On a Range of Exponents for Absence of Lavrentiev Phenomenon for Double Phase Functionals

MIROSLAV BULÍČEK, PIOTR GWIAZDA & JAKUB SKRZECZKOWSKI

Abstract

For a class of functionals having the \((p, q)\)-growth, we establish an improved range of exponents \(p, q\) for which the Lavrentiev phenomenon does not occur. The proof is based on a standard mollification argument and Young convolution inequality. Our contribution is two-fold. First, we observe that it is sufficient to regularise only bounded functions. Second, we exploit the \(L^\infty\) bound on the function rather than the \(L^p\) estimate on the gradient. Our proof does not rely on the properties of minimizers to variational problems but it is rather a consequence of the underlying Musielak–Orlicz function spaces. Moreover, our method works for unbounded boundary data, the variable exponent functionals and vectorial problems. In addition, the result seems to be optimal for \(p \leq d\).

1. Introduction

We consider a class of functionals with the so-called \((p, q)\)-growth. The prominent example we have in mind is

\[
G(u) := \int_{\Omega} |\nabla u(x)|^p \, dx + \int_{\Omega} a(x) |\nabla u(x)|^q \, dx. \tag{1.1}
\]

Here, \(\Omega \subset \mathbb{R}^d\) is a bounded Lipschitz domain, \(u : \Omega \to \mathbb{R}\) is an argument of the functional \(G\), \(a : \Omega \to [0, \infty)\) is a given nonnegative function and \(1 \leq p < q < \infty\).
are given numbers. Functional $\mathcal{G}$ is the interesting toy model for studying minimization of functionals with the so-called non-standard growth. Indeed, depending on whether $a = 0$ or $a > 0$, $\mathcal{G}$ exhibits either the $p$- or the $q$-growth.

A well-known feature of functional $\mathcal{G}$ is the so-called Lavrentiev phenomenon. For instance, there exists a function $a \in C^\alpha(\Omega)$ with $\alpha \in (0, 1)$, exponents $p, q$ fulfilling $p < d < d + \alpha < q$ and boundary data $u_0 \in W^{1,q}(\Omega)$ such that

$$\inf_{u \in u_0 + W^{1,p}_0(\Omega)} \mathcal{G}(u) < \inf_{u \in u_0 + W^{1,q}_0(\Omega)} \mathcal{G}(u).$$

On the other hand, it is known that if $q \leq p + \alpha \frac{p}{d}$, the Lavrentiev phenomenon does not occur for the toy model (1.1); see [21]. Under the additional assumption $u_0 \in L^\infty(\Omega)$, the range of exponents has been improved to $q \leq p + \alpha$ [16, Proposition 3.6, Remark 5]. The latter work heavily depends on the properties of minimizers and the $L^\infty$ bound for the minimizer of the functional (1.1) form a nontrivial part of the result in [16].

In this paper we prove that neither the assumption $u_0 \in L^\infty(\Omega)$ nor any additional property of minimizer (higher integrability, continuity) is irrelevant for the absence of Lavrentiev phenomenon. More precisely, we prove that one does not observe Lavrentiev phenomenon if

$$q \leq p + \alpha \max\left(1, \frac{p}{d}\right)$$

and boundary data $u_0 \in W^{1,q}(\Omega)$. In this case, we have

$$\inf_{u \in u_0 + W^{1,p}_0(\Omega)} \mathcal{G}(u) = \inf_{u \in u_0 + W^{1,q}_0(\Omega)} \mathcal{G}(u) = \inf_{u \in u_0 + C^\infty_c(\Omega)} \mathcal{G}(u).$$

This significantly improves the available results for the case $p < d$. Moreover, our proof is elementary as it is based on a simple regularisation argument together with Young’s convolution inequality. In particular, we do not use estimates on minimizers of functional (1.1). Consequently, our method easily extends to the vector-valued maps and cover variable-exponent functionals as well (see Section 3.2.)

The question of whether (1.2) or (1.4) holds true is related to the density of $C^\infty_c(\Omega)$ in the Musielak–Orlicz–Sobolev space $W^{1,\psi}(\Omega)$ corresponding to the functional (1.1), see (4.1)–(4.3) for definitions. In this context, we prove that the density result hold true for $p, q$ satisfying (1.3) which is again better then so-far known regime of exponents announced in [1].

Let us discuss the result of the paper within the context of previous works related to this topic. The first studies concerning functionals changing their ellipticity rate at each point have been carried out by Zhikov [39–42]. In particular, in [41] he observed that it may happen that (1.4) does not hold, extending thus similar observations made by Lavrentiev [24] and Mania [26]. Another related direction of research is the regularity of minimizers. Although the fundamental results for minimizers were obtained by Marcellini [27–30] more than 20 years ago, it is in fact still an active topic of research; see for instance [2,4–7,10,12,16–19,31–33,35,38].

Going back to the functional (1.1), the available results for boundary data $u_0 \in W^{1,q}(\Omega)$ provide both positive and negative answers to the question whether (1.4)
holds true. On the one hand, if $q \leq p + \frac{p\alpha}{d}$ then (1.4) is indeed valid [20,21]. On the other hand, if $q > p + \alpha \max\left(1, \frac{p-1}{d-1}\right)$ then counterexample in [3, Theorem 34] shows that (1.4) is violated (see also [21, Lemma 7] for a weaker result concerning the case $p < d < d + \alpha < q$ obtained with more elementary methods). In this paper we establish (1.4) for $q \leq p + \alpha \max\left(1, \frac{p}{d}\right)$ which partially fills the gap between currently known positive and negative results concerning the Lavrentiev phenomenon. Moreover, in view of [3, Theorem 34], our result is sharp for $p \leq d$.

Next, we wish to address two issues that appeared in previous papers on this topic. First, in [17, Lemma 4.1] there is the following claim: for every $\varepsilon > 0$ and ball $B_r(x) \subset \Omega$, there exists $p_\varepsilon < q_\varepsilon$ satisfying

$$
\varepsilon \ p_\varepsilon > q_\varepsilon - p_\varepsilon - \alpha_\varepsilon \frac{p_\varepsilon}{d} > 0,
$$

(1.5)
a coefficient $a_\varepsilon \in C^\infty(\Omega)$ and a boundary data $u_0 \in W^{1,q}(B_r(x)) \cap L^\infty(B_r(x))$ such that

$$
\inf_{u \in u_0 + W^{1,p_\varepsilon}_0(B_r(x))} G(u) < \inf_{u \in u_0 + W^{1,p_\varepsilon}(B_r(x)) \cap W^{1,q_\varepsilon}_{\text{loc}}(B_r(x))} G(u).
$$

Although this is a very nice result, it does not prove that range of exponents $q \leq p + \alpha \frac{p}{d}$ is optimal for absence of the Lavrentiev phenomenon and it does not contradict our result about the range stated in (1.3). In fact, authors refer to the counterexample from [21] constructed for exponents satisfying $p < d < d + \alpha < q$ i.e. exponents that do not meet our range because the distance between $p$ and $q$ is greater than $\alpha$. In fact, it is shown that there exists $p_\varepsilon$ and $q_\varepsilon$ but it follows also from the proof that they are constructed in the following way: for $\delta > 0$ to be specified later, we define $p_\varepsilon := d - \delta$, $q_\varepsilon := d + \alpha + \delta$ and find a proper counterexample constructed in [21]. Then, when $p_\varepsilon \geq 1$, we have

$$
\varepsilon \ p_\varepsilon \geq \varepsilon, \quad q_\varepsilon - p_\varepsilon - \alpha_\varepsilon \frac{p_\varepsilon}{d} = 2\delta + \alpha \frac{\delta}{d} = \delta \left(2 + \frac{\alpha}{d}\right),
$$

so that (1.5) is satisfied if we let $\delta := \frac{\varepsilon}{2(2 + \alpha/d)}$. Consequently, $p_\varepsilon \to d$ as $\varepsilon \to 0$, which is in perfect coincidence with (1.3).

Second, we also want compare our result with [16], where authors proved that the Lavrentiev phenomenon is not observed for $q \leq p + \alpha$ in the particular cases when minimizers of (1.1) are bounded, but this requires an extra assumption on the boundary data, namely that the boundary data $u_0$ is bounded and apply the maximum principle [25]. In addition, reasoning in [16] is based on the so-called Morrey type estimate on the gradient of minimizer which is not an obvious result itself. Comparing to our work, we prove that the Lavrentiev phenomenon does not occur independently of the properties of minimizers or boundedness of boundary data. Our methods are elementary and are based on simple estimates on convolutions. We point out that one could naively think that our result is a consequence of [16] and a simple approximation argument (boundary data $u_0 \in W^{1,q}(\Omega)$ is approximated with a sequence $\{u_{0,n}\}_{n \in \mathbb{N}} \subset W^{1,q}(\Omega) \cap L^\infty(\Omega)$) but it is not necessarily true that sequence of minimizers has then a subsequence converging again to a minimizer of the limit problem.
Finally, we want to point out and emphasize the main novelties of the paper. Standard methods [17, 20] for proving (1.4) are based on regularization of arbitrary function $u \in W^{1,p}_0(\Omega)$ satisfying $G(u) < \infty$ with a sequence of smooth functions $u^\varepsilon = u * \eta_\varepsilon$ and passing to the limit $G(u^\varepsilon) \to G(u)$ as $\varepsilon \to 0$. The latter is not trivial because the integrand in (1.1) is $x$-dependent. More precisely, if the integrand is convex and autonomous (i.e. it does not depend on $x$) one can use Jensen’s inequality and Vitali convergence theorem to prove that $G(u^\varepsilon) \to G(u)$ whenever $G(u) < \infty$. In particular, there is no Lavrentiev phenomenon in this case; see also [9].

The strategy to deal with the non-autonomous case is to approximate locally the integrand with autonomous function that does not depend on $x$ (see Lemma 5.4) so that one can exploit Jensen’s inequality. The approximation requires good estimate on $\|\nabla u^\varepsilon\|_\infty$ which results in constraint on exponents $p$ and $q$. The estimate on gradient is obtained by writing $\nabla u^\varepsilon = \nabla u * \eta_\varepsilon$ and using the fact that $\nabla u \in L^p(\Omega)$. Our main contribution is an observation that it is sufficient to approximate only bounded functions (i.e. $u \in L^\infty(\Omega)$). It turns out that for $p < d$, it is a better strategy to write $\nabla u^\varepsilon = u * \nabla \eta_\varepsilon$ and exploit the estimate $u \in L^\infty(\Omega)$ rather that $\nabla u \in L^p(\Omega)$. We remark that these observations have been already used in our recent paper on parabolic equations [11] but at that point we did not observe that similar ideas may bring new information to analysis of the Lavrentiev phenomenon.

The structure of the paper is as follows: in Section 2 we present the main result, Theorem 2.3. The theorem holds true under rather complicated assumption so in Section 3 we discuss two representative examples. In Section 4 we review the most important properties of the Musielak–Orlicz–Sobolev spaces. We explain here why it is sufficient to approximate only bounded functions, see Lemmas 4.2 and 4.3. Then, in Section 5 we present the proof of the main result in the particular case of functional $G$ as in (1.1) and $\Omega = B$ (i.e. a unit ball). In this case we may neglect many technical difficulties and clearly present main ideas. Section 6 is devoted to the proof of Theorem 2.3 in the general case. Finally, in Section 7 we briefly discuss how to extend our work to the case of vectorial problems.

2. Main result

Let us first set notation. We always assume that $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain and $d$ is the dimension of the space. We write $B$ for the unit open ball centered at 0. For balls with radius $r$ we use $B_r$ and if the center is at some general point $x$, we write $B_r(x)$ so that $B_1(0) = B$ and $B_r(0) = B_r$. Concerning function spaces, we write $C^\infty_\text{c}(\Omega)$ for the space of smooth compactly supported functions, $W^{1,p}(\Omega)$ and $W^{1,p}_0(\Omega)$ are usual Sobolev spaces, $W^{1,\psi}(\Omega)$ and $W^{1,\psi}_0(\Omega)$ are the Musielak–Orlicz–Sobolev spaces defined in Section 4 while $C^\alpha(\overline{\Omega})$ is the space of Hölder continuous functions on $\overline{\Omega}$ with exponent $\alpha \in (0, 1]$. Finally, $\eta_\varepsilon : \mathbb{R}^d \to \mathbb{R}$ is a usual mollification kernel.

We already introduced the key motivation of the paper, i.e., the functional (1.1), but the main result concern more general cases. We focus in the paper on functionals
being of the form
\[ H(u) = \int_\Omega \psi(x, |\nabla u(x)|) \, dx, \tag{2.1} \]
where \( \psi \) is the so-called \( \mathcal{N} \)-function and it satisfies the following assumptions:

**Assumption 2.1.** We assume that an \( \mathcal{N} \)-function \( \psi : \Omega \times \mathbb{R}^+ \to \mathbb{R}^+ \) satisfies

(A1) (vanishing at 0) \( \psi(x, \xi) = 0 \) if and only if \( \xi = 0 \),

(A2) (convexity) for each \( x \), the map \( \mathbb{R}^+ \ni \xi \mapsto \psi(x, \xi) \) is convex,

(A3) \( (p - q \text{ growth}) \) there exist exponents \( 1 \leq p < q < \infty \) and \( \xi_0 \geq 1 \) and constants \( C_1 \) and \( C_2 \) such that
\[ C_1 |\xi|^p \leq \psi(x, \xi) \text{ for } |\xi| \geq \xi_0, \quad \psi(x, \xi) \leq C_2 (1 + |\xi|^q) \text{ for all } \xi \geq 0, \]

(A4) \( (\Delta_2 \text{ condition}) \) there exists a constant \( C_4 \) such that
\[ \psi(x, 2\xi) \leq C_4 \psi(x, \xi). \]

(A5) (autonomous lower-bound) there is function \( m_\psi : \mathbb{R}^+ \to \mathbb{R}^+ \) and \( \xi_0 \) such that for \( \xi \geq \xi_0 \) we have \( m_\psi(\xi) \leq \psi(x, \xi) \) and \( m_\psi(\xi) / \xi \to \infty \) as \( \xi \to \infty \).

**Assumption 2.2.** Let \( \psi \) be an \( \mathcal{N} \)-function satisfying Assumption 2.1. We assume that for all \( D > 1 \), there are constants \( M = M(p, q, D) \) and \( N = N(p, q, D) \) such that
\[ \psi(z, \xi) \leq M \psi(y, \xi) + N \]
for all balls \( B_\gamma(x) \), all \( y, z \in \overline{B_\gamma(x)} \cap \Omega \), all \( \xi \in \left[ 0, D\gamma - \min\left(1, \frac{d}{\gamma}\right) \right] \) and all \( \gamma \in (0, \frac{1}{2}) \).

Let us make few comments on Assumptions 2.1 and 2.2. Conditions (A1)–(A2) are standard in the theory of Orlicz spaces while (A3) reflects growth of the \( \mathcal{N} \)-function being trapped between \( p - \) and \( q - \) growth. Condition (A4) ensures good functional analytic properties in \( W_0^{1, \psi}(\Omega) \) cf. Lemma 4.1. Finally, condition (A5) guarantees good behaviour of second conjugate on \( \psi \) and it is only necessary for Step 4 in the proof of Lemma 6.2. In particular, this condition is necessary only for the general \( \mathcal{N} \)-functions so it is does not have to hold in the special case \( \phi(x, \xi) = |\xi|^p + a(x) |\xi|^q \).

Assumption 2.1 thus reflects the basic functional setting. The real cornerstone of the paper is however Assumption 2.2. It is in fact an abstractly formulated condition on continuity of \( \psi \). To understand it better, we note that it is always possible to estimate, for all \( x \in \Omega \),
\[ \inf_{y \in B_\gamma(x) \cap \Omega} \psi(y, \xi) \leq \psi(x, \xi). \]

Assumption 2.2 states that the above estimate can be inverted (with a suitable constant). As it seems to be hard to verify it directly, we provide two model examples of \( \mathcal{N} \)-functions \( \psi \) satisfying this condition in Section 3. Nevertheless, we would like to emphasize that the prototypic functional (1.1) satisfying (1.3) fulfils also Assumption 2.1.

The main result of this paper is
Theorem 2.3. Let $\mathcal{H}$ be a functional defined with (2.1) with $\psi$ satisfying Assumptions 2.1, 2.2. Then, for all $u_0 \in W^{1,q}_0(\Omega)$ we have

$$\inf_{u \in u_0 + W^{1,p}_0(\Omega)} \mathcal{H}(u) = \inf_{u \in u_0 + W^{1,q}_0(\Omega)} \mathcal{H}(u) = \inf_{u \in u_0 + C_c^\infty(\Omega)} \mathcal{H}(u).$$

Moreover, space $C_c^\infty(\Omega)$ is dense in the Musielak–Orlicz–Sobolev space $W^{1,\psi}_0(\Omega)$.

3. Examples of $N$-functions satisfying Assumption 2.2

3.1. Standard double phase functionals

In this section, we prove that the $N$-function

$$\varphi(x, \xi) = |\xi|^p + a(x) |\xi|^q$$

satisfies Assumption 2.2 provided that $a \in C^{\alpha}(\overline{\Omega})$ and $q \leq p + \alpha \max(1, \frac{p}{q})$. The related functional reads as

$$\mathcal{G}(u) := \int_{\Omega} |\nabla u(x)|^p + a(x) |\nabla u(x)|^q \, dx.$$ 

To show thus, we use the following lemma, whose assumptions are evidently satisfied by the example given above:

Lemma 3.1. Suppose that $\psi$ satisfies Assumption 2.1 with exponents $p$ and $q$. Moreover, assume that there is $\alpha \in (0, 1]$ and constant $C_3$ such that, for all $x_1, x_2 \in \Omega$ and $\xi \geq \xi_0$, we have that

$$|\psi(x_1, \xi) - \psi(x_2, \xi)| \leq C_3 |x_1 - x_2|^\alpha (1 + |\xi|^q). \quad (3.1)$$

Then, $\psi$ satisfies Assumption 2.2 provided that $q \leq p + \alpha \max(1, \frac{p}{q})$.

Proof of Lemma 3.1. First, we may assume that $\xi > \xi_0$ as for $\xi \in [0, \xi_0]$ we have that

$$\psi(x, \xi) \leq C_2 (1 + |\xi|^q) \leq C_2 (1 + |\xi_0|^q) + \psi(y, \xi) \quad (3.2)$$

so the assertion follows with $M = 1$ and $N = C_2 (1 + |\xi_0|^q)$. Hence, we fix $\xi > \xi_0$ and some ball $B_\rho(x)$ such that $B_\rho(x) \cap \overline{\Omega}$ is not empty. Thanks to (3.1), we have for all $y, z \in B_\rho(x) \cap \overline{\Omega}$,

$$\psi(z, \xi) \geq \psi(y, \xi) - C_3 (1 + |\xi|^q) |y - z|^\alpha \geq \psi(y, \xi) - C_3 (1 + |\xi|^q) |y - z|^\alpha \quad (3.3)$$

As $\xi \geq \xi_0 \geq 1$, we have in fact that

$$\psi(z, \xi) \geq \psi(y, \xi) - 2 C_3 |\xi|^q |y - z|^\alpha.$$ 

To bootstrap this estimate, we fix $\delta \in (0, 1)$ and write

$$\psi(z, \xi) = \delta \psi(z, \xi) + (1 - \delta) \psi(z, \xi) \geq \delta \psi(y, \xi) - \delta 2 C_3 |\xi|^q |y - z|^\alpha + (1 - \delta) C_1 |\xi|^p, \quad (3.4)$$
where we used (3.3) to estimate the first term and lower bound \( \psi(z, \xi) \geq C_1 |\xi|^p \) to estimate the second term. Now, we may write

\[
2 \delta C_3 |\xi|^q \gamma^\alpha = 2 \delta C_3 |\xi|^{q-p} |\xi|^p \gamma^\alpha \leq 2 \delta C_3 D^{q-p} \gamma^{\alpha-(q-p) \min\left(1, \frac{d}{\gamma} \right) |\xi|^p,}
\]

(3.5)

where we used \(|\xi| \leq D \gamma^{-\min\left(1, \frac{d}{\gamma} \right) \min\left(1, \frac{p}{d} \right) \min\left(1, \frac{p}{d} \right) = \alpha - \alpha = 0. \)

It follows that \( \gamma^{\alpha-(q-p) \min\left(1, \frac{d}{\gamma} \right) \leq 1 \text{ for } \gamma \in (0, \frac{1}{2}) \). Hence, coming back to (3.4) we obtain

\[
\psi(z, \xi) \geq \delta \psi(y, \xi) - \delta 2 C_3 D^{q-p} |\xi|^p + (1 - \delta) C_1 |\xi|^p
\]

\[
= \delta \psi(y, \xi) + ((1 - \delta) C_1 - \delta 2 C_3 D^{q-p}) |\xi|^p.
\]

We choose \( \delta = \frac{C_1}{C_1 + 2 C_3 D^{q-p}} \) so that \(((1 - \delta) C_2 - \delta 2 C_3 D^{q-p}) |\xi|^p = 0. \) Hence, for all \( y, z \in \Omega \),

\[
\psi(z, \xi) \geq \delta \psi(y, \xi),
\]

so combining this with (3.2), the proof is concluded with \( M = \max (1/\delta, 1) \) and \( N = C_2 (1 + |\xi_0|^q) \).

3.2. Variable exponent double phase functionals

In this section we prove that \( N \)-function

\[
\phi(x, \xi) := |\xi|^{p(x)} + a(x) |\xi|^{q(x)}
\]

satisfies our Assumption 2.2. The related functional reads as

\[
\mathcal{J}(u, \Omega) := \int_{\Omega} \left[ |\nabla u|^{p(x)} + a(x) |\nabla u|^{q(x)} \right] \, dx.
\]

3.2. Variable exponent double phase functionals

In this section we prove that \( N \)-function

\[
\phi(x, \xi) := |\xi|^{p(x)} + a(x) |\xi|^{q(x)}
\]

satisfies our Assumption 2.2. The related functional reads as

\[
\mathcal{J}(u, \Omega) := \int_{\Omega} \left[ |\nabla u|^{p(x)} + a(x) |\nabla u|^{q(x)} \right] \, dx.
\]

Assumption 3.2. We assume that

(B1) \((p - q \text{ growth}) \text{ there exist } p, q \text{ with } 1 < p \leq q \text{ such that the functions } p(x), q(x) : \Omega \rightarrow [1, \infty) \text{ satisfy } p \leq p(x) \leq q(x) \leq q. \)

(B2) \((\log\text{-Hölder continuity}) \text{ there are constants } C_p, C_q \text{ such that for all } x, y \in \Omega \text{ with } |x - y| \leq \min \left( \text{diam } \Omega, \frac{1}{2} \right) \text{ we have } \)

\[
|p(x) - p(y)| \leq -\frac{C_p}{\log |x - y|}, \quad |q(x) - q(y)| \leq -\frac{C_q}{\log |x - y|}.
\]

(B3) \((\alpha\text{-Hölder continuity}) a \in C^\alpha(\Omega) \) with constant \(|a|_\alpha.\)
Lemma 3.3. Under Assumption 3.2, $N$-function $\phi$ defined with (3.6) satisfies Assumption 2.2 for $q$ and $p$ such that $q \leq p + \alpha \max\left(1, \frac{p}{\gamma}\right)$.

**Proof.** As in the proof of Lemma 3.1, we only need to consider $\xi \geq 1$. Let us estimate $\frac{\phi(x, \xi)}{\phi(y, \xi)}$ for $x, y \in \Omega$ such that $|x - y| \leq \min\left(\text{diam } \Omega, \frac{\xi}{2}\right)$. Using $a \geq 0$, we have that

$$\frac{\phi(x, \xi)}{\phi(y, \xi)} = \left|\xi\right|^{p(x)} + a(x) \left|\xi\right|^{q(x)} \leq \left|\xi\right|^{p(y)} + a(y) + \left|\xi\right|^{q(y)} \leq \left|\xi\right|^{q(y)} + a(y)$$

Now, let $D > 1$ and $y \in (0, \frac{1}{2})$. Suppose that $q \leq p + \alpha \max\left(1, \frac{p}{\gamma}\right)$, $|x - y| \leq \gamma$ and $\xi \in \left[1, D\gamma^{-\min\left(1, \frac{d}{\gamma}\right)}\right]$. Let $C > 0$, then

$$\left|\xi\right|^{-\frac{C_p}{\log\left(\frac{x}{\gamma}\right)}} \leq D\gamma^{-\min\left(1, \frac{d}{\gamma}\right)} = D\gamma^{-\frac{C_p}{\log\left(\frac{x}{\gamma}\right)}} = D\gamma^{-\frac{C_q}{\log\left(\frac{x}{\gamma}\right)}} \leq E$$

Applying this estimate with $C = C_p, C_q$ and finding a numerical constant $E$ such that

$$D^{-\frac{C_q}{\log\left(\frac{x}{\gamma}\right)}} e^{C_q}, D^{-\frac{C_p}{\log\left(\frac{x}{\gamma}\right)}} e^{C_p} \leq E$$

for all $y \in (0, \frac{1}{2})$. It follows that

$$\left|\xi\right|^{-\frac{C_p}{\log\left(\frac{x}{\gamma}\right)}}, \left|\xi\right|^{-\frac{C_q}{\log\left(\frac{x}{\gamma}\right)}} \leq E.$$  

Using (3.8), we obtain

$$\frac{\phi(x, \xi)}{\phi(y, \xi)} \leq E \left(E^2 + |a|_{\alpha} \gamma^{\alpha} \left(D\gamma^{-\min\left(1, \frac{d}{\gamma}\right)} \right)^{\alpha \max\left(1, \frac{p}{\gamma}\right)} + 1\right)$$

$$= E \left(E^2 + D^{\alpha \max\left(1, \frac{p}{\gamma}\right)} |a|_{\alpha} + 1\right) =: M,$$

where we used $\max\left(1, \frac{p}{\gamma}\right) \min\left(1, \frac{d}{\gamma}\right) = 1$ in the last line. We deduce that

$$\phi(x, \xi) \leq M \phi(y, \xi).$$
4. Musielak–Orlicz–Sobolev spaces

Our results are based on smooth approximation in the Musielak–Orlicz spaces, so we first recall their definitions and basic properties. For more details, we refer to monographs [13,23]. We consider an $N$-function $\psi : \Omega \times \mathbb{R}^+ \to \mathbb{R}$ satisfying (A1)–(A4) in Assumption 2.1. For $f : \Omega \to \mathbb{R}^d$ such that $\int_{\Omega} \psi(x, |f(x)|) \, dx < \infty$, we define the related Luxembourg norm with

$$\| f \|_\psi = \inf \left\{ \lambda > 0 : \int_{\Omega} \psi \left( x, \frac{|f(x)|}{\lambda} \right) \, dx \leq 1 \right\}.$$  \hspace{1cm} (4.1)

Finally, the Musielak–Orlicz–Sobolev spaces are defined as

$$W^{1,\psi}(\Omega) = \{ w \in W^{1,1}(\Omega) : \| \nabla w \|_\psi < \infty \}, \quad W_0^{1,\psi}(\Omega) = W_0^{1,1}(\Omega) \cap W^{1,\psi}(\Omega),$$

the latter one corresponds to the space of functions vanishing at the boundary. These are normed spaces with norm

$$\| w \|_{1,\psi} = \| w \|_1 + \| \nabla w \|_\psi.$$  \hspace{1cm} (4.3)

One can think of $W^{1,\psi}(\Omega)$ as the space of functions having gradient integrable with $p$ or $q$ power depending on whether $a = 0$ or not.

We summarize some properties of the Musielak–Orlicz spaces in the following lemma. They are mainly consequences of (A4) in Assumption 2.1. The proof can be found in many texts on Orlicz spaces [13,23], yet for the sake of completeness, we present the proof in “Appendix A.2”.

Lemma 4.1. Let $\psi$ satisfy Assumption 2.1 and let $f, f_n : \Omega \to \mathbb{R}^d$. Then,

(C1) $\| f \|_\psi < \infty \iff \int_{\Omega} \psi(x, c|f(x)|) \, dx < \infty$ for some $c > 0$ \iff $\int_{\Omega} \psi(x, c|f(x)|) \, dx < \infty$ for all $c > 0$,
(C2) $\| f_n - f \|_\psi \to 0 \iff$ for some $c > 0$ $\int_{\Omega} \psi(x, c|f_n(x) - f(x)|) \, dx \to 0$ \iff for all $c > 0$ $\int_{\Omega} \psi(x, c|f_n(x) - f(x)|) \, dx \to 0$,
(C3) if $\| f \|_\psi < \infty$ and any of the conditions in (C2) is satisfied, we have $\int_{\Omega} \psi(x, |f_n(x)|) \, dx \to \int_{\Omega} \psi(x, |f(x)|) \, dx$,
(C4) if $f_n \to f$ a.e. on $\Omega$, $\| f \|_\psi < \infty$ and the sequence $\{ \psi(x, |f_n(x)|) \}_{n \in \mathbb{N}}$ is uniformly integrable then $\| f_n - f \|_\psi \to 0$,
(C5) if $\| f \|_\psi < \infty$ then $f \in L^1(\Omega)$.

Next two lemmas show that to prove the absence of the Lavrentiev phenomenon, it is sufficient to demonstrate that every $u \in W_0^{1,\psi}(\Omega) \cap L^\infty(\Omega)$ can be approximated in the topology of $W^{1,\psi}$ by smooth function from $C_0^\infty(\Omega)$.

First lemma shows, that it is enough to consider only bounded functions. Notice that we do not impose any specific assumption on the $N$-function $\psi$ here.

Lemma 4.2. Space $W_0^{1,\psi}(\Omega) \cap L^\infty(\Omega)$ is dense in $W_0^{1,\psi}(\Omega)$. 

Proof. Let \( u \in W^{1,\psi}_0(\Omega) \). Consider truncation of \( u \) defined as

\[
T_k(u) = \begin{cases} 
  u & \text{if } |u| \leq k, \\
  k \frac{u}{|u|} & \text{if } |u| > k.
\end{cases}
\]

Clearly, \( T_k(u) \in L^\infty(\Omega) \). Moreover, chain rule for Sobolev maps implies that

\[
\nabla T_k(u) = \nabla u \mathbb{1}_{|u| \leq k}
\]

so that \( \nabla T_k(u) \to \nabla u \) a.e. as \( k \to \infty \). As \( \psi(x, 0) = 0 \), we have

\[
0 \leq \psi(x, |\nabla T_k(u)|) = \psi(x, |\nabla u|) \mathbb{1}_{|u| \leq k} \leq \psi(x, |\nabla u|)
\]

so that the sequence \( \{\psi(x, |\nabla T_k(u)|)\}_{k \in \mathbb{N}} \) is uniformly integrable. Application of (C4) from Lemma 4.1 concludes the proof. \( \square \)

Lemma 4.3. Suppose that \( \psi \) satisfies (A1)–(A4) in Assumption 2.1. Let \( p < q \) be exponents as in (A3) in Assumption 2.1. Suppose that for every \( u \in W^{1,\psi}_0(\Omega) \cap L^\infty(\Omega) \) there exists a sequence \( \{u^n\}_{n=1}^{\infty} \subset C^\infty_c(\Omega) \) such that \( \|u^n - u\|_{1,\psi} \to 0 \) as \( n \to \infty \). Then, the space \( C^\infty_c(\Omega) \) is dense in \( W^{1,\psi}_0(\Omega) \) and the Lavrentiev phenomenon does not occur, i.e., for all \( u_0 \in W^{1,q}(\Omega) \) we have that

\[
\inf_{u \in u_0 + W^{1,p}_0(\Omega)} \mathcal{H}(u) = \inf_{u \in u_0 + W^{1,q}_0(\Omega)} \mathcal{H}(u) = \inf_{u \in u_0 + C^\infty_c(\Omega)} \mathcal{H}(u).
\]

Proof. Thanks to Lemma 4.2, \( C^\infty_c(\Omega) \) is dense in \( W^{1,\psi}_0(\Omega) \). Let \( u^* \in W^{1,p}(\Omega) \) be the minimizer of \( \mathcal{H} \) i.e.

\[
\inf_{u \in u_0 + W^{1,p}(\Omega)} \mathcal{H}(u) = \mathcal{H}(u^*).
\]

The minimizer exists by a usual application of direct method in calculus of variations, cf. [37, Theorem 2.7]. Note that we always have that

\[
\mathcal{H}(u^*) = \inf_{u \in u_0 + W^{1,p}(\Omega)} \mathcal{H}(u) \leq \inf_{u \in u_0 + W^{1,q}(\Omega)} \mathcal{H}(u) \leq \inf_{u \in u_0 + C^\infty_c(\Omega)} \mathcal{H}(u),
\]

because \( p < q \). To prove the reversed inequality, we write \( u^* = u_0 + \bar{u} \) where \( \bar{u} \in W^{1,p}_0(\Omega) \). Note that \( u_0 \in W^{1,\psi}(\Omega) \) (because \( W^{1,q}(\Omega) \subset W^{1,\psi}(\Omega) \)) and \( u^* \in W^{1,\psi}(\Omega) \) (because \( \mathcal{H}(u^*) < \infty \) cf. Lemma 4.1 (C1)). It follows that \( \bar{u} = u^* - u_0 \in W^{1,\psi}_0(\Omega) \). Now, consider the sequence \( \{u_n\}_{n \in \mathbb{N}} \subset C^\infty_c(\Omega) \) such that \( u_n \to \bar{u} \) in \( W^{1,\psi}(\Omega) \) which exists due to the assumptions. It follows that \( u_n + u_0 \to \bar{u} + u_0 = u^* \) in \( W^{1,\psi}(\Omega) \). In particular, \( \mathcal{H}(u_0 + u_n) \to \mathcal{H}(u^*) \) cf. Lemma 4.1 (C3). Note that \( u_0 + u_n \in u_0 + C^\infty_c(\Omega) \). It follows that

\[
\inf_{u \in u_0 + C^\infty_c(\Omega)} \mathcal{H}(u) \leq \mathcal{H}(u_0 + u_n) \to \mathcal{H}(u^*) \quad \text{as } n \to \infty.
\]

\( \square \)
5. Proof of Theorem 2.3 in the special case

In this section we prove Theorem 2.3 in the case when $\Omega = B$ (unit ball centered at 0) and the $N$-function is defined via the formula

$$\varphi(x, \xi) = |\xi|^p + a(x) |\xi|^q.$$  \hfill (5.1)

The corresponding functional then takes the form

$$\mathcal{G}(u) := \int_B |\nabla u(x)|^p + a(x) |\nabla u(x)|^q \, dx.$$ 

Note that, if $a \in C^\alpha(B)$ and $q \leq p + \alpha \max \left(1, \frac{p}{d}\right)$, it follows from Lemma 3.1 that $\varphi$ satisfies Assumption 2.2.

The main purpose of this section is that we avoid all technical difficulties and focus only on the main parts of the proof. More precisely, we do not need to take care of difficulties coming from

- geometric properties of general Lipschitz domain $\Omega$,
- situation when for general $N$-function $\psi$ there is no local minimizer of the map $x \mapsto \psi(x, \xi)$ valid for all values of $\xi$.

We start with introducing mollification that will be used to define the approximation.

**Definition 5.1. (Mollification with squeezing)** For $\varepsilon \in (0, 1/4)$ we set $\eta_\varepsilon(y) = \frac{1}{\varepsilon^d} \eta \left(\frac{y}{\varepsilon}\right)$ where $\eta$ is a usual mollification kernel. Then, for arbitrary $u : \mathbb{R}^d \to \mathbb{R}$, we define $u^\varepsilon : \mathbb{R}^d \to \mathbb{R}$ as

$$u^\varepsilon(x) = \int_{\mathbb{R}^d} \eta_\varepsilon(y) u \left(\frac{x}{1-2\varepsilon} - y\right) \, dy.$$ 

**Lemma 5.2.** Let $u \in W^{1,1}_0(B)$ and be extended by zero onto $\mathbb{R}^d$. Then, $u^\varepsilon \in C^\infty_c(B)$. Moreover, $x \frac{1}{1-2\varepsilon} - y \in B_{5\varepsilon}(x)$ for all $y$ such that $|y| \leq \varepsilon$.

**Proof.** Smoothness follows from standard properties of convolutions cf. [22, Appendix C.4]. To see the compact support, let $|x| \geq 1 - \varepsilon$ and $|y| \leq \varepsilon$. Then,

$$\left| \frac{x}{1-2\varepsilon} - y \right| \geq \frac{1 - \varepsilon}{1-2\varepsilon} - \varepsilon = \frac{1 - \varepsilon}{1 - 2\varepsilon} = \frac{1 - \varepsilon}{1 - 2\varepsilon} \frac{1}{1 - 2\varepsilon} = 1 + \frac{2\varepsilon^2}{1 - 2\varepsilon} > 1$$

so that $u \left(\frac{x}{1-2\varepsilon} - y\right) = 0$. It follows that $u^\varepsilon$ is supported in $B_{1-\varepsilon}$. To see the second property, we estimate

$$\left| x - \frac{x}{1-2\varepsilon} + y \right| \leq |x| \frac{2\varepsilon}{1 - 2\varepsilon} + |y| \leq 4\varepsilon + \varepsilon = 5\varepsilon,$$

where we used $\frac{1}{1-2\varepsilon} \leq 2$, i.e. $\varepsilon \leq \frac{1}{4}$.

\hfill $\Box$
Before formulating the main theorem of this section, we state and prove two results: a technical lemma concerning approximating sequence and a simple observation concerning $\mathcal{N}$-function $\varphi$.

**Lemma 5.3.** Let $u \in W^{1,1}_0(B)$ be such that $G(u) < \infty$ and consider its extension to $\mathbb{R}^d$. Then,

(D1) $\varphi \left( \frac{x}{1 - 2\varepsilon}, |\nabla u| \left( \frac{x}{1 - 2\varepsilon} \right) \right) \rightarrow \varphi(x, |\nabla u(x)|)$ in $L^1(\mathbb{R}^d)$,

(D2) $\int_{\mathbb{R}^d} \varphi \left( \frac{x}{1 - 2\varepsilon} - y, |\nabla u| \left( \frac{x}{1 - 2\varepsilon} - y \right) \right) \eta_\varepsilon(y) \, dy \rightarrow \varphi(x, |\nabla u(x)|)$ in $L^1(\mathbb{R}^d)$.

**Proof.** To see (D1), we note that the convergence holds in the pointwise sense. Moreover, the considered sequence is supported only for $x \in B_{1-2\varepsilon}$. Therefore, to establish convergence in $L^1(\mathbb{R}^d)$, it is sufficient to prove equiintegrability of the sequence $\left\{ \varphi \left( \frac{x}{1 - 2\varepsilon}, |\nabla u| \left( \frac{x}{1 - 2\varepsilon} \right) \right) \right\}$ and apply the Vitali convergence theorem. To this end, we need to prove that

$$\forall \eta > 0 \exists \delta > 0 \forall A \subset B, |A| \leq \delta \int_A \varphi \left( \frac{x}{1 - 2\varepsilon}, |\nabla u| \left( \frac{x}{1 - 2\varepsilon} \right) \right) \, dx \leq \eta.$$

We fix $\eta$ and arbitrary $A \subset B$. Using change of variables we have that

$$\int_A \varphi \left( \frac{x}{1 - 2\varepsilon}, |\nabla u| \left( \frac{x}{1 - 2\varepsilon} \right) \right) \, dx = (1 - 2\varepsilon)^d \int_{A/(1-2\varepsilon)} \varphi(x, |\nabla u(x)|) \, dx \leq \int_{2A} \varphi(x, |\nabla u(x)|) \, dx,$$

where for $c \in \mathbb{R}^+$, $cA$ denotes a usual scaled set. By assumption, we have that $G(u) < \infty$, so if we set

$$\omega(\tau) := \sup_{C \subset \mathbb{R}^d : |C| \leq \tau} \int_C \varphi(x, |\nabla u(x)|) \, dx,$$

then $\omega(\tau)$ is a non-decreasing function, continuous at 0. Therefore, we may find $\tau$ such that $\omega(\tau) \leq 2^{-d} \eta$. Then, we choose $\delta = 2^{-d} \tau$ to conclude the proof of (D1). Finally, the convergence result (D2) follows from Young’s convolutional inequality and (D1).

**Lemma 5.4.** Let $\varphi$ be given by (5.1). Then for all balls $B_\gamma(x)$ such that $\overline{B_\gamma(x)} \cap \overline{B}$ is nonempty, there exists $x^* \in \overline{B_\gamma(x)} \cap \overline{B}$ such that, for all $\xi$,

$$\inf_{y \in B_\gamma(x) \cap \overline{B}} \varphi(y, \xi) = \varphi(x^*, \xi).$$

**Proof.** Using continuity of $a$ and compactness of $\overline{B_\gamma(x)} \cap \overline{B}$ we have that

$$\inf_{y \in \overline{B_\gamma(x)} \cap \overline{B}} \varphi(y, \xi) = \inf_{y \in \overline{B_\gamma(x)} \cap \overline{B}} \left[ |\xi|^p + a(y) |\xi|^q \right] = |\xi|^p + |\xi|^q \inf_{y \in \overline{B_\gamma(x)} \cap \overline{B}} a(y),$$

and we choose $y^*$ such that $\inf_{y \in \overline{B_\gamma(x)} \cap \overline{B}} a(y) = a(y^*)$. \qed
Theorem 5.5. (Theorem 2.3 in the special case) Let \( u \in W^{1,\varphi}_0 (B) \cap L^\infty (B) \) with \( a \in C^\alpha (\overline{B}) \). Suppose that
\[
1 \leq p < q \leq p + \alpha \max \left( 1, \frac{P}{d} \right).
\]
Consider sequence \( u^\varepsilon \) as in Definition 5.1 with \( \varepsilon \in (0, \frac{1}{4}) \). Then,

1. \( u^\varepsilon \in C^\infty_c (B) \),
2. \( G (u^\varepsilon) \to G (u) \) as \( \varepsilon \to 0 \),
3. \( u^\varepsilon \to u \) in \( W^{1,\varphi} (B) \) as \( \varepsilon \to 0 \),
4. \( C^\infty_c (B) \) is dense in \( W^{1,\varphi}_0 (B) \) and Lavrentiev phenomenon does not occur, i.e.
\[
\inf_{u \in \mathcal{G} \cap \mathcal{L} \supseteq \mathcal{G}} u_0 = \inf_{u \in \mathcal{G} \cap \mathcal{L} \supseteq \mathcal{G}} u_0 = \inf_{u \in \mathcal{G} \cap \mathcal{L} \supseteq \mathcal{G}} u_0.
\]

Proof. The first property follows from construction. To prove convergence, we note that
\[
G (u^\varepsilon) = \int_B \varphi (x, |\nabla u^\varepsilon| (x)) \, dx.
\]

We would like to take mollification out of the function \( \varphi \) using its convexity and Jensen’s inequality. However, this is not possible as function \( \varphi \) depends also on \( x \) explicitly. To overcome this problem, we apply Assumption 2.2, which allows us to approximate the function \( \varphi (x, \xi) \) locally by a function depending only on \( \xi \). Notice that \( \varphi \) satisfies Assumption 2.2 thanks to Lemma 3.1 and the structural assumption (5.1).

Case 1 \( p \leq d \). In this case we have \( q \leq p + \alpha \). Using Young’s convolution inequality, we obtain
\[
\| \nabla u^\varepsilon \|_\infty \leq \| u \|_\infty \| \nabla \eta_\varepsilon \|_1 \leq D (5\varepsilon)^{-1},
\]
where we choose \( D := 5 \| u \|_\infty \| \nabla \eta \|_1 \). Let \( x \in B \). Applying Assumption 2.2 with \( \gamma = 5 \varepsilon \) and Lemma 5.4 we obtain \( x^* \in B_{5\varepsilon} (x) \cap \overline{B} \) and constants \( M, N \) such that
\[
\varphi (x, |\nabla u^\varepsilon| (x)) \leq M \varphi (x^*, |\nabla u^\varepsilon| (x^*)) + N.
\]
Note that
\[
\varphi (x^*, |\nabla u^\varepsilon (x)|) = \varphi \left( x^*, \frac{1}{1 - 2\varepsilon} \left| \int_{B_\varepsilon} \nabla u \left( \frac{x}{1 - 2\varepsilon} - y \right) \eta_\varepsilon (y) \, dy \right| \right) \leq \left( \frac{1}{1 - 2\varepsilon} \right)^q \varphi \left( x^*, \int_{B_\varepsilon} |\nabla u| \left( \frac{x}{1 - 2\varepsilon} - y \right) \eta_\varepsilon (y) \, dy \right) \leq 2^q \varphi \left( x^*, \int_{B_\varepsilon} |\nabla u| \left( \frac{x}{1 - 2\varepsilon} - y \right) \eta_\varepsilon (y) \, dy \right),
\]
where we used that \( \varphi \) is of the form (5.1). Then, Jensen’s inequality implies that
\[
\varphi \left( x^*, \int_{B_\varepsilon} |\nabla u| \left( \frac{x}{1 - 2\varepsilon} - y \right) \eta_\varepsilon(y) \, dy \right) \\
\leq \int_{B_\varepsilon} \varphi \left( x^*, |\nabla u| \left( \frac{x}{1 - 2\varepsilon} - y \right) \right) \eta_\varepsilon(y) \, dy.
\]

If \( \frac{x}{1 - 2\varepsilon} - y \) does not belong to \( \overline{B} \) then \( \varphi \left( x^*, |\nabla u| \left( \frac{x}{1 - 2\varepsilon} - y \right) \right) = 0 \). Otherwise, Lemma 5.2 implies \( \frac{x}{1 - 2\varepsilon} - y \in \overline{B} \cap B_{5\varepsilon}(x) \), so that
\[
\varphi \left( x^*, |\nabla u| \left( \frac{x}{1 - 2\varepsilon} - y \right) \right) \leq \varphi \left( \frac{x}{1 - 2\varepsilon} - y, |\nabla u| \left( \frac{x}{1 - 2\varepsilon} - y \right) \right),
\]
due to the minimality of \( x^* \) and nonnegativity of \( a \). As \( x \in B \) was fixed, we obtain inequality
\[
\varphi(x, |\nabla u^\varepsilon|(x)) \leq 2^q \, M \int_{B_\varepsilon} \varphi \left( \frac{x}{1 - 2\varepsilon} - y, |\nabla u| \left( \frac{x}{1 - 2\varepsilon} - y \right) \right) \eta_\varepsilon(y) \, dy + N,
\]
valid for all \( x \in B \). Now, we observe that \( \varphi(x, |\nabla u^\varepsilon|(x)) \) converges a.e. to \( \varphi(x, |\nabla u|(x)) \). Moreover, the (RHS) of (5.4) is convergent in \( L^1(B) \) cf. Lemma 5.3 (D2) so that \( \{ \varphi(x, |\nabla u^\varepsilon|(x)) \}_\varepsilon \) is uniformly integrable in \( L^1(B) \). Therefore, Vitali convergence theorem implies
\[
\varphi(x, |\nabla u^\varepsilon|(x)) \rightarrow \varphi(x, |\nabla u|(x)) \quad \text{in } L^1(B) \text{ as } \varepsilon \rightarrow 0.
\]

Thanks to triangle inequality we obtain (E2). To see (E3), we note a simple estimate
\[
|a + b|^q \leq 2^{q-1} (|a|^q + |b|^q)
\]
so that
\[
\varphi \left( x, |\nabla u(x) - \nabla u^\varepsilon(x)| \right) \leq 2^{q-1} \varphi \left( x, |\nabla u|(x) \right) + 2^{q-1} \varphi \left( x, |\nabla u^\varepsilon|(x) \right).
\]

It follows that the sequence \( \{ \varphi \left( x, |\nabla u(x) - \nabla u^\varepsilon(x)| \right) \}_\varepsilon \) is again uniformly integrable and Vitali convergence theorem yields that
\[
\varphi \left( x, |\nabla u(x) - \nabla u^\varepsilon(x)| \right) \rightarrow 0 \quad \text{in } L^1(B) \text{ as } \varepsilon \rightarrow 0,
\]
concluding the proof of (E3). This shows that any bounded function in \( W_{0,1}^q(B) \) can be approximated with smooth compactly supported functions so that (E4) follows from Lemma 4.3.

**Case 2** \( p > d \). In this case we have \( q \leq p + \alpha \frac{p}{d} \). Note that
\[
\nabla u^\varepsilon(x) = \frac{1}{1 - 2\varepsilon} \int_{B_\varepsilon} \nabla u \left( \frac{x}{1 - 2\varepsilon} - y \right) \eta_\varepsilon(y) \, dy.
\]

Therefore, instead of (5.2), we can compute that
\[
\|\nabla u^\varepsilon\|_\infty \leq \frac{1}{1 - 2\varepsilon} \left\| \nabla u \left( \frac{\cdot}{1 - 2\varepsilon} \right) \right\|_p \|\eta_\varepsilon\|_{p'} \leq 2 \left\| \nabla u \left( \frac{\cdot}{1 - 2\varepsilon} \right) \right\|_p \|\eta_\varepsilon\|_{p'}, \tag{5.5}
\]
where $p'$ is the usual Hölder conjugate exponent. Using change of variables we obtain

$$
\|\eta_\varepsilon\|_{p'} = \int_{B_\varepsilon} \frac{1}{\varepsilon^d p'} |\eta\left(\frac{x}{\varepsilon}\right)|^{p'} dx = \varepsilon^{d(1-p')} \int_B |\eta(x)|^{p'} dx = \varepsilon^{-p' \frac{d}{p}} \|\eta\|_{p'},
$$

so that $\|\eta_\varepsilon\|_{p'} = \varepsilon^{-\frac{d}{p}} \|\eta\|_{p'}$. Using a change of variables again,

$$
\left\| \nabla u \left( \frac{\cdot}{1-2\varepsilon} \right) \right\|_p \leq \|\nabla u\|_p,
$$

which is finite as $G(u) < \infty$. Therefore, (5.5) boils down to

$$
\|\nabla u^\varepsilon\|_\infty \leq D \left( 5\varepsilon \right)^{-\frac{d}{p}},
$$

where $D := 5^d \|\nabla u\|_p \|\eta\|_{p'}$. Using Assumption 2.2 we obtain estimate (5.3). The rest of the proof is exactly the same. $\square$

6. Proof of Theorem 2.3 in the general case

In this section we generalize construction from Section 5 to prove Theorem 2.3 in the general case.

6.1. Second convex conjugate function

For general $\mathcal{N}$-function $\psi$ satisfying Assumption 2.1, Lemma 5.4 is not necessarily true. Therefore, to control mollifications, we need a different method to approximate $\psi(x, \xi)$ with a function depending only on $\xi$. The construction below is somehow standard and has appeared in many works before; see [14,15]. We start more generally. Let $f : \mathbb{R} \rightarrow \mathbb{R}$. We define convex conjugate $f^* : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ of $f$ as

$$
f^*(\eta) = \sup_{\xi \in \mathbb{R}} (\xi \cdot \eta - f(\xi)).
$$

Moreover, the second convex conjugate of $f^{**}$ is defined as

$$
f^{**}(\xi) = \sup_{\eta \in \mathbb{R}} (\xi \cdot \eta - f^*(\eta)).
$$

We now list some basic properties of the convex conjugates cf. [37, Propositions 2.21, 2.28].

**Lemma 6.1.** Let $f, g : \mathbb{R} \rightarrow \mathbb{R}$. Then, the following holds true:

(F1) $f^*$ and $f^{**}$ are convex functions,

(F2) if $f \leq g$ on $\mathbb{R}$, then $g^* \leq f^*$ on $\mathbb{R}$,

(F3) if $f \leq g$ on $\mathbb{R}$, then $f^{**} \leq g^{**}$ on $\mathbb{R}$,

(F4) if $f$ is convex then $f^{**} = f$ on $\mathbb{R}$.
(F5) \( f^{**} \) is the greatest convex minorant of \( f \).

Now, we apply these notions to \( \mathcal{N} \)-functions. Given \( \mathcal{N} \)-function \( \psi(x, \xi) \) satisfying Assumption 2.1, we extend it by 0 for \( \xi < 0 \) (hence this extension is surely convex), we consider a ball \( B_x(x) \) such that \( B_x(x) \cap \Omega \) is nonempty and we define

\[
\psi_x, \gamma(\xi) : \mathbb{R} \to \mathbb{R}, \quad \psi_x, \gamma(\xi) := \inf_{y \in B_x(x) \cap \Omega} \psi(y, \xi). \quad (6.1)
\]

**Lemma 6.2.** Let \( \psi \) be as in Assumptions 2.1, 2.2 and \( \psi_x, \gamma \) be as in (6.1).

(G1) Let \( \mathcal{D} > 1 \). Then, there are constants \( \mathcal{M} = \mathcal{M}(p, q, \mathcal{D}), \mathcal{N} = \mathcal{N}(p, q, \mathcal{D}) \) such that

\[
\psi(y, \xi) \leq \mathcal{M} \psi^{**}_{x, \gamma}(\xi) + \mathcal{N} \quad (6.2)
\]

for all balls \( B_x(x) \), all \( y \in B_x(x) \cap \Omega \), all \( \xi \) such that \( \xi \leq \mathcal{D} \gamma - \min\left(1, \frac{q}{p}\right) \) and all \( \gamma \in (0, \frac{1}{2}) \).

(G2) It holds \( 0 \leq \psi^{**}_{x, \gamma}(\xi) \leq \psi(y, \xi) \) for all balls \( B_x(x) \), all \( y \in B_x(x) \cap \Omega \) and all \( \xi \in \mathbb{R} \).

One could try to prove Lemma 6.2 by applying property (F3) from Lemma 6.1 to the estimate appearing in Assumption 2.2. However, this estimate is valid only on some bounded interval rather than the whole real line. The correct argument is presented in [13] but since it contains some imperfections, we present it below.

**Proof of Lemma 6.2.** The proof of (G2) follows easily from (F3) and (F4) stated in Lemma 6.1. For (G1) we split the proof into several steps. Recall that a convex function has a supporting line so that for all \( \eta \in \mathbb{R} \), there exists supporting line \( h_\eta \) such that \( \psi^{**}_{x, \gamma}(\xi) \geq h_\eta(\xi) \) and \( \psi^{**}_{x, \gamma}(\eta) = h_\eta(\eta) \).

**Step 1.** The map \( \mathbb{R} \ni \xi \mapsto \psi_x, \gamma(\xi) \) is locally Lipschitz continuous.

**Proof.** Fix \( y \in B_x(x) \) and interval \([-R, R] \subset \mathbb{R} \). The map \( \mathbb{R} \ni \xi \mapsto \psi(y, \xi) \) is convex so its difference quotients are monotone. Hence, for \( \xi_1, \xi_2 \in [-R, R] \) with \( \xi_1 < \xi_2 \) we have

\[
\frac{\psi(y, \xi_2) - \psi(y, -R - 1)}{\xi_2 - (-R - 1)} \leq \frac{\psi(y, \xi_2) - \psi(y, \xi_1)}{\xi_2 - \xi_1} \leq \frac{\psi(y, R + 1) - \psi(y, \xi_1)}{R + 1 - \xi_1}
\]

Since \( |\psi(y, R + 1)| \leq C_2 (1 + (R + 1))^q \) cf. Assumption 2.1 (A3), the map \( \xi \mapsto \psi(y, \xi) \) is Lipschitz continuous with constant \( 2 C_2 (1 + (R + 1))^q \). By triangle inequality we have that

\[
\psi(y, \xi_1) \leq |\psi(y, \xi_1) - \psi(y, \xi_2)| + \psi(y, \xi_2) \\
\leq 2 C_2 (1 + (R + 1))^q |\xi_1 - \xi_2| + \psi(y, \xi_2).
\]

Finding a sequence \( \{y_k\} \) such that \( \psi(y_k, \xi_2) \to \psi_x, \gamma(\xi_2) \) we obtain

\[
\psi_x, \gamma(\xi_1) \leq \liminf_{k \to \infty} \psi(y_k, \xi_1) \leq 2 C_2 (1 + (R + 1))^q |\xi_1 - \xi_2| + \psi_x, \gamma(\xi_2).
\]

As \( \xi_1 \) and \( \xi_2 \) can be interchanged, the conclusion follows. \( \square \)
Step 2. For \( \xi \leq 0 \) we have \( \psi_{x, y}^{**}(\xi) = 0 \) and for \( \xi > 0 \) we have \( \psi_{x, y}^{**}(\xi) \geq 0 \). In particular, estimate (6.2) is satisfied for function \( \psi \) and \( \xi \leq 0 \).

**Proof.** First, note that the function being identically zero is convex. As \( 0 \leq \psi_{x, y}(\xi) \), we deduce

\[
0 \leq \psi_{x, y}^{**}(\xi). \tag{6.3}
\]

Finally, as \( \psi_{x, y}^{**}(\xi) \leq \psi_{x, y}(\xi) \), we deduce \( \psi_{x, y}^{**}(\xi) = 0 \) for \( \xi \leq 0 \). \( \square \)

Step 3. Fix \( \eta \) such that \( 0 \leq \eta \leq D \gamma^{-\min\left(1, \frac{d}{p}\right)} \) and assume that \( h_\eta(\xi) = \psi_{x, y}^{**}(\xi) \) only for \( \xi = \eta \). Then, \( \psi_{x, y}^{**}(\eta) = \psi_{x, y}(\eta) \) and estimate (6.2) is satisfied for \( \xi = \eta \).

**Proof.** Suppose that \( \psi_{x, y}^{**}(\eta) < \psi_{x, y}(\eta) \) (we always have \( \psi_{x, y}^{**}(\eta) \leq \psi_{x, y}(\eta) \)). Using Lipschitz continuity from Step 1, we find two lines such that \( \psi_{x, y} \) is above them (see dotted lines in Fig. 1). Hence, we observe that \( \psi_{x, y} \) is not the largest convex minorant of \( \psi_{x, y} \), see Fig. 1. It follows that \( \psi_{x, y}^{**}(\eta) = \psi_{x, y}(\eta) \) and estimate (6.2) follows directly from Assumption 2.2. \( \square \)

Step 4. Fix \( \eta \) such that \( 0 \leq \eta \leq D \gamma^{-\min\left(1, \frac{d}{p}\right)} \) and assume that \( h_\eta(\xi) = \psi_{x, y}^{**}(\xi) \) for some interval \([a, b]\) containing \( \eta \) (so that \( h_\eta \) and \( \psi_{x, y}^{**} \) have joint line interval). Then, estimate (6.2) is satisfied for \( \xi = \eta \).

**Proof.** First, from Step 2 we may assume that \( a \geq 0 \) and by assumption (A5) we can assume \( b < \infty \) (as \( \psi(x, \xi) \) is superlinear as \( \xi \to \infty \)). Second, the reasoning from Step 4 shows that

\[
\psi_{x, y}(a) = \psi_{x, y}^{**}(a), \quad \psi_{x, y}(b) = \psi_{x, y}^{**}(b).
\]

Moreover, by the assumption, there exists \( t \in [0, 1] \) such that

\[
\psi_{x, y}(\eta) = t \psi_{x, y}^{**}(a) + (1 - t) \psi_{x, y}(b) = t \psi_{x, y}(a) + (1 - t) \psi_{x, y}(b).
\]

By definition of \( \psi_{x, y} \), there exist sequences \( \{x_n^a\}_{n \in \mathbb{N}}, \{x_n^b\}_{n \in \mathbb{N}} \subset B_\gamma(x) \) such that

\[
\psi_{x, y}^{**}(\eta) \geq t \psi(x_n^a, a) + (1 - t) \psi(x_n^b, b) - \frac{1}{n}. \tag{6.4}
\]

With these at hand, we proceed to the final proof. By definition and convexity,

\[
\psi_{x, y}(\eta) \leq \psi(x_n^b, \eta) \leq t \psi(x_n^b, a) + (1 - t) \psi(x_n^b, b) \tag{6.5}
\]

To apply (6.4), we have to replace \( \psi(x_n^b, a) \) with \( \psi(x_n^a, a) \). This can be done with Assumption 2.2: we note that \( |x_n^a - x_n^b| = 2 \gamma \) so if we let \( D := 2^{-\min\left(1, \frac{d}{p}\right)} \) we have

\[
|\eta| \leq D \gamma^{-\min\left(1, \frac{d}{p}\right)} = D (2 \gamma)^{-\min\left(1, \frac{d}{p}\right)}
\]

and Assumption 2.2 implies existence of constants \( M(D), N(D) \) (we skip dependence of these constants on \( p \) and \( q \) as these exponents are fixed) such that

\[
\psi(x_n^b, \eta) \leq M(D) \psi(x_n^a, a) + N(D).
\]
It follows from (6.5) that
\[ \psi_{x,\gamma}(\eta) \leq t \left( M(D) \psi(x_n^a, a) + N(D) \right) + (1 - t) \psi(x_n^b, b) \]

Letting \( \tilde{M}(D) := \max(M(D), 1) \) and exploiting (6.4) we have that
\[ \psi_{x,\gamma}(\eta) \leq \tilde{M}(D) \left( t \psi(x_n^a, a) + (1 - t) \psi(x_n^b, b) \right) + N(D) \]
\[ \leq \tilde{M}(D) \psi^{**}_{x,\gamma}(\eta) + \frac{\tilde{M}(D)}{n} + N(D). \]

Sending \( n \to \infty \), we deduce that
\[ \psi_{x,\gamma}(\eta) \leq \tilde{M}(D) \psi^{**}_{x,\gamma}(\eta) + N(D). \]

Exploiting Assumption 2.2 once again, we obtain for all \( y \in B_{\gamma}(x) \)
\[ \psi(y, \eta) \leq M(D) \psi_{x,\gamma}(\eta) + N(D) \leq M(D) \tilde{M}(D) \psi^{**}_{x,\gamma}(\eta) + N(D) + N(D). \]

The conclusion follows with \( M := M(D) \tilde{M}(D) \) and \( N = N(D) + N(D) \).

Step 5 Cases considered in Steps 2-4 are the only possible ones.

\textbf{Proof.} Clearly, the tangent line \( h_\eta \) touches the epigraph of \( \psi^{**}_{x,\gamma} \) at least in one point. The case where it is touched exactly at one point was studied in Step 3 while the situation when it is touched along some interval \([a, b]\) was analyzed in Step 4. Now, suppose that there are \( \eta < \eta_1 < \eta_2 \) such that
\[ \psi^{**}_{x,\gamma}(\eta) = h_\eta(\eta), \quad \psi^{**}_{x,\gamma}(\eta_2) = h_\eta(\eta_2), \quad \psi^{**}_{x,\gamma}(\eta_1) > h_\eta(\eta_1). \]

Then, \( \psi^{**}_{x,\gamma} \) is not convex raising contradiction.

6.2. Geometric issues

As \( \Omega \) is not a ball in general, we cannot define compactly supported approximation by retracting the function to the interior part of \( \Omega \) as in Definition 5.1. However, one can still do that for star-shaped domains.

\textbf{Definition 6.3.}

1. A bounded domain \( U \subset \mathbb{R}^d \) is said to be star-shaped with respect to \( \vec{x} \) if every ray starting from \( \vec{x} \) intersects with \( \partial U \) at one and only one point.
2. A bounded domain \( U \subset \mathbb{R}^d \) is said to be star-shaped with respect to the ball \( B_{\gamma}(x_0) \) if \( U \) is star-shaped with respect to all \( y \in B_{\gamma}(x_0) \).

The following lemma shows that star-shaped domains can be uniformly shrinked which allows for defining compactly supported approximations:
Fig. 1. We assume that there is \( \eta > 0 \) such that functions \( \psi_{x, \gamma}^{**, \eta} \) (black line) and \( \psi_{x, \gamma} \) (grey line) satisfy \( \psi_{x, \gamma}^{**, \eta} (\eta) < \psi_{x, \gamma} (\eta) \) and tangent line \( h_\eta \) touches \( \psi_{x, \gamma}^{**, \eta} \) only at \( \eta \). As \( \psi_{x, \gamma} \) is Lipschitz continuous, we can estimate it from below (dotted lines). Then, the function obtained by combining \( \psi_{x, \gamma}^{**, \eta} \) and the dashed line is convex. It lies below \( \psi_{x, \gamma} \) and above \( \psi_{x, \gamma}^{**, \eta} \) raising contradiction with Lemma 6.1 (F5)

Lemma 6.4. Let \( U \subset \mathbb{R}^d \) be a star-shaped domain with respect to the ball \( B_R \). Let \( \kappa_\varepsilon = 1 - \frac{4\varepsilon}{R} \). Then, \( \text{dist}(\kappa_\varepsilon U, \partial U) \geq 2\varepsilon \). In particular,

\[
\kappa_\varepsilon U + \varepsilon B \subset U.
\]

More generally, if \( U \) is star-shaped with respect to the ball \( B_R(x_0) \),

\[
\kappa_\varepsilon (U - x_0) + \varepsilon B \subset (U - x_0).
\]

Proof. Let \( b \in \partial U \) and let \( c \in \partial (\kappa_\varepsilon U) \) such that \( c \) lies on the interval \([0, b]\). Let \( T \) be a sphere of radius \( R \) perpendicular to the interval \([0, b]\) and let \( S \) be the cone with base \( T \) and apex \( b \), see Fig. 2. First, we have

\[
\text{dist}(\partial U, c) \geq \text{dist}(\partial S, c).
\]

Let \( \alpha \) be a half of an apex angle of the cone \( C \), see Fig. 2. It follows that

\[
\text{dist}(\partial S, c) = \sin(\alpha) |b - c| = \sin(\alpha) (1 - \kappa_\varepsilon) |b| = \sin(\alpha) (1 - \kappa_\varepsilon) |b| \geq \sin(\alpha) \frac{4\varepsilon}{R} |b|.
\]

so it is sufficient to estimate \( \sin(\alpha) \) from below. Using notation from Fig. 2, the length of interval \([d, b]\) equals \( \frac{b}{\cos(\alpha)} \). Therefore,

\[
\sin(\alpha) = \frac{R}{|b|/\cos(\alpha)} \implies \tan(\alpha) = \frac{R}{|b|}.
\]
Fig. 2. Three dimensional adaptation of the construction performed in Lemma 6.4. Point $b$ belongs to the boundary of star-shaped domain $U$ while point $c = \kappa \epsilon b$ belongs to the boundary of the rescaled set $\kappa \epsilon U$. Sphere $T$ is perpendicular to the interval $[0, b]$ and since $U$ is star-shaped with respect to the ball $B_R$, the cone $S$ with base $T$ and apex $b$ lies inside $U$.

As $\sin^2(\alpha) = \frac{\tan^2(\alpha)}{1 + \tan^2(\alpha)}$ we have that

$$\sin^2(\alpha) = \frac{R^2}{R^2 + b^2} \geq \frac{R^2}{4b^2} \implies \sin(\alpha) \geq \frac{R}{2|b|},$$

where we used $R^2 \leq |b|^2 \leq 3|b|^2$. We conclude that $\text{dist}(\partial U, c) \geq 2\epsilon$. As this argument can be repeated for all $c \in \partial (\kappa \epsilon U)$, we obtain $\text{dist}(\partial U, \kappa \epsilon U) \geq 2 \epsilon$. The second statement follows from observation that the set $U - x_0$ is star-shaped with respect to the ball $B_R$.

On star-shaped domain we can define mollification with squeezing as in Definition 5.1.

**Definition 6.5.** *(Mollification with squeezing on star-shaped domain)* Let $U$ be a star-shaped domain with respect to the ball $B_R(x_0)$. Given $u \in W^{1,1}_{0}(U)$ we extend it with 0 to $\mathbb{R}^d$ and define

$$S^\epsilon_U u(x) := \int_{\mathbb{R}^d} u \left( x_0 + \frac{x - x_0 - y}{\kappa \epsilon} \right) \eta_\epsilon(y) \, dy,$$

where $\kappa \epsilon = 1 - \frac{4\epsilon}{R}$.

The reader may think about the case $x_0 = 0$ first.

**Lemma 6.6.** Function $S^\epsilon_U u$ from Definition 6.5 belongs to $C^\infty_c(U)$. 


Proof. The smoothness is clear from standard properties of convolutions. Concerning compact support, we claim that $S_{U}^{\varepsilon}u$ is supported in $x_{0} + \kappa_{\varepsilon}(U - x_{0}) + \varepsilon B$ which is a compact subset of $U$ due to Lemma 6.4. Indeed, let $x \notin x_{0} + \kappa_{\varepsilon}(U - x_{0}) + \varepsilon B$ and suppose that there is $y$ with $|y| \leq \varepsilon$ such that $x_{0} + \frac{x - x_{0} - y}{\kappa_{\varepsilon}} \in U$. Then, we can write

$$x = x_{0} + \kappa_{\varepsilon}\left(x_{0} + \frac{x - x_{0} - y}{\kappa_{\varepsilon}} - x_{0}\right) + y$$

so that $x \in x_{0} + \kappa_{\varepsilon}(U - x_{0}) + \varepsilon B$ raising contradiction. It follows that for $x \in x_{0} + \kappa_{\varepsilon}(U - x_{0}) + \varepsilon B$ we have either

$$x_{0} + \frac{x - x_{0} - y}{\kappa_{\varepsilon}} \in U \text{ and } |y| > \varepsilon \quad \text{or} \quad x_{0} + \frac{x - x_{0} - y}{\kappa_{\varepsilon}} \notin U$$

so that the integral $\int_{\mathbb{R}^{d}} u\left(x_{0} + \frac{x - x_{0} - y}{\kappa_{\varepsilon}}\right) \eta_{\varepsilon}(y) \, dy = 0$. ∎

To move from star-shaped domains to Lipschitz ones we will use the following decomposition cf. [34, Lemma 3.14]:

Lemma 6.7. Suppose that $\Omega \subset \mathbb{R}^{d}$ is a bounded Lipschitz domain. Then, there exist domains $\{U_{i}\}_{i=1,...,n}$ such that

$$\overline{\Omega} \subset \bigcup_{i=1}^{n} U_{i},$$

and $\Omega \cap U_{i}$ is star-shaped with respect to some ball $B_{R_{i}}(x_{i})$.

6.3. Approximating sequence and proof of Theorem 2.3

We are in position to define the approximating sequence. Let $\Omega$ be a Lipschitz bounded domain. From Lemma 6.7 we obtain a family of domains such that $\overline{\Omega} \subset \bigcup_{i=1}^{n} U_{i}$ where $\{\Omega \cap U_{i}\}_{i=1,...,n}$ are star-shaped domains with respect to balls $B_{R}(x_{i})$ (without loss of generality, we may assume that the radii of the balls are the same by taking $R := \min_{i=1,...,n} R_{i}$). In particular, $\{U_{i}\}_{i=1,...,n}$ is an open covering of $\overline{\Omega}$ so there exists partition of unity related to this covering: family of functions $\{\theta_{i}\}_{i=1,...,n}$ such that

$$\theta_{i} \in C_{c}^{\infty}(U_{i}), \quad 0 \leq \theta_{i} \leq 1, \quad \sum_{i=1}^{n} \theta_{i} = 1 \text{ on } \overline{\Omega}.$$

Given $u \in W_{0}^{1,1}(\Omega) \cap L^{\infty}(\Omega)$ we extend it with 0 as above and we set

$$S_{U}^{\varepsilon}u := \sum_{i=1}^{n} S_{U_{i}}^{\varepsilon}(u \theta_{i}) = \sum_{i=1}^{n} \int_{B_{\varepsilon}} (u \theta_{i}) \left(x_{i} + \frac{x - x_{i} - y}{\kappa_{\varepsilon}}\right) \eta_{\varepsilon}(y) \, dy, \quad (6.6)$$

where $\kappa_{\varepsilon} = 1 - \frac{4\varepsilon}{R}$. We note that since $u$ vanishes outside of $\Omega$, function $u \theta_{i}$ is supported in $\Omega \cap U_{i}$ which is star-shaped.

Before formulating the main result of this section, we will state and prove two technical lemmas concerning approximating sequence.
Lemma 6.8. Let $\kappa_e = 1 - \frac{4\varepsilon}{R}$, $x \in \Omega$ and $|y| \leq \varepsilon$. Then, there exists a constant $C_{\Omega, R}$ such that for $\varepsilon \leq \frac{R}{8}$ we have $x_i + \frac{x - x_i - y}{\kappa_e} \in B_\varepsilon c_{\Omega, R}(x)$.

Proof. Note that for $\varepsilon \leq \frac{R}{8}$, we have $\frac{1}{\kappa_e} \leq 2$. We compute that

$$
|x_i - x| + \frac{x - x_i - y}{\kappa_e} - x| = \left|(x_i - x) \left(1 - \frac{1}{\kappa_e} \right) - \frac{y}{\kappa_e} \right| \\
\leq |x_i - x| + \frac{\varepsilon}{\kappa_e} \leq 8 |x_i - x| + 2 \varepsilon.$$

As $|x_i - x| \leq \text{diam}(\Omega)$ (the diameter of $\Omega$), we choose $C_{\Omega, R} := \frac{8 \text{diam}(\Omega)}{R} + 2$. □

Lemma 6.9. Let $u \in W_0^{1,1}(\Omega)$ be such that $\mathcal{H}(u) < \infty$ and consider its extension to $\mathbb{R}^d$. Then,

(H1) $\psi \left( x_i + \frac{x - x_i}{\kappa_e}, (|\nabla u|, \theta_i) \left( x_i + \frac{x - x_i}{\kappa_e} \right) \right) \to \psi(x, |\nabla u|, \theta_i)$ in $L^1(\mathbb{R}^d)$,

(H2) $\int_{B_e} \psi \left( x_i + \frac{x - x_i}{\kappa_e}, (|\nabla u|, \theta_i) \left( x_i + \frac{x - x_i}{\kappa_e} \right) \right) \eta_e(y) \, dy \to \psi(x, |\nabla u|, \theta_i)$ in $L^1(\mathbb{R}^d)$.

Proof. Concerning (H1), we note that the convergence holds in the pointwise sense. Moreover, the considered sequence is supported on $\Omega \cap U_i$. Therefore, to establish convergence in $L^1(\mathbb{R}^d)$, it is sufficient to prove equiintegrability of the sequence $\left\{ \psi \left( x_i + \frac{x - x_i}{\kappa_e}, (|\nabla u|, \theta_i) \left( x_i + \frac{x - x_i}{\kappa_e} \right) \right) \right\}_e$ and apply Vitali convergence theorem. To this end, we need to prove

$$
\forall \eta > 0 \exists \delta > 0 \forall A \subset \Omega \cap U_i, |A| \leq \delta \\
\int_A \psi \left( x_i + \frac{x - x_i}{\kappa_e}, (|\nabla u|, \theta_i) \left( x_i + \frac{x - x_i}{\kappa_e} \right) \right) \, dx \leq \eta.
$$

We fix $\eta$ and arbitrary $A \subset \Omega \cap U_i$. Using convexity,

$$
\psi \left( x_i + \frac{x - x_i}{\kappa_e}, (|\nabla u|, \theta_i) \left( x_i + \frac{x - x_i}{\kappa_e} \right) \right) \\
\leq \psi \left( x_i + \frac{x - x_i}{\kappa_e}, |\nabla u| \left( x_i + \frac{x - x_i}{\kappa_e} \right) \right)
$$

as $0 \leq \theta_i \leq 1$. Second, using change of variables we have that

$$
\int_A \psi \left( x_i + \frac{x - x_i}{\kappa_e}, |\nabla u| \left( x_i + \frac{x - x_i}{\kappa_e} \right) \right) \, dx \\
= (\kappa_e)^d \int_{\tilde{A}} \psi(x, |\nabla u| (x)) \, dx \leq \int_{\tilde{A}} \psi(x, |\nabla u| (x)) \, dx,
$$

where $\tilde{A}$ is a set obtained from $A$ after the performed change of variables. Note that measures of these sets satisfy $|\tilde{A}| \leq \frac{1}{\kappa_e} |A| \leq 2^d |A|$. Having this in mind, we let

$$
\omega(\tau) := \sup_{C \subset \mathbb{R}^d : |C| \leq \tau} \int_{C} \psi(x, |\nabla u| (x)) \, dx.
$$
Function $\omega(\tau)$ is a non-decreasing function, continuous at 0 because $\mathcal{H}(u) < \infty$. Therefore, we may find $\tau$ such that $\omega(\tau) \leq \eta$. Then, we choose $\delta = 2^{-d} \tau$ to conclude the proof of (H1). Finally, (H2) follows from Young’s convolutional inequality and (H1).

$\square$

**Theorem 6.10.** (Theorem 2.3 in the general case) Let $u \in W_0^{1,\psi}(\Omega) \cap L^\infty(\Omega)$ where $\psi$ satisfies Assumptions 2.1 and 2.2. Let $\mathcal{H}$ be given with (2.1). Suppose that $q \leq p + \alpha \max\left(1, \frac{p}{d}\right)$.

Consider sequence $\{S^\varepsilon u\}_{\varepsilon > 0}$ as in (6.6) with $\varepsilon \leq \frac{R}{8}$. Then,

(I1) $S^\varepsilon u \in C^\infty_c(\Omega)$,

(I2) $\mathcal{H}(S^\varepsilon u) \to \mathcal{H}(u)$ as $\varepsilon \to 0$,

(I3) $S^\varepsilon u \to u$ in $W^{1,\psi}(\Omega)$ as $\varepsilon \to 0$,

(I4) space $C^\infty_c(\Omega)$ is dense in $W^{1,\psi}(\Omega)$ and Lavrentiev phenomenon does not occur, i.e. for all boundary data $u_0 \in W^{1,\psi}(\Omega)$:

$$\inf_{u \in u_0 + W^{1,p}_0(\Omega)} \mathcal{H}(u) = \inf_{u \in u_0 + W^{1,q}_0(\Omega)} \mathcal{H}(u) = \inf_{u \in u_0 + C^\infty_c(\Omega)} \mathcal{H}(u).$$

**Proof.** The first property follows from Lemma 6.6. To prove (I2), we note that

$$\mathcal{H}(S^\varepsilon u) = \int_{\Omega} \psi(x, \nabla S^\varepsilon u(x)) \, dx.$$  

To take mollification out of the function $\psi$ we want to use Jensen’s inequality and Lemma 6.2. The latter requires estimate on $\|\nabla S^\varepsilon u\|_\infty$

**Case 1:** $p \leq d$. In this case we have $q \leq p + \alpha$. Using Young’s convolution inequality we obtain

$$\|\nabla S^\varepsilon u\|_\infty \leq \sum_{i=1}^n \|u \theta_i\|_\infty \|\nabla \eta_\varepsilon\|_1 \leq \sum_{i=1}^n \|u\|_\infty \|\nabla \eta_\varepsilon\|_1 \leq D (\varepsilon C_{\Omega,R})^{-1},$$

where we choose $D := n \|u\|_\infty \|\nabla \eta_\varepsilon\|_1 C_{\Omega,R}$ and $C_{\Omega,R}$ is a constant from Lemma 6.8.

Let $x \in \Omega$ be fixed. Applying Lemma 6.2 with $\gamma = \varepsilon C_{\Omega,R}$ we obtain constants $\mathcal{M}, \mathcal{N}$ such that

$$\psi_x(\nabla S^\varepsilon u) \leq \mathcal{M} \psi^{**}_{x,\gamma}(|\nabla S^\varepsilon u|) + \mathcal{N},$$

where function $\psi^{**}_{x,\gamma}$ is the second convex conjugate of the function defined in (6.1). Now, we want to estimate $\psi^{**}_{x,\gamma}(|\nabla S^\varepsilon u|)$. Due to its convexity, Jensen’s inequality
implies
\[
\psi_{x,y}^{**} \left( \left\lVert \nabla S^\varepsilon u \right\rVert \right) = \psi_{x,y}^{**} \left( \sum_{i=1}^{n} \int_{B_\varepsilon} \nabla_x (u \theta_i) \left( x_i + \frac{x - x_i - y}{\kappa_\varepsilon} \right) \eta_\varepsilon(y) \, dy \right)
\]
\[
\leq \int_{B_\varepsilon} \psi_{x,y}^{**} \left( \sum_{i=1}^{n} \nabla_x (u \theta_i) \left( x_i + \frac{x - x_i - y}{\kappa_\varepsilon} \right) \right) \eta_\varepsilon(y) \, dy
\]
\[
\leq \frac{1}{2} \int_{B_\varepsilon} \psi_{x,y}^{**} \left( \frac{2}{\kappa_\varepsilon} \sum_{i=1}^{n} (\nabla u \theta_i) \left( x_i + \frac{x - x_i - y}{\kappa_\varepsilon} \right) \right) \eta_\varepsilon(y) \, dy
\]
\[
+ \frac{1}{2} \int_{B_\varepsilon} \psi_{x,y}^{**} \left( \frac{2}{\kappa_\varepsilon} \sum_{i=1}^{n} (u \nabla \theta_i) \left( x_i + \frac{x - x_i - y}{\kappa_\varepsilon} \right) \right) \eta_\varepsilon(y) \, dy
\]
\[
=: X + Y.
\]
Concerning term \( Y \), using upper bound from Lemma 6.2 and \( q \)-growth cf. Assumption 2.1 (A3) we can estimate as follows:
\[
\psi_{x,y}^{**} \left( \frac{2}{\kappa_\varepsilon} \sum_{i=1}^{n} (u \nabla \theta_i) \left( x_i + \frac{x - x_i - y}{\kappa_\varepsilon} \right) \right)
\]
\[
\leq C_2 \left( 1 + \frac{2n}{\kappa_\varepsilon} \left\| u \right\|_\infty \sup_{i=1,\ldots,n} \left\| \nabla \theta_i \right\|_\infty^q \right)
\]
\[
\leq C_2 \left( 1 + \frac{4n}{\kappa_\varepsilon} \left\| u \right\|_\infty \sup_{i=1,\ldots,n} \left\| \nabla \theta_i \right\|_\infty^q \right).
\]
Here we used \( \frac{1}{\kappa_\varepsilon} \leq 2 \) for \( \varepsilon \leq \frac{R}{8} \) in the second inequality. Therefore, using
\[
\int_{B_\varepsilon} \eta_\varepsilon(y) \, dy = 1,
\]
\[
Y \leq M C_2 \left( 1 + \frac{4n}{\kappa_\varepsilon} \left\| u \right\|_\infty \sup_{i=1,\ldots,n} \left\| \nabla \theta_i \right\|_\infty^q \right):= \mathcal{Y} < \infty.
\]
Concerning term \( X \), we use convexity so get that
\[
X \leq \frac{M}{2n} \sum_{i=1}^{n} \int_{B_\varepsilon} \psi_{x,y}^{**} \left( \frac{2n}{\kappa_\varepsilon} \left( \nabla u \theta_i \right) \left( x_i + \frac{x - x_i - y}{\kappa_\varepsilon} \right) \right) \eta_\varepsilon(y) \, dy,
\]
so that we can study each summand independently. If \( x_i + \frac{x - x_i - y}{\kappa_\varepsilon} \) does not belong to \( \Omega \) then \( \psi_{x,y}^{**} \left( \frac{2n}{\kappa_\varepsilon} \left( \nabla u \theta_i \right) \left( x_i + \frac{x - x_i - y}{\kappa_\varepsilon} \right) \right) = 0 \) because \( \nabla u \) vanishes at this point cf. Lemma 6.2 (G2). Otherwise, \( x_i + \frac{x - x_i - y}{\kappa_\varepsilon} \in \Omega \cap B_{C_\varepsilon,\varepsilon}(x) \) cf. Lemma 6.8 so that
\[
\psi_{x,y}^{**} \left( \frac{2n}{\kappa_\varepsilon} \left( \nabla u \theta_i \right) \left( x_i + \frac{x - x_i - y}{\kappa_\varepsilon} \right) \right)
\]
\[
\leq \psi \left( x_i + \frac{x - x_i - y}{\kappa_\varepsilon}, \frac{2n}{\kappa_\varepsilon} \left( \nabla u \theta_i \right) \left( x_i + \frac{x - x_i - y}{\kappa_\varepsilon} \right) \right),
\]
due to Lemma 6.2. Applying Assumption 2.1 (A4) iteratively, we obtain
\[
\psi \left( x_i + \frac{x - x_i - y}{\kappa \varepsilon}, \frac{2 n}{\kappa \varepsilon} \left| (\nabla u \theta_i) \left( x_i + \frac{x - x_i - y}{\kappa \varepsilon} \right) \right| \right) \leq C_k^4 \psi \left( x_i + \frac{x - x_i - y}{\kappa \varepsilon}, \frac{2 n}{\kappa \varepsilon} \left| (\nabla u \theta_i) \left( x_i + \frac{x - x_i - y}{\kappa \varepsilon} \right) \right| \right),
\]
where \( k \) is the smallest natural number such that \( \frac{2 n}{\kappa \varepsilon} \leq 2^k \) (see (A.2) in the proof of Lemma A.2). Using (6.8) and the fact that \( x \in \Omega_1 \) was arbitrary, we obtain the inequality
\[
\psi \left( x, \left| \nabla S^\varepsilon u \right| \right) \leq N + Y + \frac{C_k^4}{2n} \sum_{i=1}^n \int_{B_i} \psi \left( x_i + \frac{x - x_i - y}{\kappa \varepsilon}, \frac{1}{\kappa \varepsilon} \left| (\nabla u \theta_i) \left( x_i + \frac{x - x_i - y}{\kappa \varepsilon} \right) \right| \right) \eta(\varepsilon(y)) \, dy,
\]
valid for all \( x \in \Omega \). Now, we observe that \( \psi \left( x, \left| \nabla S^\varepsilon u \right| \right) \) converges a.e. to \( \psi \left( x, \left| \nabla u \right| \right) \). Moreover, the last term on (RHS) of (6.9) is convergent in \( L^1(\Omega) \) cf. Lemma 6.9 (H2). Therefore, Corollary A.2 (Vitali convergence theorem) implies
\[
\psi \left( x, \left| \nabla S^\varepsilon u \right| \right) \rightarrow \psi \left( x, \left| \nabla u \right| \right) \text{ in } L^1(\Omega) \text{ as } \varepsilon \rightarrow 0.
\]

Thanks to triangle inequality we obtain (I2). Now, (I3) follows from Lemma 4.1 (C4) while property (I4) follows from Lemma 4.3.

**Case 2:** \( p > d \). In this case we have \( q \leq p + \alpha \frac{d}{2} \). In this situation, instead of (6.7), we compute using change of variables
\[
\left\| \nabla S^\varepsilon u \right\|_\infty \leq \frac{1}{\kappa \varepsilon} \sum_{i=1}^n \left\| (\nabla u \theta_i) \right\|_p \left\| \eta \right\|_p \leq 2 \sum_{i=1}^n \left\| (\nabla u \theta_i) \right\|_p \left\| \eta \right\|_p',
\]
where \( p' \) is the usual Hölder conjugate exponent. Using change of variables we obtain
\[
\left\| \eta \right\|_{p'} = \int_{B_{\varepsilon \varepsilon}} \frac{1}{\varepsilon^{d \varepsilon}} \left| \eta \left( \frac{x}{\varepsilon} \right) \right|^{p'} \, dx = \varepsilon^{d \left( 1 - \frac{d}{p'} \right)} \int_B \left| \eta(x) \right|^{p'} \, dx = \varepsilon^{- d \frac{d}{p}} \left\| \eta \right\|_{p'},
\]
so that \( \left\| \eta \right\|_{p'} = \varepsilon^{- d \frac{d}{p}} \left\| \eta \right\|_{p'} \). Concerning the term with function \( u \),
\[
\left\| (\nabla u \theta_i) \right\|_p \leq \left\| \nabla u \right\|_p \left\| \theta_i \right\|_\infty + \left\| u \right\|_p \left\| \nabla \theta_i \right\|_\infty,
\]
which is finite as \( \mathcal{H}(u) < \infty \) and \( u \in L^\infty(\Omega) \). Therefore, (5.5) boils down to
\[
\left\| \nabla S^\varepsilon u \right\|_\infty \leq D \left( \varepsilon C_{\Omega, R} \right)^{- d \frac{d}{p}}, \quad D := 2 C_{\Omega, R}^d \left\| \eta \right\|_p \sum_{i=1}^n \left( \left\| \nabla u \right\|_p \left\| \theta_i \right\|_\infty + \left\| u \right\|_p \left\| \nabla \theta_i \right\|_\infty \right).
\]
Now, we can apply Lemma 6.2 (G1) to obtain estimate (6.8). The rest of the proof is exactly the same. □
7. Extension of Theorem 2.3 to vector-valued maps

Many authors consider variational problems with vector-valued functions. However, in our work functionals depend only on the length of the gradient so there is almost no difficulty in extending our result to the vector case setting. In this section, we write \( u = (u^1, \ldots, u^n) \) for the map \( u : \Omega \to \mathbb{R}^n \). For simplicity, we use the same notation for spaces of vector-valued functions as for spaces of scalar-valued ones.

The main point that needs explanation is a generalisation of Lemma 4.2, where we applied truncation to approximate functions from \( W^{1,\psi} (\Omega) \) by functions from \( W^{1,\psi} (\Omega) \cap L^\infty (\Omega) \).

Lemma 7.1. Let \( u : \Omega \to \mathbb{R}^n \), \( u = (u^1, \ldots, u^n) \) where \( n \in \mathbb{N} \). Suppose that \( u \in W^{1,\psi} (\Omega) \). Then, \( u^i \in W^{1,\psi} (\Omega) \). Moreover, suppose that for each \( i = 1, \ldots, n \) we have \( u^i_k \to u^i \) in \( W^{1,\psi} (\Omega) \). Let \( u_k := (u^1_k, \ldots, u^n_k) \). Then, \( u_k \to u \) in \( W^{1,\psi} (\Omega) \).

Proof. We observe that if we interpret \( |\nabla u| \) component-wisely, we have \( |\nabla u^i| \leq |\nabla u| \). By convexity of \( \xi \mapsto \psi (x, \xi) \), we have that
\[
0 \leq \psi (x, |\nabla u^i|) \leq \psi (x, |\nabla u|).
\]
To see the second statement, we note that
\[
|\nabla u - \nabla u_k| \leq \sum_{i=1}^n |\nabla u^i - \nabla u^i_k|,
\]
so that by convexity of the mapping \( \xi \mapsto \psi (x, \xi) \),
\[
0 \leq \psi (x, |\nabla u - \nabla u_k|) \leq \frac{1}{n} \sum_{i=1}^n \psi \left( x, n |\nabla u^i - \nabla u^i_k| \right).
\]
Using Lemma 4.1 (C2), we conclude that \( u_k \to u \) in \( W^{1,\psi} (\Omega) \).

Theorem 7.2. Suppose that \( p \leq q + \alpha \max \left( 1, \frac{d}{n} \right) \). Let \( \mathcal{H} \) be a functional defined by (2.1) with \( \psi \) satisfying Assumptions 2.1 and 2.2. Then, for all \( u_0 \in W^{1,q} (\Omega) \) we have that
\[
\inf_{u \in u_0 + W^{1,p}_0 (\Omega)} \mathcal{H}(u) = \inf_{u \in u_0 + W^{1,q}_0 (\Omega)} \mathcal{H}(u) = \inf_{u \in u_0 + C_0^\infty (\Omega)} \mathcal{H}(u).
\]
Moreover, space \( C_0^\infty (\Omega) \) is dense in the Musielak–Orlicz–Sobolev space \( W^{1,\psi}_0 (\Omega) \).

Proof. We first prove that \( C_0^\infty (\Omega) \) is dense in the Musielak–Orlicz–Sobolev space \( W^{1,\psi}_0 (\Omega) \). This follows from the following facts:

- \( W^{1,\psi}_0 (\Omega) \cap L^\infty (\Omega) \) is dense in \( W^{1,\psi}_0 (\Omega) \). Indeed, let \( u \in W^{1,\psi}_0 (\Omega) \) and \( u_k := (T_k(u^1), \ldots, T_k(u^n)) \) where \( T_k \) was defined in (4.4). It follows from Lemmas 4.2 and 7.1 that \( u_k \to u \) in \( W^{1,\psi} (\Omega) \).
\* \( C_c^\infty(\Omega) \) is dense in \( W_0^{1,\psi}(\Omega) \cap L^\infty(\Omega) \). Indeed, let \( \mathbf{u} \in W_0^{1,\psi}(\Omega) \cap L^\infty(\Omega) \). Then, each \( u^i \in W_0^{1,\psi}(\Omega) \cap L^\infty(\Omega) \). By Theorem 2.3, we have a sequence \( \{u^i_k\}_{k \in \mathbb{N}} \subset C_c^\infty(\Omega) \) such that \( u^i_k \to u^i \) in \( W^{1,\psi}(\Omega) \). Let \( \mathbf{u}_k := (u^1_k, \ldots, u^n_k) \). By Lemma 7.1, \( \mathbf{u}_k \to \mathbf{u} \) in \( W^{1,\psi}(\Omega) \).

Having density of \( C_c^\infty(\Omega) \) in \( W_0^{1,\psi}(\Omega) \) in hand, the absence of Lavrentiev phenomenon follows as in the proof of Lemma 4.3.

\[ \square \]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix A

A.1. Vitali convergence theorem

In this section we recall a convergence result that is used several times in this paper. For the proof, see [8, Theorem 4.5.4].

**Theorem A.1.** Let \((X, \mathcal{F}, \mu)\) be a finite measure space (i.e. \( \mu(X) < \infty \)). Let \( \{f_n\}_{n \in \mathbb{N}} \subset L^1(X, \mathcal{F}, \mu) \) and \( f \) be an \( \mathcal{F} \)-measurable function. Then, \( f_n \to f \) in \( L^1(X, \mathcal{F}, \mu) \) if and only if \( f_n \to f \) in measure and \( \{f_n\}_{n \in \mathbb{N}} \) is uniformly integrable, i.e.

\[
\forall \varepsilon > 0 \exists \delta > 0 \forall A \in \mathcal{F} \quad \mu(A) < \delta \Rightarrow \sup_{n \in \mathbb{N}} \int_A |f_n| \, d\mu < \varepsilon.
\]

In the proof of the main results, we have applied the following corollary.

**Corollary A.2.** Let \((X, \mathcal{F}, \mu)\) be a finite measure space (i.e. \( \mu(X) < \infty \)). Let \( \{f_n\}_{n \in \mathbb{N}} \subset L^1(X, \mathcal{F}, \mu) \) be a nonnegative sequence and \( f \) be an \( \mathcal{F} \)-measurable function. Suppose that

(J1) \( f_n \to f \) in measure,

(J2) there exists a sequence of functions \( \{g_n\}_{n \in \mathbb{N}} \) convergent in \( L^1(X, \mathcal{F}, \mu) \) and function \( h \in L^1(X, \mathcal{F}, \mu) \) such that

\[
0 \leq f_n \leq g_n + h.
\]

Then, \( f_n \to f \) in \( L^1(X, \mathcal{F}, \mu) \).
Proof. In view of Theorem A.1, it is sufficient to prove that \( \{ f_n \}_{n \in \mathbb{N}} \) is uniformly integrable. To this end, for an arbitrary set \( A \), we have

\[
\int_A |f_n| \, d\mu = \int_A f_n \, d\mu \leq \int_A g_n \, d\mu + \int_A h \, d\mu \leq \int_A |g_n + h| \, d\mu.
\]

Let \( \varepsilon > 0 \). As \( \{ g_n \}_{n \in \mathbb{N}} \) is convergent in \( L^1(X, \mathcal{F}, \mu) \), the same is true for \( \{ g_n + h \}_{n \in \mathbb{N}} \). It follows that \( \{ g_n + h \}_{n \in \mathbb{N}} \) is uniformly integrable. Therefore, there exists \( \delta > 0 \) such that if \( \mu(A) < \delta \), we have \( \int_A |g_n + h| \, d\mu < \varepsilon \). It follows that

\[
\int_A |f_n| \, d\mu < \varepsilon.
\]

\( \square \)

Lemma A.3. Let \( \{ f_n \} \) be a sequence of functions and \( \psi \) be an \( \mathcal{N} \)-function satisfying (A1), (A2) and first inequality in (A3) in Assumption 2.1. Assume that

\[
\int_{\Omega_1} \psi(x, |f_n - f|) \rightarrow 0.
\]

Then, \( f_n \rightarrow f \) in measure.

Proof. Let \( \{ \mu_x \} \) be the Young measure generated by sequence \( \{ f_n - f \} \) cf. [36, Theorem 6.2] (note that with \( p \geq 1 \)

\[
\sup_{n \in \mathbb{N}} \int_{\Omega} g(|f_n - f|) < \infty \quad \text{for} \quad g(x) = \begin{cases} C_1 |\xi_0|^p & \text{if} \ x < \xi_0, \\ C_1 |\xi|^p & \text{if} \ x > \xi_0. \end{cases}
\]

due to (A3) in Assumption 2.1 so that assumptions of [36, Theorem 6.2] are satisfied). Hence,

\[
0 = \int_{\Omega} \int_{\mathbb{R}^+} \psi(x, \lambda) \, d\mu_x(\lambda) \, dx
\]

so that for a.e. \( x \in \Omega, \mu_x = \delta_0 \). We conclude that \( f_n - f \rightarrow 0 \) in measure [37, Lemma 4.12]. \( \square \)

A.2. Proof of Lemma 4.1

Proof. The first equivalence in (C1) follows directly from definition of the norm (4.1) so in fact it is sufficient to prove that if \( \int_{\Omega} \psi(x, c|f(x)|) \, dx < \infty \) for some \( c > 0 \) then \( \int_{\Omega} \psi(x, d|f(x)|) \, dx < \infty \) for all \( d > 0 \). First, if \( d < c \), this follows by convexity and Jensen’s inequality:

\[
\int_{\Omega} \psi(x, d|f(x)|) \, dx = \int_{\Omega} \psi \left( x, \frac{d}{c} |f(x)| + 0 \right) \, dx \leq \frac{d}{c} \int_{\Omega} \psi(x, c|f(x)|) \, dx. \quad (A.1)
\]

If \( d > c \), we find \( k \in \mathbb{N} \) such that \( d \leq 2^k c \) and apply (A4) in Assumption 2.1:

\[
\int_{\Omega} \psi(x, d|f(x)|) \, dx \leq C_4^k \int_{\Omega} \psi \left( x, \frac{d}{2^k} |f(x)| \right) \, dx \leq C_4^k \frac{d}{2^k c} \int_{\Omega} \psi(x, c|f(x)|) \, dx. \quad (A.2)
\]

where we used the first part.
Concerning (C2), we first prove equivalence:

\[ \|f_n - f\|_\psi \to 0 \iff \int_\Omega \psi(x, c | f_n - f|) \, dx \to 0 \text{ for all } c > 0. \]

To prove \( \Rightarrow \) we fix \( c > 0 \) and we note that there exists \( n_c \) such that for all \( n \geq n_c \) we have \( c \| f_n - f \|_\psi < 1. \) By definition (4.1), there exists a sequence \( \{\delta_k\}_{k \in \mathbb{N}} \) convergent to 0 such that \( c \| f_n - f \|_\psi + \delta_k < 1 \) and

\[ \int_\Omega \psi \left( x, \frac{c | f_n - f |}{c \| f_n - f \|_\psi + \delta_k} \right) \leq 1. \]

Using convexity of \( \psi \) and equality \( \psi(x, 0) = 0 \) we obtain

\[ \int_\Omega \psi(x, c | f_n - f|) \leq (\|f_n - f\|_\psi + \delta_k) \int_\Omega \psi \left( x, \frac{c | f_n - f |}{\|f_n - f\|_\psi + \delta_k} \right) \leq c \|f_n - f\|_\psi + \delta_k. \]

Letting \( k \to \infty \) (so that \( \delta_k \to 0 \)) and \( n \to \infty \) we conclude the proof. For \( \Leftarrow \), we note that for each \( c > 0 \), there exists \( n_c \) such that for all \( n \geq n_c \) we have \( \int_\Omega \psi(x, c | f_n - f|) \, dx \leq 1 \), i.e. \( \|f_n - f\|_\psi \leq \frac{1}{c} \). The conclusion follows by letting \( c \to \infty \). We are left to prove equivalence

\[ \int_\Omega \psi(x, c | f_n - f|) \, dx \to 0 \text{ for all } c > 0 \iff \int_\Omega \psi(x, c | f_n - f|) \, dx \to 0 \text{ for some } c > 0. \]

This follows from (A.1) and (A.2).

To prove (C3) we assume that \( \int_\Omega \psi(x, | f_n - f|) \, dx \to 0 \) and \( \|f\|_\psi < \infty \) which implies \( \int_\Omega \psi(x, | f|) \, dx < \infty \). First, thanks to Lemma A.3, we deduce that \( f_n \to f \) in measure. Second, we can estimate

\[ 0 \leq \psi(x, |f_n|) \leq \psi \left( x, \frac{1}{2} |f_n - f| + \frac{1}{2} |f| \right) \leq \frac{1}{2} \psi(x, 2 |f_n - f|) + \frac{1}{2} \psi(x, 2 |f|) \leq \frac{C_4}{2} \psi(x, |f_n - f|) + \frac{C_4}{2} \psi(x, |f|). \]

Corollary A.2 implies that \( \psi(x, |f_n|) \to \psi(x, |f|) \) in \( L^1(\Omega) \) so in particular, \( \int_\Omega \psi(x, |f_n|) \, dx \to \int_\Omega \psi(x, |f|) \, dx \).

Concerning (C4), in view of Vitali convergence theorem cf. Theorem A.1, it is sufficient to prove that the sequence \( \{\psi(x, f_n - f)\}_n \) is uniformly integrable. Using convexity and \( \Delta_2 \) condition we obtain

\[ 0 \leq \psi(x, |f_n - f|) \leq \psi \left( x, \frac{1}{2} |f_n| + \frac{1}{2} |f| \right) \leq \frac{1}{2} \psi(x, 2 |f_n|) + \frac{1}{2} \psi(x, 2 |f|) \leq \frac{C_4}{2} \psi(x, |f_n|) + \frac{C_4}{2} \psi(x, |f|). \]

It follows that \( \psi(x, |f_n - f|) \to 0 \) in \( L^1(\Omega) \) and the conclusion follows from (C2).
Finally, to show (C5), we have to prove that if \( \int_{\Omega} \psi(x, |f(x)|) \, dx < \infty \) then \( \int_{\Omega} |f(x)|^p \, dx < \infty \). To this end, we estimate
\[
\int_{\Omega} |f(x)|^p \, dx = \int_{\Omega \cap \{|f| \leq \xi_0\}} |f(x)|^p \, dx + \int_{\Omega \cap \{|f| \geq \xi_0\}} |f(x)|^p \, dx \\
\leq |\xi_0|^p + \int_{\Omega} \psi(x, |f(x)|) \, dx
\]
where we used (A3) in Assumption 2.1.

References

1. Ahmida, Y., Chlebicka, I., Gwiazda, P., Youssfi, A.: Gossez’s approximation theorems in Musielak–Orlicz–Sobolev spaces. *J. Funct. Anal.* **275**(9), 2538–2571, 2018
2. Alberti, G., Serra Cassano, F.: Non-occurrence of gap for one-dimensional autonomous functionals. In *Calculus of Variations, Homogenization and Continuum Mechanics (Marseille, 1993)*, Volume 18 of *Ser. Adv. Math. Appl. Sci.*, pp. 1–17. World Scientific Publishing, River Edge, 1994
3. Balci, A.K., Diening, L., Surnachev, M.: New examples on Lavrentiev gap using fractals. *Calc. Var. Partial Differ. Equ.*, **59**(5): Paper No. 180, 34, 2020
4. Balci, A.K., Surnachev, M.: Lavrentiev gap for some classes of generalized Orlicz functions. *Nonlinear Anal.*, 207: Paper No. 112329, 22, 2021
5. Baroni, P., Colombo, M., Mingione, G.: Harnack inequalities for double phase functionals. *Nonlinear Anal.* **121**, 206–222, 2015
6. Baroni, P., Colombo, M., Mingione, G.: Regularity for general functionals with double phase. *Calc. Var. Partial Differ. Equ.*, **57**(2): Paper No. 62, 48, 2018
7. Bella, P., Schäffner, M.: On the regularity of minimizers for scalar integral functionals with \((p, q)\)-growth. *Anal. PDE* **13**(7), 2241–2257, 2020
8. Bogachev, V.I.: Measure Theory, vol. I, II. Springer, Berlin, 2007
9. Bouguet, P.: Non occurrence of the Lavrentiev gap for multidimensional autonomous problems. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **2022** (to appear)
10. Breit, D.: New regularity theorems for non-autonomous variational integrals with \((p, q)\)-growth. *Calc. Var. Partial Differ. Equ.* **44**(1–2), 101–129, 2012
11. Bulíček, M., Gwiazda, P., Skrzeczkowski, J.: Parabolic equations in Musielak–Orlicz spaces with discontinuous in time \(N\)-function. *J. Differ. Equ.* **290**, 17–56, 2021
12. Byun, S.-S., Oh, J.: Regularity results for generalized double phase functionals. *Anal. PDE* **13**(5), 1269–1300, 2020
13. Chlebicka, I., Gwiazda, P., Wróblewska-Kamińska, A., Świernczewska-Gwiazda, A.: Partial Differential Equations in anisotropic Musielak–Orlicz spaces. Springer Monographs in Mathematics. Springer, Cham (2021)
14. Chlebicka, I., Gwiazda, P., Zatorska-Goldstein, A.: Parabolic equation in time and space dependent anisotropic Musielak–Orlicz spaces in absence of Lavrentiev’s phenomenon. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **36**(5), 1431–1465, 2019
15. Chlebicka, I., Gwiazda, P., Zatorska-Goldstein, A.: Renormalized solutions to parabolic equations in time and space dependent anisotropic Musielak–Orlicz spaces in absence of Lavrentiev’s phenomenon. *J. Differ. Equ.* **267**(2), 1129–1166, 2019
16. Colombo, M., Mingione, G.: Bounded minimisers of double phase variational integrals. *Arch. Ration. Mech. Anal.* **218**(1), 219–273, 2015
17. Colombo, M., Mingione, G.: Regularity for double phase variational problems. *Arch. Ration. Mech. Anal.* **215**(2), 443–496, 2015
18. De Filippis, C., Mingione, G.: On the regularity of minima of non-autonomous functionals. *J. Geom. Anal.* **30**(2), 1584–1626, 2020
19. De Filippis, C., Palatucci, G.: Hölder regularity for nonlocal double phase equations. *J. Differ. Equ.* **267**(1), 547–586, 2019
20. **Esposito, A., Leonetti, F., Vincenzo Petricca, P.**: Absence of Lavrentiev gap for non-autonomous functionals with \((p, q)\)-growth. *Adv. Nonlinear Anal.* **8**(1), 73–78, 2019

21. **Esposito, L., Leonetti, F., Mingione, G.**: Sharp regularity for functionals with \((p, q)\) growth. *J. Differ. Equ.* **204**(1), 5–55, 2004

22. **Evans, L.C.**: *Partial Differential Equations. Graduate Studies in Mathematics*. American Mathematical Society, Providence, 1998

23. **Harjulehto, P., Hästö, P.**: *Orlicz Spaces and Generalized Orlicz Spaces. Lecture Notes in Mathematics*, vol. 2236. Springer, Cham, 2019

24. **Lavrentieff, M.**: Sur quelques problèmes du calcul des variations. *Ann. Mat. Pura Appl.* **4**(1), 7–28, 1927

25. **Leonetti, F., Siepe, F.**: Maximum principle for vector valued minimizers. *J. Convex Anal.* **12**(2), 267–278, 2005

26. **Mania, B.**: Sopra un esempio di Lavrentieff. *Boll. Un. Mat. Ital.* **13**, 147–153, 1934

27. **Marcellini, P.**: Regularity of minimizers of integrals of the calculus of variations with nonstandard growth conditions. *Arch. Ration. Mech. Anal.* **105**(3), 267–284, 1989

28. **Marcellini, P.**: Regularity and existence of solutions of elliptic equations with \(p, q\)-growth conditions. *J. Differ. Equ.* **90**(1), 1–30, 1991

29. **Marcellini, P.**: Regularity for elliptic equations with general growth conditions. *J. Differ. Equ.* **105**(2), 296–333, 1993

30. **Marcellini, P.**: Everywhere regularity for a class of elliptic systems without growth conditions. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)** **23**(1), 1–25, 1996

31. **Mariconda, C., Treu, G.**: Non-occurrence of a gap between bounded and Sobolev functions for a class of nonconvex Lagrangians. *J. Convex Anal.* **27**(4), 1247–1259, 2020

32. **Mariconda, C., Treu, G.**: Non-occurrence of the Lavrentiev phenomenon for a class of convex nonautonomous Lagrangians. *Open Math.* **18**(1), 1–9, 2020

33. **Mingione, G.**: Regularity of minima: an invitation to the dark side of the calculus of variations. *Appl. Math.* **51**(4), 355–426, 2006

34. **Novotný, A., Straškraba, I.**: *Introduction to the Mathematical Theory of Compressible Flow. Oxford Lecture Series in Mathematics and Its Applications*, vol. 27. Oxford University Press, Oxford, 2004

35. **Ok, J.**: Regularity for double phase problems under additional integrability assumptions. *Nonlinear Anal.* **194**, 111408, 13, 2020

36. **Pedregal, P.**: *Parametrized Measures and Variational Principles. Progress in Nonlinear Differential Equations and their Applications*, vol. 30. Birkhäuser Verlag, Basel, 1997

37. **Rindler, F.**: *Calculus of Variations. Universitext*. Springer, Cham, 2018

38. **Schmidt, T.**: Regularity of minimizers of \(W^{1,p}\)-quasiconvex variational integrals with \((p, q)\)-growth. *Calc. Var. Partial Differ. Equ.* **32**(1), 1–24, 2008

39. **Zhikov, V.**: Lavrentiev phenomenon and homogenization for some variational problems. *C. R. Acad. Sci. Paris Sér. I Math.* **316**(5), 435–439, 1993

40. **Zhikov, V.V.**: Problems of convergence, duality, and averaging for a class of functionals of the calculus of variations. *Dokl. Akad. Nauk SSSR* **267**(3), 524–528, 1982

41. **Zhikov, V.V.**: On Lavrentiev’s phenomenon. *Russ. J. Math. Phys.* **3**(2), 249–269, 1995

42. **Zhikov, V.V.**: On some variational problems. *Russ. J. Math. Phys.* **5**(1), 105–116, 1997
