Differential and Integral Equations Volume 10, Number 4, July 1997, pp. 687 – 716.

VARIATIONAL INTEGRALS OF NEARLY LINEAR GROWTH

Luigi Greco*
Dipartimento di Matematica e Applicazioni, "R. Caccioppoli", Università,
Via Cintia, Complesso Monte S. Angelo, 80126 Napoli, Italia

Tadeusz Iwaniec**
Department of Mathematics, Syracuse University, Syracuse, NY 13244, USA

Carlo Sbordone*
Dipartimento di Matematica e Applicazioni, "R. Caccioppoli", Università,
Via Cintia, Complesso Monte S. Angelo, 80126 Napoli, Italia

(Submitted by: Reza Aftabizadeh)

Abstract. We study variational integrals and related equations whose integrand grows almost linearly with respect to the gradient. A prototype of such functionals is

$$I[u] = \int_\Omega |\nabla u|\, A\big(|\nabla u|\big)\, dx,$$

where A is slowly increasing to ∞. For instance, $A(t) = \log^\alpha(1+t)$, $\alpha > 0$, or $A(t) = \log\log(e+t)$, etc. We show that the minimizer u subject to the Dirichlet data v satisfies the estimate

$$\int_\Omega |\nabla u|\, A^{1\pm\varepsilon}\big(|\nabla u|\big)\, dx \le C \int_\Omega |\nabla v|\, A^{1\pm\varepsilon}\big(|\nabla v|\big)\, dx$$

at least for some small $\varepsilon > 0$. This extends previous results [12], [16] on integrals with power growth.

0. Introduction. This paper is concerned with variational functionals of the form

$$I[u] = \int_\Omega W(x, \nabla u)\, dx \qquad \text{in } \Omega \subset \mathbb{R}^n, \tag{0.1}$$

for a mapping $u\colon \Omega \to \mathbb{R}^m$ with prescribed Dirichlet boundary data. Here the integrand $W\colon \Omega \times \mathbb{R}^{n\times m} \to \mathbb{R}_+$ is strictly convex with respect to the variable $\xi \in \mathbb{R}^{n\times m}$. The existence and uniqueness of the minimizer is, therefore, established by the direct method. In many classical situations the integrand grows as a power function $|\xi|^p$ with some $p > 1$. This type of growth was crucial for applying various tools of

Received for publication May 1996.

* Supported by M.U.R.S.T. 40 % (1994).

** Supported in part by NSF Grant DMS-9401104 and in part by IAM-CNR, Napoli, via P. Castellino.

AMS Subject Classifications: 35J60, 35J99, 49N60.

real analysis to establish regularity theory of the minimizers. Here we will not be concerned with such case in any essential way. Our special emphasis will be placed on the integrands which are not too far from being linear in $|\xi|$, that is

$$\lim_{\xi\to\infty} |\xi|^{-p} W(x,\xi) = 0 \qquad \text{for each } p > 1.$$

We are therefore led to study the stationary points of (0.1) in the Orlicz-Sobolev space $W^1 L_N(\Omega, \mathbb{R}^m)$ for an Orlicz function N that grows almost linearly; see Section 1 for the notation. This case presents a more delicate problem because the usual tools such as maximal operators and reverse Hölder inequalities are not well adapted to such spaces. Thus, we will need to refine those tools. Let us first look at some typical examples. A prototype of our functionals is

$$I[u] = \int_\Omega F\big(|\nabla u|\big)\,dx, \qquad F(t) = \int_0^t A(\tau)\,d\tau, \tag{0.2}$$

where $A\colon [0,\infty[\ \to [a,\infty[$, $A(0) = a \geq 0$, is a strictly increasing function of class $C^1\big([0,\infty[\big)$. Our basic assumption on A is formulated in terms of the inverse function, denoted by $B(s) = A^{-1}(s)$, $s \geq a$. It is required that

$$l^{-1}\, B(s) \leq s\, B'(s) \leq B(l\, s), \tag{0.3}$$

for all $s \geq a$, where $l > 1$ is a constant. Unless otherwise stated, no further condition on A will be imposed. Here are some illustrative examples:

$$A(t) = t^\alpha, \qquad \log^\alpha(1+t), \qquad \log^\alpha(e+t), \qquad \text{with } \alpha > 0$$

and

$$A(t) = \big[\log\log(e+t)\big]^\alpha, \qquad \big[\log\log\log(e^e+t)\big]^\alpha, \qquad \text{etc.}$$

Roughly speaking, the function A may approach infinity with increasing t as slowly as one wants.

Variational functionals whose integrands depend on $x \in \Omega$ typically arise in the study of a deformation of nonhomogeneous material body occupying the region $\Omega \subset \mathbb{R}^n$. Also, a functional (0.2) for a function u on a Riemannian manifold (with measurable metric tensor) can be written in local coordinates as

$$I[u] = \int_\Omega F\big(\langle G(x)\,\nabla u, \nabla u\rangle^{1/2}\big)\,J(x)\,dx, \tag{0.4}$$

where $G(x)$ is a symmetric matrix satisfying a uniform ellipticity condition

$$K^{-2}|\xi|^2 \leq \langle G(x)\,\xi,\xi\rangle \leq K^2|\xi|^2, \tag{0.5}$$

for almost every $x \in \Omega$ with K independent of x. Also, the Jacobian determinant satisfies

$$K^{-1} \le J(x) \le K \qquad \text{a.e. in } \Omega. \tag{0.6}$$

We shall now briefly discuss the Euler-Lagrange equation of the functionals in question. It has the form

$$\operatorname{div} \mathcal{A}(x, \nabla u) = 0, \tag{0.7}$$

where $\mathcal{A}\colon \Omega \times \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m}$ is the derivative of $W(x,\xi)$ with respect to ξ. For the functional (0.4) we have

$$\mathcal{A}(x,\xi) = J(x)\, \|\xi\|_x^{-1}\, A\big(\|\xi\|_x\big) \cdot G(x)\, \xi,$$

where we denote by $\|\xi\|_x = \langle G(x)\,\xi, \xi\rangle^{1/2}$ for $\xi \in \mathbb{R}^n$. In particular, if $G(x) \equiv \mathrm{Id}$ and $J(x) \equiv 1$ we obtain the Euler-Lagrange equation of our prototype functional

$$\operatorname{div} \mathcal{A}(\nabla u) = 0, \qquad \mathcal{A}(\xi) = |\xi|^{-1} A\big(|\xi|\big)\, \xi. \tag{0.8}$$

It will be shown that the mapping $\mathcal{A}\colon \Omega \times \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m}$ (in our examples) satisfies the three conditions

$$\big|\mathcal{A}(x,\xi)\big| \le \beta\, A\big(|\xi|\big), \tag{0.9}$$

$$\big|\mathcal{A}(x,\xi+\zeta) - \mathcal{A}(x,\xi)\big| \le \frac{\beta\,|\zeta|}{|\xi|+|\zeta|}\, \Big[A\big(|\xi|\big) + A\big(|\zeta|\big)\Big], \tag{0.10}$$

$$\langle \xi, \mathcal{A}(x,\xi)\rangle \ge \alpha\, |\xi|\, A\big(|\xi|\big), \tag{0.11}$$

where α and β are positive numbers depending only on l in (0.3) and K in (0.5)–(0.6). The above examples are merely illustrative rather than representative of the variety of applications of the $\mathcal{A}$-harmonic equation (0.7). Our arguments will strongly depend on the assumptions (0.9)–(0.11). We will not really need that $\mathcal{A}$ is the derivative of a variational integrand. All we will require are conditions (0.9)–(0.11), where A satisfies inequalities (0.3). These inequalities yield

$$F(t) \le t\, A(t) \le 2^{l+1}\, F(t), \tag{0.12}$$

where F is a convex function given by $F(t) = \int_0^t A(\tau)\, d\tau$; see Section 1. Therefore, the natural domain of the $\mathcal{A}$-harmonic operator is the Orlicz-Sobolev space $W^1 L_F(\Omega, \mathbb{R}^m)$. Equation (0.7) simply means that

$$\int_\Omega \langle \mathcal{A}(x, \nabla u), \nabla \eta\rangle\, dx = 0, \tag{0.13}$$

for all test functions $\eta \in W_0^1 L_F(\Omega, \mathbb{R}^m)$. Given $v \in W^1 L_F(\Omega, \mathbb{R}^m)$, one may consider the Dirichlet problem

$$\begin{cases} \operatorname{div} \mathcal{A}(x, \nabla u) = 0 \\ u - v \in W_0^1 L_F(\Omega, \mathbb{R}^m). \end{cases} \tag{0.14}$$

A routine computation, see (6.4), reveals that

$$\int_\Omega F\big(|\nabla u|\big)\,dx \le C(\mathcal{A}) \int_\Omega F\big(|\nabla v|\big)\,dx. \tag{0.15}$$

Here and subsequently, $C(\mathcal{A})$ stands for a number which depends on the structural constants of the $\mathcal{A}$-harmonic operator, but neither on the solution u nor on the Dirichlet data v. Estimates in the natural spaces such as (0.15) involve no regularity of the domain Ω. However, we wish to investigate the Dirichlet problem for the $\mathcal{A}$-harmonic equation in the Orlicz-Sobolev spaces $W^1L_N(\Omega,\mathbb{R}^m)$ not too far from the natural one. For this, it will be necessary to put some restrictions on Ω. We want to know whether ∇u enjoys the same degree of integrability as ∇v. The affirmative solution is contained in two main theorems of this paper. First we shall be concerned with estimates below the natural degree of integrability, that is, for very weak solutions of (0.14). This concept goes back at least as far as [11], [12] and [16]. A very weak solution of the $\mathcal{A}$-harmonic equation is simply a locally integrable function $u\colon\Omega\to\mathbb{R}^m$ whose distributional gradient ∇u belongs to the Orlicz space $L_N(\Omega,\mathbb{R}^{n\times m})$, $N(t)=t\,A^{1-\varepsilon}(t)$ for some $0<\varepsilon<1$ and the integral identity (0.13) remains valid for all test functions $\eta\in \mathrm{Lip}\,(\Omega,\mathbb{R}^m)$ with compact support. For certain domains $\Omega\subset\mathbb{R}^n$ such as the half space, cubes, balls or bilipschitz images of a cube, we succeed in proving the following uniform estimate.

Theorem 1. *For each $\mathcal{A}$-harmonic operator there exist an exponent $0<\varepsilon<1$ and a constant $C(\mathcal{A})$ with the following properties: if we set $N(t)=t\,A^{1-\varepsilon}(t)$ and consider $u\in W^1L_N(\Omega,\mathbb{R}^m)$ a very weak solution of* (0.7) *with the Dirichlet condition $u-v\in W^1_0L_N(\Omega,\mathbb{R}^m)$, then*

$$\int_\Omega N\big(|\nabla u|\big)\,dx \le C(\mathcal{A}) \int_\Omega N\big(|\nabla v|\big)\,dx. \tag{0.16}$$

Thus very weak solutions are quasiminima of the variational integral $\int_\Omega N\big(|\nabla u|\big)$.

For Ω a regular domain, see Definition 6.1, we obtain the following higher integrability result for the solutions of (0.14).

Theorem 2. *For each $\mathcal{A}$-harmonic operator in a regular domain there exist an exponent $0<\varepsilon<1$ and a constant $C(\mathcal{A})$ such that, if $v\in W^1L_N(\mathbb{R}^n,\mathbb{R}^m)$ with $N(t)=t\,A^{1+\varepsilon}(t)$, then also $u\in W^1L_N(\Omega,\mathbb{R}^m)$ and we have the uniform bound*

$$\int_\Omega N\big(|\nabla u|\big)\,dx \le C(\mathcal{A}) \int_{\mathbb{R}^n} N\big(|\nabla v|\big)\,dx.$$

We give the proof only for the case when the function $t\mapsto A^n(t)$ is concave in $[0,\infty[$. However with a slight modification our proof still goes when we drop this assumption. The arguments establishing Theorem 2 strongly depend on an extension of Gehring's lemma [7] to the Orlicz space $L_F(\mathbb{R}^n)$ which might be of independent interest. Using the standard notation of this theory we formulate the result as

Lemma 1. *Let $F(t) = t\,A(t)$ be strictly convex, where $A\colon \mathbb{R}_+ \to \mathbb{R}_+$, $A(0) = 0$ is continuously increasing and satisfies the Δ_2-condition. Suppose that nonnegative functions g, $h \in L_F(\mathbb{R}^n)$ satisfy*

$$F^{-1}\Big[\fint_Q F(g)\,dx\Big] \le a \ \fint_{4Q} g\,dx + b\,F^{-1}\Big[\fint_{4Q} F(h)\,dx\Big], \tag{0.17}$$

for every cube $Q \subset \mathbb{R}^n$ with constants $a > 1$ and $b > 0$ independent of the cube. Then there exist $\varepsilon = \varepsilon(n, a, b, F) > 0$ and $C = C(n, a, b, F)$ such that

$$\int_{\mathbb{R}^n} gA^{1+\varepsilon}(g)\,dx \le C \int_{\mathbb{R}^n} hA^{1+\varepsilon}(h)\,dx. \tag{0.18}$$

It is worth pointing out that for this lemma condition (0.3) is not required. What we really need is condition (0.12) for the convex function $F(t) = \int_0^t A(\tau)\,d\tau$, which obviously follows from (0.3). See Section 5 for more details. One more point of emphasis is that the functional $g \mapsto F^{-1}\big[\fint_Q F(g)\big]$ need not be a norm and, in the cases of most interest to us, may not even be equivalent (comparable) with any norm in $L_F(\mathbb{R}^n)$. Such a situation always occurs when F has nearly linear growth. That is why this case needs handling with greater care.

1. Notation and preliminaries. In this section we introduce the notation about Orlicz and Orlicz-Sobolev spaces we will follow through the paper. For more information on these spaces see [2], [4], [9], [10], [14], [15], [18].

Let Ω be an open subset of $\mathbb{R}^n$. The Orlicz space corresponding to a function F is denoted by $L_F(\Omega)$ and the closure in $L_F(\Omega)$ of the subspace of bounded measurable functions with compact support is denoted by $E_F(\Omega)$. If F satisfies the Δ_2-condition we have $E_F(\Omega) = L_F(\Omega)$. The Orlicz-Sobolev space of functions in $L_F(\Omega)$ with first distributional derivatives in $L_F(\Omega)$ is denoted by $W^1L_F(\Omega)$ and the closure of $C_0^\infty(\Omega)$ in $W^1L_F(\Omega)$ is denoted by $W_0^1L_F(\Omega)$.

Now we prepare some estimates for the function $A = A(t)$ which we will need in the sequel. Recall that $A\colon [0,\infty[\to [a,\infty[$ is strictly increasing and its inverse function $B\colon [a,\infty[\to [0,\infty[$ satisfies the condition

$$l^{-1}\,B(s) \le s\,B'(s) \le B(l\,s), \qquad s \ge a, \tag{1.1}$$

where $l \ge 1$ is a constant independent of s. The first of the above inequalities can be rephrased as

$$t\,A'(t) \le l\,A(t), \qquad t \ge 0, \tag{1.2}$$

which simply means that the function $t \mapsto t^{-l}A(t)$ is decreasing. It is worth noting that

$$\text{the function } t\,A^{-\varepsilon}(t) \ \text{ is increasing} \tag{1.3}$$

for all $0 < \varepsilon \le 1/l$. Another consequence of (1.2) is the inequality

$$A(\lambda\, t) \le \lambda^l\, A(t) \qquad \text{for all } t \ge 0 \text{ and } \lambda \ge 1. \tag{1.4}$$

Indeed, as $\lambda \ge 1$, we have $A(\lambda t)/(\lambda t)^l \le A(t)/t^l$ and therefore the conclusion. This inequality is usually stated for $\lambda = 2$ and in such form is referred to as the Δ_2-property of A. The proof of (0.12) now proceeds

$$F(t) = \int_0^t A(\tau)\, d\tau \le t\, A(t) \le 2^l\, t\, A(t/2) \le 2^{l+1} \int_{t/2}^t A(\tau)\, d\tau \le 2^{l+1} F(t).$$

Note that we actually have

$$A(t+s) \le 2^l \big(A(t) + A(s)\big) \qquad \text{for all } t, s \ge 0\,. \tag{1.5}$$

Indeed, for $t \le s$ we write $A(t+s) \le A(2s) \le 2^l A(s) \le 2^l\big(A(t) + A(s)\big)$; the case $s \le t$ is similar. Next we shall show that for each $M \ge 0$,

$$A(M) \le l\, A\Big[\int_0^M \frac{t\, A'(t)}{A(t)}\, dt\Big]. \tag{1.6}$$

To this effect we make the substitution $t = B(ls)$ and then use the inequality in the right hand side of (1.1) to obtain

$$\begin{aligned}\int_0^M \frac{t\, A'(t)}{A(t)}\, dt \ge \int_{B(la)}^M \frac{t\, A'(t)}{A(t)}\, dt &= \int_a^{A(M)/l} B(ls)\, \frac{ds}{s} \\ &\ge \int_a^{A(M)/l} B'(s)\, ds = B\big(A(M)/l\big),\end{aligned}$$

which is the same as (1.6). For use in the forthcoming section we define

$$\varphi_\varepsilon(t) = \frac{A'(t)}{A^{1+\varepsilon}(t)}\,, \qquad \varepsilon > 0 \tag{1.7}$$

and

$$\Phi_\varepsilon(t) = A(t) \int_0^t s\, \varphi_\varepsilon(s)\, ds. \tag{1.8}$$

We see at once that

$$\int_M^\infty \varphi_\varepsilon(t)\, dt = \frac{1}{\varepsilon\, A^\varepsilon(M)}\,, \qquad \text{for } M > 0. \tag{1.9}$$

Furthermore, for $0 < \varepsilon < 1/l$ we have

$$\int_0^t s\, \varphi_\varepsilon(s)\, ds \le \frac{l}{1 - l\varepsilon}\, t\, A^{-\varepsilon}(t). \tag{1.10}$$

Indeed, in view of (1.3),

$$\int_0^t s\,\varphi_\varepsilon(s)\,ds = \int_0^t s\,A^{-1/l}(s)\,A^{1/l-\varepsilon-1}(s)\,A'(s)\,ds$$
$$\leq t\,A^{-1/l}(t)\int_0^t A^{1/l-\varepsilon-1}(s)\,A'(s)\,ds = \frac{l}{1-l\varepsilon}\,t\,A^{-\varepsilon}(t).$$

To estimate the function Φ_ε, $0<\varepsilon<1/l$, we deduce from (1.10) that

$$\Phi'_\varepsilon(t) = A'(t)\int_0^t s\,\varphi_\varepsilon(s)\,ds + t\,A'(t)\,A^{-\varepsilon}(t) \leq \big(\frac{l}{1-l\varepsilon}+1\big)t\,A'(t)\,A^{-\varepsilon}(t).$$

Hence

$$\int_0^s \frac{\Phi'_\varepsilon(t)}{t}\,dt \leq \big(\frac{l}{1-l\varepsilon}+1\big)\frac{1}{1-\varepsilon}\big[A^{1-\varepsilon}(s)-A^{1-\varepsilon}(0)\big].$$

Now, in view of (1.4), it is obvious that for each $h\geq 0$,

$$\int_0^{2h}\frac{\Phi'_\varepsilon(t)}{t}\,dt \leq 8^l A^{1-\varepsilon}(h)\,, \qquad \text{if } 0<\varepsilon\leq 1/(3l). \tag{1.11}$$

Having disposed of these preliminary steps, we can now prove the following crucial estimate.

Lemma 1.1. *Suppose* $0\leq h\leq M$ *and* $0<\varepsilon\leq 1$. *Then for each* $0<\sigma\leq 1$,

$$h\,A^{1-\varepsilon}(h) \leq \sigma\,\Phi_\varepsilon(M) + l\,\sigma^{-l}h\,A(h)\,A^{-\varepsilon}(M). \tag{1.12}$$

Proof. We begin with the inequality

$$h \leq \sigma\int_0^M \frac{t\,A'(t)}{A(t)}\,dt + l\,\sigma^{-l}\,h\big[\frac{A(h)}{A(M)}\big]^\varepsilon. \tag{1.13}$$

To see this, we need only to consider the case $\int_0^M \big[t\,A'(t)/A(t)\big]\,dt < h/\sigma$. Then by (1.6),

$$h\leq \frac{l\,h}{A(M)}\,A\big[\int_0^M \frac{t\,A'(t)}{A(t)}\,dt\big] \leq \frac{l\,h}{A(M)}\,A\big(h/\sigma\big) \leq \frac{l\,h\,\sigma^{-l}A(h)}{A(M)} \leq l\,\sigma^{-l}\,h\big[\frac{A(h)}{A(M)}\big]^\varepsilon$$

as desired. Next we multiply (1.13) by $A^{1-\varepsilon}(h)$ to obtain

$$h\,A^{1-\varepsilon}(h) \leq \sigma\,A^{1-\varepsilon}(M)\int_0^M \frac{t\,A'(t)}{A(t)}\,dt + l\,\sigma^{-l}\,h\,A(h)\,A^{-\varepsilon}(M)$$
$$\leq \sigma\,A(M)\int_0^M \frac{t\,A'(t)}{A^{1+\varepsilon}(t)}\,dt + l\,\sigma^{-l}\,h\,A(h)\,A^{-\varepsilon}(M),$$

which is exactly inequality (1.12). □

Finally, we need to examine the function $\Gamma\colon \mathbb{R}_+\times\mathbb{R}_+\to\mathbb{R}_+$ given by

$$\Gamma(g,h) = g\,\frac{A(g)+A(h)}{g+h} \leq A(g)+A(h). \tag{1.14}$$

Lemma 1.2. *Suppose* $0 < \varepsilon \leq 1/l$ *and* $0 < \sigma \leq 1$*; then*

$$h\, A^{-\varepsilon}(h)\, \Gamma(g,h) \leq 2\sigma\, h\, A^{1-\varepsilon}(h) + 2\sigma^{-l-1} g\, A^{1-\varepsilon}(g).$$

Proof. The case $g \leq \sigma\, h$ is easy, because $\Gamma(g,h) \leq \sigma\, \big[A(g) + A(h)\big] \leq 2\, \sigma\, A(h)$. On the other hand, if $h < g/\sigma$, then we can write

$$\Gamma(g,h) \leq A(g) + A(g/\sigma) \leq (1 + \sigma^{-l})\, A(g)$$

and as the function $t\, A^{-\varepsilon}(t)$ is increasing, we also have

$$h \leq (g/\sigma)\, A^{-\varepsilon}(g/\sigma)\, A^{\varepsilon}(h) \leq (g/\sigma)\, A^{-\varepsilon}(g)\, A^{\varepsilon}(h).$$

Hence

$$h\, A^{-\varepsilon}(h)\, \Gamma(g,h) \leq (1 + \sigma^{-l})\, (g/\sigma)\, A^{1-\varepsilon}(g) \leq 2\sigma^{-l-1} g\, A^{1-\varepsilon}(g).$$

Lemma 1.3. *Suppose* $0 < \varepsilon < 1/l$, $0 \leq h \leq M$ *and* $g \geq 0$*; then*

$$\Gamma(g,h) \int_0^M t\, \varphi_\varepsilon(t)\, dt \leq 2\Phi_\varepsilon(M) + \frac{2l}{1 - l\varepsilon}\, g\, A^{1-\varepsilon}(g).$$

Proof. We consider two cases. If $g \leq M$, then $\Gamma(g,h) \leq A(g) + A(h) \leq 2A(M)$. Thus

$$\Gamma(g,h) \int_0^M t\, \varphi_\varepsilon(t)\, dt \leq 2A(M) \int_0^M t\, \varphi_\varepsilon(t)\, dt = 2\Phi_\varepsilon(M).$$

Now suppose that $g > M \geq h$. Then estimate (1.10) yields

$$\begin{aligned}\Gamma(g,h) \int_0^M t\, \varphi_\varepsilon(t)\, dt &\leq \frac{l}{1 - l\varepsilon}\, M\, A^{-\varepsilon}(M)\, \Gamma(g,h) \leq \frac{l}{1 - l\varepsilon}\, g\, A^{-\varepsilon}(g)\, \Gamma(g,h) \\ &\leq \frac{l}{1 - l\varepsilon}\, g\, A^{-\varepsilon}(g)\, 2A(g) = \frac{2l}{1 - l\varepsilon}\, g\, A^{1-\varepsilon}(g).\end{aligned}$$

Here we have exploited again the fact that the function $t\, A^{-\varepsilon}(t)$ is decreasing; see (1.3). □

We finish this section by proving condition (0.10) for the operator

$$\mathcal{A}(\xi) = A\big(|\xi|\big)\, \xi/|\xi|.$$

The remaining conditions (0.9) and (0.11) are easy to check.

Case 1. $|\xi| \le 3|\zeta|$. From (1.5) it follows that

$$\begin{aligned}\big|\mathcal{A}(\xi+\zeta)-\mathcal{A}(\xi)\big| &\le A\big(|\xi+\zeta|\big)+A\big(|\xi|\big) \le (2^l+1)\Big[A\big(|\xi|\big)+A\big(|\zeta|\big)\Big]\\ &\le \frac{\beta\,|\zeta|}{|\xi|+|\zeta|}\Big[A\big(|\xi|\big)+A\big(|\zeta|\big)\Big]\end{aligned}$$

with $\beta = 4(2^l+1)$.

Case 2. $|\xi| \ge 3|\zeta|$. We can write

$$\big|\mathcal{A}(\xi+\zeta)-\mathcal{A}(\xi)\big| = \Big|\int_0^1 \frac{d}{dt}\,\mathcal{A}(\xi+t\zeta)\,dt\Big| \le |\zeta|\int_0^1 \big|\mathcal{A}'(\xi+t\zeta)\big|\,dt.$$

Here $\mathcal{A}' \colon \mathbb{R}^n \to \mathbb{R}^{n\times m}$ is easily computed to be

$$\mathcal{A}'(\xi) = \frac{A\big(|\xi|\big)}{|\xi|}\Big[\mathrm{Id}-\frac{\xi\otimes\xi}{|\xi|^2}\Big] + A'\big(|\xi|\big)\,\frac{\xi\otimes\xi}{|\xi|^2}.$$

Hence, in view of (1.2), we find that $\big|\mathcal{A}'(\xi)\big| \le (l+2)\,A\big(|\xi|\big)/|\xi|$ and continue the chain of inequalities started above:

$$\begin{aligned}&\le (l+2)\,|\zeta|\int_0^1 \frac{A\big(|\xi+t\zeta|\big)}{|\xi+t\zeta|}\,dt \le (l+2)\,\frac{|\zeta|}{|\xi-\zeta|}\,A\big(|\xi|+|\zeta|\big)\\ &\le (2l+4)\,\frac{|\zeta|}{|\xi|+|\zeta|}\,2^l\Big[A\big(|\xi|\big)+A\big(|\zeta|\big)\Big].\end{aligned}$$

In the last step we have used (1.5). The proof of condition (0.10) is complete.

2. Maximal inequalities in Orlicz spaces. This section is devoted to a precise formulation and a complete proof of the maximal theorem in Orlicz spaces and a few related estimates of the Hardy-Littlewood maximal operator

$$\mathcal{M}h(x) = \sup\Big\{ ⨍_Q |h| : x\in Q\subset\mathbb{R}^n\Big\}\,, \qquad h\in L^1_{\mathrm{loc}}(\mathbb{R}^n).$$

Here the supremum is taken over all cubes Q containing x and, as always, the symbol $⨍_Q$ stands for the integral mean over the cube.

Let λ denote the distribution function of $\mathcal{M}h$, that is, $\lambda(t) = \big|\{x : \mathcal{M}h(x) > t\}\big|$. Let $\Phi\colon [0,\infty[\to [0,\infty[$, $\Phi(0)=0$, be continuously differentiable with $\Phi'(t)\ge 0$ and assume that

$$\Psi(s) = \int_0^s \frac{\Phi'(t)}{t}\,dt < \infty \qquad \text{for all } s \ge 0.$$

Note that $\Phi(a) \le a\int_0^a \Phi'(t)/t\,dt = a\Psi(a)$ for all $a > 0$.

Proposition 2.1. *Suppose h is a nonnegative function in $\mathbb{R}^n$ such that $h\,\Psi(2h) \in L^1(\mathbb{R}^n)$; then*

$$\int_{\mathbb{R}^n} \Phi(\mathcal{M}h)\,dx \le 4\cdot 3^n \int_{\mathbb{R}^n} h\,\Psi(2h)\,dx. \tag{2.1}$$

Proof. The following estimate for the distribution function is a very well known consequence of Vitali's covering lemma:

$$\lambda(t) \le \frac{2\cdot 3^n}{t}\int_{2h>t} h(x)\,dx \qquad \text{for all } t>0.$$

Note that the integral in the right hand side is finite. Indeed,

$$\int_{2h>t} h\,dx \le \frac{1}{\Psi(t)}\int_{2h>t} h\,\Psi(2h)\,dx < \infty.$$

Now, for each $0<a<b<\infty$ we may integrate by parts to obtain

$$\begin{aligned}
\int_{a<\mathcal{M}h<b} \Phi(\mathcal{M}h)\,dx &= -\int_a^b \Phi(t)\,d\lambda(t) = -\lambda(t)\,\Phi(t)\Big|_a^b + \int_a^b \Phi'(t)\,\lambda(t)\,dt \\
&\le \lambda(a)\,\Phi(a) + \int_0^\infty \Phi'(t)\,\lambda(t)\,dt \\
&\le 2\cdot 3^n\Big[\Psi(a)\int_{2h>a} h\,dx + \int_0^\infty \frac{\Phi'(t)}{t}\Big(\int_{2h>t} h\,dx\Big)dt\Big] \\
&\le 2\cdot 3^n\Big[\int_{2h>a} h\,\Psi(2h)\,dx + \int_{\mathbb{R}^n} h\Big(\int_0^{2h}\frac{\Phi'(t)}{t}\,dt\Big)dx\Big] \\
&\le 4\cdot 3^n \int_{\mathbb{R}^n} h\,\Psi(2h)\,dx.
\end{aligned}$$

Letting a go to zero and b go to ∞, inequality (2.1) follows. $\square$

Clearly, our result applies to the function Φ_ε given by (1.8), which in view of estimate (1.11) yields

$$\int_{\mathbb{R}^n} \Phi_\varepsilon(\mathcal{M}h)\,dx \le 4\cdot 3^n\cdot 8^l \int_{\mathbb{R}^n} h\,A^{1-\varepsilon}(h)\,dx\,, \qquad 0<\varepsilon\le 1/(3l), \tag{2.2}$$

for each nonnegative function h of Orlicz class $L_N(\mathbb{R}^n)$ with $N(t) = t\,A^{1-\varepsilon}(t)$. Combining it with Lemma 1.1 yields

$$\int_{\mathbb{R}^n} h\,A^{1-\varepsilon}(h)\,dx \le 10^{2+nl+l^2}\int_{\mathbb{R}^n} h\,A(h)\,A^{-\varepsilon}(\mathcal{M}h)\,dx, \tag{2.3}$$

for all $0 < \varepsilon \le 1/(3l)$. To see this, we apply inequality (1.12) to the functions h and $M = \mathcal{M}h$ and then integrate over $\mathbb{R}^n$:

$$\begin{aligned}\int_{\mathbb{R}^n} h\, A^{1-\varepsilon}(h)\, dx &\le \sigma \int_{\mathbb{R}^n} \Phi_\varepsilon(\mathcal{M}h)\, dx + l\, \sigma^{-l} \int_{\mathbb{R}^n} h\, A(h)\, A^{-\varepsilon}(\mathcal{M}h)\, dx \\ &\le 4 \cdot 3^n \cdot 8^l\, \sigma \int_{\mathbb{R}^n} h\, A^{1-\varepsilon}(h)\, dx + l\, \sigma^{-l} \int_{\mathbb{R}^n} h\, A(h)\, A^{-\varepsilon}(\mathcal{M}h)\, dx.\end{aligned}$$

Putting $\sigma = 3^{-n} 8^{-l-1}$, inequality (2.3) follows.

3. Estimates below the natural exponent. We shall examine a nonhomogeneous $\mathcal{A}$-harmonic equation

$$\operatorname{div} \mathcal{A}(x, f + \nabla w) = 0 \qquad \text{in } \mathbb{R}^n, \tag{3.1}$$

for $w \in W_0^1 L_N(\mathbb{R}^n, \mathbb{R}^m)$ and f of the Orlicz class $L_N(\mathbb{R}^n, \mathbb{R}^{n\times m})$, $N(t) = t\, A^{1-\varepsilon}(t)$. Our goal is to prove the following uniform estimate.

Proposition 3.1. *To each $\mathcal{A}$-harmonic operator there correspond an exponent $0 < \varepsilon(\mathcal{A}) < 1$ and a constant $C(\mathcal{A}) > 0$ such that*

$$\int_{\mathbb{R}^n} |\nabla w|\, A^{1-\varepsilon}\big(|\nabla w|\big)\, dx \le C(\mathcal{A}) \int_{\mathbb{R}^n} |f|\, A^{1-\varepsilon}\big(|f|\big)\, dx, \tag{3.2}$$

for all f, $\nabla w \in L_N(\mathbb{R}^n, \mathbb{R}^{n\times m})$ satisfying (3.1) with $N(t) = t\, A^{1-\varepsilon}(t)$ and $0 < \varepsilon \le \varepsilon(\mathcal{A})$.

Remark 3.1. Without loss of generality, it can be assumed that $w \in C_0^\infty(\mathbb{R}^n, \mathbb{R}^m)$. For, if not, we approximate w by functions $w_j \in C_0^\infty(\mathbb{R}^n, \mathbb{R}^m)$ such that $\int N\big(|\nabla w_j - \nabla w|\big) \to 0$. Consider the equation $\operatorname{div} \mathcal{A}(x, f_j + \nabla w_j) = 0$ with $f_j = \nabla w - \nabla w_j + f$. Having disposed of the inequality

$$\int_{\mathbb{R}^n} N\big(|\nabla w_j|\big)\, dx \le C(\mathcal{A}) \int_{\mathbb{R}^n} N\big(|f_j|\big)\, dx,$$

we may let j go to ∞ to conclude with (3.2) for the function w.

The idea of the proof is similar to that of John Lewis [16]. Some innovations, however, are in order. The crucial fact is the following extension lemma; see [1], [3].

Lemma 3.1. *Given $w \in C_0^\infty(\mathbb{R}^n, \mathbb{R}^m)$ and $t > 0$, there exists a Lipschitz function $\eta \in \operatorname{Lip}(\mathbb{R}^n, \mathbb{R}^m)$ with compact support such that*

$$\big|\nabla \eta(x)\big| \le C(n,m)\, t \qquad \textit{for all } x \in \mathbb{R}^n$$

and

$$\eta(x) = w(x) \qquad \textit{if } \mathcal{M}\big(\nabla w\big)(x) \le t.$$

Here we recall that $\mathcal{M}$ stands for the maximal operator in $\mathbb{R}^n$.

Proof of Proposition 3.1. For notational simplicity, we denote $h = |\nabla w|$, $g = |f|$ and $\mathcal{M} = \mathcal{M}h$. For fixed $t > 0$ we then apply the identity

$$\int_{\mathbb{R}^n} \langle \mathcal{A}(x, f + \nabla w), \nabla \eta \rangle = 0 \tag{3.3}$$

to the function $\eta = \eta_t$ from Lemma 3.1. This identity splits as

$$\begin{aligned} &\int_{\mathcal{M} \le t} \langle \mathcal{A}(x, \nabla w), \nabla w \rangle \, dx \\ &= - \int_{\mathcal{M} > t} \langle \mathcal{A}(x, \nabla w), \nabla \eta \rangle \, dx + \int_{\mathcal{M} \le t} \langle \mathcal{A}(x, \nabla w) - \mathcal{A}(x, f + \nabla w), \nabla w \rangle \, dx \\ &\quad + \int_{\mathcal{M} > t} \langle \mathcal{A}(x, \nabla w) - \mathcal{A}(x, f + \nabla w), \nabla \eta \rangle \, dx. \end{aligned}$$

Using inequalities (0.9)-(0.11) and the definition of the function $\Gamma(g, h)$, see (1.14), we obtain

$$\begin{aligned} \alpha \int_{\mathcal{M} \le t} h \, A(h) \, dx &\le \beta \, C(n, m) \, t \int_{\mathcal{M} > t} A(h) \, dx \\ &+ \beta \int_{\mathcal{M} \le t} h \, \Gamma(g, h) \, dx + \beta \, C(n, m) \, t \int_{\mathcal{M} > t} \Gamma(g, h) \, dx. \end{aligned}$$

Next, we multiply this inequality by $\varphi_\varepsilon(t)$, see formula (1.7), and integrate with respect to t:

$$\begin{aligned} &\alpha \int_0^\infty \varphi_\varepsilon(t) \Big[\int_{\mathcal{M} \le t} h \, A(h) \, dx \Big] dt \le \beta \, C(n, m) \int_0^\infty t \, \varphi_\varepsilon(t) \Big[\int_{\mathcal{M} > t} A(h) \, dx \Big] dt \\ &+ \beta \int_0^\infty \varphi_\varepsilon(t) \Big[\int_{\mathcal{M} \le t} h \Gamma(g, h) \, dx \Big] dt + \beta \, C(n, m) \int_0^\infty t \varphi_\varepsilon(t) \Big[\int_{\mathcal{M} > t} \Gamma(g, h) \, dx \Big] dt. \end{aligned}$$

By Fubini's theorem we change the order of integration and then multiply by ε:

$$\begin{aligned} &\varepsilon \int_{\mathbb{R}^n} h \, A(h) \int_{\mathcal{M}}^\infty \varphi_\varepsilon(t) \, dt \, dx \le C \, \varepsilon \int_{\mathbb{R}^n} A(h) \int_0^{\mathcal{M}} t \, \varphi_\varepsilon(t) \, dt \, dx \\ &+ C \, \varepsilon \int_{\mathbb{R}^n} h \, \Gamma(g, h) \int_{\mathcal{M}}^\infty \varphi_\varepsilon(t) \, dt \, dx + C \, \varepsilon \int_{\mathbb{R}^n} \Gamma(g, h) \int_0^{\mathcal{M}} t \, \varphi_\varepsilon(t) \, dt \, dx, \end{aligned} \tag{3.4}$$

where $C = C(n, m, \alpha, \beta)$. Observe that $h \le \mathcal{M}$ pointwise. We shall now examine each of the four terms in (3.4). From (1.9) we get

$$\varepsilon \int_{\mathbb{R}^n} h \, A(h) \int_{\mathcal{M}}^\infty \varphi_\varepsilon(t) \, dt \, dx = \int_{\mathbb{R}^n} h \, A(h) \, A^{-\varepsilon}(\mathcal{M}) \, dx.$$

By $A(h) \le A(\mathcal{M})$ and the definition of Φ_ε, see (1.8), it follows that

$$C\,\varepsilon \int_{\mathbb{R}^n} A(h) \int_0^{\mathcal{M}} t\,\varphi_\varepsilon(t)\,dt\,dx \le C\,\varepsilon \int_{\mathbb{R}^n} \Phi_\varepsilon(\mathcal{M})\,dx.$$

Lemma 1.2 together with formula (1.9) yield

$$\begin{aligned} &C\,\varepsilon \int_{\mathbb{R}^n} h\,\Gamma(g,h) \int_{\mathcal{M}}^{\infty} \varphi_\varepsilon(t)\,dt\,dx \\ &\quad \le 2\sigma\, C \int_{\mathbb{R}^n} h\,A(h)\,A^{-\varepsilon}(\mathcal{M})\,dx + 2\sigma^{-l-1} C \int_{\mathbb{R}^n} g\,A^{1-\varepsilon}(g)\,dx, \end{aligned}$$

for all $0 < \sigma \le 1$. The last integral in (3.4) is estimated by using Lemma 1.3,

$$C\,\varepsilon \int_{\mathbb{R}^n} \Gamma(g,h) \int_0^{\mathcal{M}} t\varphi_\varepsilon(t)\,dt\,dx \le 2C\varepsilon \int_{\mathbb{R}^n} \Phi_\varepsilon(\mathcal{M})\,dx + \frac{2Cl\varepsilon}{1-l\varepsilon} \int_{\mathbb{R}^n} gA^{1-\varepsilon}(g)\,dx.$$

This holds for all $0 < \varepsilon < 1/(3l)$. Put $\sigma = 1/(4C)$ to conclude from (3.4) that

$$\begin{aligned} &\frac{1}{2} \int_{\mathbb{R}^n} h\,A(h)\,A^{-\varepsilon}(\mathcal{M})\,dx \\ &\quad \le 3C\,\varepsilon \int_{\mathbb{R}^n} \Phi_\varepsilon(\mathcal{M})\,dx + \big[C + 2C(4C)^{l+1}\big] \int_{\mathbb{R}^n} g\,A^{1-\varepsilon}(g)\,dx. \end{aligned}$$

The last step in our proof is to apply maximal inequalities (2.3) and (2.2):

$$\begin{aligned} &\int_{\mathbb{R}^n} h\,A^{1-\varepsilon}(h)\,dx \le 10^{2+nl+l^2} \int_{\mathbb{R}^n} h\,A(h)\,A^{-\varepsilon}(\mathcal{M})\,dx \\ &\le 6C \cdot 10^{2+nl+l^2}\,\varepsilon \int_{\mathbb{R}^n} \Phi_\varepsilon(\mathcal{M})\,dx + 2\big[C + 2C(4C)^{l+1}\big]10^{2+nl+l^2} \int_{\mathbb{R}^n} gA^{1-\varepsilon}(g)\,dx \\ &\le \varepsilon\,C_1 \int_{\mathbb{R}^n} h\,A^{1-\varepsilon}(h)\,dx + C_1 \int_{\mathbb{R}^n} g\,A^{1-\varepsilon}(g)\,dx, \end{aligned}$$

where $C_1 = C_1(n,m,l,\alpha,\beta)$. Finally, we choose $\varepsilon = 1/(2C_1)$ to conclude with (3.2), where $C(\mathcal{A}) = C_1(n,m,l,\alpha,\beta)$. This completes the proof of Proposition 3.1.

Proof of Theorem 1 for $\Omega = \mathbb{R}^n_+$. With the aid of a reflection we shall now carry over inequality (3.2) to $\Omega = \mathbb{R}^n_+$. For notational simplicity, we confine ourselves to the scalar case, that is, $m = 1$. Denoting $w = u - v \in W^1_0 L_N(\mathbb{R}^n_+, \mathbb{R})$, equation (0.7) takes the form

$$\operatorname{div} \mathcal{A}(x, f + \nabla w) = 0, \tag{3.5}$$

where $f = \nabla v \in L_N(\mathbb{R}^n_+, \mathbb{R}^n)$. Let $\rho\colon \mathbb{R}^n \to \mathbb{R}^n$ denote the reflection about the hyperplane $x_n = 0$, that is $\rho(x_1, \dots, x_{n-1}, x_n) = (x_1, \dots, x_{n-1}, -x_n)$. We then extend equation (3.5) to $\mathbb{R}^n$ by setting

$$\widetilde{w}(x) = \begin{cases} w(x) & \text{in } \mathbb{R}^n_+ \\ -w(\rho x) & \text{in } \mathbb{R}^n_-. \end{cases}$$

Clearly, $\nabla\widetilde{w} \in L_N(\mathbb{R}^n, \mathbb{R}^n)$ and we have the formula

$$\nabla\widetilde{w}(x) = \begin{cases} \nabla w(x) & \text{in } \mathbb{R}^n_+ \\ -\rho\nabla w(\rho x) & \text{in } \mathbb{R}^n_-. \end{cases}$$

This also suggest how to extend the vector field f. Accordingly,

$$\widetilde{f}(x) = \begin{cases} f(x) & \text{in } \mathbb{R}^n_+ \\ -\rho f(\rho x) & \text{in } \mathbb{R}^n_-. \end{cases}$$

By change of variables we obviously have

$$\int_{\mathbb{R}^n} N(\widetilde{f})\,dx = 2\int_{\mathbb{R}^n_+} N(f)\,dx \quad \text{and} \quad \int_{\mathbb{R}^n} N(\nabla\widetilde{w})\,dx = 2\int_{\mathbb{R}^n_+} N(\nabla w)\,dx. \tag{3.6}$$

Now we extend $\mathcal{A}(x,\xi)$ to all $x \in \mathbb{R}^n$ and $\xi \in \mathbb{R}^n$,

$$\widetilde{\mathcal{A}}(x,\xi) = \begin{cases} \mathcal{A}(x,\xi) & \text{if } x \in \mathbb{R}_+ \\ -\rho\mathcal{A}(\rho x, -\rho\xi) & \text{if } x \in \mathbb{R}_-. \end{cases}$$

It is easily seen that the mapping $\widetilde{\mathcal{A}}\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ satisfies conditions (0.9)-(0.11) with the same constants α and β. Consider the vector field F in $\mathbb{R}^n$ given by

$$F(x) = \widetilde{\mathcal{A}}\big(x, \widetilde{f}(x) + \nabla\widetilde{w}(x)\big).$$

Clearly, $\operatorname{div} F(x) = 0$ in $\mathbb{R}^n_+$ in the sense of distributions. For $x \in \mathbb{R}^n_-$ we find that $F(x) = -\rho F(\rho x)$, thus F is divergence free in $\mathbb{R}^n_-$ as well. It is an easy exercise to verify that F is in fact divergence free on the entire space $\mathbb{R}^n$, that is

$$\int_{\mathbb{R}^n} \langle F(x), \nabla\varphi(x)\rangle\,dx = 0,$$

for all Lipschitz functions φ with compact support. Indeed, this integral splits as

$$\begin{aligned} &\int_{\mathbb{R}^n_+} \langle F(x), \nabla\varphi(x)\rangle\,dx - \int_{\mathbb{R}^n_-} \langle \rho F(\rho x), \nabla\varphi(x)\rangle\,dx \\ &\quad = \int_{\mathbb{R}^n_+} \langle F(x), \nabla\varphi(x)\rangle\,dx - \int_{\mathbb{R}^n_+} \langle F(x), \rho\nabla\varphi(\rho x)\rangle\,dx \\ &\quad = \int_{\mathbb{R}^n_+} \langle F(x), \nabla\psi(x)\rangle\,dx = 0, \end{aligned}$$

where we notice that $\psi(x) = \varphi(x) - \varphi(\rho x)$ vanishes on $\partial\mathbb{R}^n_+ = \mathbb{R}^{n-1}$. In other words, we arrive at the $\widetilde{\mathcal{A}}$-harmonic equation on $\mathbb{R}^n$,

$$\operatorname{div} \widetilde{\mathcal{A}}(x, \widetilde{f} + \nabla\widetilde{w}) = 0.$$

We refer the reader to [13] for more details about reflection principle for nonlinear $\mathcal{A}$-harmonic equations. Now, by Proposition 3.1 and identities (3.6), we obtain

$$\int_{\mathbb{R}^n_+} |\nabla w|\, A^{1-\varepsilon}\big(|\nabla w|\big)\, dx \le C(\mathcal{A}) \int_{\mathbb{R}^n_+} |\nabla v|\, A^{1-\varepsilon}\big(|\nabla v|\big)\, dx.$$

Finally, since $\nabla u = \nabla v - \nabla w$, a pointwise estimate holds:

$$|\nabla u|\, A^{1-\varepsilon}\big(|\nabla u|\big) \le 2^{l+1} |\nabla v|\, A^{1-\varepsilon}\big(|\nabla v|\big) + 2^{l+1} |\nabla w|\, A^{1-\varepsilon}\big(|\nabla w|\big).$$

Therefore,

$$\int_{\mathbb{R}^n_+} |\nabla u|\, A^{1-\varepsilon}\big(|\nabla u|\big)\, dx \le \big[1 + C(\mathcal{A})\big]\, 2^{l+1} \int_{\mathbb{R}^n_+} |\nabla v|\, A^{1-\varepsilon}\big(|\nabla v|\big)\, dx$$

as desired.

4. Proof of Theorem 1. Our next objective is to prove estimate (0.16) for Ω a cube in $\mathbb{R}^n$. We confine ourselves to the scalar case $m = 1$, the vectorial case being left to the reader. There is no loss of generality in assuming that Ω is the unit cube $[0,1]^n$. As before, Theorem 1 will be proved once we establish the estimate

$$\int_{\Omega} |\nabla w|\, A^{1-\varepsilon}\big(|\nabla w|\big)\, dx \le C(\mathcal{A}) \int_{\Omega} |f|\, A^{1-\varepsilon}\big(|f|\big)\, dx, \tag{4.1}$$

for all $w \in C_0^\infty(\Omega)$ and $f \in L_N(\Omega, \mathbb{R}^n)$, $N(t) = t\, A^{1-\varepsilon}(t)$, which satisfy the nonhomogeneous $\mathcal{A}$-harmonic equation

$$\operatorname{div} \mathcal{A}(x, f + \nabla w) = 0 \qquad \text{in } \Omega. \tag{4.2}$$

The proof falls naturally into three steps.

Step 1. We extend equation (4.2) to the cube $\mathbf{Q} = [-1,1]^n$ by reflecting it through the hyperplanes $\{x : x_i = 0\}$, $i = 1, \dots, n$. This is done in much the same method as in the case of the half space $\Omega = \mathbb{R}^n_+$, the only difference being in the choice of the coordinate axes. For simplicity of notation, we use the same letters f, w and $\mathcal{A}$ for the extended functions. Thus

$$\operatorname{div} \mathcal{A}(x, f + \nabla w) = 0 \qquad \text{in } \mathbf{Q} = [-1,1]^n. \tag{4.3}$$

We see at once that

$$\int_{\mathbf{Q}} N(\nabla w)\, dx = 2^n \int_{\Omega} N(\nabla w)\, dx \quad \text{and} \quad \int_{\mathbf{Q}} N(f)\, dx = 2^n \int_{\Omega} N(f)\, dx.$$

This is only a preliminary extension. Continuing in this fashion we extend equation (4.3) to the entire space $\mathbb{R}^n$ by reflecting it through the faces of the cube $\mathbf{Q}$. Actually, this leads to a periodic extension. To be more precise, for each n-tuple of integers $\mathbf{m} = (m_1, \dots, m_n)$ we denote

$$\mathbf{Q_m} = \mathbf{Q} + 2\mathbf{m} = \big\{x : \ \|x - 2\mathbf{m}\| \leq 1\big\}.$$

These cubes are congruent to $\mathbf{Q}_0 = \mathbf{Q}$ and are centered at $2\mathbf{m}$. Hereafter we use the max-norm $\|x\| = \max\limits_{1\leq i\leq n} |x_i|$. As before, we continue to write f, w and $\mathcal{A}$ for the functions extended to $\mathbb{R}^n$. Thus, for each lattice point $\mathbf{m} = \pm e_i$, $i = 1, \dots, n$, we have

$$w(x + 2\mathbf{m}) = w(x)\,, \quad f(x + 2\mathbf{m}) = f(x) \quad \text{and} \quad \mathcal{A}(x + 2\mathbf{m}, \xi) = \mathcal{A}(x, \xi).$$

Moreover,

$$\operatorname{div} \mathcal{A}(x, f + \nabla w) = 0 \qquad \text{in } \mathbb{R}^n.$$

Step 2. We shall prepare to use arguments from the proof of Proposition 3.1. Fix an arbitrary integer $k > 2$ and set

$$U_k = \bigcup_{\|\mathbf{m}\|\leq k} \mathbf{Q_m}.$$

Of course U_k is a closed cube composed of $(2k + 1)^n$ cubes congruent to $\mathbf{Q}$. Hence

$$\int_{U_k} N(\nabla w)\, dx = (2k + 1)^n \int_{\mathbf{Q}} N(\nabla w)\, dx = (4k + 2)^n \int_{\Omega} N(\nabla w)\, dx.$$

Similarly,

$$\int_{U_k} N(f)\, dx = (4k + 2)^n \int_{\Omega} N(f)\, dx.$$

We shall now truncate the functions w and f to be supported in U_k,

$$w_k = \begin{cases} w & \text{in } U_k \\ 0 & \text{in } \mathbb{R}^n \setminus U_k, \end{cases} \qquad f_k = \begin{cases} f & \text{in } U_k \\ 0 & \text{in } \mathbb{R}^n \setminus U_k. \end{cases}$$

Clearly, $f_k \in L_N(\mathbb{R}^n, \mathbb{R}^n)$ and $w_k \in C_0^\infty(\mathbb{R}^n)$. Actually, the support of w_k lies in the interior of U_k. To be more specific denote $\delta = \operatorname{dist}(\partial\Omega, \operatorname{supp} w)$. Then, because of periodicity of the extension, we have

$$\delta = \operatorname{dist}(\partial U_k, \operatorname{supp} w) > 0, \tag{4.4}$$

where we emphasize that δ is independent of the integer k. Now we shall mimic the notation of the proof of Proposition 3.1 by setting $h_k = |\nabla w_k|$ and $g_k = |f_k|$. Note that h_k is $2\mathbf{m}$-periodic in the region U_k for all $\mathbf{m} = \pm e_i$ $(i = 1, \dots, n)$, but not in $\mathbb{R}^n$. As before, we shall apply Lemma 3.1 with w_k in place of w. Accordingly, for each k and $t > 0$ there exists a Lipschitz function $\eta_{k,t}$ on $\mathbb{R}^n$ with compact support, such that

$$\left|\nabla\eta_{k,t}(x)\right| \le C(n)\, t \qquad \text{for all } x \in \mathbb{R}^n.$$

Moreover, $\eta_{k,t}(x) = w_k(x)$ whenever $\mathcal{M}h_k(x) \le t$ and $x \in \mathbb{R}^n$. However, as things develop afterwards this latter condition will be required to hold only for $x \in U_k$. Therefore, we may redefine some of the functions $\eta_{k,t}$, simply putting $\eta_{k,t} \equiv 0$ if $\mathcal{M}h_k(x) > t$ for all $x \in U_k$. In conclusion, we obtain that for each $k > 2$ and $t > 0$,

$$\begin{aligned} &\eta_{k,t}(x) = w(x) && \text{if } x \in U_k \text{ with } \mathcal{M}h_k(x) \le t, \\ &\eta_{k,t} \equiv 0 && \text{if the set } \left\{x \in U_k : \mathcal{M}h_k(x) \le t\right\} \text{ is empty.} \end{aligned}$$

Lemma 4.1. *The maximal function $\mathcal{M}h_k$ is $2\mathbf{m}$-periodic on U_{k-1} for all $\mathbf{m} = \pm e_i$, $i = 1, \dots, n$.*

Proof. We want to show that $\mathcal{M}h_k(x) = \mathcal{M}h_k(x + 2\mathbf{m})$ whenever x and $x + 2\mathbf{m}$ belong to U_{k-1}. Since h_k vanishes outside U_k, it is clear that

$$\mathcal{M}h_k(x) = \sup\Big\{ ⨍_D h; x \in D \subset U_k \Big\},$$

where the supremum is taken only over the cubes $D \subset U_k$. Fix such a cube D. Two cases are possible.

Case 1. $x + 2\mathbf{m} \in D$. Then by trivial means we have

$$⨍_D h_k \le \mathcal{M}h_k(x + 2\mathbf{m}).$$

Case 2. $x + 2\mathbf{m} \notin D$. A simple geometric observation shows that $D + 2\mathbf{m}$ is also contained in U_k. Then, because of $2\mathbf{m}$-periodicity of h, we obtain

$$⨍_D h_k = ⨍_D h = ⨍_{D+2\mathbf{m}} h = ⨍_{D+2\mathbf{m}} h_k \le \mathcal{M}h_k(x + 2\mathbf{m}).$$

In conclusion, $\mathcal{M}h_k(x) \le \mathcal{M}h_k(x + 2\mathbf{m})$. The opposite inequality follows because we can interchange the points x and $x+2\mathbf{m}$. This completes the proof of the lemma.

Lemma 4.2. *There exist constants T and L, independent of k, such that*

$$\eta_{k,t} \textit{ is supported in } U_k \textit{ if } t \ge T, \quad |\eta_{k,t}| \le L \textit{ in } U_k \textit{ if } t < T. \tag{4.5}$$

Proof. Let us fix an arbitrary point $x \in \mathbb{R}^n \backslash U_k$ and a cube $D \subset \mathbb{R}^n$ containing x. Since h_k vanishes outside U_k, we have

$$⨍_D h_k = 0 \qquad \text{if } l(D) \le \delta,$$

where $l(D)$ denotes the length of the edges of D and δ is given by (4.4). On the other hand, because of the periodicity of h, we find that for $l(D) > \delta$,

$$\fint_D h_k \leq \fint_D h \leq 2^n \delta^{-n} \fint_\Omega h.$$

In either case we conclude with the estimate

$$\mathcal{M}h_k(x) \leq 2^n \delta^{-n} \fint_\Omega |\nabla w| = T,$$

the latter being the definition of T. Now, for $t \geq T$ we have $\eta_{k,t}(x) = 0$, because it either coincides with $w_k(x)$ which vanishes outside U_k or is zero identically.

In order to prove (4.5), we fix $t < T$ and assume that the set $U_k \cap \big\{x : \mathcal{M}h_k(x) \leq t\big\}$ is not empty, otherwise $\eta_{k,t} \equiv 0$ and there is nothing to prove. Because of periodicity of $\mathcal{M}h_k$ in U_{k-1}, see Lemma 4.1, there are congruent points, say, $x_{\mathbf{2m}} \in \mathbf{Q}_{\mathbf{2m}} \subset U_{k-1}$, for all $\|\mathbf{m}\| \leq k-1$, such that $\mathcal{M}h_k(x_{\mathbf{2m}}) \leq t$. This means that $\eta_{k,t}(x_{\mathbf{2m}}) = w_k(x_{\mathbf{2m}}) = w(x_{\mathbf{2m}}) = w(x_{\mathbf{o}})$. Finally, for each $x \in U_k$ one can find at least one of the points $x_{\mathbf{2m}}$ to be close enough to x so that $\|x - x_{\mathbf{2m}}\| \leq 4$. Hence

$$\big|\eta_{k,t}(x)\big| \leq \big|\eta_{k,t}(x_{\mathbf{2m}})\big| + \|x - x_{\mathbf{2m}}\| \, \|\nabla \eta_{k,t}\|_\infty \leq \big|w(x_{\mathbf{o}})\big| + C(n)\, T = L.$$

Step 3. We shall mimic the proof of Proposition 3.1. Instead of identity (3.3), we have the following inequality for all t:

$$\int_{U_k} \langle \mathcal{A}(x, f + \nabla w), \nabla \eta_{k,t} \rangle \, dx \leq C\, k^{n-1} \chi_T(t), \tag{4.6}$$

where C is a constant independent of k and t and χ_T denotes the characteristic function of the interval $(0, T)$; see Lemma 4.2 for the definition of T.

Indeed, the case $t \geq T$ is obvious, because $\eta_{k,t}$ is compactly supported in U_k and the vector field $\mathcal{F}(x) = \mathcal{A}(x, f + \nabla w)$ is divergence free. Thus in this case the integral in (4.6) vanishes, as in identity (3.3). A difference is forthcoming when $t < T$. In this case $\big|\eta_{k,t}(x)\big| \leq L$ for all $x \in U_k$; see Lemma 4.2. Let $\varphi \in C_0^\infty(U_k)$ be a nonnegative function such that $\varphi \equiv 1$ on U_{k-1} and $\big|\nabla \varphi(x)\big| \leq C(n)$. By using Green's formula, we then compute

$$\begin{aligned}
\int_{U_k} \langle \mathcal{F}, \nabla \eta_{k,t} \rangle \, dx &= \int_{U_k} \langle \varphi\, \mathcal{F}, \nabla \eta_{k,t} \rangle \, dx + \int_{U_k} \langle (1 - \varphi)\, \mathcal{F}, \nabla \eta_{k,t} \rangle \, dx \\
&\leq -\int_{U_k} \eta_{k,t} \operatorname{div} (\varphi\, \mathcal{F}) \, dx + \int_{U_k \setminus U_{k-1}} |\mathcal{F}| \, |\nabla \eta_{k,t}| \, dx \\
&\leq -\int_{U_k} \eta_{k,t} \langle \nabla \varphi, \mathcal{F} \rangle \, dx + C(n)\, t \int_{U_k \setminus U_{k-1}} |\mathcal{F}| \, dx \\
&\leq C(n)\, L \int_{U_k \setminus U_{k-1}} |\mathcal{F}| \, dx + C(n)\, T \int_{U_k \setminus U_{k-1}} |\mathcal{F}| \, dx.
\end{aligned}$$

The number of cubes $\mathbf{Q_m}$ which are contained in $U_k \backslash U_{k-1}$ is equal to $(2k+1)^n - (2k-1)^n \leq C(n) k^{n-1}$. Since $\mathcal{F}$ is $2\mathbf{m}$-periodic, the integrals of $|\mathcal{F}|$ over such cubes are the same. Hence

$$\int_{U_k} \langle \mathcal{F}, \nabla \eta_{k,t} \rangle \, dx \leq C(n)\, (L+T)\, K^{n-1} \int_{\mathbf{Q}} |\mathcal{F}| \, dx = C\, K^{n-1},$$

which implies inequality (4.6).

Now the rest of the proof of estimate (4.1) runs in much the same way as the proof of Proposition 3.1. Inequality (3.4) takes the form

$$\begin{aligned} \varepsilon \int_{U_k} h\, A(h) \int_{\mathcal{M}h_k}^{\infty} \varphi_e(t)\, dt\, dx &\leq C\, \varepsilon \int_{U_k} A(h) \int_0^{\mathcal{M}h_k} t\, \varphi_\varepsilon(t)\, dt\, dx \\ &+ C\, \varepsilon \int_{U_k} h\, \Gamma(g,h) \int_{\mathcal{M}h_k}^{\infty} \varphi_\varepsilon(t)\, dt\, dx + C\, \varepsilon \int_{U_k} \Gamma(g,h) \int_0^{\mathcal{M}h_k} t\, \varphi_\varepsilon(t)\, dt\, dx \\ &+ C\, K^{n-1} \varepsilon \int_0^{\infty} \varphi_\varepsilon(t)\, \chi_T(t)\, dt. \end{aligned}$$

In this inequality we may replace h and g by h_k and g_k, respectively. The inequality also remains valid if we replace U_k by the entire space $\mathbb{R}^n$. The integrals in the right hand side will possibly increase, while the integral in the left hand side will not change, because h_k vanishes outside U_k. Having disposed with this inequality we can now proceed analogously to the proof of Proposition 3.1 to conclude

$$\int_{\mathbb{R}^n} h_k\, A^{1-\varepsilon}(h_k)\, dx \leq C(\mathcal{A}) \int_{\mathbb{R}^n} g_k\, A^{1-\varepsilon}(g_k)\, dx + C\, K^{n-1}.$$

Here we notice that

$$\int_{\mathbb{R}^n} h_k\, A^{1-\varepsilon}(h_k)\, dx = \int_{U_k} N(h)\, dx = (4k+2)^n \int_{\Omega} N\big(|\nabla w|\big)\, dx.$$

Similarly,

$$\int_{\mathbb{R}^n} g_k\, A^{1-\varepsilon}(g_k)\, dx = \int_{U_k} N\big(|f|\big)\, dx = (4k+2)^n \int_{\Omega} N\big(|f|\big)\, dx.$$

Finally, dividing by $(4k+2)^n$ and letting k go to ∞ we obtain

$$\int_{\Omega} |\nabla w|\, A^{1-\varepsilon}\big(|\nabla w|\big)\, dx \leq C(\mathcal{A}) \int_{\Omega} |f|\, A^{1-\varepsilon}\big(|f|\big)\, dx$$

with the same constant $C(\mathcal{A})$ as in Proposition 3.1.

Remark 4.1. The proof above gives more than Theorem 1, as f need not be a gradient field.

Changing variables in (0.14) leads to further extensions of the result to domains such as balls in $\mathbb{R}^n$ or other type of domains which are obtained from a cube via a bilipschitz transformation.

5. Gehring's Lemma in Orlicz spaces. Let $F\colon \mathbb{R}_+ \to \mathbb{R}_+$ be a convex function of the form $F(t) = t\,A(t)$, where A is an increasing function with $A(0) = 0$. Throughout we assume that A satisfies the Δ_2-condition, that is,

$$A(2t) \le K\,A(t), \tag{5.1}$$

for some $K \ge 1$ and all $t \ge 0$. Notice the inequality for the inverse function,

$$F^{-1}(x+y) \le F^{-1}(x) + F^{-1}(y)\,, \qquad x, y \in \mathbb{R}_+. \tag{5.2}$$

We also have

$$x\,A(y) \le F(x) + F(y)\,, \qquad x, y \in \mathbb{R}_+. \tag{5.3}$$

The latter is an immediate consequence of the monotonicity of A. This implies another useful inequality,

$$F^{-1}(x) \le t + \frac{x}{A(t)}, \tag{5.4}$$

for every $t > 0$ and all $x \ge 0$. The Orlicz space $L_F(\mathbb{R}^n)$ consists of all functions $g \in L^1_{\text{loc}}(\mathbb{R}^n)$ such that

$$[g]_F = F^{-1}\Big[\int_{\mathbb{R}^n} F\big(|g|\big)\,dx\Big] < \infty.$$

Proof of Lemma 1. The general idea of the proof is in the same spirit as in the celebrated work of F.W. Gehring [7]. Fix an arbitrary positive number t and define

$$s = 2(a+b)t. \tag{5.5}$$

We then use the Calderòn-Zygmund decomposition

$$\mathbb{R}^n = \Big[\bigcup_{j=1}^{\infty} Q_j\Big] \cup \Big[\mathbb{R}^n \setminus \bigcup_{j=1}^{\infty} Q_j\Big],$$

where Q_j are pairwise disjoint cubes such that

$$F(s) \le \fint_{Q_j} F(g)\,dx \le 2^n F(s)\,, \qquad j = 1, 2, \ldots \tag{5.6}$$

and

$$\big|g(x)\big| \le s \qquad \text{for almost every } x \in \mathbb{R}^n \setminus \bigcup_{j=1}^{\infty} Q_j. \tag{5.7}$$

Denote the distribution sets of g and h by

$$G_s = \{x \in \mathbb{R}^n :\ g(x) > s\} \qquad \text{and} \qquad H_s = \{x \in \mathbb{R}^n :\ h(x) > s\}.$$

Thus (5.7) implies that $G_s \subset \cup_j Q_j$. While, in view of (5.6), we obtain

$$\int_{G_s} F(g)\,dx \le \sum_{j=1}^{\infty} \int_{Q_j} F(g)\,dx \le 2^n F(s) \sum_{j=1}^{\infty} |Q_j|. \tag{5.8}$$

Next we are going to estimate the measure of the cube Q_j by using inequalities (5.6) and (0.17). Accordingly,

$$s = F^{-1}\big(F(s)\big) \le F^{-1}\big[\fint_{Q_j} F(g)\,dx\big] \le a \fint_{4Q_j} g\,dx + b\,F^{-1}\big[\fint_{4Q_j} F(h)\,dx\big]. \tag{5.9}$$

Obviously, we have

$$\fint_{4Q_j} g\,dx \le \frac{1}{|4Q_j|} \int_{4Q_j \cap G_t} g\,dx + t. \tag{5.10}$$

Similarly we shall estimate the second term of the right hand side of (5.9). Using inequalities (5.2) and (5.4) we obtain

$$\begin{aligned} F^{-1}\big[\fint_{4Q_j} F(h)\,dx\big] &\le F^{-1}\big[\frac{1}{|4Q_j|} \int_{4Q_j \cap H_t} F(h)\,dx + F(t)\big] \\ &\le F^{-1}\big[\frac{1}{|4Q_j|} \int_{4Q_j \cap H_t} F(h)\,dx\big] + t \le t + \frac{1}{|4Q_j|\,A(t)} \int_{4Q_j \cap H_t} F(h)\,dx + t. \end{aligned} \tag{5.11}$$

Estimates (5.10) and (5.11) combined with (5.9) yield

$$s \le (a+2b)t + \frac{a}{|4Q_j|} \int_{4Q_j \cap G_t} g\,dx + \frac{b}{|4Q_j|\,A(t)} \int_{4Q_j \cap H_t} F(h)\,dx,$$

which in view of (5.5) implies

$$at|4Q_j| \le a \int_{4Q_j \cap G_t} g\,dx + \frac{b}{A(t)} \int_{4Q_j \cap H_t} F(h)\,dx. \tag{5.12}$$

Now, by the well known Vitali covering argument, we can extract from $\{4Q_j\}_{j\in\mathbb{N}}$ pairwise disjoint cubes, say $\{4Q_{j_\nu}\}$, such that

$$\big|\bigcup_{j=1}^{\infty} 4Q_j\big| \le 3^n \sum_{\nu} |4Q_{j_\nu}|.$$

Summing up inequalities (5.12) yields

$$at\sum_{j=1}^{\infty}|Q_j| \leq at\Big|\bigcup_{j=1}^{\infty}4Q_j\Big| \leq 3^n at\sum_{\nu}\big|4Q_{j_\nu}\big|$$
$$\leq 3^n a\sum_{\nu}\int_{4Q_{j_\nu}\cap G_t} g\,dx + \frac{3^n b}{A(t)}\sum_{\nu}\int_{4Q_{j_\nu}\cap H_t} F(h)\,dx.$$

Finally, we obtain an estimate for the measure of the set $\cup_j Q_j$,

$$\sum_{j=1}^{\infty}|Q_j| \leq \frac{3^n}{t}\int_{G_t} g\,dx + \frac{3^n b}{a\,F(t)}\int_{H_t} F(h)\,dx.$$

We then substitute into inequality (5.8),

$$\int_{G_s} F(g)\,dx \leq \frac{6^n F(s)}{t}\int_{G_t} g\,dx + \frac{6^n b\,F(s)}{a\,F(t)}\int_{H_t} F(h)\,dx. \tag{5.13}$$

On the other hand we have the trivial estimate

$$\int_{G_t\setminus G_s} F(g)\,dx = \int_{G_t\setminus G_s} g\,A(g)\,dx \leq A(s)\int_{G_t} g\,dx.$$

Adding to (5.13), in view of (5.5) and the formula $F(t) = t\,A(t)$, we obtain

$$\int_{G_t} F(g)\,dx$$
$$\leq \big(1 + 2\cdot 6^n a + 2\cdot 6^n b\big)A(s)\int_{G_t} g\,dx + \big(2\cdot 6^n b + 2\cdot 6^n b/a\big)\frac{A(s)}{A(t)}\int_{H_t} F(h)\,dx.$$

Next observe that the Δ_2-condition (5.1) implies

$$A(s) = A\big[(2a+2b)t\big] \leq C(a,b,K)\,A(t).$$

We then estimate

$$\int_{G_t} g\,A(g)\,dx \leq \alpha\,A(t)\int_{G_t} g\,dx + \beta\int_{H_t} h\,A(h)\,dx, \tag{5.14}$$

where $\alpha = \alpha(n,a,b,K) > 1$ and $\beta = \beta(n,a,b,K)$ are independent of t and the functions g and h.

Let us introduce a continuously decreasing function

$$\varphi(t) = \int_{G_t} g\,dx \leq \frac{1}{A(t)}\int_{\mathbb{R}^n} F(g)\,dx < \infty.$$

The formula

$$\int_{G_t} g\, A^p(g)\, dx = -\int_t^\infty A^p(s)\, d\varphi(s) \tag{5.15}$$

holds at least for $p = 1$. The integral in the right hand side is of Stieltjes type. Inequality (5.14) reads

$$-\int_t^\infty A(s)\, d\varphi(s) \le \alpha\, A(t)\, \varphi(t) + \beta \int_{H_t} h\, A(h)\, dx. \tag{5.16}$$

Our next step is to deduce from (5.16) that

$$-\int_0^\infty A^p(s)\, d\varphi(s) \le \frac{p\beta}{\alpha - \alpha p + p} \int_0^\infty A^{p-1}(t) \big[-\int_{H_t} h\, A(h)\, dx\big] dt, \tag{5.17}$$

for all $p \in [1, \alpha/(\alpha-1)[$. We shall integrate by parts. Some difficulty arises, since we do not know if the integral $\int A^p d\varphi$ is actually converging. To avoid this problem, we first replace φ by its cut-off function

$$\varphi_T(s) = \begin{cases} \varphi(s) & \text{if } 0 \le s \le T \\ 0 & \text{if } s > T. \end{cases}$$

Notice that inequality (5.16) is preserved, namely

$$-\int_t^\infty A(s)\, d\varphi_T(s) \le \alpha A(t)\, \varphi_T(t) + \beta \int_{H_t} h\, A(h)\, dx, \tag{5.16$_T$}$$

for all $t \ge 0$. Now the integral $\int_0^\infty A^p(s)\, d\varphi_T(s)$ is converging for all $p \ge 1$, so we can perform the computations

$$\begin{aligned} -\int_0^\infty A^p(s)\, d\varphi_T(s) &= \int_0^\infty A^{p-1}(t) \big[\int_t^\infty A(s)\, d\varphi_T(s)\big]' dt \\ &= (p-1) \int_0^\infty A^{p-2}(t)\, A'(t) \big[-\int_t^\infty A(s)\, d\varphi_T(s)\big] dt \\ &\le \alpha\, (p-1) \int_0^\infty A^{p-1}(t)\, A'(t)\, \varphi_T(t)\, dt \\ &\quad + \beta\, (p-1) \int_0^\infty A^{p-2}(t)\, A'(t) \big[\int_{H_t} h\, A(h)\, dx\big] dt \\ &= -\frac{\alpha(p-1)}{p} \int_0^\infty A^p(t)\, d\varphi_T(t) - \beta \int_0^\infty A^{p-1}(t) \big[\int_{H_t} h\, A(h)\, dx\big]' dt. \end{aligned}$$

Hence for $p \in [1, \alpha/(\alpha - 1)[$,

$$-\int_0^\infty A^p(s)\, d\varphi_T(s) \le \frac{p\,\beta}{\alpha - \alpha\, p + p} \int_0^\infty A^{p-1}(t) \big[-\int_{H_t} h\, A(h)\, dx\big]' dt.$$

Inequality (5.17) follows letting T go to ∞. Now formula (5.15) can be applied giving the estimate

$$\int_{\mathbb{R}^n} g\, A^p(g)\, dx \leq \frac{p\,\beta}{\alpha - \alpha\, p + p} \int_{\mathbb{R}^n} h\, A^p(h)\, dx.$$

This completes the proof.

Remark 5.1. Inequality (0.18) can be generalized in the following way. Consider a mapping $\mathcal{G}\colon [0,\infty[\, \to [0,\infty[$, with $\mathcal{G}(0) = 0$ and $\mathcal{G}(t)/t$ increasing and assume that there exists $p \geq 1$ such that

$$\int_0^t \frac{\mathcal{G}(\tau)}{\tau}\, d\tau \geq \frac{1}{p}\, \mathcal{G}(t), \qquad \forall t \geq 0.$$

Then, by similar calculations as above, if $p \leq \alpha/(\alpha - 1)$ we see that

$$\int_{\mathbb{R}^n} g\, \mathcal{G}\big(A(g)\big)\, dx \leq \frac{p\,\beta}{\alpha - \alpha\, p + p} \int_{\mathbb{R}^n} h\, \mathcal{G}\big(A(h)\big)\, dx.$$

6. Higher integrability of the gradient. In this section we prove Theorem 2. To be more specific, we examine the $\mathcal{A}$-harmonic equation (0.7) in a domain $\Omega \subset \mathbb{R}^n$, where $\mathcal{A}\colon \Omega \times \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m}$ satisfies the conditions

$$\big|\mathcal{A}(x,\xi)\big| \leq \beta A\big(|\xi|\big) \tag{6.1}$$

and

$$\langle \mathcal{A}(x,\xi), \xi\rangle \geq \alpha\, |\xi|\, A\big(|\xi|\big). \tag{6.2}$$

Here, as in Gehring's lemma, A is continuously increasing on $\mathbb{R}_+$, $A(0) = 0$ and $F(t) = t\, A(t)$ is convex. Also, A satisfies the Δ_2-condition (1.4), that is,

$$A(\lambda t) \leq \lambda^l A(t), \tag{6.3}$$

for all $t \geq 0$ and $\lambda \geq 1$. We give the proof only under the additional assumption that the function $t \mapsto A^n(t)$ is concave. This condition is not particularly restrictive as we are interested in no power growth of A anyway.

We begin with estimate (0.15), that is, with the case $\varepsilon = 0$. Indeed, putting $\eta = u - v$ into (0.13) yields

$$\alpha \int_\Omega F\big(|\nabla u|\big)\, dx \leq \beta \int_\Omega |\nabla v|\, A\big(|\nabla u|\big)\, dx \leq \frac{\beta}{\lambda} \int_\Omega \Big[F\big(\lambda|\nabla v|\big) + F\big(|\nabla u|\big)\Big]\, dx,$$

for all $\lambda \geq 1$. Here we applied inequality (5.3). Put $\lambda = 2\beta/\alpha$ to obtain

$$\int_\Omega F\big(|\nabla u|\big)\, dx \leq \int_\Omega F\big(2\beta\, |\nabla v|/\alpha\big)\, dx \leq \big(2\beta/\alpha\big)^{l+1} \int_\Omega F\big(|\nabla v|\big)\, dx, \tag{6.4}$$

where we used Δ_2-condition (6.3). This estimate holds without any regularity assumption on the domain Ω. To deal with the case $\varepsilon > 0$ we make the following definition.

Definition 6.1. A bounded open set $\Omega \subset \mathbb{R}^n$ is said to be regular if there is $\mu > 0$ such that for each cube $D \not\subset \Omega$ we have

$$\mu|D| \leq |2D \backslash \Omega|.$$

Observe that bounded Lipschitz domains are regular.

The proof consists of several steps. First we extend u to $\mathbb{R}^n$ by letting it be equal to v outside Ω.

Lemma 6.1. (Reverse Jensen Inequality) *There are constants $a = a > 1$ and $b = b > 0$ depending on n, l, α, β and μ such that*

$$F^{-1}\big[\fint_Q F\big(|\nabla u|\big)\,dx\big] \leq a \fint_{4Q} |\nabla u|\,dx + b\,F^{-1}\big[\fint_{4Q} F\big(|\nabla v|\big)\,dx\big],$$

for every cube $Q \subset \mathbb{R}^n$.

Proof. We distinguish two cases.

Case 1. $2Q \subset \Omega$. Take $\varphi \in C_0^\infty(2Q)$ to be a nonnegative function such that $\varphi \equiv 1$ on Q and $|\nabla\varphi| \leq C(n)/R$, with $R = \operatorname{diam} Q$. Then the function $\eta = \varphi\,(u - u_{2Q}) \in W_0^1 L_F(\Omega)$ can be used as a test function in (0.13). Here, as usual, u_{2Q} denotes the integral mean of u over the double cube $2Q$. This yields

$$\int_\Omega \varphi\,\langle \nabla u, \mathcal{A}(x, \nabla u)\rangle\,dx = -\int_\Omega (u - u_{2Q})\,\langle \nabla\varphi, \mathcal{A}(x, \nabla u)\rangle\,dx.$$

With the aid of inequalities (6.1) and (6.2) we obtain

$$\begin{aligned}\alpha \fint_Q F\big(|\nabla u|\big)\,dx &\leq \frac{\beta C(n)}{R} \fint_{2Q} |u - u_{2Q}|\,A\big(|\nabla u|\big)\,dx \\ &\leq \frac{\beta C(n)}{R}\big(\fint_{2Q} |u - u_{2Q}|^{\frac{n}{n-1}}\,dx\big)^{\frac{n-1}{n}}\big(\fint_{2Q} A^n\big(|\nabla u|\big)\,dx\big)^{\frac{1}{n}}.\end{aligned}$$

The first integral in the last term will be estimated by Sobolev-Poincaré inequality, the second one by using the concavity of the function $t \mapsto A^n(t)$ and Jensen's inequality

$$\fint_Q F\big(|\nabla u|\big)\,dx \leq C(n)\frac{\beta}{\alpha}\big[\fint_{2Q} |\nabla u|\,dx\big] A\big(\fint_{2Q} |\nabla u|\,dx\big) = C(n, \beta/\alpha)\,F\big(\fint_{2Q} |\nabla u|\,dx\big).$$

Finally, in view of the convexivity of F we conclude that

$$F^{-1}\big[\fint_Q F\big(|\nabla u|\big)\,dx\big] \leq C(n, \beta/\alpha) \fint_{2Q} |\nabla u|\,dx$$

as desired.

Case 2. Assume that $2Q \not\subset \Omega$. According to the definition of regular domain,

$$|4Q\backslash\Omega| \geq \mu|2Q|. \tag{6.6}$$

We shall test identity (0.13) with the function $\eta = \varphi\,(u-v)$, where $\varphi \in C_0^\infty(4Q)$, $0 \leq \varphi \leq 1$, $\varphi \equiv 1$ on Q and $|\nabla\varphi| \leq C(n)/R$. Notice that $u-v$ vanishes on $4Q\backslash\Omega$ and, therefore, $\eta \in W_0^1 L_F(\Omega)$ and

$$\nabla\eta = \varphi\,(\nabla u - \nabla v) + (u-v)\,\nabla\varphi.$$

Substituting into (0.13) yields

$$\int_{4Q} \varphi\,\langle \mathcal{A}(x,\nabla u), \nabla u - \nabla v\rangle\,dx = -\int_{4Q} \langle \mathcal{A}(x,\nabla u), \nabla\varphi\rangle\,(u-v)\,dx.$$

Hence, by (5.3),

$$\begin{aligned}
\alpha \int_{4Q} \varphi\,F\big(|\nabla u|\big)\,dx &\leq \int_{4Q} \varphi\,\langle \mathcal{A}(x,\nabla u), \nabla u\rangle\,dx \\
&= \int_{4Q} \varphi\,\langle \mathcal{A}(x,\nabla u), \nabla v\rangle\,dx - \int_{4Q} \langle \mathcal{A}(x,\nabla u), \nabla\varphi\rangle\,(u-v)\,dx \\
&\leq \beta \int_{4Q} \varphi\,|\nabla v|\,A\big(|\nabla u|\big)\,dx + \frac{C(n)\beta}{R} \int_{4Q} |u-v|\,A\big(|\nabla u|\big)\,dx \\
&\leq \frac{\beta}{\lambda} \int_{4Q} \varphi\,\Big[F\big(|\nabla u|\big) + F\big(\lambda|\nabla v|\big)\Big]\,dx + \frac{C(n)\beta}{R} \int_{4Q} |u-v|\,A\big(|\nabla u|\big)\,dx.
\end{aligned}$$

Letting $\lambda = 2\beta/\alpha$, in view of Δ_2-property of F, we obtain

$$\begin{aligned}
\int_{4Q} \varphi\,F\big(|\nabla u|\big)\,dx \leq C(n,l,\alpha,\beta) \int_{4Q} \varphi\,F\big(|\nabla v|\big)\,dx \\
+ \frac{C(n,l,\alpha,\beta)}{R} \int_{4Q} |u-v|\,A\big(|\nabla u|\big)\,dx.
\end{aligned} \tag{6.7}$$

Then, by Hölder's inequality,

$$\fint_Q F\big(|\nabla u|\big)\,dx \leq C \fint_{4Q} F\big(|\nabla v|\big)\,dx + C\Big[\fint_{4Q} \Big|\frac{u-v}{R}\Big|^{\frac{n}{n-1}}dx\Big]^{\frac{n-1}{n}} \Big[\fint_{4Q} A^n\big(|\nabla u|\big)\,dx\Big]^{\frac{1}{n}}, \tag{6.8}$$

where, in view of the concavity of A^n, we have

$$\Big[\fint_{4Q} A^n\big(|\nabla u|\big)\,dx\Big]^{\frac{1}{n}} \leq A\Big[\fint_{4Q} |\nabla u|\,dx\Big]. \tag{6.9}$$

What remains is to estimate the second integral in the right hand side. To this effect we appeal to the following form of the Sobolev-Poincaré inequality.

Let $w \in W^{1,p}(4Q)$, $1 \le p < n$, and let E be an arbitrary measurable subset of $4Q$ with positive measure. Then

$$\Big[\int_{4Q} \big|w(x) - w_E\big|^{\frac{np}{n-p}}\, dx\Big]^{\frac{n-p}{np}} \le C\big(n, p, |Q|/|E|\big)\Big[\int_{4Q} |\nabla w|^p\, dx\Big]^{\frac{1}{p}},$$

where w_E stands for the integral mean of w over the set E.

Take $w = u - v$ and $E = 4Q\backslash\Omega$. In view of (6.6), $|E| \ge \mu|2Q|$, thus $w_E = 0$ and

$$\frac{1}{R}\Big[\fint_{4Q}|u - v|^{\frac{n}{n-1}}\, dx\Big]^{\frac{n-1}{n}} \le C(n,\mu)\fint_{4Q}|\nabla u - \nabla v|\, dx. \tag{6.10}$$

Insert (6.10) and (6.9) into (6.8) to obtain

$$\fint_Q F\big(|\nabla u|\big)\, dx \le C\fint_{4Q} F\big(|\nabla v|\big)\, dx + C\Big[\fint_{4Q}|\nabla u|\, dx + \fint_{4Q}|\nabla v|\, dx\Big]A\Big(\fint_{4Q}|\nabla u|\, dx\Big).$$

Using inequality (5.3) yields

$$\fint_Q F\big(|\nabla u|\big)\, dx \le C\, F\Big(\fint_{4Q}|\nabla u|\, dx\Big) + C\fint_{4Q} F\big(|\nabla v|\big)\, dx.$$

In the last step we apply (5.2) and concavity of the inverse function, which gives $F^{-1}(Cx) \le C\, F^{-1}(x)$ for $C \ge 1$. Therefore,

$$F^{-1}\Big[\fint_Q F\big(|\nabla u|\big)\, dx\Big] \le C\fint_{4Q}|\nabla u|\, dx + C\, F^{-1}\Big[\fint_{4Q} F\big(|\nabla v|\big)\, dx\Big].$$

The proof of Lemma 6.1 is complete.

Proof of Theorem 2. It is an immediate consequence of Gehring's lemma and Reverse Jensen Inequality that $|\nabla u| \in L_N(\mathbb{R}^n)$ with $N(t) = t\, A^{1+\varepsilon}(t)$. Moreover

$$\int_\Omega N\big(|\nabla u|\big)\, dx \le \int_{\mathbb{R}^n} N\big(|\nabla u|\big)\, dx \le C(\mathcal{A})\int_{\mathbb{R}^n} N\big(|\nabla v|\big)\, dx.$$

Note that we did not really have to begin with the Dirichlet data v in $W^1L_N(\mathbb{R}^n, \mathbb{R}^m)$ on the entire space $\mathbb{R}^n$; we could have applied Withney type extension of a function $v \in W^1L_N(\Omega, \mathbb{R}^m)$ to $\mathbb{R}^n$ such that

$$\int_{\mathbb{R}^n} N\big(|\nabla v|\big)\, dx \le C_\Omega(n, N)\int_\Omega N\big(|\nabla v|\big)\, dx.$$

This would give the estimate

$$\int_\Omega N\big(|\nabla u|\big)\,dx \le C(\mathcal{A},\Omega)\int_\Omega N\big(|\nabla v|\big)\,dx$$

at least for some domains, such as a cube or a ball. It is of some interest to identify such domains. Unfortunately, this topic exceeds the scope of this paper.

7. Existence of weak solutions. In this paragraph, we give an example of existence theorem of solutions for the boundary value problem

$$\begin{cases} \operatorname{div}\mathcal{A}(x,f+\nabla u)=h & \text{in } \Omega \\ u\in W_0^1L_F(\Omega,\mathbb{R}), \end{cases} \tag{7.1}$$

where Ω is a bounded open subset of $\mathbb{R}^n$, the function $\mathcal{A}$ satisfies condition (0.9) and (0.11), $F(t)=t\,A(t)$, $f\in L_F(\mathbb{R}^n,\mathbb{R}^n)$ and h is a distribution on Ω. Note that if (7.1) is the Euler equation of a functional of the Calculus of Variations, then the existence follows easily by direct methods; see e.g. [6]. In the general case, we will use an abstract existence result due to Gossez and Mustonen [10].

Let X be a Banach space and X^*, X^{**} its dual and second dual. As usual, we identify X with a (norm) closed subspace of X^{**}. Let $T\colon X^*\to X^{**}$ be a mapping with the following properties.

(1) (i) (*finite continuity*) T is continuous from each finite dimensional subspace of X^* to X^{**}, for the $\sigma(X^{**},X^*)$ topology.
(2) (ii) (*sequential pseudo-monotonicity*) For any sequence (u_n) in X^*, if $u_n\rightharpoonup u$ for $\sigma(X^*,X)$, $Tu_n\overset{*}{\rightharpoonup}\chi$ for $\sigma(X^{**},X^*)$ and $\limsup_n\langle u_n,Tu_n\rangle\le\langle u,\chi\rangle$, then $Tu=\chi$ and $\langle u_n,Tu_n\rangle\to\langle u,\chi\rangle$.
(3) (iii) (*boundedness*) T maps bounded subsets into bounded subsets.
(4) (iv) (*weak coercivity*) $\langle u,Tu-h\rangle>0$ if u has sufficiently large norm.

Theorem 7.1. *If X^* is separable and T satisfies conditions* (i), (ii), (iii) *and* (iv) *with $h\in X$, then the equation*

$$Tu=h$$

has at least one solution.

Now, in order to define a suitable mapping T between Orlicz-Sobolev spaces having the properties required in Theorem 7.1, we impose besides (0.9) and (0.11) one more condition on the function $\mathcal{A}$; for almost every $x\in\Omega$, for all $\xi,\,\zeta\in\mathbb{R}^n$ with $\xi\ne\zeta$,

$$\langle\xi-\zeta,\mathcal{A}(x,\xi)-\mathcal{A}(x,\zeta)\rangle>0. \tag{7.2}$$

We denote by $\widetilde{F}$ the function conjugate to F and by $E_{\widetilde{F}}(\Omega)$ the closure in $L_{\widetilde{F}}(\Omega)$ of the set of bounded functions with compact support. Also, we consider the spaces of distributions

$$W^{-1}L_{\widetilde{F}}(\Omega) = \Big\{\ h \in \mathcal{D}'(\Omega) : h = \varphi_0 - \sum_{i=1}^{n} \frac{\partial \varphi_i}{\partial x_i}\ ,\ \varphi_i \in L_{\widetilde{F}}(\Omega)\ \Big\},$$

$$W^{-1}E_{\widetilde{F}}(\Omega) = \Big\{\ h \in \mathcal{D}'(\Omega) : h = \varphi_0 - \sum_{i=1}^{n} \frac{\partial \varphi_i}{\partial x_i}\ ,\ \varphi_i \in E_{\widetilde{F}}(\Omega)\ \Big\},$$

endowed with their usual norm as quotient spaces of $L_{\widetilde{F}}(\Omega, \mathbb{R}^n)$ and $E_{\widetilde{F}}(\Omega, \mathbb{R}^n)$, respectively. Then (see [9]), if we define the pairing between $u \in W_0^1 L_F(\Omega)$ and $h = \varphi_0 - \sum_i \partial\varphi_i/\partial x_i \in W^{-1}L_{\widetilde{F}}(\Omega)$ by

$$\langle u, \varphi\rangle = \int_\Omega \big(u\,\varphi_0 + \sum_{i=1}^{n} \frac{\partial u}{\partial x_i}\,\varphi_i\big)\,dx,$$

$W_0^1 L_F(\Omega)$ can be identified with the dual space to $W^{-1}E_{\widetilde{F}}(\Omega)$ and $W^{-1}L_{\widetilde{F}}(\Omega)$ with the dual space to $W_0^1 L_F(\Omega)$.

Consider the mapping

$$T\colon W_0^1 L_F(\Omega) \to W^{-1}L_{\widetilde{F}}(\Omega) \tag{7.3}$$

defined by

$$\langle v, Tu\rangle = \int_\Omega \langle \mathcal{A}(x, f + \nabla u), \nabla v\rangle\,dx. \tag{7.4}$$

Now we can state the existence result.

Theorem 7.2. *Let Ω be a bounded open subset of $\mathbb{R}^n$ and assume conditions* (0.9), (0.11) *and* (7.2) *for the function $\mathcal{A}$. Set $F(t) = t\,A(t)$, where A satisfies* (0.3). *Then problem* (7.1) *has at least one solution.*

Proof. The existence follows from Theorem 7.1, since for $X = W^{-1}E_{\widetilde{F}}(\Omega)$ the mapping T defined by (7.3) and (7.4) has the properties requested there. Properties (i) and (ii) follow from Lemma 4.3 and Theorem 5.1 of [9], respectively. The conditions (iii) and (iv) follow easily from our assumptions on $\mathcal{A}$ and A.

REFERENCES

[1] E. Acerbi and N. Fusco, *Semicontinuity problems in the calculus of variations,* Arch. Rational Mech. Anal., 86 (1984), 125–145.

[2] R. Adams, "Sobolev Spaces," Academic Press, New York, 1975.

[3] A. Dolcini, *A uniqueness result for very weak solutions of p-harmonic type equations,* Boll. UMI, Serie VII, Vol. X-A, N 1, 71–84.

[4] T.K. Donaldson and N.S. Trudinger, *Orlicz-Sobolev spaces and imbedding theorems,* J. Functional Analysis, 8 (1971), 52–75.

[5] A. Elcrat and N. Meyers, *Some results on regularity for solutions of nonlinear elliptic systems and quasiregular functions,* Duke Math. J., 42(1) (1975), 121–136.

[6] I. Ekeland and R. Temam, "Convex Analysis and Variational Problems," North-Holland, 1978.

[7] F.W. Gehring, *The L^p-integrability of the partial derivatives of a quasiconformal mapping,* Acta Math., 130 (1973), 265–277.

[8] M. Giaquinta and E. Giusti, *Quasi minima,* Ann. Inst. H. Poin. An. Non Lineaire I (1984), 79–107.

[9] J.-P. Gossez, *Nonlinear elliptic boundary problems for equations with rapidly (or slowly) increasing coefficients,* Trans. Am. Math. Soc., 190 (1974), 163–205.

[10] J.-P. Gossez and V. Mustonen, *Variational inequalities in Orlicz-Sobolev spaces*, Nonlinear Anal., Theory, Methods & Applications, 11(3) (1987), 379–392.

[11] T. Iwaniec, *p-Harmonic tensors and quasiregular mappings,* Annals of Mathematics, 136 (1992), 589–624.

[12] T. Iwaniec and C. Sbordone, *Weak minima of variational integrals*, J. Reine Angew Math., 454 (1994), 143–161.

[13] T. Iwaniec, C. Scott and B. Stroffolini, *Nonlinear Hodge theory on manifolds with boundary,* to appear.

[14] A. Kufner, O. John and S. Fučik, "Function Spaces," Noordhoff, 1977.

[15] M.A. Krasnosel'skii and Ya.B. Rutickii, "Convex Functions and Orlicz Spaces," Noordhoff, Groningen, 1961.

[16] J. Lewis, *On very weak solutions of certain elliptic systems,* Comm. Part. Diff. Equ., 18(9&10) (1993), 1515–1537.

[17] R. Landes and V. Mustonen, *Pseudo-monotone mappings in Orlicz-Sobolev spaces and nonlinear boundary value problems on unbounded domains*, J. Math. Analysis Appl., 88 (1982), 25–36.

[18] M.M. Rao and Z.D. Ren, "Theory of Orlicz spaces," Marcel Dekker, 1991.

[19] C. Sbordone, *Maximal inequalities and applications to regularity problems in the calculus of variations,* in Progress in PDE: the Metz Survey 3, ed. by Chipot, Pitman Res. Notes in Math. Series, 1995, 230–244.