A HIGHER ORDER MODEL FOR IMAGE RESTORATION: THE ONE DIMENSIONAL CASE

G. DAL MASO, I. FONSECA, G. LEONI, AND M. MORINI

Abstract. The one-dimensional version of the higher order total variation-based model for image restoration proposed by Chan, Marquina, and Mulet in [4] is analyzed. A suitable functional framework in which the minimization problem is well posed is being proposed and it is proved analytically that the higher order regularizing term prevents the occurrence of the *staircase effect*.

Keywords: image segmentation, total variation models, staircase effect, higher order regularization, relaxation
2000 Mathematics Subject Classification: 49J45 (26A45, 65K10, 68U10)

Contents

1. Introduction

Deblurring and denoising of images are fundamental problems in image processing and gave rise in the past few years to a vast variety of techniques and methods touching different fields of mathematics. Among them, variational methods based on the minimization of some energy functional have been successfully employed to treat a fairly general class of image restoration problems. Typically, such functionals present a *fidelity term*, which penalizes the distance between the reconstructed image u and the noisy image g with respect to a suitable metric, and a regularizing term, which makes high frequency noise energetically unfavorable.

When the fidelity term is given by the squared L^2 distance multiplied by a parameter $\lambda > 0$ and the regularizing term is represented by the total variation, we are led to the following minimization problem

$$\min\left\{|Du|(\Omega) + \lambda\int_\Omega |u-g|^2\,dx : u \in BV(\Omega)\right\}, \tag{1.1}$$

which was proposed by Rudin, Osher, and Fatemi in [11]. Here Ω is an open bounded domain in one or two dimensions, $BV(\Omega)$ denotes the space of functions of bounded variations in Ω, and $|Du|(\Omega)$ stands for the total variation of u in Ω. The main feature of the total variation-based image restoration is perhaps represented by the tendency to yield (almost) piecewise-constant solutions or, in other words, "blocky" images. Typically one observes that *ramps* (i.e., affine regions) in the original image give rise to staircase-like structures in the reconstructed image, a phenomenon which is often referred to as the *staircase effect*. This means that the original edges are well preserved by this method, but also that many artificial discontinuities can be generated by the presence of noise, while the finer details of the objects contained in the image may not be properly recovered.

Several variants of (1.1) have been subsequently proposed in order to fix these drawbacks. In this paper we follow the approach of Chan, Marquina, and Mulet [4]: Since the total variation does not distinguish between jumps and smooth transitions their idea is to consider an additional penalization of the discontinuities by taking second derivatives into account. More precisely, they propose a regularizing term of the form

$$\int_\Omega |\nabla u|\,dx + \int_\Omega \psi(|\nabla u|)h(\Delta u)\,dx\,, \tag{1.2}$$

where ψ is a function that must satisfy suitable conditions at infinity in order to allow jumps.

In this paper we consider the following 1-D version of (1.2):

$$\mathcal{F}_p(u) := \int_a^b |u'|\,dx + \int_a^b \psi(|u'|)|u''|^p\,dx\,, \tag{1.3}$$

where $a < b$ are real numbers and $p \in [1, +\infty)$. Our main analytical objective is twofold:

(i) to set up a proper functional framework where the minimization problem corresponding to
$$\mathcal{F}_p(u) + \lambda \int_\Omega |u-g|^2\,dx$$
is well posed;

(ii) to give an analytical proof of the fact the higher order regularizing term eliminates the staircase effect.

We point out here that we carry out the first part of this program by using the theory of relaxation (see [5] for a general introduction): We regard $\mathcal{F}_p$ as defined for all functions in the Sobolev space $W^{2,p}(]a,b[)$, we extend it to $L^1(]a,b[)$ by setting $\mathcal{F}_p(u) := +\infty$ if $u \in L^1(]a,b[) \setminus W^{2,p}(]a,b[)$, and then we identify its lower semicontinuous envelope with respect to the strong L^1 convergence. The extension of our results to higher dimensions will be the subject of a subsequent paper.

For completeness we conclude by mentioning that other approaches have been considered to avoid staircasing: The works by Geman and Reynolds [7] and Chambolle and Lions [3] contain a different use of higher order derivatives as regularizing terms; in [2], Blomgren, Chan, and Mulet propose a BV-H^1 interpolation approach, while Kindermann, Osher, and Jones avoid in [9] the use of second derivatives by considering a sort of nonlocal total variation.

The plan of the paper is the following: In Section 2 we consider the case $p = 1$; i.e., we identify the relaxation $\overline{\mathcal{F}}_1$ of $\mathcal{F}_1$, while in Section 3 we deal with the case $p > 1$. The analysis turns out to be considerably more delicate in the former case. Moreover, the domains of the relaxed functionals are quite peculiar (see Definitions 2.1 and 3.1) and display properties which are qualitatively different in the two cases. In particular, it turns out that piecewise constant functions corresponding to images with genuine edges are approximable by sequences with bounded energy only for $p = 1$. Finally, in Section 4 we investigate the staircase effect. After exhibiting an analytical example of staircasing for the Rudin-Osher-Fatemi model (Theorem 4.3), we prove that the new model does indeed prevent the occurrence of this phenomenon. More precisely, we show that whenever the datum g is of the form $g = g_1 + h$, with g_1 a regular image and h a highly oscillating noise, the reconstructed image is regular as well (Theorems 4.5 and 4.8).

2. The case $p = 1$

We start by studying the compactness properties and the relaxation of (1.3) in the case $p = 1$. Throughout this section $\psi\colon \mathbb{R} \to]0, +\infty[$ will be a bounded Borel function such that

$$M := \int_{-\infty}^{+\infty} \psi(t)\, dt < +\infty \tag{2.1}$$

and

$$\inf_{t\in K} \psi(t) > 0 \quad \text{for every compact set } K \subset \mathbb{R}\,. \tag{2.2}$$

Let $\Psi_1 \colon \overline{\mathbb{R}} \to [0, M]$ be the increasing function defined by

$$\Psi_1(t) := \int_{-\infty}^{t} \psi(s)\, ds$$

and let $\Psi_1^{-1} \colon [0, M] \to \overline{\mathbb{R}}$ be its inverse function.

Given a bounded open interval $]a, b[$ in $\mathbb{R}$, we consider the functional $\mathcal{F}_1 \colon L^1(]a, b[) \to [0, +\infty]$ defined by

$$\mathcal{F}_1(u) := \begin{cases} \displaystyle\int_a^b |u'|\, dx + \int_a^b \psi(u')|u''|\, dx & \text{if } u \in W^{2,1}(]a, b[)\,, \\ +\infty & \text{otherwise.} \end{cases} \tag{2.3}$$

The first step in the study of (2.3) will consist in identifying the subspace of L^1 functions which can be approximated by energy bounded sequences. In order to do so we need to introduce some notation and recall some basic facts about BV functions of one variable. This will be the content of the next subsection.

2.1. BV functions of one variable.

We recall that a function $u \in L^1(]a, b[)$ belongs to $BV(]a, b[)$ if and only if

$$\sup\Big\{ \int_a^b u\varphi'\, dx : \varphi \in C^1_c(]a, b[)\,,\ |\varphi| \le 1 \Big\} < +\infty\,. \tag{2.4}$$

Note that this implies that the distributional derivative u' of u is a bounded Radon measure in $]a, b[$. We will often consider the Lebesgue decomposition

$$u' = (u')^a \mathcal{L}^1 + (u')^s$$

where $(u')^a$ is the density of the absolutely continuous part of u' with respect to the Lebesgue measure $\mathcal{L}^1$ on $]a, b[$, while $(u')^s$ is its singular part. We will denote the total variation measure of u' by $|u'|$. In particular, $|u'|(]a, b[)$ equals the value of the supremum in (2.4). For every function $u \in BV(]a, b[)$ the following left and right approximate limits

$$u_-(y) := \lim_{\varepsilon\to 0+} \frac{1}{\varepsilon} \int_{y-\varepsilon}^{y} u(x)\, dx\,, \qquad u_+(y) := \lim_{\varepsilon\to 0+} \frac{1}{\varepsilon} \int_{y}^{y+\varepsilon} u(x)\, dx$$

are well defined at every point $y \in]a, b[$. In fact, $u_-(y)$ is well defined also at $y = b$ while $u_+(y)$ exists also at $y = a$. The functions u_- and u_+ coincide $\mathcal{L}^1$-a.e. with u and are left and right continuous, respectively. Moreover, it turns out that the set $S_u := \{y \in]a, b[: u_-(y) \neq u_+(y)\}$ is at most countable. The set S_u is often referred to as the set of essential discontinuities or *jump points* of u.

It is well known that, in turn, the singular part $(u')^s$ splits into the sum of an atomic measure concentrated on S_u and a singular diffuse measure $(u')^c$, called the *Cantor part* of u':

$$(u')^s = [u]\mathcal{H}^0 \llcorner S_u + (u')^c \,,$$

where we set $[u] := u_+ - u_-$ and $\mathcal{H}^0$ stands for the counting measure. Finally, we recall that every $u \in BV(]a,b[)$ is differentiable at $\mathcal{L}^1$-a.e. y in $]a,b[$ with derivative given by $(u')^a(y)$. In this case we will often write, with a slight abuse of notation, $u'(y)$ instead of $(u')^a(y)$.

We say that a sequence $\{u_k\}$ of functions in $BV(]a,b[)$ *weakly star converges in* $BV(]a,b[)$ to a function $u \in BV(]a,b[)$ if $u_n \to u$ in $L^1(]a,b[)$ and $u'_k \rightharpoonup u'$ weakly* in $M_b(]a,b[)$, where $M_b(]a,b[)$ is the space of bounded Radon measures.

We will also need sometimes the notion of total variation for a function defined everywhere. We recall that $u\colon]a,b[\to \mathbb{R}$ has bounded *pointwise total variation* over the interval $]c,d[\subset]a,b[$ if

$$\mathrm{Var}\,(u;]c,d[) := \sup \sum_{i=1}^{k} |u(y_i) - u(y_{i-1})| < +\infty \,,$$

where the supremum is taken over all finite families $y_0, y_1, \ldots, y_k$ such that $c < y_0 < y_1 < \cdots < y_k < d$, $k \in \mathbb{N}$. It is easy to see that if u has bounded pointwise total variation in $]a,b[$, then it admits left and right limits at every point, it belongs to $BV(]a,b[)$, and $|u'|(]c,d[) \le \mathrm{Var}\,(u;]c,d[)$ for every interval $]c,d[\subset]a,b[$. Conversely, if $u \in BV(]a,b[)$, the *precise representatives* u_- and u_+ have bounded pointwise total variation and satisfy

$$|u'|(]c,d[) = \mathrm{Var}\,(u_-;]c,d[) = \mathrm{Var}\,(u_+;]c,d[)$$

for every interval $]c,d[\subset]a,b[$.

Finally, we recall the Helly theorem: For every bounded sequence of functions $u_k :]a,b[\to \mathbb{R}$ such that $\sup_k \mathrm{Var}\,(u_k;]a,b[) < +\infty$, there exist u, with pointwise total variation in $]a,b[$, and a subsequence (not relabeled) such that $u_k \to u$ pointwise.

We refer to [12] and [8] for an exhaustive exposition of the properties of BV functions of one variable.

2.2. **Compactness.** To define the subspace of L^1 functions that can be approximated by energy bounded sequences, for every function $u \in BV\,(]a,b[)$ we consider the sets

$$Z^+[(u')^a] := \left\{ x \in]a,b[\, : \lim_{\varepsilon \to 0+} \frac{1}{2\varepsilon} \int_{x-\varepsilon}^{x+\varepsilon} (u')^a \, dx = +\infty \right\}, \tag{2.5}$$

$$Z^-[(u')^a] := \left\{ x \in]a,b[\, : \lim_{\varepsilon \to 0+} \frac{1}{2\varepsilon} \int_{x-\varepsilon}^{x+\varepsilon} (u')^a \, dx = -\infty \right\}. \tag{2.6}$$

It is also convenient to define

$$Z[(u')^a] := Z^+[(u')^a] \cup Z^-[(u')^a] \,.$$

Definition 2.1. Let $X^1_\psi(]a,b[)$ be the set of all functions $u \in BV(]a,b[)$ such that $v := \Psi_1 \circ (u')^a$ belongs to $BV(]a,b[)$ and the positive part $((u')^c)^+$ and the negative part $((u')^c)^-$ of the measure $(u')^c$ are concentrated on $Z^+[(u')^a]$ and $Z^-[(u')^a]$, respectively.

Remark 2.2. Note that if $u \in X^1_\psi(]a,b[)$ then the limits

$$(u')^a_-(y) := \lim_{\varepsilon \to 0+} \frac{1}{\varepsilon} \int_{y-\varepsilon}^{y} (u')^a \, dx \,, \qquad (u')^a_+(y) := \lim_{\varepsilon \to 0+} \frac{1}{\varepsilon} \int_{y}^{y+\varepsilon} (u')^a \, dx \tag{2.7}$$

exist in $\overline{\mathbb{R}}$ for every y. More precisely, $(u')^a_-$ exists also at $y = b$ while $(u')^a_+$ is well defined also at $y = a$. Indeed, since $v = \Psi_1 \circ (u')^a$ is a BV function, it admits a precise representative $\tilde{v}$ such that the right and left limits exist at every point, and the same property holds for $\Psi_1^{-1}(\tilde{v})$. As $\Psi_1^{-1}(\tilde{v}) = (u')^a$ $\mathcal{L}^1$-a.e. in $]a,b[$, the limits considered in (2.7) are everywhere well-defined. Moreover the set $S_{(u')^a} := S_v$ is at most countable and

$$(u')^a_- = (u')^a_+ \quad \text{on }]a,b[\setminus S_{(u')^a}. \tag{2.8}$$

We also remark that $(u')^a_-$ and $(u')^a_+$ are left and right continuous, which, in turn, implies that the functions defined by

$$(u')^a_\vee(x) := \max\left\{(u')^a_+(x),(u')^a_-(x)\right\}, \ (u')^a_\wedge(x) := \min\left\{(u')^a_+(x),(u')^a_-(x)\right\} \text{ if } x\in]a,b[,$$
$$(u')^a_\vee(a) = (u')^a_\wedge(a) := (u')^a_+(a), \quad \text{and } (u')^a_\vee(b) = (u')^a_\wedge(b) := (u')^a_-(b)$$

are upper and lower semicontinuous in $[a,b]$, respectively. By (2.8) we have

$$Z^+[(u')^a]\setminus S_{(u')^a} = \{x\in]a,b[: (u')^a_\wedge(x) = +\infty\}\setminus S_{(u')^a},$$
$$Z^-[(u')^a]\setminus S_{(u')^a} = \{x\in]a,b[: (u')^a_\vee(x) = -\infty\}\setminus S_{(u')^a}.$$

Therefore $((u')^c)^+$ is concentrated on the set $\{x\in]a,b[: (u')^a_\wedge(x) = +\infty\}$ and $((u')^c)^-$ is concentrated on the set $\{x\in]a,b[: (u')^a_\vee(x) = -\infty\}$.

Before we proceed we show that the space $X^1_\psi(]a,b[)$ contains functions with nontrivial Cantor part when ψ satisfies suitable decay estimates at infinity.

Proposition 2.3. *Assume that* $\psi\colon \mathbb{R}\to]0,+\infty[$ *is a bounded Borel function satisfying* (2.1), (2.2), *and*

$$\psi(t) \le \frac{c}{t^\alpha} \tag{2.9}$$

for all $t\ge 1$ *and for some* $c>0$, $\alpha>1$. *Then there exists* $u\in X^1_\psi(]a,b[)$ *with* $(u')^c\ne 0$.

Proof. For simplicity we take $]a,b[=]0,1[$.

Step 1: We start by recalling the definition of the generalized Cantor set $\mathbb{D}_\delta$, where $\delta\in]0,\frac{1}{2}[$ (see for instance [6, Chapter 1, Section 2.4]). The construction is entirely similar to the one of the (ternary) Cantor set with the only difference that the middle intervals removed at each step have length $1-2\delta$ times the length of the intervals remaining from the previous step. To be more precise, remove from $[0,1]$ the interval $I_{11} := (\delta, 1-\delta)$. At the second step remove from each of the remaining closed intervals $[0,\delta]$ and $[1-\delta,1]$ the middle intervals, denoted by I_{12} and I_{22}, of length $\delta(1-2\delta)$. Continuing in this fashion at each step n we remove 2^{n-1} middle intervals $I_{1n}, \ldots, I_{2^{n-1}n}$, each of length $\delta^{n-1}(1-2\delta)$. The *generalized Cantor set* $\mathbb{D}_\delta$ is defined as

$$\mathbb{D}_\delta := [0,1]\setminus \bigcup_{n=1}^{\infty}\bigcup_{k=1}^{2^{n-1}} I_{kn}.$$

The set $\mathbb{D}_\delta$ is closed (since its complement is given by a family of open intervals) and

$$\mathcal{L}^1(\mathbb{D}_\delta) = 1-\sum_{n=1}^{\infty}\sum_{k=1}^{2^{n-1}}\mathcal{L}^1(I_{kn}) = 1-\sum_{n=1}^{\infty}\sum_{k=1}^{2^{n-1}}\delta^{n-1}(1-2\delta) = 1-(1-2\delta)\sum_{n=1}^{\infty}(2\delta)^{n-1} = 0.$$

Next we recall the definition of the corresponding Cantor function f_δ. Set

$$g_n := \frac{1}{(2\delta)^n}\Big(1-\sum_{j=1}^{n}\sum_{k=1}^{2^{j-1}}\chi_{I_{kj}}\Big),$$

and define $f_n(x) := \int_0^x g_n(t)\,dt$. It can be shown that $\{f_n\}$ converges uniformly to a continuous nondecreasing function f_δ such that $f_\delta(0)=0$, $f_\delta(1)=1$, and $f'_\delta = (f'_\delta)^c$ is supported on $\mathbb{D}_\delta$.

Step 2: We claim that it is enough to find a constant $\delta\in]0,\frac{1}{2}[$ for which it is possible to construct a continuous integrable function $w_\delta :]0,1[\to [0,+\infty]$ such that $\Psi_1\circ w_\delta \in BV(]0,1[)$ and $w_\delta(x) = +\infty$ if and only if $x\in\mathbb{D}_\delta$. Indeed, setting $u_\delta(x) := \int_0^x w_\delta(t)\,dt +$

$f_\delta(x)$, we have that $u_\delta \in BV(]0,1[)$, u_δ is continuous, $(u'_\delta)^a = w_\delta$ so that $Z^+[(u'_\delta)^a] = Z[(u'_\delta)^a] = \mathbb{D}_\delta$ and $\Psi_1 \circ (u'_\delta)^a \in BV(]0,1[)$. Moreover, $(u'_\delta)^c = (f'_\delta)^c$ is supported on $\mathbb{D}_\delta = Z^+[(u'_\delta)^a]$. Hence u_δ belongs to $X^1_\psi(]a,b[)$.

Step 3: It remains to construct w_δ for a suitable $\delta \in]0, \frac{1}{2}[$. Consider a convex function $\phi :]0,1[\to [0,+\infty)$ such that

$$\lim_{x\to 0^+} \phi(x) = \lim_{x\to 1^-} \phi(x) = +\infty, \quad \phi(\tfrac{1}{2}) = 0\,, \tag{2.10}$$

and

$$\int_0^1 \phi(x)\,dx = 1\,. \tag{2.11}$$

Choose $s > 0$ so large that

$$\alpha > \frac{s+1}{s}\,. \tag{2.12}$$

For $x \in I_{kn}$ (see Step 1) define

$$\phi_{kn}(x) := 2^{sn} + \phi\left(\tfrac{x-a_{kn}}{\delta^{n-1}(1-2\delta)} + \tfrac{1}{2}\right)\,, \tag{2.13}$$

where a_{kn} is the mid point of the interval I_{kn}. Finally set

$$w_\delta := \sum_{n=1}^{\infty}\sum_{k=1}^{2^{n-1}} \phi_{kn}\chi_{I_{kn}} + I_{\mathbb{D}_\delta}\,,$$

where $I_{\mathbb{D}_\delta}$ is the indicator function of the set $\mathbb{D}_\delta$, that is,

$$I_{\mathbb{D}_\delta}(x) := \begin{cases} +\infty & \text{if } x \in \mathbb{D}_\delta\,, \\ 0 & \text{otherwise.} \end{cases}$$

Using the fact that

$$\int_{I_{kn}} \phi_{kn}\,dx = (2^{sn}+1)\,\delta^{n-1}\,(1-2\delta)\,,$$

which follows from (2.11) and a change of variables, we have

$$\int_0^1 w_\delta\,dx = \sum_{n=1}^{\infty}\sum_{k=1}^{2^{n-1}} (2^{sn}+1)\,\delta^{n-1}\,(1-2\delta) < \infty$$

for $\delta < \frac{1}{2^{s+1}}$. To estimate the total variation of $v := \Psi_1 \circ w_\delta$ we consider the approximating sequence

$$v_m(x) := \begin{cases} \Psi_1 \circ \phi_{kn}(x) & \text{if } x \in I_{kn},\ 1 \le k \le 2^{n-1},\ 1 \le n \le m, \\ M & \text{otherwise.} \end{cases}$$

By (2.9), (2.10), (2.13), and the convexity of ϕ it can be seen that

$$\mathrm{Var}\,(v_m; I_{kn}) = 2(M - \Psi_1(2^{sn})) = 2\int_{2^{sn}}^{+\infty} \psi(t)\,dt \le \frac{2c}{\alpha-1}\frac{1}{2^{sn(\alpha-1)}}\,.$$

It follows that

$$\mathrm{Var}\,(v_m;]0,1[) \le \frac{2c}{\alpha-1}\sum_{n=1}^{m}\sum_{k=1}^{2^{n-1}} \frac{1}{2^{sn(\alpha-1)}} \le \frac{2c}{\alpha-1}\sum_{n=1}^{\infty} \frac{1}{2^{sn(\alpha-1)-n+1}}\,.$$

The last series is finite thanks to (2.12). Therefore the v_m's have equibounded total variations and, since $v_m \to v$ in $L^1(]0,1[)$, we conclude that $v \in BV(]0,1[)$. □

Energy bounded sequences are compact in $X^1_\psi(]a,b[)$, as made precise by the following theorem.

Theorem 2.4. *Let* $\{u_k\}$ *be a sequence of functions bounded in* $L^1(]a,b[)$ *such that*

$$C := \sup_k \mathcal{F}_1(u_k) < +\infty\,. \tag{2.14}$$

Then there exist a subsequence (not relabeled) $\{u_k\}$ *and a function* $u \in X^1_\psi(]a,b[)$ *such that*

$$u_k \rightharpoonup u \quad \text{weakly}^* \text{ in } BV(]a,b[)\,, \tag{2.15}$$

$$\Psi_1 \circ u_k' \rightharpoonup \Psi_1 \circ (u')^a \quad \text{weakly}^* \text{ in } BV(]a,b[)\,, \tag{2.16}$$

$$u_k' \to (u')^a \quad \text{pointwise } \mathcal{L}^1\text{-a.e. in }]a,b[\,. \tag{2.17}$$

Proof. By (2.3) and (2.14) we have that each u_k belongs to $W^{2,1}(]a,b[)$ and

$$C_1 := \sup_k \int_a^b [\,|u_k| + |u_k'| + \psi(u_k')|u_k''|\,]\,dx < +\infty\,. \tag{2.18}$$

Let us define

$$v_k := \Psi_1 \circ u_k'\,. \tag{2.19}$$

As Ψ_1 is Lipschitz in $\mathbb{R}$, the functions v_k belong to $W^{1,1}(]a,b[)$ and

$$v_k' = \psi(u_k')u_k'' \qquad \mathcal{L}^1\text{-a.e. on }]a,b[\,. \tag{2.20}$$

It follows from (2.1) and (2.14) that

$$\int_a^b [\,|v_k| + |v_k'|\,]\,dx \le M(b-a) + C\,. \tag{2.21}$$

By (2.18) and (2.21) and the Helly theorem, passing to a subsequence if necessary, we may assume that

$$u_k \rightharpoonup u \quad \text{weakly}^* \text{ in } BV(]a,b[)$$

and

$$v_k\,(x) \to v\,(x) \quad \text{for all } x \in\,]a,b[\tag{2.22}$$

for some $u \in BV(]a,b[)$ and $v\colon]a,b[\, \to [0,M]$ with pointwise bounded variation. Note that (2.22) determines the values of v at every $x \in\,]a,b[$.

Since Ψ_1^{-1} is continuous, we obtain

$$u_k' \to w := \Psi_1^{-1}(v) \quad \text{pointwise in }]a,b[\,. \tag{2.23}$$

Moreover w has left and right limits in $\overline{\mathbb{R}}$ at each point $x \in\,]a,b[$, denoted by $w_-(x)$ and $w_+(x)$ respectively, and

$$w(x) = w_-(x) = w_+(x) \quad \text{except for a countable set of points } x\,. \tag{2.24}$$

We now split the remaining part of the proof into two steps.

Step 1: We prove that

$$w = (u')^a \quad \mathcal{L}^1\text{-a.e. in }]a,b[\,. \tag{2.25}$$

If not, we have $\mathcal{L}^1(\{w \neq (u')^a\}) > 0$. By (2.2) the function Ψ_1^{-1} is locally Lipschitz and so $w = \Psi_1^{-1}(v)$ is finite $\mathcal{L}^1$-a.e. since $v \in L^1(]a,b[)$. Hence there exists $t_0 > 0$ such that

$$\mathcal{L}^1(\{w \neq (u')^a\} \cap \{|w| < t_0\}) > 0$$

and, in particular, we may find an infinite number of disjoint open intervals I such that

$$\mathcal{L}^1(\{w \neq (u')^a\} \cap \{|w| < t_0\} \cap I) > 0\,. \tag{2.26}$$

By a change of variables we obtain

$$\int_I \psi(u_k')|u_k''|\,dx \ge \int_{m_k}^{M_k} \psi(t)\,dt\,, \tag{2.27}$$

where

$$m_k := \inf_I u_k' \quad \text{and} \quad M_k := \sup_I u_k'\,.$$

We claim that at least one of the two sequences $\{m_k\}$ and $\{M_k\}$ is divergent. Indeed, if not, a subsequence of $\{u'_k\}$ would be bounded in $L^\infty(I)$. This implies that $u' \in L^\infty(I)$ and that $u'_k \rightharpoonup u'$ weakly* in $L^\infty(I)$. As $u'_k \to w$ pointwise $\mathcal{L}^1$-a.e. in I, we deduce that $u' = w$ $\mathcal{L}^1$-a.e. in I, which contradicts (2.26). Hence the claim holds. If

$$\lim_{k\to\infty} M_k = +\infty\,, \tag{2.28}$$

then by (2.23) and (2.26)

$$\limsup_{k\to\infty} m_k < t_0\,. \tag{2.29}$$

From (2.27), (2.29), and (2.28) we obtain

$$\liminf_{k\to\infty} \int_I \psi(u'_k)|u''_k|\,dx \geq \int_{t_0}^{+\infty} \psi(t)\,dt > 0\,.$$

Analogously, if $\lim_k m_k = -\infty$ then

$$\liminf_{k\to\infty} \int_I \psi(u'_k)|u''_k|\,dx \geq \int_{-\infty}^{-t_0} \psi(t)\,dt > 0\,.$$

In any case we can choose an arbitrarily large number m of disjoint intervals I satisfying (2.26). Adding the contributions of each interval we obtain

$$\liminf_{k\to\infty} \int_a^b \psi(u'_k)|u''_k|\,dx \geq m \min\Big\{ \int_{t_0}^{+\infty} \psi(t)\,dx, \int_{-\infty}^{-t_0} \psi(t)\,dt \Big\}\,,$$

which contradicts (2.18) for m large enough. This concludes the proof of (2.25) .

Step 2: To prove that $u \in X^1_\psi(]a,b[)$ it remains to show that the positive part $((u')^c)^+$ and the negative part $((u')^c)^-$ of the measure $(u')^c$ are concentrated on $Z^+[(u')^a]$ and $Z^-[(u')^a]$ respectively, that is

$$((u')^c)^\pm\,(]a,b[\setminus Z^\pm[(u')^a]) = 0\,. \tag{2.30}$$

To this purpose we introduce the sets

$$E^+[u'] := \Big\{ x \in\,]a,b[\, : \lim_{\varepsilon\to 0+} \frac{(u')^+\,(]x-\varepsilon, x+\varepsilon[)}{2\varepsilon} = +\infty \Big\}\,, \tag{2.31}$$

$$E^-[u'] := \Big\{ x \in\,]a,b[\, : \lim_{\varepsilon\to 0+} \frac{(u')^-\,(]x-\varepsilon, x+\varepsilon[)}{2\varepsilon} = +\infty \Big\}\,, \tag{2.32}$$

$$E[u'] := \Big\{ x \in\,]a,b[\, : \lim_{\varepsilon\to 0+} \frac{|u'|(]x-\varepsilon, x+\varepsilon[)}{2\varepsilon} = +\infty \Big\}\,.$$

Since $((u')^s)^+ = ((u')^+)^s$ is concentrated on $E^+[u']$ and $((u')^s)^- = ((u')^-)^s$ is concentrated on $E^-[u']$ (see, e.g., [1, Theorem 2.22]), to prove (2.30) it is enough to show that

$$E^+[u'] \setminus Z^+[(u')^a] \text{ and } E^-[u'] \setminus Z^-[(u')^a] \text{ are at most countable.} \tag{2.33}$$

We only show that $E^+[u'] \setminus Z^+[(u')^a]$ is at most countable, since the other property can be proved in a similar way. Assume by contradiction that $E^+[u'] \setminus Z^+[(u')^a]$ is not countable. Since by (2.5) and (2.25)

$$Z^+[(u')^a] \subset \{x \in\,]a,b[\, : \max\{w_-(x), w_+(x)\} = +\infty\}\,,$$

by (2.24) there exists $t_0 > 0$ such that

$$(E^+[u'] \setminus Z^+[(u')^a]) \cap \{w < t_0\} \quad \text{is uncountable.}$$

Fix $t_1 > t_0$ and let $x_1, \ldots, x_m$ be m distinct points in $(E^+[u'] \setminus Z^+[(u')^a]) \cap \{w < t_0\}$. By (2.31) there exists $\varepsilon > 0$ such that the intervals $I_j := \,]x_j - \varepsilon, x_j + \varepsilon[$ are pairwise disjoint and

$$\frac{(u')^+\,(]x_j - \varepsilon, x_j + \varepsilon[)}{2\varepsilon} > t_1 \quad \text{for } i = 1, \ldots, m\,. \tag{2.34}$$

By a change of variables we obtain

$$\int_{I_j} \psi(u_k')|u_k''|\,dx \geq \int_{m_{kj}}^{M_{kj}} \psi(t)\,dt\,, \tag{2.35}$$

where

$$m_{kj} := \inf_{I_j} u_k' \quad \text{and} \quad M_{kj} := \sup_{I_j} u_k'\,.$$

By (2.23) and the fact that $w\,(x_j) < t_0$ we deduce that

$$\limsup_{k\to\infty} m_{kj} < t_0 \tag{2.36}$$

for $j = 1, \dots, m$. On the other hand, (2.15) and (2.34) yield

$$\liminf_{k\to\infty} \frac{1}{2\varepsilon} \int_{x_j-\varepsilon}^{x_j+\varepsilon} (u_k')^+\,dx \geq \frac{(u')^+\,(]x_j-\varepsilon, x_j+\varepsilon[)}{2\varepsilon} > t_1$$

(this can be seen as a particular case of the Reshetnyak lower semicontinuity theorem, with $f = (\cdot)^+$). This implies that $\liminf_{k\to\infty} M_{kj} > t_1$ for $j = 1, \dots, m$. Hence, also by (2.35) and (2.36), we obtain

$$\liminf_{k\to\infty} \sum_{j=1}^{m} \int_{I_j} \psi(u_k')|u_k''|\,dx \geq \sum_{j=1}^{m} \liminf_{k\to\infty} \int_{I_j} \psi(u_k')|u_k''|\,dx \geq m \int_{t_0}^{t_1} \psi(t)\,dt\,,$$

which contradicts (2.18) for m large enough. This shows (2.33) and concludes the proof of the theorem. □

2.3. **Relaxation.** The following theorem, which is the main result of the section, is devoted to the characterization of the relaxation of $\mathcal{F}_1$ with respect to strong convergence in $L^1(]a,b[)$.

Theorem 2.5. *Let $\overline{\mathcal{F}}_1\colon L^1(]a,b[) \to [0,+\infty]$ be defined by:*

$$\overline{\mathcal{F}}_1\,(u) := \inf\left\{ \liminf_{k\to\infty} \mathcal{F}_1\,(u_k) :\ u_k \to u \ \text{in}\ L^1(]a,b[) \right\}$$

for every $u \in L^1(]a,b[)$. Then

$$\overline{\mathcal{F}}_1(u) = \begin{cases} |u'|(]a,b[) + |v'|(]a,b[\setminus S_u) + \sum_{x\in S_u} \Phi(\nu_u, (u')^a_-, (u')^a_+) & \text{if } u \in X^1_\psi(]a,b[)\,, \\ +\infty & \text{otherwise,} \end{cases} \tag{2.37}$$

where $v := \Psi_1 \circ (u')^a$, $\nu_u := \operatorname{sign}(u_+ - u_-)$, and

$$\begin{aligned} \Phi(1,t_1,t_2) &:= \int_{t_1}^{+\infty} \psi(t)\,dt + \int_{t_2}^{+\infty} \psi(t)\,dt\,, \\ \Phi(-1,t_1,t_2) &:= \int_{-\infty}^{t_1} \psi(t)\,dt + \int_{-\infty}^{t_2} \psi(t)\,dt\,. \end{aligned} \tag{2.38}$$

Remark 2.6. For every $x \in S_u$ we have

$$\Phi(\nu_u(x), (u')^a_-(x), (u')^a_+(x)) = |v'|(\{x\}) + \hat{\Phi}(\nu_u(x), (u')^a_-(x), (u')^a_+(x))\,,$$

where

$$\hat{\Phi}(1,t_1,t_2) := \int_{\max\{t_1,t_2\}}^{+\infty} \psi(t)\,dt \qquad \text{and} \qquad \hat{\Phi}(-1,t_1,t_2) := \int_{-\infty}^{\min\{t_1,t_2\}} \psi(t)\,dt\,.$$

In particular, for every Borel set $B \subset\,]a,b[$

$$|v'|(B\setminus S_u) + \sum_{x\in S_u\cap B} \Phi(\nu_u, (u')^a_-, (u')^a_+) = |v'|(B) + \sum_{x\in S_u\cap B} \hat{\Phi}(\nu_u, (u')^a_-, (u')^a_+) \geq |v'|(B)\,.$$

Proof of Theorem 2.5. Let $\mathcal{G}$ be the functional defined by the right hand side of (2.37). We prove that for every $u_k \to u$ in $L^1(]a,b[)$ we have

$$\mathcal{G}(u) \le \liminf_{k\to\infty} \mathcal{F}_1(u_k)\,. \tag{2.39}$$

It is enough to consider sequences $\{u_k\}$ for which the liminf is a limit and has a finite value and $u_k \to u$ pointwise $\mathcal{L}^1$-a.e. in $]a,b[$. Then u_k belongs to $W^{2,1}(]a,b[)$ and (2.14) is satisfied. This implies that

$$|u'|(]a,b[) \le \liminf_{k\to\infty} \int_a^b |u_k'|\,dx\,. \tag{2.40}$$

Moreover, it follows from Theorem 2.4 that $u \in X^1_\psi(]a,b[)$ and that, up to a subsequence, $\{u_k'\}$ converges to $(u')^a$ pointwise $\mathcal{L}^1$-a.e. in $]a,b[$.

Let F be a finite subset of S_u. We want to prove that

$$|v'|(]a,b[\setminus F) + \sum_{x\in F} \Phi(\nu_u, (u')^a_-, (u')^a_+) \le \liminf_{k\to\infty} \int_a^b \psi(u_k')|u_k''|\,dx\,. \tag{2.41}$$

We write F as $\{x_1,\dots,x_m\}$, with $a < x_1 < \dots < x_m < b$. For every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) \in\,]0,\varepsilon[$ such that $a < x_1 - \delta < x_1 + \delta < x_2 - \delta < x_2 + \delta < \dots < x_{m-1} - \delta < x_{m-1} + \delta < x_m - \delta < x_m + \delta < b$ and

$$|u(x_j-\delta) - u_-(x_j)| < \varepsilon\,, \quad |u(x_j+\delta) - u_+(x_j)| < \varepsilon\,, \tag{2.42}$$

$$|(u')^a(x_j-\delta) - (u')^a_-(x_j)| < \varepsilon\,, \quad |(u')^a(x_j+\delta) - (u')^a_+(x_j)| < \varepsilon\,, \tag{2.43}$$

$$u_k(x_j-\delta) \to u(x_j-\delta)\,, \quad u_k(x_j+\delta) \to u(x_j+\delta) \quad \text{as} \quad k\to\infty\,, \tag{2.44}$$

$$u_k'(x_j-\delta) \to (u')^a(x_j-\delta)\,, \quad u_k'(x_j+\delta) \to (u')^a(x_j+\delta) \quad \text{as} \quad k\to\infty\,, \tag{2.45}$$

$$|(u')^a(x_j-\delta)| + |(u')^a(x_j+\delta)| + \varepsilon < \frac{|[u](x_j)| - 4\varepsilon}{2\delta}\,,$$

for $j = 1,\dots,m$.

Since $v_k \to v$ pointwise $\mathcal{L}^1$-a.e. in $]a,b[$ and $v_k' = \psi(u_k')u_k''$ $\mathcal{L}^1$-a.e. in $]a,b[$, we obtain

$$|v'|(]x_j+\delta, x_{j+1}-\delta[) \le \liminf_{k\to\infty} \int_{x_j+\delta}^{x_{j+1}-\delta} \psi(u_k')|u_k''|\,dx$$

for $j = 1,\dots,m-1$. A similar result holds for the intervals $]a, x_1-\delta[$ and $]x_m+\delta, b[$. Let F_δ be the union of the intervals $[x_j-\delta, x_j+\delta]$ for $j=1,\dots,m$. Summing with respect to j, and adding the contributions of the intervals $]a, x_1-\delta[$ and $]x_m+\delta, b[$, we obtain

$$|v'|(]a,b[\setminus F_\delta) \le \liminf_{k\to\infty} \int_{]a,b[\setminus F_\delta} \psi(u_k')|u_k''|\,dx\,. \tag{2.46}$$

We consider now the interval $I^\delta_j := [x_j-\delta, x_j+\delta]$, assuming that $[u](x_j) = u_+(x_j) - u_-(x_j) > 0$. By the mean value theorem there exists $y^\delta_{kj} \in\,]x_j-\delta, x_j+\delta[$ such that

$$u_k'(y^\delta_{kj}) = \frac{u_k(x_j+\delta) - u_k(x_j-\delta)}{2\delta} \ge \frac{[u](x_j) - 4\varepsilon}{2\delta}\,, \tag{2.47}$$

where the last inequality follows from (2.42) and (2.44) for k sufficiently large. By a change of variables we obtain

$$\int_{u_k'(x_j-\delta)}^{\frac{[u](x_j)-4\varepsilon}{2\delta}} \psi(t)\,dt \le \int_{x_j-\delta}^{y^\delta_{kj}} \psi(u_k')|u_k''|\,dx\,,$$

$$\int_{u_k'(x_j+\delta)}^{\frac{[u](x_j)-4\varepsilon}{2\delta}} \psi(t)\,dt \le \int_{y^\delta_{kj}}^{x_j+\delta} \psi(u_k')|u_k''|\,dx\,.$$

Adding these inequalities and taking the limit as $k \to \infty$ we obtain, thanks to (2.45) ,

$$\int_{(u')^a(x_j-\delta)}^{\frac{[u](x_j)-4\varepsilon}{2\delta}} \psi(t)\,dt + \int_{(u')^a(x_j+\delta)}^{\frac{[u](x_j)-4\varepsilon}{2\delta}} \psi(t)\,dt \le \liminf_{k\to\infty} \int_{x_j-\delta}^{x_j-\delta} \psi(u_k')|u_k''|\,dx\,. \tag{2.48}$$

Similarly, if $[u](x_j) < 0$, then we have

$$\int_{\frac{[u](x_j)+4\varepsilon}{2\delta}}^{(u')^a(x_j-\delta)} \psi(t)\,dt + \int_{\frac{[u](x_j)+4\varepsilon}{2\delta}}^{(u')^a(x_j+\delta)} \psi(t)\,dt \le \liminf_{k\to\infty} \int_{x_j-\delta}^{x_j-\delta} \psi(u_k')|u_k''|\,dx\,. \tag{2.49}$$

From (2.46), (2.48), and (2.49) we deduce that

$$\begin{aligned} |v'|(]a,b[\setminus F_\delta) &+ \sum_{[u](x_j)>0} \Big(\int_{(u')^a(x_j-\delta)}^{\frac{[u](x_j)-4\varepsilon}{2\delta}} \psi(t)\,dt + \int_{(u')^a(x_j+\delta)}^{\frac{[u](x_j)-4\varepsilon}{2\delta}} \psi(t)\,dt \Big) \\ &+ \sum_{[u](x_j)<0} \Big(\int_{\frac{[u](x_j)+4\varepsilon}{2\delta}}^{(u')^a(x_j-\delta)} \psi(t)\,dt + \int_{\frac{[u](x_j)+4\varepsilon}{2\delta}}^{(u')^a(x_j+\delta)} \psi(t)\,dt \Big) \\ &\le \liminf_{k\to\infty} \int_a^b \psi(u_k')|u_k''|\,dx\,. \end{aligned}$$

Taking the limit as $\varepsilon \to 0$ (which implies $\delta(\varepsilon) \to 0$) we obtain (2.41) thanks to (2.43) .

Since S_u is at most countable, (2.39) can be obtained from (2.41) by taking the supremum over all finite sets F contained in S_u.

Conversely, let $u \in X^1_\psi(]a,b[)$. We claim that there exists a sequence $\{u_k\}$ in $W^{2,1}(]a,b[)$ such that $u_k \to u$ in $L^1(]a,b[)$ and

$$\mathcal{G}(u) \ge \limsup_{k\to\infty} \mathcal{F}_1(u_k)\,. \tag{2.50}$$

It is clearly enough to consider the case $\mathcal{G}(u) < +\infty$.

We divide the proof into three steps.

Step 1: We prove (2.50) under the additional assumptions that $(u')^a$ is bounded and that $S_u = \{x_1, \dots, x_m\}$, with $x_1 < \dots < x_m$. Note that in this case $Z[(u')^a] = \emptyset$, hence $(u')^c = 0$.

Construct a sequence $\{v_k\}$ in $W^{1,1}(]a,b[)$ such that $v_k \to v = \Psi_1 \circ (u')^a$ pointwise $\mathcal{L}^1$-a.e. in $]a,b[$, $\Psi_1\left(-\left\|(u')^a\right\|_\infty\right) \le v_k \le \Psi_1\left(\left\|(u')^a\right\|_\infty\right)$, and

$$\int_a^b |v_k'(x)|\,dx \to |v'|(]a,b[)\,.$$

Setting $w_k := \Psi_1^{-1}(v_k)$, we have $w_k \in W^{1,1}(]a,b[)$ thanks to (2.2),

$$w_k \to (u')^a \quad \text{pointwise } \mathcal{L}^1\text{-a.e. in }]a,b[\,, \tag{2.51}$$

and $\|w_k\|_\infty \le \left\|(u')^a\right\|_\infty$. Find $\delta_k \to 0^+$ such that

$$w_k(x_j - \delta_k) \to (u')^a_-(x_j)\,, \quad w_k(x_j + \delta_k) \to (u')^a_+(x_j) \quad \text{for } j = 1, \dots, m\,, \tag{2.52}$$

and

$$\begin{gathered} \int_{x_{j-1}+\delta_k}^{x_j-\delta_k} |v_k'|\,dx \to |v'|(]x_{j-1},x_j[) \quad \text{for } j = 2, \dots, m, \\ \int_a^{x_1-\delta_k} |v_k'|\,dx \to |v'|(]a,x_1[)\,, \ \int_{x_m+\delta_k}^b |v_k'|\,dx \to |v'|(]x_m+\delta_k,b[)\,. \end{gathered} \tag{2.53}$$

By (2.51) and by the dominated convergence theorem we have

$$u_+(x_{j-1}) + \int_{x_{j-1}+\delta_k}^{x_j-\delta_k} w_k(s)\,ds \longrightarrow u_+(x_{j-1}) + \int_{x_{j-1}}^{x_j} (u')^a\,ds = u_-(x_j) \tag{2.54}$$

for $j = 2, \dots, m$, with the obvious changes for $j = 1$ and $j = m+1$.

To deal with the jump point x_j, assume first that

$$u_+ (x_j) - u_- (x_j) > 0\,. \tag{2.55}$$

In this case we need to construct functions $f_{kj} \in C^2([x_j - \delta_k, x_j + \delta_k])$ that satisfy the following properties: there exist $y_{kj} \in \,]x_j - \delta_k, x_j + \delta_k[$ such that

$$f_{kj} (x_j - \delta_k) = u_+ (x_{j-1}) + \int_{x_{j-1}+\delta_k}^{x_j-\delta_k} w_k (s)\, ds, \quad f_{kj} (x_j + \delta_k) = u_+ (x_j)\,, \tag{2.56}$$

$$f'_{kj} (x_j - \delta_k) = w_k (x_j - \delta_k)\,, \quad f'_{kj} (x_j + \delta_k) = w_k (x_j + \delta_k)\,, \tag{2.57}$$

$$f''_{kj}(x) > 0 \text{ for } x \in \,]x_j - \delta_k, y_{kj}[\,, \quad f''_{kj}(x) < 0 \text{ for } x \in \,]y_{kj}, x_j + \delta_k[\,, \tag{2.58}$$

$$|f_{kj} (x_j - \delta_k) - \min_{[x_j-\delta_k, y_{kj}]} f_{kj}| \le \tfrac{1}{k}\,, \quad |f_{kj} (x_j + \delta_k) - \max_{[y_{kj}, x_j+\delta_k]} f_{kj}| \le \tfrac{1}{k}\,, \tag{2.59}$$

where we replace x_{j-1} and $x_{j-1} - \delta_k$ by a in the case $j = 1$.

We now discuss briefly the existence of such functions. We observe that the latter conditions in equations (2.56)–(2.58) imply that the graph of f_{kj} in the interval $[y_{kj}, x_j + \delta_k[$ lies below the straight line passing through the point $(x_j + \delta_k, u_+(x_j))$ with slope $w_k(x_j + \delta_k)$, i.e.,

$$f_{kj}(x) \le u_+(x_j) + w_k(x_j + \delta_k)(x - x_j - \delta_k)$$

for $x \in [y_{kj}, x_j + \delta_k[$. It is then easy to see that the inequality

$$u_+ (x_j) - 2w_k (x_j + \delta_k)\, \delta_k - u_+ (x_{j-1}) - \int_{x_{j-1}+\delta_k}^{x_j-\delta_k} w_k (s)\, ds > 0\,, \tag{2.60}$$

allows to fulfill also the former conditions in equations (2.56)–(2.58), as well as (2.59). By (2.52), (2.54), and (2.55), inequality (2.60) is satisfied when δ_k is small enough.

If the left-hand side of (2.55) is negative then we choose f_{kj} so that (2.56) and (2.57) hold, and there exists $y_{kj} \in \,]x_j - \delta_k, x_j + \delta_k[$ such that

$$f''_{kj}(x) < 0 \text{ for } x \in \,]x_j - \delta_k, y_{kj}[\,, \quad f''_{kj}(x) > 0 \text{ for } x \in \,]y_{kj}, x_j + \delta_k[\,,$$
$$|f_{kj} (x_j - \delta_k) - \max_{[x_j-\delta_k, x_j+\delta_k]} f_{kj}| \le \tfrac{1}{k}\,, \quad |f_{kj} (x_j + \delta_k) - \min_{[x_j-\delta_k, x_j+\delta_k]} f_{kj}| \le \tfrac{1}{k}\,.$$

In the same way the construction is possible if δ_k is small enough.

We are now ready to define the approximating sequence

$$u_k (x) := \begin{cases} u_+ (a) + \int_a^x w_k (s)\, ds & \text{if } a \le x < x_1 - \delta_k\,, \\ f_{kj} (x) & \text{if } x_j - \delta_k \le x < x_j + \delta_k,\ j = 1, \dots, m\,, \\ u_+ (x_{j-1}) + \int_{x_{j-1}+\delta_k}^x w_k (s)\, ds & \text{if } x_{j-1} + \delta_k \le x < x_j - \delta_k,\ j = 2, \dots, m\,, \\ u_+ (x_m) + \int_{x_m+\delta_k}^x w_k (s)\, ds & \text{if } x_m + \delta_k \le x < b\,. \end{cases}$$

Let us define $x_0 := a$ and $x_{m+1} := b$. Since $w_k \to (u')^a$ in $L^1 (]a, b[)$, we have

$$u_k (x) \to u_+ (x_{j-1}) + \int_{x_{j-1}}^x (u')^a (s)\, ds = u (x)$$

for every $x \in \,]x_{j-1}, x_j[$ and $j = 1, \dots, m + 1$ and, in turn, $u_k \to u$ in $L^1 (]a, b[)$. As

$$\begin{aligned} \int_{x_{j-1}+\delta_k}^{x_j-\delta_k} |u'_k|\, dx + \int_{x_{j-1}+\delta_k}^{x_j-\delta_k} \psi (u'_k)\, |u''_k|\, dx &= \int_{x_{j-1}+\delta_k}^{x_j-\delta_k} |w_k|\, dx + \int_{x_{j-1}+\delta_k}^{x_j-\delta_k} \psi (w_k)\, |w'_k|\, dx \\ &\le \int_{x_{j-1}}^{x_j} |w_k|\, dx + \int_{x_{j-1}+\delta_k}^{x_j-\delta_k} |v'_k|\, dx\,, \end{aligned}$$

by (2.53) and the fact that $w_k \to (u')^a$ in $L^1(]a,b[)$ we have

$$\begin{aligned}&\limsup_{k\to\infty}\Big(\int_{x_{j-1}+\delta_k}^{x_j-\delta_k}|u_k'|\;dx+\int_{x_{j-1}+\delta_k}^{x_j-\delta_k}\psi(u_k')\,|u_k''|\;dx\Big)\\&\le\int_{x_{j-1}}^{x_j}\big|(u')^a\big|\;dx+|v'|\,(]x_{j-1},x_j[)\,.\end{aligned}\tag{2.61}$$

Similarly,

$$\limsup_{k\to\infty}\Big(\int_a^{x_1-\delta_k}|u_k'|\;dx+\int_a^{x_1-\delta_k}\psi(u_k')\,|u_k''|\;dx\Big)\le\int_a^{x_1}\big|(u')^a\big|\;dx+|v'|\,(]a,x_1[)\ ,\tag{2.62}$$

$$\limsup_{k\to\infty}\Big(\int_{x_m+\delta_k}^{b}|u_k'|\;dx+\int_{x_m+\delta_k}^{b}\psi(u_k')\,|u_k''|\;dx\Big)\le\int_{x_m}^{b}\big|(u')^a\big|\;dx+|v'|\,(]x_m,b[)\ .\tag{2.63}$$

Assume that $[u](x_j)=u_+(x_j)-u_-(x_j)>0$. Then (2.55) holds for k sufficiently large. By (2.57), (2.58), (2.59), and a change of variables we obtain

$$\begin{aligned}&\int_{x_j-\delta_k}^{x_j+\delta_k}|u_k'|\;dx+\int_{x_j-\delta_k}^{x_j+\delta_k}\psi(u_k')\,|u_k''|\;dx=\int_{x_j-\delta_k}^{x_j+\delta_k}|f_{kj}'|\,dx+\int_{x_j-\delta_k}^{x_j+\delta_k}\psi\left(f_{kj}'\right)\left|f_{kj}''\right|\,dx\\&\le f_{kj}(x_j+\delta_k)-f_{kj}(x_j-\delta_k)+\int_{w_k(x_j-\delta_k)}^{f_{kj}'(y_{kj})}\psi(t)\;dt+\int_{w_k(x_j+\delta_k)}^{f_{kj}'(y_{kj})}\psi(t)\;dt+\frac{2}{k}\,.\end{aligned}\tag{2.64}$$

By (2.58) we have

$$f_{kj}'(y_{kj})=\max_{[x_j-\delta_k,x_j+\delta_k]}f_{kj}'\ge\frac{1}{2\delta_k}\left[f_{kj}(x_j+\delta_k)-f_{kj}(x_j-\delta_k)\right]\,.\tag{2.65}$$

By (2.56) and the fact that $w_k\to(u')^a$ in $L^1(]a,b[)$ we obtain

$$f_{kj}(x_j+\delta_k)-f_{kj}(x_j-\delta_k)\to u_+(x_j)-\Big(u_+(x_{j-1})+\int_{x_{j-1}}^{x_j}(u')^a\;ds\Big)=[u](x_j)\ .$$

In turn, using (2.65), we get that $f_{kj}'(y_{kj})\to\infty$. Thus, letting $k\to\infty$ in (2.64) and using (2.52), we infer

$$\begin{aligned}&\limsup_{k\to\infty}\Big(\int_{x_j-\delta_k}^{x_j+\delta_k}|u_k'|\;dx+\int_{x_j-\delta_k}^{x_j+\delta_k}\psi(u_k')\,|u_k''|\;dx\Big)\\&\le[u](x_j)+\int_{(u')_-^a(x_j)}^{+\infty}\psi(t)\;dt+\int_{(u')_+^a(x_j)}^{+\infty}\psi(t)\;dt\,.\end{aligned}\tag{2.66}$$

Similarly, if $[u](x_j)=u_+(x_j)-u_-(x_j)<0$, we find

$$\begin{aligned}&\limsup_{k\to\infty}\Big(\int_{x_j-\delta_k}^{x_j+\delta_k}|u_k'|\;dx+\int_{x_j-\delta_k}^{x_j+\delta_k}\psi(u_k')\,|u_k''|\;dx\Big)\\&\le|[u](x_j)|+\int_{-\infty}^{(u')_-^a(x_j)}\psi(t)\;dt+\int_{-\infty}^{(u')_+^a(x_j)}\psi(t)\;dt\,.\end{aligned}\tag{2.67}$$

Summing over j in (2.61), (2.66), (2.67) and combining with (2.62), (2.63), inequality (2.50) follows.

Step 2: Assume only that $u\in X_\psi^1(]a,b[)$ and that S_u is finite. We claim that there exists a sequence $\{u_k\}$ such that $u_k\to u$ in $L^1(]a,b[)$, each u_k satisfies the hypotheses of Step 1, and

$$\mathcal{G}(u)\ge\limsup_{k\to\infty}\mathcal{G}(u_k)\,.\tag{2.68}$$

Note that if (2.68) holds then, by applying Step 1 to each u_k we may find a sequence $u_{km} \in W^{2,1}(]a,b[)$ converging to u_k in $L^1(]a,b[)$ and satisfying

$$\mathcal{G}(u_k) \geq \limsup_{m\to\infty} \mathcal{F}_1(u_{km}) .$$

By (2.68) we then have

$$\mathcal{G}(u) \geq \limsup_{k\to\infty} \limsup_{m\to\infty} \mathcal{F}_1(u_{km})$$

and a standard diagonalization argument now yields the existence of a sequence $m_k \to \infty$ such that $u_{km_k} \to u$ in $L^1(]a,b[)$ and

$$\mathcal{G}(u) \geq \limsup_{k\to\infty} \mathcal{F}_1(u_{km_k}) .$$

In the construction of the sequence satisfying (2.68) we need to consider the precise representatives $(u')^a_\vee$ and $(u')^a_\wedge$ defined in Remark 2.2. We recall that $(u')^a_\vee$ is upper semicontinuous while $(u')^a_\wedge$ is lower semicontinuous, and so for each $k \in \mathbb{N}$ we may decompose the open sets $\{(u')^a_\wedge > k\}$ and $\{(u')^a_\vee < -k\}$ into the union of two finite sequences of pairwise disjoint open sets U^+_{kj} and U^-_{kj}, that is,

$$\bigcup_j U^+_{kj} = \{(u')^a_\wedge > k\} , \quad \bigcup_j U^-_{kj} = \{(u')^a_\vee < -k\} ,$$

such that

$$\operatorname{diam}(U^+_{kj}) \leq \mathcal{L}^1(\{(u')^a_\wedge > k\}) , \ \operatorname{diam}(U^-_{kj}) \leq \mathcal{L}^1(\{(u')^a_\vee < -k\}) \quad \text{for every } j . \tag{2.69}$$

Note that, setting $v_\vee := \Psi_1 \circ (u')^a_\vee$ and $v_\wedge := \Psi_1 \circ (u')^a_\wedge$, we have

$$|v'|(]c,d[) = \operatorname{Var}(v_\vee;]c,d[) = \operatorname{Var}(v_\wedge;]c,d[) \tag{2.70}$$

for every interval $]c,d[\subset]a,b[$.

For every set $U^\pm_{kj}$ we fix a nonnegative function $g^\pm_{kj} \in C^1_c(U^\pm_{kj})$ such that

$$\int_{U^\pm_{kj}} g^\pm_{kj}(x)\, dx = ((u')^c)^\pm (U^\pm_{kj}) , \tag{2.71}$$

and $(g^\pm_{kj})'$ has only one zero in the interior of the support of $g^\pm_{kj}$. Then we define

$$g^+_k := \sum\nolimits_j g^+_{kj} , \quad g^-_k := \sum\nolimits_j g^-_{kj} , \quad g_k := g^+_k - g^-_k , \quad w_k := T^k_{-k} \circ (u')^a + g_k , \tag{2.72}$$

where for any pair of constants $h < k$ the truncation function T^k_h is defined by

$$T^k_h(t) := \begin{cases} h & \text{for } t \leq h , \\ t & \text{for } h \leq t \leq k , \\ k & \text{for } t \geq k . \end{cases}$$

We claim that

$$w_k \mathcal{L}^1 \rightharpoonup (u')^a \mathcal{L}^1 + (u')^c \ \text{ weakly}^* \text{ in } M_b(]a,b[) . \tag{2.73}$$

Define

$$A_k := \left\{(u')^a_\wedge > k\right\} \cup \left\{(u')^a_\vee < -k\right\} .$$

Since by the Chebychev inequality

$$k\mathcal{L}^1(A_k) \to 0 , \tag{2.74}$$

it suffices to show that

$$\Big(\sum_j g^\pm_{kj}\Big) \mathcal{L}^1 \rightharpoonup \big((u')^c\big)^\pm \ \text{ weakly}^* \text{ in } M_b(]a,b[) . \tag{2.75}$$

Let $\varphi \in C_0(]a,b[)$ and $\varepsilon > 0$. By uniform continuity there exists $\delta = \delta(\varepsilon) > 0$ such that $|\varphi(x) - \varphi(y)| \leq \varepsilon$ for all $x, \ y \in]a,b[$ with $|x - y| \leq \delta$. In view of (2.69) and (2.74), for all

k sufficiently large and for all j we have that $\operatorname{diam}(U_{kj}^{\pm}) \le \delta$. Let us fix $y_{kj}^{\pm} \in U_{kj}^{\pm}$. Then, by (2.71),

$$\begin{aligned}\Big| \int_{U_{kj}^{\pm}} \varphi(x)\, g_{kj}^{\pm}(x)\, dx - \int_{U_{kj}^{\pm}} \varphi(x)\, d\left((u')^c\right)^{\pm}(x) \Big| \\ = \Big| \int_{U_{kj}^{\pm}} \left[\varphi(x) - \varphi(y_{kj}^{\pm})\right] g_{kj}^{\pm}(x)\, dx - \int_{U_{kj}^{\pm}} \left[\varphi(x) - \varphi(y_{kj}^{\pm})\right] d\left((u')^c\right)^{\pm}(x) \Big| \\ \le \varepsilon \Big(\int_{U_{kj}^{\pm}} g_{kj}^{\pm}(x)\, dx + \left((u')^c\right)^{\pm}(U_{kj}^{\pm}) \Big) \le 2\varepsilon \left((u')^c\right)^{\pm}(U_{kj}^{\pm}).\end{aligned}$$

Summing over j and using the fact that the measures $\left(\sum_j g_{kj}^{+}\right)\mathcal{L}^1$ and $\left((u')^c\right)^{+}$ are concentrated on $\left\{(u')^a_\wedge > k\right\}$, while the measures $\left(\sum_j g_{kj}^{-}\right)\mathcal{L}^1$ and $\left((u')^c\right)^{-}$ are concentrated on $\left\{(u')^a_\vee < -k\right\}$ (see Remark 2.2), we obtain (2.75).

Moreover, we claim that

$$\lim_{k\to\infty} \int_a^b |w_k|\, dx = \int_a^b |(u')^a|\, dx + |(u')^c|(]a,b[). \tag{2.76}$$

Indeed, using (2.71), (2.72), and Remark 2.2, we deduce that

$$\begin{aligned}\int_a^b |w_k|\, dx &\le \int_{\{|(u')^a| \le k\}} \left|(u')^a\right| dx + k\mathcal{L}^1(A_k) + \sum_j \int_{U_{kj}^{+}} g_{kj}^{+}\, dx + \sum_j \int_{U_{kj}^{-}} g_{kj}^{-}\, dx \\ &\le \int_a^b |(u')^a|\, dx + k\mathcal{L}^1(A_k) + \sum_j \left((u')^c\right)^{+}(U_{kj}^{+}) + \sum_j \left((u')^c\right)^{-}(U_{kj}^{-}) \\ &\le \int_a^b |(u')^a|\, dx + k\mathcal{L}^1(A_k) + |(u')^c|(]a,b[),\end{aligned}$$

and the limit superior inequality follows from (2.74). The limit inferior inequality follows from (2.73) and the lower semicontinuity of the total variation.

Set

$$u_k(x) := u_+(a) + \int_a^x w_k(s)\, ds + \sum_{x_j < x,\, x_j \in S_u} [u](x_j) \tag{2.77}$$

and $v_k := \Psi_1 \circ (u_k')^a = \Psi_1 \circ w_k$.

We claim that $u_k \to u$ in $L^1(]a,b[)$. For $x \in\,]a,b[$ by (2.73) and (2.76) it follows that

$$\int_a^x w_k\, dy \to \int_a^x (u')^a\, dy + (u')^c(]a,x[),$$

and so u_k converges to u pointwise $\mathcal{L}^1$-a.e. and, in turn, in $L^1(]a,b[)$.

Next we show that

$$\limsup_{k\to\infty} \mathcal{G}(u_k) \le \mathcal{G}(u). \tag{2.78}$$

From (2.76) we get

$$|u_k'|(]a,b[) \to |u'|(]a,b[). \tag{2.79}$$

Moreover, as $v_k = T_{\Psi_1(-k)}^{\Psi_1(k)} \circ v$ $\mathcal{L}^1$-a.e. in the open set $V_k :=]a,b[\setminus \operatorname{supp} g_k$, we have $|v_k'| \le |v'|$ as measures in V_k. In particular, this yields $|v_k'|(]a,b[\setminus (A_k \cup S_u)) \le |v'|(]a,b[\setminus (A_k \cup S_u))$ and hence

$$|v_k'|(]a,b[\setminus (A_k \cup S_u)) \le |v'|(]a,b[\setminus (A_\infty \cup S_u)), \tag{2.80}$$

where

$$A_\infty := \bigcap_k A_k = \left\{(u')^a_\wedge = +\infty\right\} \cup \left\{(u')^a_\vee = -\infty\right\}.$$

Using the properties of g_{kj}^{+} we have

$$|v_k'|(\{(u')^a_\wedge > k\} \setminus S_u) = \sum_j \int_{U_{kj}^+} \psi(k + g_{kj}^+)|(g_{kj}^+)'|\,dx$$

$$= 2\sum_j \int_k^{k+\sup g_{kj}^+} \psi(t)\,dt \le 2\mathcal{H}^0(\{j : ((u')^c)^+(U_{kj}^+) > 0\}) \int_k^\infty \psi\,(t)\;dt\,. \tag{2.81}$$

We claim that

$$2\mathcal{H}^0(\{j : ((u')^c)^+(U_{kj}^+) > 0\}) \int_k^\infty \psi\,(t)\;dt$$
$$\le |v'|\left(\{(u')^a_\wedge > k\} \setminus S_u\right) + 4\int_k^\infty \psi\,(t)\;dt\,.$$

Indeed, if $\left((u')^c\right)^+(U_{kj}^+) > 0$, then there exists a connected component $I_{kj}^+ =]a_{kj}, b_{kj}[$ of $U_{kj}^+ \setminus S_u$ such that $\left((u')^c\right)^+(I_{kj}^+) > 0$. Assume that $I_{kj}^+ \subset\subset]a,b[$. Then by Remark 2.2 we may find $c_{kj} \in I_{kj}^+$ such that $(u')^a_\wedge(c_{kj}) = +\infty$, while $(u')^a_\wedge(a_{kj})$, $(u')^a_\wedge(b_{kj}) \le k$. Hence by (2.70)

$$|v'|(U_{kj}^+ \setminus S_u) \ge |v'|(I_{kj}^+) \ge 2\int_k^\infty \psi\,(t)\;dt\,.$$

Summing over all such intervals and adding the possible contribution of the intervals I_{kj}^+ with at least one endpoint in $\{a,b\}$ we obtain the claim. In turn, by (2.81) we have

$$|v_k'|\left(\{(u')^a_\wedge > k\} \setminus S_u\right) \le |v'|\left(\{(u')^a_\wedge > k\} \setminus S_u\right) + 4\int_k^\infty \psi\,(t)\;dt\,.$$

A similar estimate holds for the set $\{(u')^a_\vee < -k\} \setminus S_u$ thus yielding

$$\limsup_{k\to\infty} |v_k'|\,(A_k \setminus S_u) \le |v'|\,(A_\infty \setminus S_u)\,. \tag{2.82}$$

Combining (2.80) with (2.82) we obtain

$$\limsup_{k\to\infty} |v_k'|\,(]a,b[\setminus S_u) \le |v'|\,(]a,b[\setminus S_u)\,.$$

Next we show that

$$\lim_{k\to\infty} \sum_{x\in S_{u_k}} \Phi\left(\nu_{u_k}, (u_k')^a_-, (u_k')^a_+\right) = \sum_{x\in S_u} \Phi\left(\nu_u, (u')^a_-, (u')^a_+\right)\,. \tag{2.83}$$

Note that $S_{u_k} = S_u$ and $\nu_{u_k}(x) = \nu_u(x)$ for all k by (2.77). Moreover, for every $x \in S_u$ if $(u')^a_+(x) \in \mathbb{R}$ then $\left|(u')^a_+(y)\right| \le k_0$ for all y in a right neighborhood of x and for some integer k_0. Thus, by (2.72) and (2.77) we have that $(u_k')^a(y) = (u')^a(y)$ for $k \ge k_0$ and for $\mathcal{L}^1$-a.e. y in the same right neighborhood. In turn, by (2.7) we infer $(u_k')^a_+(x) = (u')^a_+(x)$ for all $k \ge k_0$. If $(u')^a_+(x) = \infty$, then for all k we have $(u')^a_+ > k$ in a right neighborhood of x by right continuity (see Remark 2.2). By construction this implies that $(u_k')^a = w_k \ge k$ $\mathcal{L}^1$-a.e. in the same right neighborhood. Thus, $(u_k')^a_+(x) \ge k \to (u')^a_+(x)$. Similarly $(u_k')^a_-(x) \to (u')^a_-(x)$, so that

$$\Phi\left(\nu_{u_k}(x), (u_k')^a_-(x), (u_k')^a_+(x)\right) \to \Phi\left(\nu_u(x), (u')^a_-(x), (u')^a_+(x)\right)\,.$$

Hence (2.83) follows. This, together with (2.79) and (2.82), yields (2.78).

Step 3: Let now u be an arbitrary function in $X^1_\psi(]a,b[)$ such that $\mathcal{G}(u) < +\infty$. As in the previous step it suffices to construct $u_k \in X^1_\psi(]a,b[)$ satisfying the hypotheses of Step

2, converging to u in $L^1(]a,b[)$ and such that (2.78) holds. Write $S_u = \{x_j\}$ and for each k define $S_u^k := \{x_j : j \le k\}$ and

$$u_k(x) = u_+(a) + \int_a^x (u')^a \, dt + (u')^c (]a,x[) + \sum_{x_j<x,\, x_j \in S_u^k} [u](x_j) \, .$$

It is clear that $\{u_k\}$ converges to u in $L^1(]a,b[)$ and that $|u_k'|(]a,b[) \to |u'|(]a,b[)$. Moreover, $|v_k'|(]a,b[\setminus S_u) = |v'|(]a,b[\setminus S_u)$ and

$$\begin{aligned}\lim_{k\to\infty} \sum_{x \in S_{u_k}} \Phi\left(\nu_{u_k}, (u_k')^a_-, (u_k')^a_+\right) &= \lim_{k\to\infty} \sum_{x \in S_u^k} \Phi\left(\nu_u, (u')^a_-, (u')^a_+\right) \\ &= \sum_{x \in S_u} \Phi\left(\nu_u, (u')^a_-, (u')^a_+\right) \, .\end{aligned}$$

This concludes the proof of the theorem. □

We end the section with a compactness result for energy bounded sequences in $X^1_\psi(]a,b[)$.

Corollary 2.7. *Let $\{u_k\}$ be a sequence of functions in $X^1_\psi(]a,b[)$ bounded in $L^1(]a,b[)$ and such that*

$$C := \sup_k \overline{\mathcal{F}}_1(u_k) < +\infty \, . \tag{2.84}$$

Then there exist a subsequence (not relabeled) $\{u_k\}$ and a function $u \in X^1_\psi(]a,b[)$ such that

$$u_k \rightharpoonup u \quad \text{weakly}^* \text{ in } BV(]a,b[) \, , \tag{2.85}$$

$$\Psi_1 \circ (u_k')^a \rightharpoonup \Psi_1 \circ (u')^a \quad \text{weakly}^* \text{ in } BV(]a,b[) \, , \tag{2.86}$$

$$(u_k')^a \to (u')^a \quad \text{pointwise } \mathcal{L}^1\text{-a.e. in }]a,b[\, .$$

Proof. It is well known that convergence in measure is metrizable with the following metric

$$d(u_1,u_2) := \int_a^b \frac{|u_1 - u_2|}{1 + |u_1 - u_2|} \, dx \, ,$$

where u_1 and u_2 are (equivalent classes of) measurable functions.

By Theorems 2.4 and 2.5, for every $k \in \mathbb{N}$ we may find $w_k \in W^{2,1}(]a,b[)$ such that

$$\int_a^b |u_k - w_k| \, dx \le \frac{1}{k} \, , \quad d\left((u_k')^a, w_k'\right) \le \frac{1}{k} \, , \tag{2.87}$$

and

$$\mathcal{F}_1(w_k) \le C + 1 \, .$$

By Theorem 2.4 we may find a subsequence (not relabeled) of $\{w_k\}$ and a function $u \in X^1_\psi(]a,b[)$ such that (2.15), (2.16), (2.17) hold (with w_k in place of u_k). It now follows from (2.87) that $u_k \to u$ in $L^1(]a,b[)$ and $(u_k')^a \to (u')^a$ in measure and hence pointwise $\mathcal{L}^1$ a.e. in $]a,b[$, up to a further subsequence. From the bound (2.84), the uniqueness of the limit, and the invertibility of Ψ_1, we deduce (2.85) and (2.86). □

3. The case $p > 1$

In this section we analyze the functional (1.3) in the case $p > 1$.

Let us state precisely the standing assumptions. Throughout this section p denotes any exponent in $]1,+\infty[$, $\psi\colon \mathbb{R} \to]0,+\infty[$ is a bounded Borel function satisfying

$$M := \int_{-\infty}^{+\infty} (\psi(t))^{1/p} \, dt < +\infty \tag{3.1}$$

in addition to (2.2), and $\Psi_p\colon \overline{\mathbb{R}} \to [0,M]$ denotes the antiderivative of $\psi^{1/p}$ defined by

$$\Psi_p(t) := \int_{-\infty}^{t} (\psi(s))^{1/p}\, ds\,. \tag{3.2}$$

The function $\Psi_p^{-1}\colon [0,M] \to \overline{\mathbb{R}}$ stands for the inverse function of Ψ_p.

We now consider the functional $\mathcal{F}_p\colon L^1(]a,b[) \to [0,+\infty]$ defined by

$$\mathcal{F}_p(u) := \begin{cases} \displaystyle\int_a^b |u'|\,dx + \int_a^b \psi(u')|u''|^p\,dx & \text{if } u \in W^{2,p}(]a,b[)\,, \\ +\infty & \text{otherwise.} \end{cases} \tag{3.3}$$

It turns out that piecewise smooth functions with bounded derivative and nonempty discontinuity set cannot be approximated by sequences with equibounded energy. This is a consequence of Remark 3.2(i) and Theorem 3.3 below, and to this end we introduce a suitable space of functions. Recall that $Z^{\pm}[(u')^a]$ are the sets defined in (2.5) and (2.6), while $(u')^s$ denotes the singular part of the gradient measure u'.

Definition 3.1. Let $X^p_\psi(]a,b[)$ be the set of all functions $u \in BV(]a,b[)$ such that $v := \Psi_p \circ (u')^a$ belongs to $W^{1,p}(]a,b[)$ and the positive part $((u')^s)^+$ and the negative part $((u')^s)^-$ of the measure $(u')^s$ are concentrated on $Z^+[(u')^a]$ and $Z^-[(u')^a]$, respectively.

Remark 3.2. (i) It follows immediately from the definition that if $u \in X^p_\psi(]a,b[)$ then $(u')^a = \Psi_p^{-1}(v)$ is continuous on $[a,b]$ with values in $\overline{\mathbb{R}}$. In particular, it turns out that

$$Z^{\pm}[(u')^a] = \{x \in]a,b[\,:\, (u')^a = \pm\infty\}\,.$$

By the assumption on the support of the singular part $(u')^s$, we have $\lim_{x\to x_0}(u')^a(x) = +\infty$ for every jump point x_0 with $u_+(x_0) - u_-(x_0) > 0$ and $\lim_{x\to x_0}(u')^a(x) = -\infty$ for every jump point x_0 with $u_+(x_0) - u_-(x_0) < 0$. This means that if S_u is nonempty then u cannot have bounded derivative outside the discontinuity set. In particular, piecewise constant functions are not included in the class $X^p_\psi(]a,b[)$.

(ii) We observe that the function $(u')^a$ is differentiable $\mathcal{L}^1$ a.e. in $]a,b[$ with

$$v' = \psi^{\frac{1}{p}}\left((u')^a\right)\left((u')^a\right)'\,. \tag{3.4}$$

To see this, we consider the open set

$$A_k := \{x \in]a,b[\,:\, -k < (u')^a < k\}\,.$$

Since by (2.2) the function Ψ_p^{-1} is Lipschitz continuous in the interval $[\Psi_p(-k), \Psi_p(k)]$ and $v \in W^{1,p}(]a,b[)$, by the chain rule we have that $(u')^a = \Psi_p^{-1} \circ v \in W^{1,p}(A_k)$ and, in particular, it is differentiable $\mathcal{L}^1$-a.e. in A_k and (3.4) holds. Since $(u')^a$ is integrable we have that

$$\mathcal{L}^1\Big(]a,b[\setminus \bigcup_k A_k\Big) = 0$$

and the conclusion follows.

(iii) It is easy to check that $X^p_\psi(]a,b[)$ may contain discontinuous functions. An example is given by the following construction: Let $\psi\colon \mathbb{R} \to]0,+\infty[$ be defined by

$$\psi(t) := \begin{cases} 1 & \text{if } |t| \le 1\,, \\ \frac{1}{|t|^\alpha} & \text{if } |t| > 1\,, \end{cases}$$

where α is any number in $]1,+\infty[$, and let $p \in]1, \frac{\alpha+1}{2}[$. Consider now the discontinuous functions $u\colon]-1,1[\to \mathbb{R}$ given by

$$u(x) := \begin{cases} -|x|^\beta & \text{if } x \le 0\,, \\ 1 + x^\beta & \text{if } x > 0\,, \end{cases}$$

with

$$0 < \beta < 1 - \frac{p-1}{\alpha - p}\,.$$

A straightforward computation shows that the function $\Psi_p \circ (u')^a$ belongs to $W^{1,p}(]-1,1[)$, which in turn implies that $u \in X^p_\psi(]-1,1[)$.

(iv) Finally, the same construction of Proposition 2.3 shows that for every admissible ψ satisfying (2.9) the space $X^p_\psi(]a,b[)$ contains a function with nontrivial Cantor part, if p is sufficiently close to 1. We omit the details of this fact which can be easily checked following step by step the proof of Proposition 2.3.

The next theorem is the counterpart of Theorem 2.4 for the case $p > 1$. It establishes that energy bounded sequences are relatively compact in $X^p_\psi(]a,b[)$. The proof is similar to the one of Theorem 2.4, nevertheless since this is the main result of this section we reproduce it in full detail for the reader's convenience.

Theorem 3.3. *Let $\{u_k\}$ be a sequence of functions bounded in $L^1(]a,b[)$ and such that*

$$C := \sup_k \mathcal{F}_p(u_k) < +\infty\,. \tag{3.5}$$

Then there exist a subsequence (not relabeled) $\{u_k\}$ and a function $u \in X^p_\psi(]a,b[)$ such that

$$u_k \rightharpoonup u \quad \text{weakly}^* \text{ in } BV(]a,b[)\,, \tag{3.6}$$

$$\Psi_p \circ u'_k \rightharpoonup \Psi_p \circ (u')^a \quad \text{weakly in } W^{1,p}(]a,b[)\,,$$

$$u'_k \to (u')^a \quad \text{pointwise in }]a,b[\,. \tag{3.7}$$

Proof. By (3.3) and (3.5) we may assume that each u_k belongs to in $W^{2,p}(]a,b[)$ and that

$$C_1 := \sup_k \int_a^b [\,|u_k| + |u'_k| + \psi(u'_k)|u''_k|^p\,]\,dx < +\infty\,. \tag{3.8}$$

Let us define

$$v_k := \Psi_p \circ u'_k\,. \tag{3.9}$$

As Ψ_p is Lipschitz in $\mathbb{R}$, the functions v_k belong to $W^{1,p}(]a,b[)$ and

$$v'_k = (\psi(u'_k))^{1/p} u''_k \qquad \mathcal{L}^1\text{-a.e. on }]a,b[\,. \tag{3.10}$$

It follows from (3.1) and (3.5) that

$$\int_a^b [\,|v_k|^p + |v'_k|^p\,]\,dx \le M^p(b-a) + C_1\,. \tag{3.11}$$

By (3.8) and (3.11), passing to a subsequence (not relabeled), we may assume that

$$u_k \rightharpoonup u \quad \text{weakly}^* \text{ in } BV(]a,b[)$$

and

$$v_k \rightharpoonup v \quad \text{weakly in } W^{1,p}(]a,b[) \tag{3.12}$$

for some functions $u \in BV(]a,b[)$ and $v \in W^{1,p}(]a,b[;[0,M])$.

Since Ψ_p^{-1} is continuous, we obtain

$$u'_k = \Psi_p^{-1} \circ v_k \to w := \Psi_p^{-1} \circ v \quad \text{pointwise in }]a,b[\,. \tag{3.13}$$

Note also that w is continuous with values in $\overline{\mathbb{R}}$.

We now split the remaining part of the proof into two steps.

Step 1: We prove that

$$w = (u')^a \quad \mathcal{L}^1\text{-a.e. on }]a,b[\,. \tag{3.14}$$

If not, arguing as for (2.26), we may find $t_0 > 0$ and an infinite number of disjoint open intervals I such that

$$\mathcal{L}^1(\{w \neq (u')^a\} \cap \{|w| < t_0\} \cap I) > 0\,. \tag{3.15}$$

By Hölder's inequality and a change of variables we obtain

$$\begin{aligned}\int_I \psi(u_k')|u_k''|^p\,dx &\geq \frac{1}{\mathcal{L}^1(I)^{p-1}}\Big(\int_I (\psi(u_k'))^{1/p}|u_k''|\,dx\Big)^p\\ &\geq \frac{1}{(b-a)^{p-1}}\Big(\int_{m_k}^{M_k}(\psi(t))^{1/p}\,dt\Big)^p,\end{aligned}\tag{3.16}$$

where $m_k := \inf_I u_k'$ and $M_k := \sup_I u_k'$.

Reasoning as in the first step of the proof of Theorem 2.4, we can show that at least one of the two sequences $\{m_k\}$ and $\{M_k\}$ is divergent. If $\lim_k M_k = +\infty$ then by (3.13) $\limsup_k m_k < t_0$ and, in turn, from (3.16) we obtain

$$\liminf_{k\to\infty}\int_I \psi(u_k')|u_k''|^p\,dx \geq \frac{1}{(b-a)^{p-1}}\Big(\int_{t_0}^{+\infty}(\psi(t))^{1/p}\,dt\Big)^p > 0\,.$$

Analogously, if $\lim_k m_k = -\infty$ then

$$\liminf_{k\to\infty}\int_I \psi(u_k')|u_k''|^p\,dx \geq \frac{1}{(b-a)^{p-1}}\Big(\int_{-\infty}^{t_0}(\psi(t))^{1/p}\,dt\Big)^p > 0\,.$$

In any case for an arbitrarily large number m of disjoint intervals I satisfying (3.15) , adding the contributions of each interval we obtain

$$\liminf_{k\to\infty}\int_a^b \psi(u_k')|u_k''|^p\,dx \geq \frac{m}{(b-a)^{p-1}}\min\Big\{\Big(\int_{t_0}^{+\infty}(\psi(t))^{1/p}\,dt\Big)^p, \Big(\int_{-\infty}^{t_0}(\psi(t))^{1/p}\,dt\Big)^p\Big\}\,,$$

which contradicts (3.8) for m large enough. This concludes the proof of (3.14) and, in turn, of (3.7) .

Step 2: To prove that $u \in X^p_\psi(]a,b[)$ it remains to show that the positive part $((u')^s)^+$ and the negative part $((u')^s)^-$ of the measure $(u')^s$ are concentrated on $Z^+[(u')^a]$ and $Z^-[(u')^a]$ respectively.

Arguing as in Step 2 of the proof of Theorem 2.4, one can see that it is enough to show

$$E^+[u'] \setminus Z^+[(u')^a] \text{ and } E^-[u'] \setminus Z^-[(u')^a] \text{ are empty,}\tag{3.17}$$

where $E^+[u']$ and $E^-[u']$ are the sets introduced in (2.31) and (2.32) . We only show that $E^+[u'] \setminus Z^+[(u')^a]$ is empty, since the other property can be proved in the same way.

Assume by contradiction that $E^+[u'] \setminus Z^+[(u')^a]$ contains a point x_0. Denote $t_0 := 2|(w(x_0)|$, fix any $t_1 > t_0$, and choose $\varepsilon_0 > 0$ such that

$$\frac{1}{(2\varepsilon_0)^{p-1}}\Big(\int_{t_0}^{t_1}(\psi(t))^{1/p}\,dt\Big)^p > C\,,\tag{3.18}$$

where C is the constant appearing in (3.5). By (2.31) there exists $0 < \varepsilon < \varepsilon_0$ such that

$$\frac{(u')^+\,(]x_0-\varepsilon, x_0+\varepsilon[)}{2\varepsilon} > t_1\,.\tag{3.19}$$

Set $I :=]x_0-\varepsilon, x_0+\varepsilon[$. By Hölder's inequality and a change of variables (see (3.16)) we obtain

$$\int_I \psi(u_k')|u_k''|^p\,dx \geq \frac{1}{(2\varepsilon_0)^{p-1}}\Big(\int_{m_k}^{M_k}(\psi(t))^{1/p}\,dt\Big)^p,\tag{3.20}$$

where $m_k := \inf_I u_k'$ and $M_k := \sup_I u_k'$. By (3.13) and the fact that $w(x_0) < t_0$, we deduce that

$$\limsup_{k\to\infty} m_k < t_0\,.\tag{3.21}$$

On the other hand, reasoning as at the end of the proof of Theorem 2.4, we deduce from (3.6) and (3.19) that

$$\liminf_{k\to\infty}\frac{1}{2\varepsilon}\int_{x_0-\varepsilon}^{x_0+\varepsilon}(u_k')^+\,dx \geq \frac{(u')^+\,(]x_0-\varepsilon, x_0+\varepsilon[)}{2\varepsilon} > t_1\,,$$

which implies that

$$\liminf_{k\to\infty} M_k > t_1 \,. \tag{3.22}$$

From (3.18), (3.20), (3.21), and (3.22) we obtain

$$\liminf_{k\to\infty} \int_I \psi(u_k')|u_k''|^p\,dx \geq \frac{1}{(2\varepsilon_0)^{p-1}}\Big(\int_{t_0}^{t_1} (\psi(t))^{1/p}\,dt\Big)^p > C\,,$$

which contradicts (3.8) . This shows (3.17) and concludes the proof of the theorem. □

We next identify the relaxation of $\mathcal{F}_p$ with respect to strong convergence in $L^1(]a,b[)$.

Theorem 3.4. *Let $\overline{\mathcal{F}}_p\colon L^1(]a,b[) \to [0,+\infty]$ be defined by*

$$\overline{\mathcal{F}}_p\,(u) := \inf\Big\{\liminf_{k\to\infty} \mathcal{F}_p\,(u_k) :\ u_k \to u \ \text{in}\ L^1(]a,b[)\Big\} \tag{3.23}$$

for every $u \in L^1(]a,b[)$. Then

$$\overline{\mathcal{F}}_p\,(u) = \begin{cases} |u'|(]a,b[) + \displaystyle\int_a^b |v'|^p\,dx & \text{if } u \in X^p_\psi(]a,b[)\,, \\ +\infty & \text{otherwise,} \end{cases} \tag{3.24}$$

where $v := \Psi_p \circ (u')^a$.

Proof. We sketch the proof focusing only on the main changes with respect to the proof of Theorem 2.5. Let $\mathcal{G}_p$ be the functional defined by the right hand side of (3.24).

We start by showing that

$$\mathcal{G}_p(u) \leq \liminf_{k\to\infty} \mathcal{F}_p(u_k)\,. \tag{3.25}$$

whenever $u_k \to u$ in $L^1(]a,b[)$. It is enough to consider sequences $\{u_k\}$ for which the liminf is a limit and has a finite value. Then u_k belongs to $W^{2,1}(]a,b[)$ and (3.5) is satisfied. Setting $v_k := \Psi_p \circ u_k'$, by Theorem 3.3 we have $v_k \rightharpoonup v$ weakly in $W^{1,p}(]a,b[)$. Using the fact that $|v_k'|^p = \psi(u_k')|u_k''|^p$, we deduce that

$$\int_a^b |v'|^p\,dx \leq \liminf_{k\to\infty} \int_a^b \psi(u_k')|u_k''|^p\,dx\,. \tag{3.26}$$

Inequality (3.25) follows now from (3.26) and the lower semicontinuity of the total variation.

We split the proof of the limsup inequality into several steps.

Step 1: Let $u \in X^p_\psi(]a,b[)$ be such that $(u')^s = 0$. We claim that there exists a sequence $\{u_k\}$ in $W^{2,p}(]a,b[)$ such that $u_k \to u$ in $L^1(]a,b[)$ and

$$\limsup_{k\to\infty} \mathcal{F}_p(u_k) \leq \mathcal{G}_p(u)\,. \tag{3.27}$$

Define $w_k := ((u')^a \vee -k) \wedge k$. Using the fact that $(u')^a \in W^{1,p}(A_{2k})$, where

$$A_{2k} := \{x \in\,]a,b[\,:\ -2k < (u')^a < 2k\}$$

as observed in Remark 3.2-(ii), one sees that $w_k \in W^{1,p}(]a,b[)$. Define

$$u_k(x) := u_+(a) + \int_a^x w_k(y)\,dy\,.$$

It is easy to see that $u_k \to u$ in $L^1(]a,b[)$ and (3.27) holds.

Step 2: Assume that $u \in X^p_\psi(]a,b[)$, $(u')^c = 0$, and S_u is finite. We claim that there exists a sequence $\{u_k\}$ of functions in $X^p_\psi(]a,b[)$, with $(u_k')^s = 0$, such that $u_k \to u$ in $L^1(]a,b[)$ and

$$\limsup_{k\to\infty} \mathcal{G}_p(u_k) \leq \mathcal{G}_p(u)\,. \tag{3.28}$$

Since the construction is local, it is enough to consider the case $S_u = \{x_0\}$ for some $x_0 \in]a,b[$ with $[u](x_0) > 0$. By the properties of $X^p_\psi(]a,b[)$ we can find two sequences $x_k \nearrow x_0$ and $y_k \searrow x_0$ such that

$$u(x_k) \to u_-(x_0)\,, \quad u(y_k) \to u_+(x_0)\,, \quad \text{and} \quad (u')^a(x_k) = (u')^a(y_k) \to (u')^a(x_0) = +\infty\,.$$

Consider the affine functions $h_k(x) := u(x_k) + (u')^a(x_k)(x - x_k)$. For every k sufficiently large there exists $z_k \in]x_k, b[$ such that $h_k(z_k) = u_+(x_0)$. Since $(u')^a(x_k) \to +\infty$ and $x_k \to x_0$, we have that $z_k \to x_0$ as $k \to \infty$. Define

$$u_k(x) := \begin{cases} u(x) & \text{if } a < x \le x_k\,, \\ h_k(x) & \text{if } x_k < x \le z_k\,, \\ u(x + y_k - z_k) + u_+(x_0) - u(y_k) & \text{if } z_k < x < b\,. \end{cases}$$

Using the fact that $(u')^a(x_k) = (u')^a(y_k)$, it is easy to check that $u_k \in X^p_\psi(]a,b[)$, with $(u'_k)^s = 0$, $u_k \to u$ in $L^1(]a,b[)$, and (3.28) holds.

Step 3: Assume that $u \in X^p_\psi(]a,b[)$ and $(u')^c = 0$. We claim that there exists a sequence of functions u_k in $X^p_\psi(]a,b[)$, with $(u'_k)^c = 0$ and S_{u_k} finite, such that $u_k \to u$ in $L^1(]a,b[)$ and (3.28) holds.

To see this, it is enough to consider the same approximation constructed in Step 3 of the proof of Theorem 2.5.

Step 4: Assume that $u \in X^p_\psi(]a,b[)$. We claim that there exists a sequence of functions u_k in $X^p_\psi(]a,b[)$, with $(u'_k)^c = 0$, such that $u_k \to u$ in $L^1(]a,b[)$ and (3.28) holds.

Since $(u')^a$ is continuous from $]a,b[$ into $\overline{\mathbb{R}}$ and integrable (see Remark 3.2), we have that $K := \{x \in]a,b[\, : |(u')^a| = +\infty\}$ is relatively closed in $]a,b[$ with zero $\mathcal{L}^1$ measure. Hence, we may find a sequence of open sets $A_k \subset]a,b[$ such that $A_k \searrow K$. Let $\{I^k_j\}_j$ be the collection of all connected components of A_k intersecting K. Let $c^k_j := (u')^s(I^k_j) > 0$. By the properties of $X^p_\psi(]a,b[)$ for every j we may choose $x^k_j \in I^k_j \cap K$ such that $(u')^a(x^k_j) = +\infty$ if $c^k_j > 0$ and $(u')^a(x^k_j) = -\infty$ if $c^k_j < 0$. Define

$$u_k(x) := u_+(a) + \int_a^x (u')^a(y)\,dy + \sum_{j:\, x^k_j \le x} c^k_j\,.$$

Using the definition of $X^p_\psi(]a,b[)$ one can check that

$$\sum_j c^k_j \delta_{x^k_j} \rightharpoonup (u')^s \quad \text{weakly}^* \text{ in } M_b(]a,b[)$$

and $|\sum_j c^k_j \delta_{x^k_j}|(]a,b[) \to |(u')^s|(]a,b[)$ as $k \to \infty$. Using this fact it is easy to see that the sequence $\{u_k\}$ meets all the requirements.

By combining Steps 1-4 with a diagonal argument one can finally prove that (3.27) holds for every u in $X^p_\psi(]a,b[)$. □

Corollary 3.5. *Let $\{u_k\}$ be a sequence of functions in $X^p_\psi(]a,b[)$ bounded in $L^1(]a,b[)$ and such that*

$$C := \sup_k \overline{\mathcal{F}}_p(u_k) < +\infty\,. \tag{3.29}$$

Then there exists a subsequence (not relabeled) $\{u_k\}$ and a function $u \in X^p_\psi(]a,b[)$ such that

$$u_k \rightharpoonup u \quad \textit{weakly}^* \textit{ in } BV(]a,b[)\,, \tag{3.30}$$

$$\Psi_p \circ (u'_k)^a \rightharpoonup \Psi_p \circ (u')^a \quad \textit{weakly in } W^{1,p}(]a,b[)\,, \tag{3.31}$$

$$(u'_k)^a \to (u')^a \quad \textit{pointwise in }]a,b[\,. \tag{3.32}$$

Proof. With an argument entirely similar to the one used in the proof of Corollary 2.7 we can extract a subsequence $\{u_k\}$ which satisfies (3.30) and (3.31). In turn (3.31) and the continuity of Ψ_p^{-1} in $\overline{\mathbb{R}}$ imply (3.32). □

4. The staircase effect

The purpose of this section is to show analytically that the presence of the higher order term in the functional $\overline{\mathcal{F}}$ prevents the occurrence of the so-called *staircase effect* as opposed to what happens in image reconstructions based on the total variation functional.

4.1. **The Rudin-Osher-Fatemi model.** We start by showing that staircase-like structures do appear in solutions to the Rudin-Osher-Fatemi problem; i.e., in minimizers for the functional $\mathrm{ROF}_{\lambda,g} : BV(]a,b[) \to \mathbb{R}$ defined by

$$\mathrm{ROF}_{\lambda,g}(w) := |w'|(]a,b[) + \lambda \int_a^b (w-g)^2 \, dx \,,$$

where $\lambda > 0$ is the *fidelity parameter* and $g \in L^2(]a,b[)$ is the given "signal" to be processed. This fact is well known and numerically observed in many situations. We provide here a simple analytical example. A different example can be found in [10]. It will be constructed by means of the following proposition which deals with minimizers of $\mathrm{ROF}_{\lambda,g}$ when g is a monotone function.

Proposition 4.1. *Let $g : [a,b] \to [0,1]$ be a nondecreasing function such that $g_+(a) = 0$ and $g_-(b) = 1$. Let g^{-1} denote the left-continuous generalized inverse of g, defined by*

$$g^{-1}(c) := \inf\{x \in [a,b] : \, g(x) \geq c\} \tag{4.1}$$

for every $c \in [0,1]$ and assume that there exist $0 < c_1 < c_2 < 1$ such that

$$2\lambda \int_a^{g^{-1}(c_1)} (c_1 - g(x)) \, dx = 1 \qquad \text{and} \qquad 2\lambda \int_{g^{-1}(c_2)}^b (g(x) - c_2) \, dx = 1 \,. \tag{4.2}$$

Then the function u, defined by

$$u(x) := \begin{cases} c_1 & \text{if } a \leq x \leq g^{-1}(c_1) \,, \\ g(x) & \text{if } g^{-1}(c_1) < x \leq g^{-1}(c_2) \,, \\ c_2 & \text{if } g^{-1}(c_2) < x \leq b \,, \end{cases}$$

is the unique minimizer of $\mathrm{ROF}_{\lambda,g}$ in $BV(]a,b[)$.

Remark 4.2. Since

$$\int_a^{g^{-1}(c)} (c - g(x)) \, dx = \int_0^c g^{-1}(y) \, dy \,, \qquad \int_{g^{-1}(c)}^b (g(x) - c) \, dx = \int_c^1 g^{-1}(y) \, dy$$

for all $c \in [0,1]$, the continuity of the integral implies that condition (4.2) is satisfied for every λ sufficiently large.

Proof of Proposition 4.1. We split the proof into two steps.
Step 1. We assume first that u is absolutely continuous. In order to prove the minimality of u, by density it suffices to show that $\mathrm{ROF}_{\lambda,g}(u+\varphi) \geq \mathrm{ROF}_{\lambda,g}(u)$ for every $\varphi \in C^1([a,b])$, which, in turn, due to the convexity of $\mathrm{ROF}_{\lambda,g}$, is equivalent to proving that

$$\frac{d^+}{d\varepsilon} \mathrm{ROF}_{\lambda,g}(u + \varepsilon\varphi)\bigg|_{\varepsilon=0} \geq 0 \quad \text{for every } \varphi \in C^1([a,b]) \,, \tag{4.3}$$

where $\frac{d^+}{d\varepsilon}$ denotes the right derivative. By a straightforward computation we have

$$\frac{d^+}{d\varepsilon} \mathrm{ROF}_{\lambda,g}(u + \varepsilon\varphi)\bigg|_{\varepsilon=0} = \int_{\{u'=0\}} |\varphi'| \, dx + \int_{\{u'>0\}} \varphi' \, dx + 2\lambda \int_a^b (u-g)\varphi \, dx \,. \tag{4.4}$$

Consider now the function $\theta : [a,b] \to [0,1]$ defined by $\theta(x) := 2\lambda \int_a^x (u-g)\,dt$. Using (4.2) and the definition of u one can check that $\theta(a) = \theta(b) = 0$, $0 \le \theta \le 1$, and $\theta \equiv 1$ in $[g^{-1}(c_1), g^{-1}(c_2)]$. In particular, $\{u' > 0\} \subset [g^{-1}(c_1), g^{-1}(c_2)] \subset \{\theta = 1\}$ so that by (4.4)

$$\frac{d^+}{d\varepsilon}\mathrm{ROF}_{\lambda,g}(u+\varepsilon\varphi)\bigg|_{\varepsilon=0} \ge \int_a^b \varphi'\theta\,dx + 2\lambda\int_a^b (u-g)\varphi\,dx = 0\,,$$

where the last equality is obtained by integrating by parts and by using the fact that $\theta' = 2\lambda(u-g)$ and $\theta(a) = \theta(b) = 0$. This shows (4.3) and concludes the proof of Step 1.

Step 2. In the general case, we construct a sequence $\{g_k\} \subset AC([g^{-1}(c_1), g^{-1}(c_2)])$ of nondecreasing functions such that $g_k(g^{-1}(c_1)) = c_1$, $g_k(g^{-1}(c_2)) = c_2$, and $g_k \to g$ in $L^2([g^{-1}(c_1), g^{-1}(c_2)])$. Let $\tilde g_k$ be the function that coincides with g_k in $[g^{-1}(c_1), g^{-1}(c_2)]$ and with g elsewhere in $[a,b]$ and, analogously, set u_k to be equal to g_k in $[g^{-1}(c_1), g^{-1}(c_2)]$ and to u elsewhere. For any $v \in BV(]a,b[)$, by applying the previous step we obtain

$$\mathrm{ROF}_{\lambda,\tilde g_k}(v) \ge \mathrm{ROF}_{\lambda,\tilde g_k}(u_k) = \mathrm{ROF}_{\lambda,g}(u)\,.$$

The minimality of u follows by letting $k \to \infty$. Finally, uniqueness is a consequence of the strict convexity of $\mathrm{ROF}_{\lambda,g}$. □

As a corollary of the previous result we can prove analytically the occurrence of the staircase effect in a very simple case. Let $g(x) := x$, $x \in [0,1]$, be the original 1D image to which we add the "noise"

$$h_n(x) := \frac{i}{n} - x \qquad \text{if } \frac{i-1}{n} \le x < \frac{i}{n}\,,\ i = 1,\dots,n\,,$$

where $n \in \mathbb{N}$, so that the resulting degraded 1D image is given by the staircase function

$$g_n(x) := \frac{i}{n} \qquad \text{if } \frac{i-1}{n} \le x < \frac{i}{n}\,,\ i = 1,\dots,n. \tag{4.5}$$

Note that, even though $h_n \to 0$ uniformly, the reconstructed image u_n preserves the staircase structure of g_n. Indeed, we show that there exists a non degenerate interval $I \subset [0,1]$ such that each u_n coincides with the degraded 1D image g_n in I for all $n \in \mathbb{N}$. More precisely we have the following theorem.

Theorem 4.3 (Staircase effect)**.** *Let $\lambda > 4$, let g_n be as in (4.5), and let u_n be the unique minimizer of $\mathrm{ROF}_{\lambda,g_n}$ in $BV(]0,1[)$. Then for all n sufficiently large there exist $0 < a_n < b_n < 1$, with*

$$a_n \to \frac{1}{\sqrt{\lambda}}, \quad b_n \to 1 - \frac{1}{\sqrt{\lambda}}$$

as $n \to \infty$, such that $u_n = g_n$ on $[a_n, b_n]$ and u_n is constant on each interval $[0,a_n)$ and $(b_n, 1]$.

Proof. Let g_n^{-1} denote the generalized inverse function of g_n defined by (4.1) with g replaced by g_n. As both $\{g_n\}$ and $\{g_n^{-1}\}$ converge uniformly to $g(x) = x$ and since $\lambda > 4$, one can check that for n large enough there exist $0 < c_1^{(n)} < c_2^{(n)} < 1$ satisfying

$$2\lambda\int_0^{g_n^{-1}(c_1^{(n)})} (c_1^{(n)} - g_n)\,dx = 1 \qquad \text{and} \qquad 2\lambda\int_{g_n^{-1}(c_2^{(n)})}^1 (g_n - c_2^{(n)})\,dx = 1$$

with $c_1^{(n)} \to c_1$ and $c_2^{(n)} \to c_2$ as $n \to \infty$, where c_1 and c_2 are defined by

$$2\lambda\int_0^{c_1} (c_1 - x)\,dx = 1 \qquad \text{and} \qquad 2\lambda\int_{c_2}^1 (x - c_2)\,dx = 1\,. \tag{4.6}$$

By Proposition 4.1 the unique minimizer u_n of $\mathrm{ROF}_{\lambda,g_n}$ in $BV(]0,1[)$ takes the form

$$u_n(x) = \begin{cases} c_1^{(n)} & \text{if } 0 \le x \le g_n^{-1}(c_1^{(n)})\,, \\ g_n(x) & \text{if } g_n^{-1}(c_1^{(n)}) < x \le g_n^{-1}(c_2^{(n)})\,, \\ c_2^{(n)} & \text{if } g_n^{-1}(c_2^{(n)}) < x \le 1\,. \end{cases}$$

The conclusion follows by observing that $a_n := g_n^{-1}(c_1^{(n)}) \to c_1\,,\ b_n := g_n^{-1}(c_2^{(n)}) \to c_2$ and that $c_1 = \frac{1}{\sqrt{\lambda}}$ and $c_2 = 1 - \frac{1}{\sqrt{\lambda}}$, thanks to (4.6) . □

4.2. **Absence of the staircase effect: The case** $p = 1$**.** Next we show that the presence of the higher order term in the functional $\overline{\mathcal{F}}_1$ prevents the occurrence of the staircase effect. We begin with the case $p = 1$. We consider the minimization problem

$$\min\left\{\overline{\mathcal{F}}_1\,(u) + \lambda \int_a^b (u-g)^2\,dx :\ u \in X^1_\psi(]a,b[)\right\}, \tag{4.7}$$

where $\overline{\mathcal{F}}_1$ is the relaxed functional given in (2.37). To prove the absence of the staircase effect we need the following auxiliary result that is of independent interest.

Proposition 4.4. *Assume that* $\psi\colon \mathbb{R} \to]0,+\infty[$ *is a bounded Borel function satisfying* (2.1) *and* (2.2)*. Let* $g : [a,b] \to \mathbb{R}$ *be Lipschitz continuous and let* $u \in X^1_\psi(]a,b[)$ *be a solution of the minimization problem* (4.7)*. Then* u *is Lipschitz continuous and* $u' \in BV(]a,b[)$*.*

Proof. The plan of the proof is the following. We will show that the discontinuity set S_u is empty and that the left and right limits $(u')^a_-$ and $(u')^a_+$, defined in (2.7), are finite everywhere on $]a,b]$ and on $[a,b[$, respectively. Note that this will imply that the sets $Z^{\pm}[(u')^a]$ (see (2.5) and (2.6)) are empty and, in turn, that $u \in W^{1,1}(]a,b[)$ by the properties of the space $X^1_\psi(]a,b[)$. Moreover, recalling that the functions $(u')^a_\vee$ and $(u')^a_\wedge$ defined in Remark 2.2 are upper and lower semicontinuous on $[a,b]$, it will also follow that both $(u')^a_-$ and $(u')^a_+$ are bounded, yielding the Lipschitz continuity of u. In turn, the fact that $u' \in BV(]a,b[)$ is a consequence of the local Lipschitz continuity of Ψ_1^{-1}.

Step 1: We start by showing that S_u is empty. We argue by contradiction, assuming that S_u contains a point x_0. Without loss of generality we may suppose that $\nu_u(x_0) = 1$; i.e., $u_+(x_0) > u_-(x_0)$. We also assume that $\frac{1}{2}(u_+(x_0) + u_-(x_0)) \ge g(x_0)$. In the following it is convenient to think of u as coinciding everywhere with its lower semicontinuous representative $u_\wedge := \min\{u_-, u_+\}$.

Find $\varepsilon > 0$ so small that

$$\sum_{\substack{x\in S_u \\ x\in]x_0,x_0+\varepsilon[}} |[u](x)| < \frac{[u](x_0)}{4} \tag{4.8}$$

and let $C > 0$ satisfy

$$C > 2\|g'\|_\infty \quad \text{and} \quad \frac{1}{2}(u_+(x_0) + u_-(x_0)) + C\varepsilon > u_-(x_0+\varepsilon)\,. \tag{4.9}$$

For $t \in [0,1]$ consider the affine function

$$h^t(x) := \tfrac{(1-t)}{2}(u_+(x_0) + u_-(x_0)) + t\big(\tfrac{1}{4}u_-(x_0) + \tfrac{3}{4}u_+(x_0)\big) + C(x - x_0)$$

and note that by (4.9) there exists $x^t \in]x_0, x_0+\varepsilon[$ such that

$$(x^t, h^t(x^t)) \in \Gamma_u \quad \text{and} \quad g < h^t < u \text{ in }]x_0, x^t[\,, \tag{4.10}$$

where Γ_u stands for the extended graph of u defined by

$$\Gamma_u := \{(x,t) \in]a,b[\times\mathbb{R} : \min\{u_-(x), u_+(x)\} \le t \le \max\{u_-(x), u_+(x)\}\}\,.$$

Let u^t be the function defined by

$$u^t(x) := \begin{cases} h^t(x) & \text{if } x \in \,]x_0, x^t[, \\ u(x) & \text{otherwise,} \end{cases} \tag{4.11}$$

and note that

$$\lambda\Big(\int_a^b |u-g|^2\,dx - \int_a^b |u^t-g|^2\,dx \Big) \geq \lambda\Big(\int_a^b |u-g|^2\,dx - \int_a^b |u^1-g|^2\,dx \Big) =: \eta > 0 \tag{4.12}$$

for every $t \in [0,1]$. Now it is convenient to approximate u with functions having only finitely many jump points. Hence the following approximation procedure is needed only when S_u is infinite. In this case write $S_u = \{x_0, x_1, \dots, x_j, \dots\}$, for each k define $S_u^k := \{x_j : 0 \leq j \leq k\}$, and for $x \in \,]a,b[$ set

$$u_k(x) = u_+(a) + \int_a^x (u')^a\,dt + (u')^c(]a,x[) + \sum_{x_j<x,\, x_j \in S_u^k} [u](x_j)\,.$$

Note that, since $u_k \to u$ in $L^\infty(]a,b[)$, for k large enough it follows from (4.9) and (4.10) that for every $t \in [0,1]$ there exists $x_k^t \in \,]x_0, x_0+\varepsilon[$ such that

$$(x_k^t, h^t(x_k^t)) \in \Gamma_{u_k} \quad \text{and} \quad g < h^t < u_k \text{ in }]x_0, x_k^t[\,,$$

where Γ_{u_k} denotes the extended graph of u_k. For all such k we consider the comparison function u_k^t defined as in (4.11), with u and x^t replaced by u_k and x_k^t, respectively. Using the uniform convergence of $\{u_k\}$ to u and (4.10), we have that $x^t \leq \liminf_k x_k^t$, which yields $u^t \geq \limsup_k u_k^t$ $\mathcal{L}^1$-a.e. on $]a,b[$. Moreover $u_k \to u$ in $\overline{\mathcal{F}}_1$ energy. Hence, also by (4.12), we may find k so large that for $t \in [0,1]$

$$\lambda\Big(\int_a^b |u_k-g|^2\,dx - \int_a^b |u_k^t-g|^2\,dx \Big) \geq \lambda\Big(\int_a^b |u_k-g|^2\,dx - \int_a^b |u_k^1-g|^2\,dx \Big) \geq \frac{\eta}{2}\,, \tag{4.13}$$

$$\overline{\mathcal{F}}_1(u_k) + \lambda \int_a^b |u_k - g|^2\,dx \leq \overline{\mathcal{F}}_1(u) + \lambda \int_a^b |u-g|^2\,dx + \frac{\eta}{4}\,. \tag{4.14}$$

Let us fix k satisfying (4.13) and (4.14). We claim that there exists $\bar{t} \in [0,1]$ such that $x_k^{\bar{t}}$ is a continuity point for u_k. Indeed, if not, then for every $t \in [0,1]$ there exists a jump point x_j, with $1 \leq j \leq k$, such that $x_k^t = x_j$ and the point $(x_k^t, h^t(x_k^t))$ belongs to the corresponding vertical segment of the extended graph of u_k. Setting $I_j := \{t \in [0,1] : x_k^t = x_j\}$ and $\sigma_j := \{(x_j, h^t(x_j)) : t \in I_j\}$, it is clear that $[0,1] = \cup_{j=1}^k I_j$ and $\mathcal{H}^1(\sigma_j) = \mathcal{H}^1(\{(x_0, h^t(x_0)) : t \in I_j\})$. Thus,

$$\sum_{\substack{x \in S_u \\ x \in]x_0, x_0+\varepsilon[}} |[u](x)| \geq \sum_{j=1}^k \mathcal{H}^1(\sigma_j) = \mathcal{H}^1\big(\{(x_0, h^t(x_0)) : t \in [0,1]\}\big) = \frac{[u](x_0)}{4}\,,$$

in contradiction with (4.8).

Since from now on $\bar{t}$ and k are fixed, to simplify the notation we set $\hat{x} := x_k^{\bar{t}}$, $\hat{u} := u_k^{\bar{t}}$, $\hat{h} := h^{\bar{t}}$, and $\hat{v} := \Psi_1 \circ (\hat{u}')^a$. By construction (see (4.11)) we have

$$|\hat{u}'|(]a,b[) \leq |u_k'|(]a,b[)\,. \tag{4.15}$$

Next we claim that

$$(u')^a_-(\hat{x}) \leq \hat{h}'(\hat{x}) = C\,. \tag{4.16}$$

If $(u')^a_-(\hat{x}) \leq 0$ there is nothing to prove. If $(u')^a_-(\hat{x}) > 0$, then by left continuity $(u')^a_-(y) > 0$ for y sufficiently close to $\hat{x}$, which, in turn, implies $(u')^c(]y, \hat{x}[) \geq 0$ by the properties of $X^1_\psi(]a,b[)$. Since S_{u_k} is finite and $\hat{x}$ is a continuity point, for y in a left neighborhood of $\hat{x}$ we can write

$$\hat{h}(\hat{x}) = u_k(\hat{x}) = u_k(y) + \int_y^{\hat{x}} (u')^a(s)\,ds + (u')^c(]y,\hat{x}[) > \hat{h}(y) + \int_y^{\hat{x}} (u')^a(s)\,ds\,,$$

where we have used the fact that $u_k(\hat{x}) = \hat{h}(\hat{x})$ and $\hat{h} < u_k$ in a left neighborhood of $\hat{x}$. Claim (4.16) follows.

Now, recalling that $\Phi(1,t_1,t_2) = 2\Psi_1(+\infty) - \Psi_1(t_1) - \Psi_1(t_2)$ for every $t_1,\ t_2 \in \overline{\mathbb{R}}$ by (2.38) and using Remark 2.6, we estimate

$$\begin{aligned}
&|v'|([x_0,\hat{x}] \setminus S_u) + \sum_{x\in S_u\cap[x_0,\hat{x}]} \Phi(\nu_u,(u')^a_-,(u')^a_+)\\
&\quad\geq |v'|(]x_0,\hat{x}]) + \Phi(1,(u')^a_-(x_0),(u')^a_+(x_0))\\
&\quad\geq |\Psi_1((u')^a_+(x_0)) - \Psi_1((u')^a_-(\hat{x}))| + |\Psi_1((u')^a_+(\hat{x})) - \Psi_1((u')^a_-(\hat{x}))|\\
&\qquad + \Phi(1,(u')^a_-(x_0),(u')^a_+(x_0))\\
&\quad= |\Psi_1((u')^a_+(x_0)) - \Psi_1((u')^a_-(\hat{x}))| + |\Psi_1((u')^a_+(\hat{x})) - \Psi_1((u')^a_-(\hat{x}))|\\
&\qquad + 2\Psi_1(+\infty) - \Psi_1((u')^a_-(x_0)) - \Psi_1((u')^a_+(x_0)) \qquad (4.17)\\
&\quad\geq -\Psi_1((u')^a_-(\hat{x})) + 2\Psi_1(+\infty) - \Psi_1((u')^a_-(x_0)) + |\Psi_1((u')^a_+(\hat{x})) - \Psi_1((u')^a_-(\hat{x}))|\\
&\quad= \Psi_1(C) - \Psi_1((u')^a_-(\hat{x})) + 2\Psi_1(+\infty) - \Psi_1((u')^a_-(x_0)) - \Psi_1(C)\\
&\qquad + |\Psi_1((u')^a_+(\hat{x})) - \Psi_1((u')^a_-(\hat{x}))|\\
&\quad\geq |\Psi_1(C) - \Psi_1((u')^a_+(\hat{x}))| + \Phi(1,(\hat{u}')^a_-(x_0),(\hat{u}')^a_+(x_0))\\
&\quad= |\hat{v}'|([x_0,\hat{x}] \setminus S_{\hat{u}}) + \sum_{x\in S_{\hat{u}}\cap[x_0,\hat{x}]} \Phi(\nu_{\hat{u}},(\hat{u}')^a_-,(\hat{u}')^a_+)\,,
\end{aligned}$$

where in the last inequality we have used (4.11) and (4.16). Collecting (4.13), (4.15), and (4.17) we deduce that

$$\overline{\mathcal{F}}_1(\hat{u}) + \lambda\int_a^b |\hat{u}-g|^2 + \frac{\eta}{2} \leq \overline{\mathcal{F}}_1(u_k) + \lambda\int_a^b |u_k - g|^2$$

and, in turn, by (4.14)

$$\overline{\mathcal{F}}_1(\hat{u}) + \lambda\int_a^b |\hat{u}-g|^2\,dx < \overline{\mathcal{F}}_1(u) + \lambda\int_a^b |u-g|^2\,dx\,, \qquad (4.18)$$

which contradicts the minimality of u.

If $\frac{1}{2}(u_+(x_0) + u_-(x_0)) < g(x_0)$ then we proceed in a similar manner: The comparison function $\hat{u}$ is now constructed by replacing u_k with an affine function (defined as before and with C and t properly chosen) in a left neighborhood of x_0. The argument is completely analogous to the previous one and we omit the details.

Step 2: We finally show that $(u')^a_-$ and $(u')^a_+$ are finite everywhere in $]a,b]$ and in $[a,b[$, respectively. We give the details only for $(u')^a_-$, since one can argue for $(u')^a_+$ in an entirely similar way.

Recall that by the previous step u is continuous. Once again we reason by contradiction by assuming that there exists $\bar{x} \in\ ]a,b]$ such that $|(u')^a_-(\bar{x})| = +\infty$. Without loss of generality we may suppose that $(u')^a_-(\bar{x}) = +\infty$. Using Remark 2.2 and the differentiability properties of BV functions we may choose a point $x_1 \in\ ]a,\bar{x}[$ such that u is differentiable at x_1 and

$$u(x_1) \neq g(x_1)\,,\ u'(x_1) = (u')^a_-(x_1) = (u')^a_+(x_1)\,,\ u'(x_1) > 2\|g'\|_\infty,\ |v'|(]a,x_1]) > 0\,. \qquad (4.19)$$

The first condition is a consequence of the fact that g is Lipschitz and u cannot be Lipschitz in any left neighborhood of $\bar{x}$, since $|(u')^a_-(\bar{x})| = +\infty$. The last condition follows easily from the fact that $(u')^a$ cannot be constant $\mathcal{L}^1$-a.e. on $]a,\bar{x}[$. Assume that $u(x_1) > g(x_1)$. Then, by (4.19) and by the previous step, we can find $\varepsilon \in [0, \frac{1}{2}[$, with

$$\Psi_1(u'(x_1)) - \Psi_1((1-\varepsilon)u'(x_1)) < |v'|([x_1, b[)\,, \tag{4.20}$$

such that the affine function $h(x) := u(x_1) + (1-\varepsilon)u'(x_1)(x - x_1)$ satisfies one of the following conditions: Either there exists a point $x_2 \in]x_1, b[$ for u such that

$$h(x_2) = u(x_2) \text{ and } g < h < u \text{ in }]x_1, x_2[\,, \tag{4.21}$$

or

$$g < h < u \text{ in }]x_1, b[\,. \tag{4.22}$$

In the latter case we set $x_2 := b$. We now consider the comparison function

$$\hat{u}(x) := \begin{cases} h(x) & \text{if } x \in]x_1, x_2[, \\ u(x) & \text{otherwise,} \end{cases}$$

and we denote $\hat{v} := \Psi_1 \circ (\hat{u}')^a$. We claim that (4.18) holds, contradicting the minimality of u. By (4.21) and (4.22) in any case we have

$$\lambda \int_a^b |\hat{u} - g|^2\, dx < \lambda \int_a^b |u - g|^2\, dx\,.$$

Moreover, if $x_2 < b$ we have by construction $|\hat{u}'|([x_1, x_2]) = u(x_2) - u(x_1) \le |u'|([x_1, x_2])$, while if $x_2 = b$ we have $|\hat{u}'|([x_1, b[) = u_-(b) - u(x_1) \le |u'|([x_1, b[)$, so that in both cases $|\hat{u}'|([a, b[) \le |u'|(]a, b[)$. Hence (4.18) will follow if we show that $|\hat{v}'|([x_1, x_2]) \le |v'|([x_1, x_2])$, where $[x_1, x_2]$ is replaced by $[x_1, b[$ if $x_2 = b$. To see this we first assume that (4.21) holds. Arguing as for (4.16), we deduce $(u')^a_-(x_2) \le h'(x_2) = (1-\varepsilon)u'(x_1)$. Therefore by (4.19) we have

$$\begin{aligned} |v'|([x_1, x_2]) =& |v'|(]x_1, x_2[) + |v'|(\{x_2\}) \\ \ge& \Psi_1(u'(x_1)) - \Psi_1((u')^a_-(x_2)) + |\Psi_1((u')^a_-(x_2)) - \Psi_1((u')^a_+(x_2))| \\ =& \Psi_1(u'(x_1)) - \Psi_1((1-\varepsilon)u'(x_1)) + \Psi_1((1-\varepsilon)u'(x_1)) - \Psi_1((u')^a_-(x_2)) \\ &+ |\Psi_1((u')^a_-(x_2)) - \Psi_1((u')^a_+(x_2))| \\ \ge& \Psi_1(u'(x_1)) - \Psi_1((1-\varepsilon)u'(x_1)) + |\Psi_1((1-\varepsilon)u'(x_1)) - \Psi_1((u')^a_+(x_2))| \\ =& |\hat{v}'|([x_1, x_2])\,. \end{aligned}$$

If (4.22) holds then, by (4.20), we obtain

$$|v'|([x_1, b[) > \Psi_1(u'(x_1)) - \Psi_1((1-\varepsilon)u'(x_1)) = |\hat{v}'|([x_1, b[)\,.$$

If $u(x_1) < g(x_1)$ we modify the previous argument in the following way. We now choose $\varepsilon \in [0, \frac{1}{2}[$ satisfying (4.20) with $|v'|([x_1, b[)$ replaced by $|v'|(]a, x_1])$ and such that the affine function $h(x)$ defined before satisfies one of the following conditions: Either there exists a point $x_2 \in]a, x_1[$ such that $h(x_2) = u(x_2)$ and $u < h < g$ in $]x_2, x_1[$, or $u < h < g$ in $]a, x_1[$. In the latter case we set $x_2 := a$. We now consider the comparison function

$$\hat{u}(x) := \begin{cases} h(x) & \text{if } x \in]x_2, x_1[, \\ u(x) & \text{otherwise,} \end{cases}$$

and we proceed exactly as before to show (4.18). □

We now turn to the main theorem of this subsection.

Theorem 4.5. *Assume that* $\psi\colon \mathbb{R} \to]0, +\infty[$ *is a bounded Borel function satisfying* (2.1) *and* (2.2)*, let* $g : [a, b] \to \mathbb{R}$ *be Lipschitz continuous, and let* $\{h_n\}$ *satisfy*

$$h_n \rightharpoonup 0 \quad \textit{weakly}^* \textit{ in } L^\infty(]a, b[)\,. \tag{4.23}$$

Define $\mathcal{A}_n$ *as the class of all solutions to* (4.7)*, with* g *replaced by* $g_n := g + h_n$*. Then for* n *large enough every solution* $u_n \in \mathcal{A}_n$ *is Lipschitz continuous. Moreover,*

$$\limsup_{n\to\infty} \sup_{w\in\mathcal{A}_n} \|w\|_{1,\infty} < +\infty \tag{4.24}$$

and for every sequence $\{u_n\} \subset \mathcal{A}_n$ *there exists a subsequence (not relabeled) and a solution* u *to* (4.7) *such that* $u_n \to u$ *in* $W^{1,p}(]a,b[)$ *for all* $p \in [1,+\infty[$*.*

Proof. It will be enough to prove that for any (sub)sequence $\{u_n\} \subset \mathcal{A}_n$ we may extract a further subsequence (not relabeled) and find a solution u to (4.7) such that u_n is Lipschitz continuous for n large enough,

$$\limsup_{n\to\infty} \|u_n\|_{1,\infty} < +\infty\,, \tag{4.25}$$

and $u_n \to u$ in $W^{1,p}(]a,b[)$ for all $p \in [1,+\infty[$. Since the sequence h_n is bounded in $L^\infty(]a,b[)$ for any $w \in X^1_\psi(]a,b[)$ we have

$$\sup_n \left(\overline{\mathcal{F}}_1(u_n) + \lambda \int_a^b (u_n - g_n)^2\,dx \right) \le \overline{\mathcal{F}}_1(w) + \lambda \int_a^b (w - g_n)^2\,dx \le C < \infty\,,$$

for a suitable constant $C > 0$ independent of n. By Corollary 2.7 there exist a subsequence not relabeled and a function $u \in X^1_\psi(]a,b[)$ such that

$$u_n \rightharpoonup u \quad \text{weakly}^* \text{ in } BV(]a,b[)\,, \tag{4.26}$$

and

$$u_n' \to (u')^a \quad \text{pointwise } \mathcal{L}^1\text{-a.e. in }]a,b[\,. \tag{4.27}$$

Moreover, since also the functions h_n^2 are equibounded, upon extracting a further subsequence we may find $f \in L^\infty(]a,b[)$ such that

$$h_n^2 \rightharpoonup f \quad \text{weakly}^* \text{ in } L^\infty(]a,b[)\,. \tag{4.28}$$

It is convenient to "localize" the functional $\overline{\mathcal{F}}_1$: For every Borel set $B \subset]a,b[$ and for $w \in X^1_\psi(]a,b[)$ we set

$$\overline{\mathcal{F}}_1(w;B) := |w'|(B) + |v'|(B \setminus S_w) + \sum_{x\in S_w\cap B} \Phi(\nu_w, (w')^a_-, (w')^a_+)\,, \tag{4.29}$$

where $v := \Psi_1 \circ (w')^a$. We divide the remaining part the proof into two steps.
Step 1: We claim that u is a solution of the minimization problem (4.7) and that for every open interval $I =]c,d[$, with $a \le c < d \le b$ and $c,\ d \in [a,b] \setminus S_{(u')^a}$,

$$\lim_{n\to\infty} \overline{\mathcal{F}}_1(u_n; I) = \overline{\mathcal{F}}_1(u; I)\,. \tag{4.30}$$

To see this, note that for each $n \in \mathbb{N}$

$$\lambda \int_I (u_n - g_n)^2\,dx = \lambda \int_I (u_n - g)^2\,dx - 2\lambda \int_I (u_n - g)\,h_n\,dx + \lambda \int_I h_n^2\,dx\,.$$

By (4.23), (4.26), and (4.28) it follows that

$$\lim_{n\to\infty} \int_I (u_n - g_n)^2\,dx = \int_I (u - g)^2\,dx + \int_I f\,dx\,. \tag{4.31}$$

Recall also that by lower semicontinuity

$$\liminf_{n\to\infty} \overline{\mathcal{F}}_1(u_n; A) \ge \overline{\mathcal{F}}_1(u; A)\,, \tag{4.32}$$

for every open set $A \subset]a,b[$.

By the minimality of u_n for every $w \in X^1_\psi(]a,b[)$ we have

$$\begin{aligned}\overline{\mathcal{F}}_1(w) &+\lambda \int_a^b (w-g)^2\,dx - 2\lambda \int_a^b (w-g)\,h_n\,dx + \lambda \int_a^b h_n^2\,dx \\ &= \overline{\mathcal{F}}_1(w) + \lambda \int_a^b (w-g_n)^2\,dx \geq \overline{\mathcal{F}}_1(u_n) + \lambda \int_a^b (u_n-g_n)^2\,dx\,.\end{aligned}$$

Using (4.32) (with $A =]a,b[$) and once again (4.23) and (4.28), we get

$$\begin{aligned}\overline{\mathcal{F}}_1(w) + \lambda \int_a^b (w-g)^2\,dx + \lambda \int_a^b f\,dx &\geq \limsup_{n\to\infty} \Big(\overline{\mathcal{F}}_1(u_n) + \lambda \int_a^b (u_n-g_n)^2\,dx\Big) \\ \geq \liminf_{n\to\infty} \Big(\overline{\mathcal{F}}_1(u_n) + \lambda \int_a^b (u_n-g_n)^2\,dx\Big) &\geq \overline{\mathcal{F}}_1(u) + \lambda \int_a^b (u-g)^2\,dx + \lambda \int_a^b f\,dx\,.\end{aligned}$$

Given the arbitrariness of $w \in X^1_\psi(]a,b[)$ this implies that u is a solution of the minimization problem (4.7). Moreover, taking $w = u$ in the previous inequalities and using (4.31) we deduce (4.30) for $I =]a,b[$; i.e.,

$$\lim_{n\to\infty} \overline{\mathcal{F}}_1(u_n) = \overline{\mathcal{F}}_1(u)\,. \tag{4.33}$$

It remains to prove (4.30) for every open interval of the form $I =]c,d[$, with $c,d \in [a,b] \setminus S_{(u')^a}$. To this end fix one such interval and assume by contradiction that

$$\limsup_{n\to\infty} \overline{\mathcal{F}}_1(u_n;I) > \overline{\mathcal{F}}_1(u;I)\,. \tag{4.34}$$

As u is continuous by Proposition 4.4, our assumption on I implies that the end points c and d do not charge $\overline{\mathcal{F}}_1(u;\cdot)$, so that $\overline{\mathcal{F}}_1(u;I) = \overline{\mathcal{F}}_1\big(u;\overline{I} \cap]a,b[\big)$. Therefore, combining (4.32), (4.33), and (4.34) we obtain

$$\begin{aligned}\overline{\mathcal{F}}_1(u) &= \overline{\mathcal{F}}_1\big(u;\overline{I}\cap]a,b[\big) + \overline{\mathcal{F}}_1\big(u;]a,b[\setminus\overline{I}\big) = \overline{\mathcal{F}}_1(u;I) + \overline{\mathcal{F}}_1\big(u;]a,b[\setminus\overline{I}\big) \\ &< \limsup_{n\to\infty} \overline{\mathcal{F}}_1(u_n;I) + \liminf_{n\to\infty} \overline{\mathcal{F}}_1\big(u_n;]a,b[\setminus\overline{I}\big) \leq \lim_{n\to\infty} \overline{\mathcal{F}}_1(u_n) = \overline{\mathcal{F}}_1(u)\,,\end{aligned}$$

which is a contradiction. This concludes the proof of (4.30).

Step 2: We now show that u_n is Lipschitz continuous for n large enough and that (4.25) holds. Note that the convergence of u_n to u in $W^{1,p}(]a,b[)$ for all $p \in [1,+\infty[$ will then easily follow from (4.25) and (4.27). Assume by contradiction that the conclusion is false. Then, arguing as at the beginning of the proof of Proposition 4.4, we may find a subsequence (not relabeled) and points $x_n \in]a,b[$ such that one of the following two cases holds:

(i) $x_n \notin S_{(u_n')^a}$ and $|(u_n')^a(x_n)| \to +\infty$;

(ii) $x_n \in S_{u_n}$ for every $n \in \mathbb{N}$.

Assume that (i) holds and, without loss of generality, that $(u_n')^a(x_n) \to +\infty$. Upon extracting a further subsequence we may also assume that $x_n \to x_0 \in [a,b]$. Recall that by Proposition 4.4 and by the previous step the function u is Lipschitz continuous. Hence there are two cases: Either

$$\overline{\mathcal{F}}_1(u;\{x_0\}\cap]a,b[) = 0 \tag{4.35}$$

or

$$x_0 \in S_{(u')^a}\,, \qquad (u')^a_\pm(x_0) \in \mathbb{R}\,, \qquad \overline{\mathcal{F}}_1(u;\{x_0\}) = |\Psi_1((u')^a_+(x_0)) - \Psi_1((u')^a_-(x_0))|\,. \tag{4.36}$$

Assume first that (4.35) holds. Set $L := \|u'\|_\infty$ and fix ε so small that,

$$\overline{\mathcal{F}}_1(u;I_\varepsilon) < \int_{L+1}^{+\infty} \psi(t)\,dt\,,$$

where $I_\varepsilon :=]x_0 - \varepsilon, x_0 + \varepsilon[\cap]a,b[$. By (4.30) we also have

$$\overline{\mathcal{F}}_1\,(u_n; I_\varepsilon) < \int_{L+1}^{+\infty} \psi(t)\,dt\,, \tag{4.37}$$

for n large enough. On the other hand by (4.27) there exists $y \in I_\varepsilon$ such that $(u_n')^a(y) < L+1$ for n large. Moreover, taking into account (i), we also have $(u_n')^a(x_n) > L+1$ for n large enough. Thus,

$$\overline{\mathcal{F}}_1\,(u_n; I_\varepsilon) \geq |v_n'|(I_\varepsilon) \geq |\Psi_1((u_n')^a(x_n)) - \Psi_1((u_n')^a(y))| \geq \Psi_1((u_n')^a(x_n)) - \Psi_1(L+1)\,,$$

for all n sufficiently large. Passing to the limit as $n \to \infty$ we then obtain

$$\liminf_{n\to\infty} \overline{\mathcal{F}}_1\,(u_n; I_\varepsilon) \geq \Psi_1(+\infty) - \Psi_1(L+1) = \int_{L+1}^{+\infty} \psi(t)\,dt\,,$$

which contradicts (4.37).

In case (4.36) holds, then $x_0 \in]a,b[$. Set

$$\eta := 2\Psi_1(+\infty) - \Psi_1((u')^a_+(x_0)) - \Psi_1((u')^a_-(x_0)) - |\Psi_1((u')^a_+(x_0)) - \Psi_1((u')^a_-(x_0))| > 0 \tag{4.38}$$

and choose ε such that both $x_0 - \varepsilon$ and $x_0 + \varepsilon$ belong to $]a,b[\setminus S_{(u')^a}$ and

$$\overline{\mathcal{F}}_1\,(u; I_\varepsilon) < |\Psi_1((u')^a_+(x_0)) - \Psi_1((u')^a_-(x_0))| + \frac{\eta}{3}\,, \tag{4.39}$$

$$|\Psi_1((u')^a_\pm(y)) - \Psi_1((u')^a_\pm(x_0))| < \frac{\eta}{4} \quad \text{for } y \in I_\varepsilon^\pm\,, \tag{4.40}$$

where $I_\varepsilon :=]x_0 - \varepsilon, x_0 + \varepsilon[$, $I_\varepsilon^+ :=]x_0, x_0 + \varepsilon[$, and $I_\varepsilon^- :=]x_0 - \varepsilon, x_0[$. Note that by (4.30) and (4.39) we have

$$\overline{\mathcal{F}}_1\,(u_n; I_\varepsilon) < |\Psi_1((u')^a_+(x_0)) - \Psi_1((u')^a_-(x_0))| + \frac{\eta}{3} \tag{4.41}$$

for n large enough. Moreover, by (4.27) and (4.40) we may find y^-, $y^+ \in I_\varepsilon$, with $y^- < x_0 < y^+$, such that

$$y^\pm \notin S_{(u_n')^a} \quad \text{and} \quad |\Psi_1((u_n')^a(y^\pm)) - \Psi_1((u')^a_\pm(x_0))| < \frac{\eta}{4} \tag{4.42}$$

for n large enough. As $y^- < x_n < y^+$ for n sufficiently large, we have

$$\begin{aligned}\overline{\mathcal{F}}_1\,(u_n; I_\varepsilon) \geq& |v_n'|(I_\varepsilon) \geq |\Psi_1((u_n')^a(x_n)) - \Psi_1((u_n')^a(y^-))| \\ &+ |\Psi_1((u_n')^a(x_n)) - \Psi_1((u_n')^a(y^+))| \\ \geq& |\Psi_1((u_n')^a(x_n)) - \Psi_1((u')^a_-(x_0))| + |\Psi_1((u_n')^a(x_n)) - \Psi_1((u')^a_+(x_0))| - \frac{\eta}{2}\,,\end{aligned} \tag{4.43}$$

where the last inequality follows from (4.42). Letting $n \to \infty$ in (4.43) and recalling (4.38) we deduce

$$\begin{aligned}\liminf_{n\to\infty} \overline{\mathcal{F}}_1\,(u_n; I_\varepsilon) &\geq 2\Psi_1(+\infty) - \Psi_1((u')^a_+(x_0)) - \Psi_1((u')^a_-(x_0)) - \frac{\eta}{2} = \\ &= |\Psi_1((u')^a_+(x_0)) - \Psi_1((u')^a_-(x_0))| + \frac{\eta}{2}\,,\end{aligned}$$

which contradicts (4.41). This concludes the proof of (4.25) if (i) holds. An entirely similar argument can be used to treat the other case. □

4.3. **Absence of the staircase effect: The case $p > 1$.** We now turn to the case $p > 1$. We consider the minimization problem

$$\min\left\{\overline{\mathcal{F}}_p\,(u) + \lambda \int_a^b |u - g|^2\,dx : \, u \in X^p_\psi(]a,b[)\right\}, \tag{4.44}$$

where $\overline{\mathcal{F}}_p$ is the relaxed functional given in (3.24). We start with two auxiliary results.

Proposition 4.6. *Let $p>1$ and assume that $\psi\colon\mathbb{R}\to\,]0,+\infty[$ is a bounded Borel function satisfying* (2.2) *and* (3.1). *Let g be Lipschitz continuous and let u_n be a sequence in $X^p_\psi(]a,b[)$ such that $\sup_n \overline{\mathcal{F}}_p(u_n)<+\infty$ and $u_n\to g$ in $L^2(]a,b[)$. Then $g\in C^1([a,b])\cap X^p_\psi(]a,b[)$. Moreover, $u_n\in C^1([a,b])$ for n large enough and $u_n\to g$ in $C^1([a,b])$.*

Proof. By the assumptions and by Corollary 3.5 we deduce that $g\in X^p_\psi(]a,b[)$. The fact that $g\in C^1([a,b])$ now follows from Remark 3.2-(i). To prove the last part of the statement we start by showing that $(u_n')^a\to g'$ uniformly in $]a,b[$. Again by Corollary 3.5 the whole sequence u_n satisfies

$$\Psi_p\circ(u_n')^a\rightharpoonup\Psi_p\circ g'\quad\text{weakly in } W^{1,p}(]a,b[)\,,\tag{4.45}$$

which implies, in particular, that

$$(\Psi_p\circ(u_n')^a)([a,b])\subset[\Psi_p(-2\|g'\|_\infty),\Psi_p(2\|g'\|_\infty)]\quad\text{for } n \text{ large enough.}\tag{4.46}$$

Since by (2.2) Ψ_p^{-1} is Lipschitz continuous on $[\Psi_p(-2\|g'\|_\infty),\Psi_p(2\|g'\|_\infty)]$, it follows from (4.45) and (4.46) that $(u_n')^a\to g'$ uniformly in $]a,b[$. In turn, by Definition 3.1 we have that $u_n'=(u_n')^a$ in $]a,b[$. In particular $u_n\in C^1([a,b])$ by Remark 3.2-(i) and $u_n\to g$ in $C^1([a,b])$. □

Proposition 4.7. *Let p and ψ be as in the previous proposition. Then for every $C>0$ there exists $\overline{\lambda}=\overline{\lambda}(C)$ with the following property: For all $g\in C^1([a,b])\cap X^p_\psi(]a,b[)$, with $\|g\|_{C^1([a,b])}\le C$ and $\overline{\mathcal{F}}_p(g)\le C$, and for all $\lambda\ge\overline{\lambda}$ every solution u to* (4.44) *belongs to $C^1([a,b])$.*

Proof. Assume by contradiction that for every $n\in\mathbb{N}$ there exist $g_n\in C^1([a,b])\cap X^p_\psi(]a,b[)$, with $\|g_n'\|_\infty\le C$ and $\overline{\mathcal{F}}_p(g_n)\le C$, and a solution u_n to

$$\min\left\{\overline{\mathcal{F}}_p(u)+n\int_a^b|u-g_n|^2\,dx:\,u\in X^p_\psi(]a,b[)\right\}$$

which does not belong to $C^1([a,b])$. Owing to Proposition 4.6 we may assume, without loss of generality, that $g_n\to g$ in $C^1([a,b])$ for a suitable function $g\in C^1([a,b])\cap X^p_\psi(]a,b[)$. Moreover, by minimality, we have

$$\overline{\mathcal{F}}_p(u_n)+n\int_a^b|u_n-g_n|^2\,dx\le\overline{\mathcal{F}}_p(g_n)\le C\,.$$

It follows in particular that $\sup_n\overline{\mathcal{F}}_p(u_n)<+\infty$ and $u_n\to g$ in $L^2(]a,b[)$. By Proposition 4.6 we conclude that $u_n\in C^1([a,b])$ for n large enough, which gives a contradiction. □

The next theorem shows that also in the case $p>1$ the staircase effect does not occur.

Theorem 4.8. *Let ψ and p be as in Proposition 4.6, let $g\in C^1([a,b])\cap X^p_\psi(]a,b[)$, and let h_n satisfy* (4.23). *For $\lambda>0$ and $n\in\mathbb{N}$ let $\mathcal{A}_{\lambda,n}\subset X^p_\psi(]a,b[)$ be the class of the solutions to the minimization problem* (4.44), *with g replaced by $g_n:=g+h_n$. Let $\overline{\lambda}$ be as in Proposition 4.7, with $C:=\max\{\|g\|_{C^1([a,b])},\overline{\mathcal{F}}_p(g)\}$. Then for all $\lambda\ge\overline{\lambda}$ we have $\mathcal{A}_{\lambda,n}\subset C^1([a,b])$ for n sufficiently large. Moreover,*

$$\lim_{\lambda\to\infty}\limsup_{n\to\infty}\sup_{u\in\mathcal{A}_{\lambda,n}}\|u-g\|_{C^1([a,b])}=0\,.\tag{4.47}$$

Proof. We start by showing the second part of the statement. Assume by contradiction that (4.47) does not hold. Then there exist $\delta>0$, a sequence of real numbers $\lambda_k\to+\infty$ and, for every k, a sequence of integers $n_j^k\to\infty$ as $j\to\infty$, such that for every k, j

$$\|u_{\lambda_k,n_j^k}-g\|_{C^1([a,b])}\ge\delta\tag{4.48}$$

for a suitable function $u_{\lambda_k,n_j^k}\in\mathcal{A}_{\lambda_k,n_j^k}$ (with the understanding that $\|u_{\lambda_k,n_j^k}-g\|_{C^1([a,b])}=+\infty$ if $u_{\lambda_k,n_j^k}\notin C^1([a,b])$). Arguing exactly as in Step 1 of the proof of Theorem 4.5 we can show that for every k there exist a subsequence (still denoted by n_j^k) and a solution u_k to (4.44) with λ replaced by λ_k, such that

$$u_{\lambda_k, n_j^k} \rightharpoonup u_k \quad \text{weakly}^* \text{ in } BV(]a,b[) \qquad \text{and} \qquad \overline{\mathcal{F}}_p(u_{\lambda_k, n_j^k}) \to \overline{\mathcal{F}}_p(u_k) \tag{4.49}$$

as $j \to \infty$. Moreover, since $g \in C^1([a,b]) \cap X^p_\psi(]a,b[)$, we have by minimality that

$$\overline{\mathcal{F}}_p(u_k) + \lambda_k \int_a^b |u_k - g|^2 \, dx \le \overline{\mathcal{F}}_p(g) , \tag{4.50}$$

which shows, in particular, that $u_k \to g$ in $L^2(]a,b[)$. Combining (4.49) and (4.50), and using a diagonal argument, we may find a subsequence $n^k_{j_k}$ such that

$$\sup_k \overline{\mathcal{F}}_p(u_{\lambda_k, n^k_{j_k}}) < +\infty \qquad \text{and} \qquad u_{\lambda_k, n^k_{j_k}} \to g \text{ in } L^2(]a,b[) .$$

Proposition 4.6 then implies that $u_{\lambda_k, n^k_{j_k}} \to g$ in $C^1([a,b])$, which contradicts (4.48).

Finally, the first part of the statement follows from a similar argument by contradiction as a consequence of Propositions 4.6 and 4.7 and of the fact that if $u_n \in \mathcal{A}_{\lambda,n}$ then, up to subsequences, u_n converges to a solution of (4.44). □

Acknowledgments. The authors thank the Center for Nonlinear Analysis (NSF Grants No. DMS-0405343 and DMS-0635983) for its support during the preparation of this paper. The research of G. Dal Maso and M. Morini was partially supported by the Projects "Calculus of Variations" 2004 and "Problemi di Calcolo delle Variazioni in Meccanica e in Scienza dei Materiali" 2006-2008, supported by the Italian Ministry of Education, University, and Research, and by the Project "Variational Problems with Multiple Scales" 2006, supported by the Italian Ministry of University and Research. The research of I. Fonseca was partially supported by the National Science Foundation under Grant No. DMS-040171 and that of G. Leoni under Grants No. DMS-0405423 and DMS-0708039.

References

[1] Ambrosio L., Fusco N., Pallara D.: Functions of bounded variation and free discontinuity problems. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[2] Blomgren P., Chan T. F., Mulet P.: Extensions to total variation denoising. *Proc. SPIE 97*, San Diego, 1997.

[3] Chambolle A., Lions P.-L.: Image recovery via total variation minimization and related problems. *Numer. Math.* **76** (1997), 167-188.

[4] Chan T., Marquina A., Mulet P.: High-order total variation-based image restoration. *SIAM J. Sci. Comput.* **22** (2000), 503-516.

[5] Dal Maso, G.: An introduction to Γ-convergence. Birkhäuser, Boston, 1993.

[6] DiBenedetto E.: Real analysis. Birkhäuser, Boston, 2002.

[7] Geman D., Reynolds G.: Constrained restoration and the recovery of discontinuities. *IEEE Trans. on Pat. An. and Mach. Intel.* **14** (1992), 367-383.

[8] Hewitt E., Stromberg K.: Real and abstract analysis. A modern treatment of the theory of functions of a real variable. Springer-Verlag, New York-Heidelberg, 1975.

[9] Kindermann S., Osher S., Jones P.W.: Deblurring and denoising of images by nonlocal functionals. *Multiscale Model. Simul.* **4** (2005), 1091-1115.

[10] Ring, W.: Stuctural properties of solutions to total variation regularization problems. *M2AN Math. Model. Numer. Anal.* **34** (2000), 799-810.

[11] Rudin L., Osher S., Fatemi E.: Nonlinear total variation based noise removal algorithms. *Physica D* **60** (1992), 259-268.

[12] Rudin W.: Real and complex analysis. Second edition. McGraw-Hill, New York-Düsseldorf-Johannesburg, 1974.

(Gianni Dal Maso) SISSA, Via Beirut 2, 34014 Trieste, Italy
E-mail address, Gianni Dal Maso: dalmaso@sissa.it

(Irene Fonseca) Carnegie Mellon University, Pittsburgh PA 15213-3890, USA
E-mail address, Irene Fonseca: fonseca@andrew.cmu.edu

(Giovanni Leoni) Carnegie Mellon University, Pittsburgh PA 15213-3890, USA
E-mail address, Giovanni Leoni: giovanni@andrew.cmu.edu

(Massimiliano Morini) SISSA, Via Beirut 2, 34014 Trieste, Italy
E-mail address, Massimiliano Morini: morini@sissa.it