A generalization of the Moreau–Yosida regularization

Aras Bacho†

Abstract

In many applications, one deals with nonsmooth functions, e.g., in nonsmooth dynamical systems, nonsmooth mechanics, or nonsmooth optimization. In order to establish theoretical results, it is often beneficial to regularize the nonsmooth functions in an intermediate step. In this work, we investigate the properties of a generalization of the Moreau–Yosida regularization on a normed space where we replace the quadratic kernel in the infimal convolution with a more general function. More precisely, for a function $f : X \to (-\infty, +\infty]$ defined on a normed space $(X, \| \cdot \|)$ and given parameters $p > 1$ and $\varepsilon > 0$, we investigate the properties of the generalized Moreau–Yosida regularization given by

$$f_\varepsilon(u) = \inf_{v \in X} \left\{ \frac{1}{p\varepsilon}\|u - v\|^p + f(v) \right\}, u \in X.$$ 

We show that the generalized Moreau–Yosida regularization satisfies the same properties as in the classical case for $p = 2$, provided that $X$ is not a Hilbert space. We further establish a convergence result in the sense of Mosco-convergence as the regularization parameter $\varepsilon$ tends to zero.

Keywords: MOREAU–YOSIDA regularization · Convex analysis · p-duality map · GâTEAUX differentiability · MOSCO-convergence · Nonsmooth analysis

Mathematics Subject Classification 34G25 · 46N10 · 49J52

1 Introduction

1.1 Preliminaries and notation

We denote $(X, \| \cdot \|)$ a normed space and $(X^*, \| \cdot \|_*)$ its topological dual space. The duality pairing between $X^*$ and $X$ is denoted by $\langle \cdot, \cdot \rangle$. For a functional $f : X \to (-\infty, +\infty]$, the effective domain is defined by $\text{dom}(f) := \{ u \in X : f(u) < +\infty \}$. The function $f$ is called proper if $\text{dom}(f) \neq \emptyset$. Furthermore, the subdifferential of $f$ in the sense of convex analysis is given by

$$\partial f(u) := \{ \xi \in X^* : f(u) - f(v) \leq \langle \xi, v \rangle \text{ for all } v \in X \}. $$

The functional $f$ is called subdifferentiable in $u \in X$ if $\partial f(u) \neq \emptyset$. The domain of the subdifferential is defined by $\text{dom}(\partial f) := \{ u \in X : \partial f(u) \neq \emptyset \}$. For a general proper, lower

---

†Ludwig-Maximilians-Universität München, Mathematisches Institut, Theresienstr. 39, 80333 München, Germany.
semicontinuous, and convex functional \( f : X \to (-\infty, +\infty] \) on a normed space \((X, \| \cdot \|)\), the classical Moreau–Yosida regularization of \( f \) is defined via
\[
f_\varepsilon(u) = \inf_{v \in X} \left\{ \frac{1}{2\varepsilon} \|u - v\|^2 + f(v) \right\}, \quad u \in X,
\]
where \( \varepsilon > 0 \) is called the regularization parameter. It is well known that the geometrical properties of the dual space \( X^* \) are intimately related to the regularity properties of the regularization \( f_\varepsilon \), see, e.g., BARBU [Bar10] and BARBU & PRECUPANU [BaP86]. Roughly speaking, the better the geometrical properties of the dual space \( X^* \) are, the better the regularization becomes. In the following, we elaborate on this in more detail. To do so, we recall the definition of the duality map \( F_X : X \rightrightarrows X^* \), which is given by the set \( F_X(v) := \{ \xi \in X^* : \langle \xi, v \rangle = \|v\|^2 = \|\xi\|^2_2 \} \). It is well known that the duality map is given by the subdifferential of the mapping \( u \mapsto \frac{1}{2} \|u\|^2 \), i.e., \( F_X(u) = \partial \left( \frac{1}{2} \|u\|^2 \right) \) for all \( u \in X \). Furthermore, it is easily checked that for all \( u \in X \), the set \( F_X(u) \) is non-empty, convex, bounded, and weak*-closed\(^1\), see, e.g., BARBU & PRECUPANU [BaP86, Section 1.2.4].

The duality map also has a geometrical interpretation: by the HAHN–BANACH theorem, see, e.g., BRÉZIS [Bré11, Theorem 1.1, p. 1], for \( u \in X \), there holds
\[
\|u\| = \max_{\xi \in X^*, \|\xi\|_* = 1} \langle \xi, u \rangle = \max_{\xi \in X^*, \|\xi\|_* \leq \|u\|} \frac{\langle \xi, u \rangle}{\|u\|} \geq \frac{\langle \xi, u \rangle}{\|u\|} \quad \text{for all } \xi \in X \text{ with } \|\xi\|_* = \|u\|.
\]
Thus, an element of the dual space belongs to the duality map \( \xi^* \in F_X(u) \) if and only if it solves the maximization problem
\[
\max_{\xi \in X^*, \|\xi\|_* = \|u\|} \frac{\langle \xi, u \rangle}{\|u\|},
\]
for which the set of maximizers is non-empty. In other words, \( \xi^* \) generates a closed supporting hyperplane to the closed ball \( \overline{B}(0, \|u\|) \).

Furthermore, we call a norm smooth if and only if the duality map is single-valued, or geometrically speaking, each supporting hyperplane which passes through a boundary point of the sphere \( S(0, \|u\|) \) with radius \( \|u\| \) is also a tangential hyperplane. We call a normed space smooth if there is an equivalent smooth norm. From (1.2), it is then readily seen that if the dual space \( X^* \) is strictly convex, i.e., the dual norm \( \| \cdot \|_* \) is strictly convex, the element which generates the supporting hyperplane is unique, meaning that the duality map \( F_X(u) \) is single-valued. In this case, the duality map is also demicontinuous \(^2\), which implies that the norm on \( X \) is Gâteaux differentiable. If the dual space \( X^* \) is uniformly convex\(^3\), then the duality map is uniformly continuous on every bounded subset of \( X \) and the norm on \( X \) is uniformly Fréchet differentiable in the sense that the limit
\[
\lim_{\lambda \to 0} \frac{\|u + \lambda v\| - 1}{\lambda}
\]

\(^1\)Therefore, \( F_X(u) \) is weak*-compact.

\(^2\)A map \( f : X \to Y \) between two normed spaces \( X \) and \( Y \) is called demicontinuous if it is strong-to-weak* continuous.

\(^3\)The normed space \( X \) is called uniformly convex if for every \( 0 < \varepsilon < 2 \) there exists \( \delta > 0 \) such that for any two vectors \( x, y \in X \) with \( \|x\| = \|y\| = 1 \) the condition \( \|x - y\| \geq \varepsilon \) implies that \( \|\frac{x - y}{\varepsilon}\| \leq 1 - \delta \). An uniformly convex space is in particular strictly convex.
A crucial assumption in all the previous results is the convexity of the functional \( \mathcal{F} \). Moreau–Yosida. The classical remarkable that similar results have been obtained for non-convex functionals \( g \). In addition, we show the convergence of the \( p \)-Moreau–Yosida regularization in the sense defined on a \( \mathcal{H} \)-space, in particular, is reflexive and that the duality map is identical with the Riesz isomorphism between the \( \mathcal{H} \)-space and its dual. For a more detailed discussion about the geometry of \( \mathcal{H} \) spaces, and in particular with regard to the duality maps, we refer the interested reader to [Bar76, Bar10, BaP86, Kie02, Byn71, Byn76, Die75, Zem91].

### 1.2 Literature review

The classical Moreau–Yosida regularization as defined in (1.1) has been studied extensively and has been employed successfully in many applications in order to circumvent the lack of regularity. The properties for the classical Moreau–Yosida regularization can, for reflexive \( \mathcal{H} \) spaces, be found in, e.g., Barbu [Bar10] and Barbu & Precupanu [BaP86] and for \( \mathcal{H} \) spaces in, e.g., Attouch [Att84] and Moreau [Mor65]. More general infimal convolutions defined by

\[
(f \boxplus g)(u) := \inf_{v \in \mathcal{H}} \{ f(u - v) + g(v) \} \quad (1.3)
\]

for proper, lower semicontinuous and convex functionals \( g \) and \( f \) defined on a \( \mathcal{H} \)-space has been studied in Bauschke & Combettes [BaC11]. In particular, the Pasch–Hausdorff envelope, i.e., \( g(v) = \beta\|v\| \), the Moreau envelope, i.e., \( g(v) = \frac{1}{\gamma^2}\|v\|^2 \), and the case \( g(v) = \frac{1}{\gamma^2}\|v\|^p, p > 1 \), have been studied. The present work generalizes the previous works by showing that the Moreau–Yosida regularization, henceforth called \( p \)-Moreau–Yosida regularization, for the kernel \( g(v) = \frac{1}{\gamma^2}\|v\|^p, p > 1 \) defined on a reflexive \( \mathcal{H} \)-space satisfies all the properties as the classical Moreau–Yosida regularization Gâteaux differentiability except from the Lipschitz continuity of the Gâteaux derivative of the regularization in the case the underlying space \( \mathcal{H} \) is a \( \mathcal{H} \)-space. In addition, we show the convergence of the \( p \)-Moerau–Yosida regularization in the sense of Mosco as the regularization parameter vanishes.

A crucial assumption in all the previous results is the convexity of the functional \( f \). It is remarkable that similar results have been obtained for non-convex functionals \( f \) that are defined on a \( \mathcal{H} \)-space via the the so-called Lions–Lasry regularization introduced by P.L. Lions and Lasry [LaL86]. The Lions–Lasry regularization of a proper function \( f : \mathcal{H} \to (-\infty, +\infty] \) that is minorized by a quadratic function is defined by

\[
(f_{\lambda})^\mu(u) := \sup_{v \in \mathcal{H}} \inf_{w \in \mathcal{H}} \left\{ f(w) + \frac{1}{2\lambda}\|v - w\|^2 - \frac{1}{2\mu}\|u - v\|^2 \right\} \quad (1.4)
\]

Similarly, one can define a regularization for a proper function \( g \) that is majorized by a quadratic function. Among other properties, it has been shown in Attouch and Aze [AtA93] that these functions are Fréchet differentiable with Lipschitz continuous derivative, weakly- or \( \lambda \)-convex (i.e. convex up to a square), satisfy \( (f_{\lambda})^\mu \leq f \), and
that \( (f_\lambda)^\mu (u) \) coincides with the MOREAU–YOSIDA regularization when \( f \) is convex. The FRÉCHET differentiability has also been shown for a more general class of kernels which includes the class of YOUNG functions. For quadratic kernels, these results have been partially extended by STRÖMBERG [Str96] to the case where \( X \) is a Banach space whose norm and dual norm are (locally) uniformly rotund, i.e., if \( \| \cdot \|^2 \) and \( \| \cdot \|^2_\ast \) are (locally) uniformly convex functions. PENOT [Pen98] has studied the FRÉCHET differentiability of the infimal convolution (1.3) in relation to the proximal mapping

\[ P_{f,g}(u) := \{ v \in X : (f \square g)(u) = f(u) + g(v) \}. \]

In particular, it has been shown that the non-emptiness of \( P_{f,g}(u) \) is related to certain properties of the (FRÉCHET or HADAMARD) subdifferential of \( f \), or under smoothness assumptions on the \( X \), the FRÉCHET derivative of \( f \). PENOT and NGAI [VaP16] have further extended the result by imposing a milder growth condition on the function \( f \). In addition, the authors studied the LIONS–LASRY regularization for more general kernels, i.e.,

\[ (f_\lambda)^\mu (u) := \sup_{v \in X} \inf_{w \in X} \left\{ f(w) + \frac{1}{\lambda} g(\|v - w\|) - \frac{1}{\mu} g(\|u - v\|) \right\} \]

for a convex, monotonically increasing, coercive, and continuously differentiable function \( h : \mathbb{R}^+ \to \mathbb{R}^+ \) with \( h(0) = 0 \).

However, for non-convex functionals \( f \), none of the previous results duplicate our results, and it is subject to future work to reproduce our results for the non-convex case with the aid of the previous results for non-convex functionals. We refer the interested reader to [Ber10, BTZ11, JTZ14] and the references therein for more results in the non-convex case. The references presented here are indeed not exhaustive.

2 Main result

The question arises: if and to what extent the properties of the duality map are related to the regularization properties of the MOREAU–YOSIDA regularization. We will see that the properties of the duality map are inherited by the subdifferential of the MOREAU–YOSIDA regularization. In fact, we will answer the question for the more general so-called \( p \)-MOREAU–YOSIDA regularization, which for \( p > 1 \), is given by

\[ f_\varepsilon (u) = \inf_{v \in X} \left\{ \frac{\varepsilon}{p} \left\| \frac{u - v}{\varepsilon} \right\|^p + f(v) \right\} , u \in X. \]  

The reason why we want to study \( p \)-MOREAU–YOSIDA regularization is simply because it maintains the growth of the functional \( f \) if it has \( p \)-growth, see, e.g., [Bac20, Bac21].

The following lemma shows some basic properties of the \( p \)-MOREAU–YOSIDA regularization on general normed spaces.

**Lemma 2.1.** Let \( f : X \to (-\infty, +\infty] \) be a proper and convex functional, and, for \( \varepsilon > 0 \) and \( p > 1 \), let \( f_\varepsilon \) be the \( p \)-MOREAU–YOSIDA regularization defined by (2.1). Then, \( f_\varepsilon : X \to \mathbb{R} \) is finite, convex, and locally LIPSCHITZ continuous. If, in addition, \( f \) is lower semicontinuous and \( X \) is a reflexive \( \text{BANACH space, then the infimum in} \)

\[ f_\varepsilon (u) = \inf_{v \in X} \left\{ \frac{\varepsilon}{p} \left\| \frac{u - v}{\varepsilon} \right\|^p + f(v) \right\} \]

is attained at every point \( u \in X \).
Proof. Let \( \tilde{u} \in \text{dom}(f) \neq \emptyset \). Then, on the one hand, there holds
\[
    f_\varepsilon(u) \leq \frac{1}{p \varepsilon^{p-1}} \|u - \tilde{u}\|^p + f(\tilde{u}) < \infty \quad \text{for every } u \in X.
\] (2.2)

On the other hand, by Ekeland & Temam [EkT99, Proposition 3.1, p. 14], there exists an affine linear minorant to \( f \), i.e., there exist \( \xi \in X^* \) and \( \alpha \in \mathbb{R} \) such that
\[
    f(v) \geq \alpha + \langle \xi, v \rangle \quad \text{for all } v \in X,
\]
so that \( f_\varepsilon(u) > -\infty \) for every \( u \in X \). This implies \( \text{dom}(f_\varepsilon) = X \). Now, for \( \lambda \in (0, 1) \) and \( u_1, u_2 \in X \), let \( (v_n^i)_{n \in \mathbb{N}} \subset X \) be a minimizing sequence for \( f_\varepsilon(u_i), i = 1, 2 \). We set \( w_n := \lambda v_n^1 + (1 - \lambda)v_n^2, n \in \mathbb{N} \). Then, by the convexity of \( f \), there holds
\[
    f_\varepsilon(\lambda u_1 + (1 - \lambda)u_2) = \inf_{v \in X} \left\{ \frac{1}{p \varepsilon^{p-1}} \|\lambda u_1 + (1 - \lambda)u_2 - v\|^p + f(v) \right\}
\[
    \leq \frac{1}{p \varepsilon^{p-1}} \|\lambda u_1 + (1 - \lambda)u_2 - w_n\|^p + f(w_n)
\[
    \leq \lambda \left( \frac{1}{p \varepsilon^{p-1}} \|u_1 - v_n^1\|^p + f(v_n^1) \right)
\[
    + (1 - \lambda) \left( \frac{1}{p \varepsilon^{p-1}} \|u_2 - v_n^2\|^p + f(v_n^2) \right)
\[
    \rightarrow \lambda f_\varepsilon(u_1) + (1 - \lambda)f_\varepsilon(u_2) \quad \text{as } n \rightarrow \infty,
\]
which shows the convexity of \( f_\varepsilon \). We note that by (2.2), \( f_\varepsilon \) is bounded on every open bounded set of \( X \). Hence, by Ekeland & Temam [EkT99, Corollary 2.4, p. 12], \( f_\varepsilon \) is locally Lipschitz continuous on \( X \). Finally, if \( X \) is a reflexive Banach space, then the infimum in \( f_\varepsilon(u) = \inf_{v \in X} \left\{ \frac{1}{p \varepsilon^{p-1}} \|\frac{v}{p}\|^p + f(v) \right\} \) is attained at every point \( u \in X \) by the direct method of calculus of variations.

In the main theorem, we will show properties of the \( p \)-Moreau–Yosida regularization under the assumption that \( X \) is reflexive such that, by the renorming theorem, \( X \) and \( X^* \) are simultaneously strictly convex and smooth. Before we progress to the next theorem, we recall that the \( p \)-duality map \( F_X^p \) is given by \( F_X^p := \partial_{\frac{1}{p}} \|\cdot\|^p \) for \( p > 1 \). Then, since the mapping \( v \mapsto \frac{1}{p} \|v\|^p \) is continuous and convex on \( X \), Ekeland & Temam [EkT99, Proposition 5.1 & 5.2, Corollary 5.1, pp. 21] ensure that \( F_X^p \) is a bounded and set-valued map such that \( F_X^p(u) \) is non-empty, convex, and weak*-closed for all \( u \in X \). Furthermore, by [EkT99, Example 4.3, pp. 19] the \( p \)-duality map is characterized by
\[
    F_X^p(u) = \{ \xi \in X^*: \langle \xi, u \rangle = \|u\|^p = \|\xi\|^{p^*} \}.
\] (2.3)

As for \( p = 2 \), if the dual space is strictly convex, then by Kien [Kie02, Proposition 2.3] and Akagi & Melchionna [AkM18, Lemma 19], the \( p \)-duality map is demi-continuous, single-valued, and monotone in the sense that
\[
    \langle F_X^p(u) - F_X^p(v), u - v \rangle \geq \left( \|u\|^{p-1} - \|v\|^{p-1} \right) (\|u\| - \|v\|) \quad \text{for all } u, v \in X.
\]

With the above-mentioned properties of the \( p \)-duality map, we are able to prove in the following theorem that the \( p \)-Moreau–Yosida regularization is, under suitable conditions, Gâteaux differentiable and has a demi-continuous Gâteaux derivative. This result generalizes and follows the proof of Barbu [Bar10, Theorem 2.58, p. 98] where the case \( p = 2 \) has been studied.
Theorem 2.2. Let $X$ be a reflexive BANACH space such that $X$ and its dual $X^*$ are strictly convex and smooth, and let $p > 1$ and $\varepsilon > 0$. Furthermore, let $f : X \rightarrow (-\infty, +\infty]$ be a proper, lower semicontinuous, and convex functional. Then, the $p$-MOREAU–YOSIDA regularization is convex and locally LIPSCHITZ continuous, and if $f$ is strictly convex, so is $f_\varepsilon$. Moreover, $f_\varepsilon(u) = \inf_{v \in X} \left\{ \frac{1}{p-1} \|u - v\|^p + f(v) \right\}$ attains at every point $u \in X$ its unique minimizer denoted by $u_\varepsilon = J_\varepsilon(u) := \arg\min_{v \in X} \left\{ \frac{1}{p-1} \|u - v\|^p + f(v) \right\}$, and $u_\varepsilon$ satisfies the EULER-LAGRANGE equation

$$0 \in F_X^p \left( \frac{u_\varepsilon - u}{\varepsilon} \right) + \partial f(u_\varepsilon).$$

Furthermore, $f_\varepsilon$ is Gâteaux-differentiable at every point $u \in X$ with the Gâteaux-derivative $A_\varepsilon : X \rightarrow X^*$ being demicontinuous on $X$ and satisfying $A_\varepsilon(u) = -F_X^p \left( \frac{u - u_\varepsilon}{\varepsilon} \right)$. If $X^*$ is uniformly convex, then $A_\varepsilon$ is continuous. Moreover, the following assertions hold:

i) $f_\varepsilon(u) = \frac{\varepsilon}{p} \|A_\varepsilon(u)\|_p^p + f(u_\varepsilon)$ for every $u \in X$,

ii) $f(u_\varepsilon) \leq f_\varepsilon(u) \leq f(u)$ for all $u \in X$ and all $\varepsilon_1 \geq \varepsilon_2 > 0$,

iii) $\lim_{\varepsilon \to 0} \|u_\varepsilon - u\| = 0$ for all $u \in \text{dom}(f)$,

iv) $\lim_{\varepsilon \to 0} f_\varepsilon(u) = f(u)$ for every $u \in X$.

v) For each $u \in \text{dom}(\partial f)$ there holds $A_\varepsilon(u) \rightharpoonup A_0(u) = \partial f(u)$ as $\varepsilon \to 0$, where $A_0(u) := \arg\min\{\|\xi\|_\varepsilon : \xi \in \partial f(u)\}$. If $X^*$ is uniformly convex, then $A_\varepsilon(u) \rightharpoonup A_0(u)$ as $\varepsilon \to 0$.

Finally, the mapping $\varepsilon \mapsto f_\varepsilon(u)$ is differentiable on $(0, +\infty)$ with

$$\frac{d}{d\varepsilon} f_\varepsilon(u) = -\frac{1}{p\varepsilon^{p-1}} \|u_\varepsilon - u\|_p^p \text{ for all } \varepsilon > 0. \tag{2.5}$$

Proof. By Lemma 2.1, the $p$-MOREAU–YOSIDA regularization is convex and locally LIPSCHITZ continuous on $X$. Now, let $f$ be strictly convex and let $u^0, u^1 \in X$ and $t \in (0, 1)$. Then, we define $u^t = tu^0 + (1 - t)u^1$ and assume

$$f_\varepsilon(u^t) = tf_\varepsilon(u^0) + (1 - t)f_\varepsilon(u^1).$$

Then, using the convexity of $\| \cdot \|_p^p$ and $f$, we obtain

$$tf_\varepsilon(u^0) + (1 - t)f_\varepsilon(u^1) = f_\varepsilon(u^t)$$

$$= \inf_{v \in X} \left\{ \frac{1}{p-1} \|u^t - v\|^p + f(v) \right\}$$

$$\leq \frac{1}{p-1} \|u^t - (tu^0 + (1 - t)u^1)\|^p + f(tu^0 + (1 - t)u^1)$$

$$\leq \frac{t}{p-1} \|u^0 - u^0\|^p + \frac{(1 - t)}{p-1} \|u^1 - u^1\|^p + tf(u^0) + (1 - t)f(u^1)$$

$$= tf_\varepsilon(u^0) + (1 - t)f_\varepsilon(u^1), \tag{2.6}$$

with $u^t = tu^0 + (1 - t)u^1$. 

Furthermore, $f_\varepsilon(u^t)$ is uniformly convex, then $A_\varepsilon(u^t) \rightharpoonup A_0(u^t)$ as $\varepsilon \to 0$, where $A_0(u^t) := \arg\min\{\|\xi\|_\varepsilon : \xi \in \partial f(u^t)\}$. If $X^*$ is uniformly convex, then $A_\varepsilon(u^t) \rightharpoonup A_0(u^t)$ as $\varepsilon \to 0$.
where $u^i_\varepsilon := \arg\min_{v \in X} \left\{ \frac{1}{p \varepsilon^{p-1}} \|u^i - v\|^p + f(v) \right\}$, $i = 0, 1$. Therefore, the inequality (2.6) becomes an equality that implies

$$\frac{1}{p \varepsilon^{p-1}} \|t(u^0_\varepsilon - u^0) + (1 - t)(u^1_\varepsilon - u^1)\|^p = \frac{1}{p \varepsilon^{p-1}} \|u^i_\varepsilon - (tu^0_\varepsilon + (1 - t)u^1_\varepsilon)\|^p$$

Then, the strict convexity of the norm $\| \cdot \|$ implies $u^0_\varepsilon - u^0_\varepsilon = u^1 - u^1_\varepsilon$ and the strict convexity of $f$ implies $u^0_\varepsilon = u^1_\varepsilon$ whence $u^0 = u^1$ and the strict convexity of $f_\varepsilon$.

The strict convexity of the norm also implies that the resolvent operator $J_\varepsilon(u) := \arg\min_{v \in X} \left\{ \frac{1}{p \varepsilon^{p-1}} \|u - v\|^p + f(v) \right\}$ is single-valued for every $u \in X$ and satisfies the inclusion (2.4) by [EkT99, Proposition 5.6, p. 26]. We define $A_\varepsilon(u) := -F_X^p \left( \frac{u - u_\varepsilon}{\varepsilon} \right)$, and note that from the characterization (2.3) of the $p$-duality map, there holds

$$f_\varepsilon(u) = \frac{\varepsilon}{p} \left\| \frac{u_\varepsilon - u}{\varepsilon} \right\|^p + f(u_\varepsilon)$$

If we show that the operator $A_\varepsilon$ is the Gâteaux derivative of $f_\varepsilon$, it follows. First, Akagi & Melchionna [AkM18, Lemma 19] have shown that the operator $A_\varepsilon : X \to X^*$ is demicontinuous, i.e., for all sequences $u_n \to u$ in $X$ as $n \to \infty$, there holds $A_\varepsilon(u_n) \rightharpoonup A_\varepsilon(u)$ in $X^*$ as $n \to \infty$. Second, we show that $A_\varepsilon(u)$ belongs to the subdifferential $\partial f_\varepsilon(u)$ for every $u \in X$. Let $u, v \in X$ and $u_\varepsilon = J_\varepsilon(u), v_\varepsilon = J_\varepsilon(v)$. Then, in view of (2.3) and the fact that $A_\varepsilon(u) = -F_X^p \left( \frac{u - u_\varepsilon}{\varepsilon} \right) \in \partial f(u_\varepsilon)$, we find

$$f_\varepsilon(u) - f_\varepsilon(v) = \frac{\varepsilon}{p} \left\| \frac{u_\varepsilon - u}{\varepsilon} \right\|^p + f(u_\varepsilon) - f(v_\varepsilon)$$

$$= \frac{\varepsilon}{p} \left\| \frac{u_\varepsilon - u}{\varepsilon} - \frac{\varepsilon}{p} \frac{v_\varepsilon - v}{\varepsilon} \right\|^p - \left\langle F_X^p \left( \frac{u_\varepsilon - u}{\varepsilon} \right), u_\varepsilon - v_\varepsilon \right\rangle$$

$$= \frac{\varepsilon}{p} \left\| \frac{u_\varepsilon - u}{\varepsilon} - \frac{\varepsilon}{p} \frac{v_\varepsilon - v}{\varepsilon} \right\|^p - \left\langle F_X^p \left( \frac{u_\varepsilon - u}{\varepsilon} \right), u_\varepsilon - u \right\rangle$$

$$- \left\langle F_X^p \left( \frac{u_\varepsilon - u}{\varepsilon} \right), v - v_\varepsilon \right\rangle - \left\langle F_X^p \left( \frac{u_\varepsilon - u}{\varepsilon} \right), v - v_\varepsilon \right\rangle$$

$$= \frac{\varepsilon}{p} \left\| \frac{u_\varepsilon - u}{\varepsilon} - \frac{\varepsilon}{p} \frac{v_\varepsilon - v}{\varepsilon} \right\|^p - \left\langle F_X^p \left( \frac{u_\varepsilon - u}{\varepsilon} \right), v - v_\varepsilon \right\rangle$$

$$= - \left\langle F_X^p \left( \frac{u_\varepsilon - u}{\varepsilon} \right), v - v_\varepsilon \right\rangle$$

for all $v \in X$.
whence \( A_\varepsilon(u) \in \partial f_\varepsilon(u) \). Subtracting each side of (2.7) by \( \langle A_\varepsilon(v), u - v \rangle \), we obtain

\[
0 \leq f_\varepsilon(u) - f_\varepsilon(v) - \langle A_\varepsilon(v), u - v \rangle \leq \langle A_\varepsilon(u) - A_\varepsilon(v), u - v \rangle \tag{2.8}
\]

for all \( \varepsilon > 0 \) and \( u, v \in X \). Choosing \( u = v + tw \), where \( t > 0 \) and \( w \in X \), and dividing (2.8) by \( t \), we obtain

\[
\lim_{t \searrow 0} \frac{f_\varepsilon(v + tw) - f_\varepsilon(v)}{t} = \langle A_\varepsilon(v), w \rangle \quad \text{for all } w \in X,
\]

where we used the demicontinuity of \( A_\varepsilon \). Hence, the functional \( f_\varepsilon \) is Gâteaux differentiable with derivative \( A_\varepsilon \). Adapting the proof of [BaP86, Proposition 1.146, p. 57], we show that \( A_\varepsilon \) is continuous, provided that \( X^* \) is a uniformly convex space: let \( u_n \to u \) and \( u_n^* = J_\varepsilon(u_n) \). Then, by the demicontinuity of \( A_\varepsilon \), there holds

\[
-A_\varepsilon(u_n) = F_X^\varepsilon \left( \frac{u_n^\varepsilon - u_n}{\varepsilon} \right) \rightharpoonup F_X^\varepsilon \left( \frac{u^\varepsilon - u}{\varepsilon} \right) = -A_\varepsilon(u) \quad \text{in } X^*,
\]

as \( n \to \infty \). Then, from (2.4) as well as the monotonicity of the duality mapping and \( \partial f \), it follows that

\[
0 \leq \left( \left\| \frac{u_n^\varepsilon - u_n}{\varepsilon} \right\|^{p-1} - \left\| \frac{u_m^\varepsilon - u_m}{\varepsilon} \right\|^{p-1} \right) \left( \| u_n^\varepsilon - u_n \| - \| u_m^\varepsilon - u_m \| \right)
\]

\[
\leq \left( F_X^\varepsilon \left( \frac{u_n^\varepsilon - u_n}{\varepsilon} \right) - F_X^\varepsilon \left( \frac{u_m^\varepsilon - u_m}{\varepsilon} \right), w_n^\varepsilon - w_m^\varepsilon - (u_m^\varepsilon - u_m) \right) \leq C \| u_n - u_m \|
\]

where in the last step we used the fact that \( A_\varepsilon \) is demicontinuous and therefore a bounded operator. By the convergence of \( (u_n)_{n \in \mathbb{N}} \), we infer that \( (u_n^\varepsilon - u_n)_{n \in \mathbb{N}} \) is convergent in the norm. Since \( \| u \|^p = \| F_X^\varepsilon(u) \|^{p^*} \), the sequence \( (F_X^\varepsilon \left( \frac{u_n^\varepsilon - u_n}{\varepsilon} \right))_{n \in \mathbb{N}} \) also converges in the norm. Since \( X^* \) is uniformly convex, norm convergence and weak convergence imply strong convergence, see, e.g., [BrÉ11, Proposition 3.32, p. 78], and thus the continuity of \( A_\varepsilon \).

We prove now the assertion ii). The chain of inequalities \( f(u_\varepsilon) \leq f_\varepsilon(u) \leq f(u) \) follows immediately from the definition of the \( p \)-MOREAU–YOSIDA regularization. To conclude ii), it remains to show that the mapping \( \varepsilon \mapsto f_\varepsilon(u) \) is monotonically decreasing on \((0, \infty)\) for every fixed \( u \in X \). Let \( u \in X \) and \( 0 < \varepsilon_2 < \varepsilon_1 \). Then, by the definition of a minimizer

\[
f_{\varepsilon_1}(u) = \frac{\varepsilon_2}{p} \left\| \frac{u_{\varepsilon_2} - u}{\varepsilon_2} \right\|^p + f(u_{\varepsilon_2})
\]

\[
\leq \frac{\varepsilon_2}{p} \left\| \frac{u_{\varepsilon_1} - u}{\varepsilon_2} \right\|^p + f(u_{\varepsilon_1})
\]

\[
= \left( \frac{1}{p \varepsilon_2^{p-1}} - \frac{1}{p \varepsilon_1^{p-1}} \right) \| u_{\varepsilon_1} - u \|^p + \frac{\varepsilon_1}{p} \left\| \frac{u_{\varepsilon_1} - u}{\varepsilon_1} \right\|^p + f(u_{\varepsilon_1})
\]

\[
= \left( \frac{1}{p \varepsilon_2^{p-1}} - \frac{1}{p \varepsilon_1^{p-1}} \right) \| u_{\varepsilon_1} - u \|^p + f_{\varepsilon_1}(u) \tag{2.9}
\]

\[
\leq f_{\varepsilon_1}(u).
\]

\(^4\)The normed space \( X \) is called uniformly convex if for each \( \varepsilon \in (0, 2) \) there exists \( \delta(\varepsilon) > 0 \), for which \( \| x \| \leq 1 \| y \| \leq 1 \) and \( \| x - y \| \geq \varepsilon \) imply \( \left\| \frac{x + y}{2} \right\| \leq 1 - \delta(\varepsilon) \).
Now, we aim to show (2.5). First, switching the roles of $\varepsilon_1$ and $\varepsilon_2$ in the inequality (2.9) and dividing both sides by $\varepsilon_1 - \varepsilon_2 > 0$, we obtain the chain of inequalities

\[
\begin{align*}
\frac{1}{p(\varepsilon_2 \varepsilon_1)^{p-1}} \left( \frac{\varepsilon_1^{p-1} - \varepsilon_2^{p-1}}{\varepsilon_2 - \varepsilon_1} \right) \|u_{\varepsilon_2} - u\|^p \\
\leq \frac{f_{\varepsilon_1}(u) - f_{\varepsilon_2}(u)}{\varepsilon_1 - \varepsilon_2} \\
\leq \frac{1}{p(\varepsilon_2 \varepsilon_1)^{p-1}} \left( \frac{\varepsilon_1^{p-1} - \varepsilon_2^{p-1}}{\varepsilon_2 - \varepsilon_1} \right) \|u_{\varepsilon_1} - u\|^p
\end{align*}
\]  
(2.10)

for all $0 < \varepsilon_2 < \varepsilon_1$. Then, (2.10) implies

\[\|u_{\varepsilon_2} - u\| \leq \|u_{\varepsilon_1} - u\| \quad \text{for all} \ 0 < \varepsilon_2 < \varepsilon_1.\]  
(2.11)

Second, since the real-valued mapping $\varepsilon \mapsto f_{\varepsilon}(u)$ is monotone for every fixed $u \in X$, it is, by LERESQUE’S differentiation theorem for monotone functions$^5$, almost everywhere differentiable and there holds

\[
\frac{d f_{\varepsilon}(u)}{d \varepsilon^+} \leq \frac{d f_{\varepsilon}(u)}{d \varepsilon^-} \quad \text{for all} \ \varepsilon > 0, u \in X,
\]

where $\frac{d f_{\varepsilon}(u)}{d \varepsilon^+}$ and $\frac{d f_{\varepsilon}(u)}{d \varepsilon^-}$ denote the right and left derivative of $\varepsilon \mapsto f_{\varepsilon}(u)$ in $\varepsilon = \varepsilon$, respectively. Let $\varepsilon > 0$ and $h > 0$ be sufficiently small. Then, choosing $\varepsilon_1 = \varepsilon + h$ and $\varepsilon_2 = \varepsilon$ in the first inequality as well as $\varepsilon_1 = \varepsilon$ and $\varepsilon_2 = \varepsilon - h$ in the second inequality of (2.10) yields

\[
\begin{align*}
\frac{1}{p((\varepsilon + h)\varepsilon)^{p-1}} \left( \frac{(\varepsilon + h)^{p-1} - \varepsilon^{p-1}}{h} \right) \|u_{\varepsilon} - u\|^p \\
\leq \frac{f_{\varepsilon + h}(u) - f_{\varepsilon}(u)}{h}
\end{align*}
\]  
(2.12)

and

\[
\begin{align*}
\frac{1}{p((\varepsilon - h)\varepsilon)^{p-1}} \left( \frac{\varepsilon^{p-1} - (\varepsilon - h)^{p-1}}{h} \right) \|u_{\varepsilon - h} - u\|^p \\
\leq \frac{f_{\varepsilon}(u) - f_{\varepsilon-h}(u)}{h}
\end{align*}
\]  
(2.13)

respectively, where we employed inequality (2.11). Finally, letting $h \to 0$ in (2.12) and (2.13) yields

\[
\frac{d f_{\varepsilon}}{d \varepsilon} = -\frac{1}{p \varepsilon^{p-1}} \|u_{\varepsilon} - u\|^p \quad \text{for all} \ \varepsilon > 0.
\]

We continue with showing assertion $iii)$. Let $u \in \text{dom}(f)$, then the first inequality of (2.10) implies

\[
\|u_{\varepsilon_2} - u\|^p \leq \left( \frac{p(\varepsilon_2 \varepsilon_1)^{p-1}}{\varepsilon_1^{p-1} - \varepsilon_2^{p-1}} \right) (f_{\varepsilon_2}(u) - f_{\varepsilon_1}(u))
\]  
(2.14)

\[
\leq \left( \frac{p(\varepsilon_2 \varepsilon_1)^{p-1}}{\varepsilon_1^{p-1} - \varepsilon_2^{p-1}} \right) (f(u) - f_{\varepsilon_1}(u))
\]

$^5$See, e.g., ELSTROD [Els05, Satz 4.5, p. 299].
for all $0 < \varepsilon_2 < \varepsilon_1$. Thus, we obtain $\lim_{\varepsilon_2 \to 0} \|u_{\varepsilon_2} - u\| = 0$. Taking into account the latter convergence and the lower semicontinuity of $f$, assertion ii) yields
\[
 f(u) \leq \liminf_{\varepsilon \to 0} f(u_\varepsilon) \\
\leq \liminf_{\varepsilon \to 0} f_\varepsilon(u) \\
\leq \limsup_{\varepsilon \to 0} f_\varepsilon(u) \leq f(u) \quad \text{for all } u \in \text{dom}(f).
\]
If $u \in X \setminus \text{dom}(f)$, we assume that there exists a sequence $(\varepsilon_n)_{n \in \mathbb{N}} \subset (0, \infty)$ with $\varepsilon_n \to 0$ as $n \to \infty$ such that $f_\varepsilon(u) \leq C$ for all $n \in \mathbb{N}$ for a constant $C > 0$. However, inequality (2.14) yields $\lim_{n \to \infty} \|u_{\varepsilon_n} - u\| = 0$, and we obtain $f(u) \leq \liminf f_\varepsilon(u) \leq C$, which is a contradiction to $u \in X \setminus \text{dom}(f)$. The assertion v) follows the exact same lines as the proof of [BaP86, Proposition 1.146 iv], p. 57.

The theorem showed us that the MOREAU–YOSIDA regularization has indeed a regularizing effect. In fact, in view of assertion iv) and (2.5), one can interpret the MOREAU–YOSIDA regularization as a regularization process described by the following HAMILTON–JACOBI equation supplemented with an initial condition
\[
\begin{cases}
\frac{\partial}{\partial t} u(t, x) + \frac{1}{p} \|d_x u(t, x)\|^p = 0, & x \in X, t > 0 \\
\ u(0+, x) = f(x), & x \in X,
\end{cases}
\tag{2.15}
\]
where a solution $u : [0, \infty) \times X \to \mathbb{R}$ is given by the so-called LAX–OLENIK formula
\[
u(t, x) = f_t(x) = \inf_{y \in X} \left\{ \frac{t}{p} \left\| \frac{x - y}{t} \right\|^p + f(y) \right\},
\]
see, e.g., LIONS [Lio81].

Moreover, we have seen to what extent these regularization and approximating properties depend on the properties of $X^\ast$. This, as previously mentioned, becomes clearer when $X = H$ is a HILBERT space. In this case, the MOREAU–YOSIDA regularization is even FRÉCHET differentiable and has a LIPSCHITZ continuous derivative with a LIPSCHITZ constant equal to the reciprocal of the regularization parameter $\varepsilon$, see, e.g., BARBU & PRECUŞANU [BaP86, Corollary 2.59, p. 99]. Thanks to these nice properties of the regularization and its derivative that are only available on a HILBERT space, the MOREAU–YOSIDA regularization is often applied to HILBERT spaces, see, e.g., BAUSCHKE & COMBETTES [BaC11] for a detailed treatise on HILBERT spaces. The MOREAU–YOSIDA regularization is related to the so-called YOSIDA approximation, which, for a given operator $A$ and $\varepsilon > 0$, refers to the operator $A_{\varepsilon} = \varepsilon^{-1}(I - S_{\varepsilon})$, which is approximative to $A$, where $S_{\varepsilon} = (I + \varepsilon A)^{-1}$. The YOSIDA approximation is successfully employed in the theory of semigroups in order to generate strongly continuous semigroups as in the eminent HILLE–YOSIDA theorem [Hil52, Yos48], the nonlinear counterpart [Dor69, CrL71], or in the theory of maximal monotone operators in BRÉZIS [Bré73].

In the next theorem, we want to show that the $p$-MOREAU–YOSIDA regularization preserves both the superlinearity or $p$–growth of a function and the MOSCO-convergence of a sequence of functions $(f_\varepsilon)_{\varepsilon \in \mathbb{N}}$ for fixed $\varepsilon > 0$. Furthermore, we show that the sequence $(f_{\varepsilon_n})_{n \in \mathbb{N}}$ converges to $f$ as $\varepsilon_n \searrow 0$ in the sense of MOSCO-convergence. Finally, we give an explicit formula for the LEGENDRE–FENCHEL transformation of the $p$-MOREAU–YOSIDA regularization of a function.
Theorem 2.3. Let $f_n : X \to (-\infty, +\infty]$ be a proper, lower semicontinuous and convex functional for each $n \in \mathbb{N}$ such that

i) for all $N > 0$, there holds
\[
\lim_{\|\xi\| \to +\infty} \frac{1}{\|\xi\|_*} \left( \inf_{n \leq N} f_n^*(\xi) \right) = \infty, \quad \lim_{\|v\| \to +\infty} \frac{1}{\|v\|} \left( \inf_{n \leq N} f_n(v) \right) = \infty.
\]

ii) the sequence $f_n$ converges to $f$ in the sense of MOSCO ($f_n \rightharpoonup^M f$), i.e., for all $u \in X$
\[
\begin{cases}
(a) \quad f(u) \leq \liminf_{n \to \infty} f_n(u_n) \quad \text{for all } u_n \rightharpoonup u \text{ in } X, \\
(b) \quad \exists \hat{u}_n \to u \text{ in } X \text{ such that } f(\hat{u}_n) \geq \limsup_{n \to \infty} f_n(\hat{u}_n).
\end{cases}
\]

Furthermore, let $\varepsilon \in (0, 1]$ and $p > 1$. Then, the $p$-MOREAU–YOSIDA regularization $f_n^\varepsilon$ satisfies i) and ii) and the convex conjugate of $f_n^\varepsilon$ is given by
\[
f_n^\varepsilon(\xi) = \frac{\varepsilon}{p^*} \|\xi\|_p^p + f_n^\varepsilon(\xi) \quad \text{for all } \xi \in X^*, n \in \mathbb{N},
\]
where $p^* > 1$ is the conjugate exponent of $p$. Moreover, $f_n^\varepsilon$ and $f_n^{\varepsilon,*}$ are uniformly superlinear with respect to $\varepsilon > 0$. Finally, for all sequences $(\varepsilon_n)_{n \in \mathbb{N}} \subset (0, 1]$ with $\varepsilon_n \to 0$ as $n \to \infty$, there holds $f_n^\varepsilon \rightharpoonup^M f$.

Proof. First, for each $\varepsilon > 0$ and $n \in \mathbb{N}$, the regularization $f_n^\varepsilon$ is a proper, lower semicontinuous, and convex functional by Lemma 2.1. The formula (2.16) follows from the calculations
\[
f_n^{\varepsilon,*}(\xi) = \sup_{v \in X} \{ \langle \xi, v \rangle - f_n^\varepsilon(v) \}
\]
\[
= \sup_{v \in X} \left\{ \langle \xi, v \rangle - \inf_{w \in X} \left\{ \frac{\varepsilon}{p} \left\| \frac{v - w}{\varepsilon} \right\|^p + f_n(w) \right\} \right\}
\]
\[
= \sup_{v \in X} \sup_{w \in X} \left\{ \langle \xi, v \rangle - \frac{\varepsilon}{p} \left\| \frac{v - w}{\varepsilon} \right\|^p - f_n(w) \right\}
\]
\[
= \sup_{v \in X} \sup_{w \in X} \left\{ \langle \xi, v \rangle - \frac{\varepsilon}{p} \left\| \frac{v - w}{\varepsilon} \right\|^p - f_n(w) \right\}
\]
\[
= \sup_{v \in X} \left\{ \sup_{w \in X} \left\{ \langle \xi, v - w \rangle - \frac{\varepsilon}{p} \left\| \frac{v - w}{\varepsilon} \right\|^p \right\} + \langle \xi, w \rangle - f_n(w) \right\}
\]
\[
= \sup_{v \in X} \left\{ \sup_{w \in X} \left\{ \langle \xi, v - w \rangle - \frac{\varepsilon}{p} \left\| \frac{v - w}{\varepsilon} \right\|^p \right\} + \langle \xi, w \rangle - f_n(w) \right\}
\]
\[
= \sup_{v \in X} \left\{ \frac{\varepsilon}{p^*} \|\xi\|_p^p + \langle \xi, w \rangle - f_n(w) \right\} = \frac{\varepsilon}{p^*} \|\xi\|_{p^*}^p + f_n^\varepsilon(\xi)
\]
for all $\xi \in X^*$ and $u \in D$, where we have used the fact $(\frac{1}{\varepsilon p^*} \| \cdot \|^p)^* = \frac{\varepsilon}{p^*} \| \cdot \|_{p^*}^{p^*}$. The expression (2.16) also shows the superlinearity of $f_n^{\varepsilon,*}$ uniformly in $\varepsilon$. We proceed by showing the superlinearity of $f_n^\varepsilon$. To do so, we note that the superlinearity of $f_n$ equivalently says that for all $N \in \mathbb{N}$ and $M > 0$, there exists a positive real number $K > 0$ such that
\[
f_n(v) \geq M \|v\|.
\]
for all \( n \geq N \) and all \( v \in X \) with \( \|v\| \geq K \). The idea is to show that for the regularization \( f_n^\varepsilon \), there exists for all \( \bar{N} \in \mathbb{N} \) and \( \bar{M} > 0 \) a positive real number \( \bar{K} > 0 \) independent of the parameter \( \varepsilon > 0 \) such that (2.17) is satisfied. So, let \( \bar{N} \in \mathbb{N} \) and \( \bar{M} > 0 \), then, for \( N = \bar{N} \) and \( M = 2\bar{M} \), there exists \( K > 0 \) such that (2.17) holds. By YOUNG’s inequality and the triangle inequality, we obtain

\[
 f_n^\varepsilon(v) = \inf_{\tilde{v} \in X} \left\{ \frac{1}{p_{\varepsilon^{-1}}} \|v - \tilde{v}\|^p + f_n(\tilde{v}) \right\}
 = \min \left\{ \inf_{\tilde{v} \in X} \left\{ \frac{1}{p_{\varepsilon^{-1}}} \|v - \tilde{v}\|^p + f_n(\tilde{v}) \right\}, \inf_{\tilde{v} \in X} \left\{ \frac{1}{p_{\varepsilon^{-1}}} \|v - \tilde{v}\|^p + f_n(\tilde{v}) \right\} \right\}
 \geq \min \inf_{\tilde{v} \in X} \left\{ \frac{1}{p_{\varepsilon^{-1}}} \|v - \tilde{v}\|^p + M\|\tilde{v}\| \right\}, \inf_{\tilde{v} \in X} \left\{ \frac{1}{p_{\varepsilon^{-1}}} \|v - \tilde{v}\|^p \right\}
 \geq \min \left\{ M\|v\| - M\varepsilon \right\}, \left\{ M\|v\| - K - M\varepsilon \right\}
 = M\|v\| - K - M\varepsilon
 \geq \frac{M}{2}\|v\| = \bar{M}\|v\|
\]

for all \( v \in X \) with \( \|v\| \geq \bar{K} := 2 \left( K + \frac{M\varepsilon}{p_{\varepsilon^{-1}}} \right) \) and \( \varepsilon \in (0, 1) \). This implies the superlinearity of \( f_n^\varepsilon \) uniformly in \( \varepsilon > 0 \), which, in turn implies the superlinearity for a fixed \( \varepsilon > 0 \). We continue by showing that \( f_n^\varepsilon \) is continuous in the sense of MOSCO-convergence. In fact, we show that for a fixed \( \varepsilon > 0 \), the regularization satisfies a stronger version of MOSCO-convergence, meaning that there not only exists a recovery sequence, but that every sequence converging against the same limit is a recovery sequence. Let \( (v_n)_{n \in \mathbb{N}} \subset X \) be a weakly convergent sequence with weak limit \( v \in X \). Now, let \( (n_k)_{k \in \mathbb{N}} \) be a subsequence such that

\[
 \liminf_{n \to \infty} f_n^\varepsilon(v_n) = \lim_{k \to \infty} f_{n_k}^\varepsilon(v_{n_k}).
\]

For each \( k \in \mathbb{N} \), we denote by \( v_k^\varepsilon \) the unique minimizer of \( v \mapsto \frac{1}{p_{\varepsilon^{-1}}} \|v - v_n\|^p + f_n(v) \) and note that thanks to the estimate

\[
 \frac{1}{p_{\varepsilon^{-1}}} \|v_n - v_k^\varepsilon\|^p \leq f_{n_k}^\varepsilon(v_{n_k}) \leq \frac{1}{p_{\varepsilon^{-1}}} \|v_n\|^p, \tag{2.18}
\]

the corresponding sequence of minimizers \( (v_k^\varepsilon)_{k \in \mathbb{N}} \) is bounded. Therefore, there exists a subsequence (labelled as before) which is weakly convergent to an element \( \tilde{v}_\varepsilon \in X \). Then, by the MOSCO-convergence \( f_n \xrightarrow{M} f \), we have

\[
 f^\varepsilon(v) \leq \frac{1}{p_{\varepsilon^{-1}}} \|v - \tilde{v}_\varepsilon\|^p + f(\tilde{v}_\varepsilon)
 \leq \liminf_{k \to \infty} \left\{ \frac{1}{p_{\varepsilon^{-1}}} \|v_n - v_k^\varepsilon\|^p + f_n(v_k^\varepsilon) \right\}
 = \lim_{k \to \infty} f_{n_k}^\varepsilon(v_{n_k}) = \liminf_{n \to \infty} f_n^\varepsilon(v_n).
\]
Now, let \( v \in X \) be arbitrary and \((v_n)_{n \in \mathbb{N}} \subset X\) any strongly convergent sequence \( v_n \to v \) as \( n \to \infty \). We extract an arbitrary subsequence \((n_k)_{k \in \mathbb{N}}\), and to each \( k \in \mathbb{N} \), we denote the minimizers of \( v \mapsto \frac{1}{p \epsilon} ||v - v_n||^p + f_{n_k}(v) \) again by \( \hat{v}_k \in X \). By \( \hat{v}_k \in X \), we denote the weak limit of a further subsequence of the very same sequence which we labelled as before. Once more, by \textit{ii}), for the minimizer \( v_\epsilon \) of \( f^\epsilon(v) \), there exists a strongly convergent recovery sequence \((\hat{v}_k)_{k \in \mathbb{N}} \subset X \) such that \( \hat{v}_k \to v_\epsilon \) and \( \lim_{k \to \infty} f_{n_k}(\hat{v}_k) = f(v_\epsilon) \). It follows

\[
\begin{align*}
f^\epsilon(v) & \leq \frac{1}{p \epsilon} ||v - \hat{v}_\epsilon||^p + f(\hat{v}_\epsilon) \\
& \leq \liminf_{k \to \infty} \left\{ \frac{1}{p \epsilon} ||v_{n_k} - \hat{v}_k||^p + \Psi_{n_k}(v_\epsilon^k) \right\} \\
& = \liminf_{k \to \infty} f^\epsilon_{n_k}(v_{n_k}) \\
& \leq \limsup_{k \to \infty} f^\epsilon_{n_k}(v_{n_k}) \\
& \leq \limsup_{k \to \infty} \left\{ \frac{1}{p \epsilon} ||v_{n_k} - \hat{v}_k||^p + f_{n_k}(\hat{v}_k) \right\} \\
& = \lim_{k \to \infty} \left\{ \frac{1}{p \epsilon} ||v_{n_k} - \hat{v}_k||^p + f_{n_k}(\hat{v}_k) \right\} \\
& = \frac{1}{p \epsilon} ||v - v_\epsilon||^p + f(v_\epsilon) = f^\epsilon(v).
\end{align*}
\]

Therefore, every subsequence \((n_k)_{k \in \mathbb{N}}\) contains a further subsequence \((n_{k_l})_{l \in \mathbb{N}}\) such that \( \lim_{l \to \infty} f^\epsilon_{n_{k_l}}(v_{n_{k_l}}) = f^\epsilon(v) \). By the subsequence principle, the convergence of the whole sequence follows. In particular, this shows \( v_\epsilon = \hat{v}_\epsilon \).

Finally, we show that the MOSCO-convergence \( f^\epsilon_n \xrightarrow{\text{M}} f \) for all sequences of regularization parameters \((\epsilon_n)_{n \in \mathbb{N}} \subset (0, 1] \) with \( \epsilon_n \to 0 \) as \( n \to \infty \). As before, let the sequence \((v_n)_{n \in \mathbb{N}} \subset X \) be given such that \( v_n \to v \in X \) as \( n \to \infty \), and let \((n_k)_{k \in \mathbb{N}}\) be a subsequence such that

\[
\liminf_{n \to \infty} f^\epsilon_n(v_n) = \lim_{k \to \infty} f^\epsilon_{n_k}(v_{n_k}).
\]

By \( \hat{v}_k \in X \), \( k \in \mathbb{N} \), we denote again the minimizer of \( f^\epsilon_{n_k}(v_{n_k}) \). Due to the same estimate as (2.18) for \((\hat{v}_k)_{k \in \mathbb{N}}\), the sequence of minimizers is bounded and therefore sequentially compact with respect to the weak topology. So, after extracting a subsequence (labelled as before), we obtain a weak limit \( \tilde{v} \in X \) such that \( \tilde{v}_k \to \tilde{v} \) as \( n \to \infty \). Now, we consider two cases:

\[i) \quad \frac{1}{p \epsilon_{n_k}} ||v_{n_k} - \hat{v}_k||^p \leq C \text{ for a constant } C > 0,\]

\[ii) \quad \frac{1}{p \epsilon_{n_k}} ||v_{n_k} - \hat{v}_k||^p \to \infty \text{ as } k \to \infty \text{ after possibly extracting a further subsequence.}\]

Ad i). We immediately find \( v = \tilde{v} \) and therefore \( \tilde{v}_k \to v \) as \( k \to \infty \). By the continuity of
f in the sense of Mosco-convergence, it follows
\[
 f(v) \leq \liminf_{k \to \infty} f_n(\tilde{v}_k) \\
 \leq \liminf_{k \to \infty} \left\{ \frac{1}{p-1} \|v_{n_k} - \tilde{v}_k\|^p + f_n(\tilde{v}_k) \right\} \\
 = \liminf_{k \to \infty} f_{n_k}(v_{n_k}) \\
 = \liminf_{n \to \infty} f_n(v_n).
\]

Ad ii). We obtain
\[
 f(v) \leq \lim_{k \to \infty} \left( \frac{1}{p-1} \|v_{n_k} - \tilde{v}_k\|^p \right) \\
 \leq \lim_{k \to \infty} \left\{ \frac{1}{p-1} \|v_{n_k} - \tilde{v}_k\|^p + f_n(\tilde{v}_k) \right\} \\
 = \liminf_{k \to \infty} f_{n_k}(v_{n_k}) \\
 = \liminf_{n \to \infty} f_n(v_n).
\]

It remains to show the existence of a recovery sequence. Let \( v \in X \) be arbitrarily chosen. Then, there exists a recovery sequence \( (v_n)_{n \in \mathbb{N}} \subset X \) for \( f \) with \( v_n \to v \) as \( n \to \infty \) such that \( \lim_{n \to \infty} f_n(v_n) = f(v) \). Proceeding as before, we take an arbitrary subsequence \( (n_k)_{k \in \mathbb{N}} \) and denote by \( (\tilde{v}_k)_{k \in \mathbb{N}} \subset X \) the minimizing sequence of \( f_{n_k}(v_{n_k}) \). Then, we consider again the two cases i) and ii).

Ad i). Since the recovery sequence is strongly convergent, it follows that \( (\tilde{v}_k)_{k \in \mathbb{N}} \) is also strongly convergent with the same limit \( v \in X \). We obtain
\[
 f(v) \leq \liminf_{k \to \infty} f_n(\tilde{v}_k) \\
 \leq \liminf_{k \to \infty} \left\{ \frac{1}{p-1} \|v_{n_k} - \tilde{v}_k\|^p + f_n(\tilde{v}_k) \right\} \\
 = \liminf_{k \to \infty} f_{n_k}(v_{n_k}) \\
 \leq \limsup_{k \to \infty} f_{n_k}(v_{n_k}) \\
 \leq \limsup_{k \to \infty} f_n(v_{n_k}) \\
 = \lim_{k \to \infty} f_n(v_{n_k}) = f(v),
\]
which by the same argument as before implies the convergence of the full sequence, i.e., \( \lim_{n \to \infty} f_n(v_n) = f(v) \).

Ad ii). Due to \( f_n(v_n) \leq f_n(v_n) \), \( n \in \mathbb{N} \), and the convergence of the right-hand side, this case cannot occur, which completes the proof.

As mentioned above, the \( p \)-MOREAU–YOSIDA regularization can be viewed as a regularization process described by the HAMILTON–JACOBI equation (2.15). However, introducing the \( p \)-MOREAU–YOSIDA regularization as a solution to the CAUCHY problem (2.15)
does not seem 'natural'. Interestingly, the regularization arises naturally when one deals with (generalized) gradient flow equations. To demonstrate this more clearly, we consider the generalized gradient flow

\[-|u'(t)|^{p-2}u'(t) \in \partial E(u(t)), \quad t > 0,\]

of a functional $E : H \to (-\infty, +\infty]$ on a Hilbert space $H$. Discretizing the equation by the implicit Euler scheme leads to

\[-\frac{U^n_\tau - U^{n-1}_\tau}{\tau} + p\frac{U^n_\tau - U^{n-1}_\tau}{\tau} \in \partial E(U^n_\tau), \quad n = 1, 2, \ldots, N,\]

where, starting with $U^n_0 = u_0 \in \text{dom}(E)$, the values $U^n_\tau, n = 1, \ldots, N$, can under certain conditions be obtained by the variational approximation scheme

$$U^n_\tau \in J_\tau(U^{n-1}) := \arg\min_{v \in H} \left\{ \frac{\tau}{p} \left| v - U^{n-1}_\tau \right|^p + E(v) \right\}, \quad n = 1, 2, \ldots, N. \tag{2.19}$$

Here, obviously the $p$-Moreau–Yosida regularization occurs naturally after discretizing the equation in time. The approximative values $U^n_\tau \in H$ are then defined by the $p$-Moreau–Yosida regularization $E_\tau$ where the regularization parameter is given by the step size $\tau$ of the time-discretization. It is also worth mentioning that the Moreau–Yosida regularization does not only regularize a function itself, but the associated resolvent operator $J_\tau(u)$ regularizes in a certain sense its arguments $u \in H$: the values $U^n_\tau \in \text{dom}(\partial E)$, which are achieved in the minimization scheme, are not only contained in the domain of the functional $E$, but also in the domain of the subdifferential $\partial E$. The latter is also referred to as the regularizing or smoothing effect of the gradient flow equation, which means that for a given initial datum $u_0 \in \text{dom}(E)$ (or in some cases even $u_0 \in \overline{\text{dom}(E)}$) the solution does not only belong to the domain of $E$ but also to the domain of its subdifferential $\partial E$ for an infinitesimal larger time step, i.e., $u(t) \in \text{dom}(\partial E)$ for every $t > 0$. It is well-known that for $p = 2$ and when $E : H \to (-\infty, +\infty]$ is a proper, lower semicontinuous, and convex functional, the subdifferential operator $\partial E$ is an infinitesimal generator of a $C_0$-semigroup such that $S(t)u_0 = u(t)$ is the unique solution to the Cauchy problem

$$\begin{cases}
u'(t) & \in -\partial E(u(t)), \quad t > 0, \\u(0) & = u_0 \in \overline{\text{dom}(E)}
\end{cases}$$

and which fulfills $S(t)u_0 = \lim_{n \to \infty} J^t_{1/n}(u_0)$, where $J^t_{1/n}$ denotes again the resolvent operator given by (2.19), see, e.g., [Bré73, Bar76]. This property even holds true in a complete metric space under slightly weaker assumptions on the functional $E$, see AMBROSIO et al. [AGS08] for a detailed discussion.

References

[AGS08] L. AMBROSIO, N. GIGLI, and G. SAVARÉ. Gradient Flows in Metric Spaces and in the Space of Probability Measures. Lectures in Mathematics ETH Zürich. Birkhäuser, Basel, 2nd edition, 2008.
[AkM18] G. Akagi and S. Melchionna. Elliptic-regularization of nonpotential perturbations of doubly-nonlinear flows of nonconvex energies: a variational approach. *J. Convex Anal.*, 25(3), 861–898, 2018.

[Asp67] E. Asplund. Averaged norms. *Israel J. Math.*, 5, 227–233, 1967.

[AtA93] H. Attouch and D. Azé. Approximation and regularization of arbitrary functions in Hilbert spaces by the Lasry-Lions method. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 10(3), 289–312, 1993.

[Att84] H. Attouch. *Variational Convergence for Functions and Operators*. Applicable Mathematics Series. Pitman, Boston, 1984.

[BaC11] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer-Verlag, New York, 2011.

[Bac20] A. Bacho. Nonsmooth analysis of doubly nonlinear evolution inclusions of second order with non-convex energy functional. *arXiv:1910.10128*, 2020.

[Bac21] A. Bacho. *On the nonsmooth analysis of doubly nonlinear evolution inclusions of first and second order with applications*. Doctoral thesis, Technische Universität Berlin, Berlin, 2021.

[BaP86] V. Barbu and T. Precupanu. *Convexity and Optimization in Banach Spaces*. Springer-Verlag, Dordrecht, 4th edition, 1986.

[Bar76] V. Barbu. *Nonlinear Semigroups and Differential Equations in Banach spaces*. Noordhoff, Leyden, 1976.

[Bar10] V. Barbu. *Nonlinear Differential Equations of Monotone Types in Banach Spaces*. Springer-Verlag, New York, 2010.

[Ber10] P. Bernard. Lasry-Lions regularization and a lemma of Ilmanen. *Rend. Semin. Mat. Univ. Padova*, 124, 221–229, 2010.

[Bré73] H. Brézis. *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*. North-Holland Publishing Co., Amsterdam, 1973.

[Bré11] H. Brézis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer-Verlag, New York, 2011.

[BTZ11] F. Bernard, L. Thibault, and N. Zlateva. Prox-regular sets and epigraphs in uniformly convex Banach spaces: various regularities and other properties. *Trans. Amer. Math. Soc.*, 363(4), 2211–2247, 2011.

[Byn71] W. L. Bynum. Characterizations of uniform convexity. *Pacific J. Math.*, 38, 577–581, 1971.

[Byn76] W. L. Bynum. Weak parallelogram laws for Banach spaces. *Canad. Math. Bull.*, 19(3), 269–275, 1976.

[CrL71] M. G. Crandall and T. M. Liggett. Generation of semi-groups of nonlinear transformations on general Banach spaces. *Amer. J. Math.*, 93, 265–298, 1971.

[Die75] J. Diestel. *Geometry of Banach Spaces-Selected Topics*. Lecture Notes in Mathematics, Vol. 485. Springer-Verlag, Berlin-New York, 1975.

[Dor69] J. R. Dorroh. A nonlinear Hille-Yosida-Phillips theorem. *J. Functional Analysis*, 3, 345–353, 1969.
[EkT99] I. EKELAND and R. TÉMAM. Convex analysis and variational problems, volume 28 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.

[Els05] J. ELSTRODT. Maß- und Integrationstheorie. Springer-Verlag, Berlin, 4th edition, 2005.

[Hil52] E. HILLE. On the generation of semi-groups and the theory of conjugate functions. Kungl. Fysiografiska Sällskapets i Lund Förhandlingar [Proc. Roy. Physiog. Soc. Lund], 21(14), 13, 1952.

[JTZ14] A. JOURANI, L. THIBAULT, and D. ZAGRODNY. Differential properties of the Moreau envelope. J. Funct. Anal., 266(3), 1185–1237, 2014.

[Kie02] B. T. KIEN. The normalized duality mapping and two related characteristic properties of a uniformly convex Banach space. Acta Math. Vietnam., 27(1), 53–67, 2002.

[LaL86] J.-M. LASRY and P.-L. LIONS. A remark on regularization in Hilbert spaces. Israel J. Math., 55(3), 257–266, 1986.

[Lio81] P.-L. LIONS. Solutions généralisées des équations de Hamilton-Jacobi du premier ordre. C. R. Acad. Sci. Paris Sér. I Math., 292(22), 953–956, 1981.

[Mor65] J.-J. MOREAU. Proximité et dualité dans un espace hilbertien. Bull. Soc. Math. France, 93, 273–299, 1965.

[Pen98] J.-P. PENOT. Proximal mappings. J. Approx. Theory, 94(2), 203–221, 1998.

[Str96] T. STRÖMBERG. On regularization in Banach spaces. Ark. Mat., 34(2), 383–406, 1996.

[VaP16] H. VAN NGAI and J.-P. PENOT. Subdifferentiation of regularized functions. Set-Valued Var. Anal., 24(1), 167–189, 2016.

[Yos48] K. YOSIDA. On the differentiability and the representation of one-parameter semi-group of linear operators. J. Math. Soc. Japan, 1, 15–21, 1948.

[Zem91] M. ZEMEK. Strong monotonicity and Lipschitz-continuity of the duality mapping. Acta Univ. Carolin. Math. Phys., 32(2), 61–64, 1991.