAM J. Control Optim.* (47), 2008, 1330-1350.

[19] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems*, Classics in Mathematics, Society for Industrial and Applied Mathematics (1999).

[20] R. Hatchi Wardrop equilibria : rigorous derivation of continuous limits from general networks models, preprint, `https://arxiv.org/abs/1506.08960`

[21] O. A. Ladyzhenskaya and N. N. Ural'tseva, *Linear and quasilinear elliptic equations, Academic Press,* New York and London, 1968.

[22] J.-M. Lasry and P.-L. Lions, Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris* 343 (10), 679–684, 2006.

[23] J.-M. Lasry and P.-L. Lions, Mean-Field Games, *Japan. J. Math.* 2, 229–260, 2007.

[24] G. Mingione, The singular set of solutions of non-differentiable elliptic systems, *Arch. Rational. Mech. Anal.*, **166** (2003), 287–301.

[25] P. Lindqvist, Notes on the p-Laplace equation, available on-line at the page `www.math.ntnu.no/~lqvist/p-laplace.pdf`.

[26] A. Prosinski and F. Santambrogio Global-in-time regularity via duality for congestion-penalized Mean Field Games. Preprint available at `cvgmt.sns.it`.

[27] F. Santambrogio *Optimal Transport for Applied Mathematicians*, book, Birkhäuser, Progress in Nonlinear Differential Equations and Their Applications 87, Birkhäuser Basel (2015)

[28] J. Simon, Régularité de la solution d'une équation non linéaire dans $\mathbb{R}^n$. In *Journées d'Analyse Non Linéaire*, P. Benilan et J. Robert éd., Lecture Notes in Mathematics, **665**, Springer (1978), 205–227.

[29] C. Villani, *Optimal Transport. Old and New.* Grundlehren der Mathematischen Wissenschaften, 338. Springer-Verlag, Berlin, 2009.

[30] K. Uhlenbeck, Regularity for a class of non-linear elliptic systems, *Acta Math.*, **138** (1977), 219–240.



Laboratoire de Mathématiques d'Orsay, Univ. Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay Cedex, France

*E-mail address*: `filippo.santambrogio@math.u-psud.fr`