Sensitivity analysis of variational inequalities via twice epi-differentiability and proto-differentiability of the proximity operator

Samir Adly, Loïc Bourdin

To cite this version:

Sensitivity analysis of variational inequalities via twice epi-differentiability and proto-differentiability of the proximity operator

Samir Adly[*], Loïc Bourdin[†]

February 25, 2018

Abstract

In this paper we investigate the sensitivity analysis of parameterized nonlinear variational inequalities of second kind in a Hilbert space. The challenge of the present work is to take into account a perturbation on all the data of the problem. This requires special adjustments in the definitions of the generalized first- and second-order differentiations of the involved operators and functions. Precisely, we extend the notions, introduced and thoroughly studied by R.T. Rockafellar, of twice epi-differentiability and proto-differentiability to the case of a parameterized lower semi-continuous convex function and its subdifferential respectively. The link between these two notions is tied to Attouch's theorem and to the new concept, introduced in this paper, of convergent supporting hyperplanes. The previous tools allow us to derive an exact formula of the proto-derivative of the generalized proximity operator associated to a parameterized variational inequality, and deduce the differentiability of the associated solution with respect to the parameter. Furthermore, the derivative is shown to be the solution of a new variational inequality involving semi- and second epi-derivatives of the data. An application is given to parameterized convex optimization problems involving the sum of two convex functions (one of them being smooth). The case of smooth convex optimization problems with inequality constraints is discussed in details. This approach seems to be new in the literature and open several perspectives towards theoretical and computational issues in nonlinear optimization.

Keywords: Sensitivity analysis; variational inequalities; twice epi-differentiability; semi-differentiability; proto-differentiability; proximity operator; Attouch's theorem; constrained convex optimization.

AMS Classification: 49J53; 49J52; 58C20; 49A50; 58C06; 65K10.

Contents

[*]Institut de recherche XLIM. UMR CNRS 7252. Université de Limoges, France. `samir.adly@unilim.fr`
[†]Institut de recherche XLIM. UMR CNRS 7252. Université de Limoges, France. `loic.bourdin@unilim.fr`

1   Introduction

In the whole paper let H be a real Hilbert space and let $\langle \cdot, \cdot \rangle$ (resp. $\| \cdot \|$) be the corresponding scalar product (resp. norm). In the sequel, $\mathbb{R}_+$ denotes the set of all nonnegative real numbers.

1.1   Context of the paper

One of the most famous optimization problem consists in finding a point y of a set $C \subset H$ that minimizes the distance to a given point $x \in H$. This historical problem can be rewritten as

$$y \in \operatorname*{argmin}_{z \in C} \| z - x \|.$$

If C is a nonempty closed convex set, it is well-known that the above problem admits a unique solution, denoted by $y = \operatorname{proj}_C(x)$, called the *metric projection* of x onto C, and characterized by the following *variational inequality*

$$\forall z \in C, \quad \langle y, z - y \rangle \geq \langle x, z - y \rangle.$$

As a first step towards the *sensitivity analysis* of the above variational inequality, one would consider a slight perturbation of the point $x \in H$, replacing it by $x(t) \in H$ where $t \in \mathbb{R}_+$ is a parameter. The behavior of $y(t) = \operatorname{proj}_C(x(t))$ was well studied in the literature and is naturally

connected to the differentiability properties (in some general sense) of the projection operator proj_C. We mention for instance the work of E.H. Zarantonello [29] who had shown the existence of directional derivatives for proj_C at the boundary of C. However, it is well-known that outside the set C, the directional differentiability of proj_C is not guaranteed (see [27] for a counter-example in $\mathbb{R}^2$). On the other hand, A. Haraux [5] and F. Mignot [11] published two fundamental papers where the conical differentiability of the projection operator is proved under a *polyhedric* (also called *density*) assumption. In that case, an asymptotic development of $y(t)$ at $t = 0$ can be obtained and the derivative $y'(0)$ can be expressed as the projection of $x'(0)$ onto a closed subset of the tangent cone of C at $y(0)$. We also refer to [4, 6, 28] for other references concerning the differentiability of the metric projection operator.

The theory of variational inequalities, introduced by G. Stampacchia and G. Fichera with motivations in unilateral mechanics, has been developed in the 1970's by the French and Italian schools. This theory has been proved to be a useful and effective tool for the study of a wide range of applications, within a unified framework, in many scientific areas such as contact mechanics, mathematical programming, operation research, economy, transportation planning, game theory, etc. In this paper we will focus on a general *nonlinear variational inequality of second kind* which consists in finding $y \in \text{H}$ such that

$$(\text{VI}(A,f,x)) \qquad \forall z \in \text{H}, \quad \langle A(y), z - y \rangle + f(z) - f(y) \geq \langle x, z - y \rangle,$$

where $A : \text{H} \to \text{H}$ is a (possibly nonlinear) operator, $f : \text{H} \to \mathbb{R} \cup \{+\infty\}$ is a proper lower semi-continuous convex function and $x \in \text{H}$ is given. Recall that $(\text{VI}(A,f,x))$ is said to be *linear* if A is linear, and that $(\text{VI}(A,f,x))$ is said to be of *first kind* if $f = \delta_\text{C}$ coincides with the indicator function of a nonempty closed convex subset $\text{C} \subset \text{H}$. In the case where $A = \text{Id}$ is the identity operator, $(\text{VI}(A,f,x))$ admits a unique solution given by $y = \text{prox}_f(x)$, where prox_f denotes the classical proximity (also well-known as proximal) operator introduced by J.-J. Moreau [13] in 1965. Note that if moreover $f = \delta_\text{C}$, then $\text{prox}_f = \text{proj}_\text{C}$. As a consequence, the Moreau's proximity operator can be viewed as a generalization of the projection operator. On the other hand, in the general case of a possibly nonlinear operator A, assumed to be Lipschitz continuous and strongly monotone, it can be proved [2, Proposition 31 p.140] that $(\text{VI}(A,f,x))$ admits a unique solution denoted by $y = \text{prox}_{A,f}(x)$, where $\text{prox}_{A,f}$ can be seen as a generalization of the Moreau's proximity operator prox_f.

A first step towards the sensitivity analysis of $(\text{VI}(A,f,x))$ is to consider a slight perturbation of the point $x \in \text{H}$, replacing it by $x(t) \in \text{H}$ where $t \in \mathbb{R}_+$ is a parameter, and to study the behavior of $y(t) = \text{prox}_{A,f}(x(t))$. The key point in the understanding of the differentiability properties of the generalized proximity operator $\text{prox}_{A,f}$ is to investigate generalized second-order differentiation theory. Precisely, R.T. Rockafellar introduced and thoroughly studied the notions of *twice epi-differentiability* [22] and *proto-differentiability* [23] of functions and set-valued maps respectively. These notions are both based on Painlevé-Kuratowski convergence of epigraphs and graphs of difference quotients. In the particular case where $A = \text{Id}$ is the identity operator and where H is finite-dimensional, thanks to Attouch's theorem [1], R.T. Rockafellar [24] proved the equivalence between the twice epi-differentiability of f, the proto-differentiability of the subdifferential operator ∂f and the proto-differentiability of the Moreau's proximity operator prox_f (see also [26, Chapter 13]). Note that C.N. Do [3] extended these results to the infinite-dimensional Hilbert setting using the concept of Mosco convergence [14]. Finally, from the twice epi-differentiability

3

of f, it can be derived that $y(t) = \mathrm{prox}_f(x(t))$ is differentiable at $t = 0$. In the general case of a possibly nonlinear operator A, the additional assumption of semi-differentiability of A (notion introduced by J.-P. Penot [16]) allows to recover easily the differentiability of $y(t) = \mathrm{prox}_{A,f}(x(t))$ at $t = 0$. Moreover $y'(0)$ can be expressed as the image of $x'(0)$ under a generalized proximity operator involving the semi-derivative of A and the second epi-derivative of f. In particular, note that $y'(0)$ is then the unique solution of the associated variational inequality.

For the calculus rules of epi-derivatives, we refer to the work of R.A. Poliquin and R.T. Rockafellar [19, 20] and references therein. In particular, it can be proved that Mignot and Haraux's *polyhedric* assumption on a closed convex set C is a sufficient condition for the twice epi-differentiability of the corresponding indicator function δ_C. We refer to the discussion in [3, Definition 2.8 and Example 2.10]. As a consequence, Rockafellar's result encompasses the one by Mignot and Haraux. Before concluding this section, we think that it is important to mention that there exist other approaches in the literature dealing with generalized second-order differentiations. It is not our aim to give here a complete list of references. Nevertheless, to mention just a few, we can refer for example to the work of B.S. Mordukhovich and R.T. Rockafellar [12] based on the dual-type constructions generated by coderivatives of first-order subdifferentials and to the paper of J.B. Hiriart-Urruty and A. Seeger [8] where calculus rules were introduced for set-valued second-order derivatives.

1.2   Contributions of the paper

The challenge of this paper is to investigate the sensitivity analysis of $(\mathrm{VI}(A, f, x))$ that takes into account a perturbation on all the data of the problem, not only on x but also on A and f. Precisely, we consider the following perturbed variational inequality which consists in finding $y(t) \in \mathrm{H}$ such that

$$(\mathrm{VI}(A(t, \cdot), f(t, \cdot), x(t))) \qquad \forall z \in \mathrm{H}, \ \ \langle A(t, y), z - y \rangle + f(t, z) - f(t, y) \geq \langle x(t), z - y \rangle,$$

where $A(t, \cdot)$, $f(t, \cdot)$ and $x(t)$ satisfy some appropriate assumptions that will be specified in Section 3. Note that the solution of $(\mathrm{VI}(A(t, \cdot), f(t, \cdot), x(t)))$ is given by

$$y(t) = \mathrm{prox}_{A(t, \cdot), f(t, \cdot)}(x(t)).$$

Our main objective in this paper is to derive sufficient conditions on $A(t, \cdot)$, $f(t, \cdot)$ and $x(t)$ under which $y(t)$ is differentiable at $t = 0$ and to provide an explicit formula for $y'(0)$. As mentioned in the previous section, R.T. Rockafellar already dealt with the *t-independent framework*, that is, with the particular case where $A(t, \cdot) = A$ and $f(t, \cdot) = f$ are t-independent. The concepts introduced by R.T. Rockafellar cannot be directly applied to the *t-dependent framework*. Therefore, we extend in Section 3 the notions of twice epi-differentiability, semi-differentiability and proto-differentiability to the case of parameterized functions, single-valued and set-valued maps respectively. These extensions are not a simple replica of the methodology developed by R.T. Rockafellar and require several adjustments in the definitions. Moreover we show explicitly in Example 31 that the situation is more complicated when dealing with the t-dependent setting. Precisely, the properness of the second epi-derivative of the one-variable function $f(x) = |x|$ is not preserved under the perturbation $f(t, x) = |x - t|$ (see Example 31), while it is preserved with the perturbation $f(t, x) = |x - t^2|$ (see Example 38). The difference between these two kinds of perturbations is the key point of the

understanding of how to handle the t-dependent situation. We refer to Section 4.1 for a detailed discussion about this issue. The new concept of *convergent supporting hyperplane*, introduced in this paper (see Definition 32), turns out to be efficient in order to prove the equivalence between the twice epi-differentiability of $f(t, \cdot)$ and the proto-differentiability of its subdifferential $\partial f(t, \cdot)$ (see Theorem 34). Actually the existence of a convergent supporting hyperplane to the second-order difference quotient is shown to be equivalent to the properness of the second epi-derivative of $f(t, \cdot)$ (see Proposition 39). Next we derive in Theorem 41 the proto-differentiability of $\mathrm{prox}_{A(t, \cdot), f(t, \cdot)}$ and we deduce in Theorem 42 that $y(t) = \mathrm{prox}_{A(t, \cdot), f(t, \cdot)}(x(t))$ is differentiable at $t = 0$. Moreover we prove that $y'(0)$ can be expressed as the image of $x'(0)$ under a generalized proximity operator involving the semi-derivative of $A(t, \cdot)$ and the second epi-derivative of $f(t, \cdot)$. In particular, note that $y'(0)$ is then the unique solution of the associated variational inequality.

Our main result (Theorem 42) encompasses and extends Rockafellar's work, and thus the results of Mignot and Haraux, mentioned in the previous section. For instance, it permits to deal with the differentiability of $t \mapsto \mathrm{proj}_{C(t)}(x(t))$ where the convex set C(t) is also perturbed in some sense (see Section 5.3). The originality of the present work is twofold. Firstly, it allows to deal with the sensitivity analysis of variational inequalities where all the data are t-dependent, via a nontrivial extension of the notion of twice epi-differentiability originally introduced by R.T. Rockafellar. Secondly, it permits to fill a gap in the literature, with respect to the works of A.J. King, A.B. Levy and R.T. Rockafellar in [9, 10]. In these references, the authors investigated the differentiability of the solutions set to parameterized generalized equations of the form

$$0 \in A(t, y) + B(t, y)$$

via the semi- and proto-derivatives of the involved single- and set-valued maps A and B respectively. However, no link has been explored with the twice epi-differentiability in the case where the set-valued map $B(t, \cdot) = \partial f(t, \cdot)$ coincides with the subdifferential of a parameterized convex function $f(t, \cdot)$.

1.3 Applications and additional comments

As an application of the theoretical results developed in this paper, we investigate the sensitivity analysis of parameterized convex optimization problems in Section 5. More precisely, we first consider in Section 5.1 the general parameterized minimization problem of the sum of two lower semi-continuous convex functions f and g (g being smooth). We show, under some appropriate assumptions, that the derivative at $t = 0$ of the perturbed solution is still a solution of a convex optimization problem involving the second epi-derivative of f and a quadratic term involving the Hessian matrix of g (see Proposition 44 for details). An illustrative example in one dimension is provided in Section 5.2 showing both the nontriviality of the t-dependent setting and the applicability of our theoretical results. Finally, the sensitivity analysis of parameterized smooth convex optimization problems with inequality constraints in a finite-dimensional setting is discussed in details in Section 5.3.

We end the paper with two discussions in Section 6. Firstly, we justify our choice of the formula of the second-order difference quotient of the twice epi-differentiability in the t-dependent setting (see Section 6.1 for details). Secondly, we show that the equivalence between the twice epi-differentiabilities of a convex function and its conjugate is not preserved in the t-dependent case

and requires some additional assumptions (that are automatically satisfied in the t-independent framework). We refer to Section 6.2 for details.

1.4 Organization of the paper

The paper is organized as follows. Section 2 is devoted to the main notations, definitions and auxiliary results from convex analysis. Section 3 is dedicated to the detailed setting of the paper and to the presentation of its main objective. Then, the extensions of the notions of generalized differentiability (semi-differentiability, proto-differentiability and twice epi-differentiability) to the t-dependent framework are specified. In Section 4, we state and prove our main results (Theorems 34, 41 and 42) where the new notion of convergent supporting hyperplane is introduced. In Section 5, as an application of our theoretical results, we investigate the sensitivity analysis of parameterized convex optimization problems. We conclude this paper with some additional comments in Section 6.

2 Recalls on convergence notions and convex analysis

In the next subsections we introduce some notations useful throughout the paper. We first give recalls about Painlevé-Kuratowski and Mosco convergences (see Section 2.1), graphical convergence (see Section 2.2) and epi-convergence (see Section 2.3). We continue with basics of convex analysis (see Section 2.4) and we conclude with recalls about Attouch's theorems (see Section 2.5). We refer to standard books like [1, 7, 17, 26] and references therein.

2.1 Two convergence modes of a parameterized family of subsets

In this paper we denote by $d(\cdot, S)$ the classical distance function to any subset S of H.

Let $(S_\tau)_{\tau>0}$ be a parameterized family of subsets of H. The *outer*, *weak-outer*, *inner* and *weak-inner limits* of $(S_\tau)_{\tau>0}$ when $\tau \to 0$ are respectively defined by

$$\limsup S_\tau := \{x \in H \mid \exists (t_n)_n \to 0, \ \exists (x_n)_n \to x, \ \forall n \in \mathbb{N}, \ x_n \in S_{t_n}\},$$

$$\text{w-}\limsup S_\tau := \{x \in H \mid \exists (t_n)_n \to 0, \ \exists (x_n)_n \rightharpoonup x, \ \forall n \in \mathbb{N}, \ x_n \in S_{t_n}\},$$

$$\liminf S_\tau := \{x \in H \mid \forall (t_n)_n \to 0, \ \exists (x_n)_n \to x, \ \exists N \in \mathbb{N}, \ \forall n \geq N, \ x_n \in S_{t_n}\},$$

$$\text{w-}\liminf S_\tau := \{x \in H \mid \forall (t_n)_n \to 0, \ \exists (x_n)_n \rightharpoonup x, \ \exists N \in \mathbb{N}, \ \forall n \geq N, \ x_n \in S_{t_n}\},$$

where $\to$ (resp. $\rightharpoonup$) denotes the strong (resp. weak) convergence in H. Note that the four following inclusions always hold true:

$$\liminf S_\tau \subset \limsup S_\tau \subset \text{w-}\limsup S_\tau \quad \text{and} \quad \liminf S_\tau \subset \text{w-}\liminf S_\tau \subset \text{w-}\limsup S_\tau.$$

In the whole paper, note that all limits with respect to τ will be considered for $\tau \to 0$. For the ease of notations, when no confusion is possible, the notation $\tau \to 0$ will be removed.

Remark 1. Let $(S_\tau)_{\tau>0}$ be a parameterized family of subsets of H. It can be shown that $\limsup S_\tau = \{x \in H \mid \liminf d(x, S_\tau) = 0\}$ and $\liminf S_\tau = \{x \in H \mid \limsup d(x, S_\tau) = 0\}$ (see, e.g., [17, Proposition 1.42 p.35]). We can deduce that $\limsup S_\tau$ and $\liminf S_\tau$ are closed subsets of H.

Remark 2. Convexity is a stable property under inner and weak-inner limits. Precisely, if $(S_\tau)_{\tau>0}$ is a parameterized family of convex subsets of H, then $\liminf S_\tau$ and w-$\liminf S_\tau$ are also convex. This result is not true in general for $\limsup S_\tau$ and w-$\limsup S_\tau$ (see, e.g., [26, p.119]).

Let us recall the two following convergence modes for a parameterized family $(S_\tau)_{\tau>0}$ of subsets of H.

Definition 3 (Painlevé-Kuratowski convergence). A parameterized family $(S_\tau)_{\tau>0}$ of subsets of H is said to be *PK-convergent* if

$$\limsup S_\tau \subset \liminf S_\tau.$$

In that case, we denote by PK-$\lim S_\tau := \liminf S_\tau = \limsup S_\tau$.

Definition 4 (Mosco convergence). A parameterized family $(S_\tau)_{\tau>0}$ of subsets of H is said to be *M-convergent* if

$$\text{w-}\limsup S_\tau \subset \liminf S_\tau.$$

In that case, we denote by M-$\lim S_\tau := \liminf S_\tau = \limsup S_\tau = \text{w-}\liminf S_\tau = \text{w-}\limsup S_\tau$.

Remark 5. (i) Let $(S_\tau)_{\tau>0}$ be a parameterized family of subsets of H. If $(S_\tau)_{\tau>0}$ M-converges, then $(S_\tau)_{\tau>0}$ PK-converges. In that case PK-$\lim S_\tau = $ M-$\lim S_\tau$. (ii) If H is finite-dimensional, the Painlevé-Kuratowski and the Mosco convergences clearly coincide.

2.2 Graphical convergence of a parameterized family of set-valued maps

For a set-valued map $A : \mathrm{H} \rightrightarrows \mathrm{H}$ on H, the *domain* of A is given by $\mathrm{Dom}(A) := \{x \in \mathrm{H} \mid A(x) \neq \emptyset\}$ and its *graph* is defined by $\mathrm{Gr}(A) := \{(x,y) \in \mathrm{H} \times \mathrm{H} \mid y \in A(x)\}$. We denote by $A^{-1} : \mathrm{H} \rightrightarrows \mathrm{H}$ the set-valued map defined by

$$A^{-1}(y) := \{x \in \mathrm{H} \mid y \in A(x)\},$$

for all $y \in \mathrm{H}$. In particular, for all x, $y \in \mathrm{H}$, it holds that

$$(x,y) \in \mathrm{Gr}(A) \Longleftrightarrow y \in A(x) \Longleftrightarrow x \in A^{-1}(y) \Longleftrightarrow (y,x) \in \mathrm{Gr}(A^{-1}).$$

Definition 6 (Graphical convergence). A parameterized family $(A_\tau)_{\tau>0}$ of set-valued maps on H is said to be *G-convergent* if $(\mathrm{Gr}(A_\tau))_{\tau>0}$ is PK-convergent. In that case, we denote by G-$\lim A_\tau : \mathrm{H} \rightrightarrows \mathrm{H}$ the set-valued map characterized by its graph as follows:

$$\mathrm{Gr}\left(\text{G-}\lim A_\tau\right) := \text{PK-}\lim \mathrm{Gr}(A_\tau).$$

Remark 7. Let $(A_\tau)_{\tau>0}$ be a parameterized family of set-valued maps on H. Then, $(A_\tau)_{\tau>0}$ is G-convergent if and only if $(A_\tau^{-1})_{\tau>0}$ is G-convergent. In that case, it holds that G-$\lim(A_\tau^{-1}) = (\text{G-}\lim A_\tau)^{-1}$.

Remark 8. Let $(A_\tau)_{\tau>0}$ be a G-convergent parameterized family of set-valued maps on H and let $A := \text{G-}\lim A_\tau$. If $(z_\tau, \xi_\tau)_{\tau>0} \to (z, \xi)$ with $(z_\tau, \xi_\tau) \in \mathrm{Gr}(A_\tau)$ for all $\tau > 0$, then $(z, \xi) \in \mathrm{Gr}(A)$.

2.3 M-convergence of a parameterized family of extended-real-valued functions

Let $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$. For an extended-real-valued function $f : \mathrm{H} \to \overline{\mathbb{R}}$ on H, the *domain* of f is defined by $\mathrm{dom}(f) := \{x \in \mathrm{H} \mid f(x) < +\infty\}$ and its *epigraph* is given by

$$\mathrm{Epi}(f) := \{(x, \lambda) \in \mathrm{H} \times \mathbb{R} \mid f(x) \leq \lambda\}.$$

We denote by $\mathcal{E}_{\mathrm{H}}$ the *set of all epigraphs* on H, that is, $S \in \mathcal{E}_{\mathrm{H}}$ if and only if there exists an extended-real-valued function $f : \mathrm{H} \to \overline{\mathbb{R}}$ such that $S = \mathrm{Epi}(f)$.

Remark 9. The set $\mathcal{E}_{\mathrm{H}}$ is stable under outer and inner limits (see, e.g., [26, p.240]). Precisely, if $(S_\tau)_{\tau>0}$ is a parameterized family of subsets of $\mathrm{H} \times \mathbb{R}$ such that $S_\tau \in \mathcal{E}_{\mathrm{H}}$ for all $\tau > 0$, then $\limsup S_\tau$ and $\liminf S_\tau$ both belong to $\mathcal{E}_{\mathrm{H}}$.

Definition 10 (M-convergence). A parameterized family $(f_\tau)_{\tau>0}$ of extended-real -valued functions on H is said to be M-convergent if $(\mathrm{Epi}(f_\tau))_{\tau>0}$ is M-convergent. In that case, we denote by E-$\lim f_\tau : \mathrm{H} \to \overline{\mathbb{R}}$ the extended-real-valued function characterized by its epigraph as follows:

$$\mathrm{Epi}(\text{E-}\lim f_\tau) := \text{M-}\lim \mathrm{Epi}(f_\tau).$$

Recall the following characterization of epi-convergence. We refer to [1, Proposition 3.19 p.297] or [26, Proposition 7.2 p.241] for details.

Proposition 11. *Let f be an extended-real-valued function on H and let $(f_\tau)_{\tau>0}$ be a parameterized family of extended-real-valued functions on H. Then $(f_\tau)_{\tau>0}$ M-converges with $f = \text{E-}\lim f_\tau$ if and only if, for all $z \in \mathrm{H}$, there exists $(z_\tau)_{\tau>0} \to z$ such that $\limsup f_\tau(z_\tau) \leq f(z)$ and, for all $(z_\tau)_{\tau>0} \rightharpoonup z$, $\liminf f_\tau(z_\tau) \geq f(z)$.*

2.4 Basics of convex analysis

A set-valued operator $A : \mathrm{H} \rightrightarrows \mathrm{H}$ is said to be *monotone* if

$$\forall (x_1, y_1), (x_2, y_2) \in \mathrm{Gr}(A), \quad \langle y_2 - y_1, x_2 - x_1 \rangle \geq 0.$$

Moreover, A is said to be *maximal monotone* if $\mathrm{Gr}(A) \subset \mathrm{Gr}(B)$ for some monotone set-valued operator $B : \mathrm{H} \rightrightarrows \mathrm{H}$ implies that $A = B$.

A set-valued map $A : \mathrm{H} \rightrightarrows \mathrm{H}$ is said to be *single-valued* if $A(x)$ is a singleton for all $x \in \mathrm{H}$. In that case, it holds in particular that $\mathrm{Dom}(A) = \mathrm{H}$ and we denote by $A : \mathrm{H} \to \mathrm{H}$ (instead of $A : \mathrm{H} \rightrightarrows \mathrm{H}$). For a single-valued map $A : \mathrm{H} \to \mathrm{H}$, we say that A is *Lipschitz continuous* if

$$\exists M \geq 0, \quad \forall x_1, x_2 \in \mathrm{H}, \quad \|A(x_2) - A(x_1)\| \leq M \|x_2 - x_1\|,$$

and we say that A is *strongly monotone* if

$$\exists \alpha > 0, \quad \forall x_1, x_2 \in \mathrm{H}, \quad \langle A(x_2) - A(x_1), x_2 - x_1 \rangle \geq \alpha \|x_2 - x_1\|^2.$$

In the whole paper, we denote by $\mathcal{A}(\mathrm{H})$ the set of all single-valued maps $A : \mathrm{H} \to \mathrm{H}$ that are Lipschitz continuous and strongly monotone.

For an extended-real-valued function $f : \mathrm{H} \to \mathbb{R} \cup \{+\infty\}$, we say that f is *proper* if $\mathrm{dom}(f) \neq \emptyset$, and that f is *lower semi-continuous* if its epigraph $\mathrm{Epi}(f)$ is closed in $\mathrm{H} \times \mathbb{R}$. Finally, recall that f is a convex function if and only if its epigraph $\mathrm{Epi}(f)$ is a convex subset of $\mathrm{H} \times \mathbb{R}$. In the whole paper, we denote by $\Gamma_0(\mathrm{H})$ the set of all extended-real-valued functions $f : \mathrm{H} \to \mathbb{R} \cup \{+\infty\}$ that are proper, lower semi-continuous and convex.

Let $f \in \Gamma_0(\mathrm{H})$. We denote by $\partial f : \mathrm{H} \rightrightarrows \mathrm{H}$ the *subdifferential operator* of f defined by

$$\partial f(x) := \{y \in \mathrm{H} \mid \forall z \in \mathrm{H}, \ \langle y, z - x \rangle \leq f(z) - f(x)\},$$

for all $x \in \mathrm{H}$. In particular, note that $x \in \mathrm{argmin}\, f$ if and only if $0 \in \partial f(x)$. Recall that ∂f is a maximal monotone operator (see [21] or [26, Theorem 12.17 p.542] for details). Moreover, it follows from the classical Brøndsted-Rockafellar theorem (see, e.g., [18, Theorem 6.5 p.333]) that $\mathrm{Dom}(\partial f) \neq \emptyset$, and thus f admits a *supporting hyperplane*, that is,

$$\exists (z, \xi, \beta) \in \mathrm{H} \times \mathrm{H} \times \mathbb{R}, \quad \begin{cases} \forall w \in \mathrm{H}, \quad f(w) \geq \langle \xi, w \rangle + \beta, \\ f(z) = \langle \xi, z \rangle + \beta. \end{cases}$$

Note that $(z, \xi, \beta) \in \mathrm{H} \times \mathrm{H} \times \mathbb{R}$ is a supporting hyperplane of f if and only if $\xi \in \partial f(z)$ and $\beta = f(z) - \langle \xi, z \rangle$.

The *generalized proximity* (also well-known as *proximal*) *operator* associated to some $(A, f) \in \mathcal{A}(\mathrm{H}) \times \Gamma_0(\mathrm{H})$ is defined by

$$\mathrm{prox}_{A,f} := (A + \partial f)^{-1}.$$

For all $x, y \in \mathrm{H}$, one can easily see that $y \in \mathrm{prox}_{A,f}(x)$ if and only if y is a solution of the *nonlinear variational inequality of second kind* given by

$(\mathrm{VI}(A, f, x))$ $$\forall z \in \mathrm{H}, \quad \langle A(y), z - y \rangle + f(z) - f(y) \geq \langle x, z - y \rangle.$$

From the contraction mapping principle, it can be proved that the above variational inequality $(\mathrm{VI}(A, f, x))$ admits a unique solution (see, e.g., [2, Proposition 31 p.140] for details) given by $\mathrm{prox}_{A,f}(x)$. We deduce that $\mathrm{prox}_{A,f} : \mathrm{H} \to \mathrm{H}$ is a single-valued map. Moreover, it can be easily proved that $\mathrm{prox}_{A,f}$ is $\frac{1}{\alpha}$-Lipschitz continuous where α denotes the strong monotonicity coefficient of A.

2.5 Recalls about Attouch's theorems

We conclude this section with two crucial results stated by H. Attouch in [1, Theorem 3.66 p.373 and Corollary 3.65 p.372]. The first result is well-known as the classical Attouch's theorem.

Theorem 12. *Let $(f_\tau)_{\tau > 0}$ be a parameterized family of functions in $\Gamma_0(\mathrm{H})$ and let $f \in \Gamma_0(\mathrm{H})$. Then, $f = \mathrm{E\text{-}lim}\, f_\tau$ if and only if the two following assertions are satisfied:*

(i) *$\partial f = \mathrm{G\text{-}lim}\, \partial f_\tau$;*

(ii) *there exists $(z_\tau, \xi_\tau)_{\tau > 0} \to (z, \xi)$ such that $(z_\tau, \xi_\tau) \in \mathrm{Gr}(\partial f_\tau)$ for all $\tau > 0$ and $f_\tau(z_\tau) \to f(z)$.*

Theorem 13. *Let $(f_\tau)_{\tau > 0}$ be a parameterized family of functions in $\Gamma_0(\mathrm{H})$. If $(\partial f_\tau)_{\tau > 0}$ is G-convergent and if $\mathrm{G\text{-}lim}\, \partial f_\tau$ is a maximal monotone operator, then there exists $f \in \Gamma_0(\mathrm{H})$ such that $\mathrm{G\text{-}lim}\, \partial f_\tau = \partial f$.*

Remark 14. Let us consider the framework of Theorem 12. If the equivalent assertions are satisfied, note that $(z, \xi) \in \mathrm{Gr}(\partial f)$ from Remark 8. In particular note that $f(z) \in \mathbb{R}$.

3   Setting, objective and generalized notions of differentiability

Before exhibiting in details the setting of the present paper and its main objective, we need to introduce first some notations.

Let $A : \mathbb{R}_+ \times \mathrm{H} \rightrightarrows \mathrm{H}$ be a parameterized set-valued map. For the simplicity of notations, we denote by $A^{-1} : \mathbb{R}_+ \times \mathrm{H} \rightrightarrows \mathrm{H}$ the parameterized set-valued map defined by

$$A^{-1}(t, y) := (A(t, \cdot))^{-1}(y)$$

for all $(t, y) \in \mathbb{R}_+ \times \mathrm{H}$.

A parameterized set-valued map $A : \mathbb{R}_+ \times \mathrm{H} \rightrightarrows \mathrm{H}$ is said to be a *parameterized single-valued map* if $A(t, \cdot) : \mathrm{H} \to \mathrm{H}$ is a single-valued map for all $t \geq 0$. In that case we denote by $A : \mathbb{R}_+ \times \mathrm{H} \to \mathrm{H}$ (instead of $A : \mathbb{R}_+ \times \mathrm{H} \rightrightarrows \mathrm{H}$).

Let us introduce:

- $\mathcal{A}(\cdot, \mathrm{H})$ the set of all parameterized single-valued maps $A : \mathbb{R}_+ \times \mathrm{H} \to \mathrm{H}$ such that $A(t, \cdot) \in \mathcal{A}(\mathrm{H})$ for all $t \geq 0$;

- $\Gamma_0(\cdot, \mathrm{H})$ the set of all parameterized extended-real-valued functions $f : \mathbb{R}_+ \times \mathrm{H} \to \mathbb{R} \cup \{+\infty\}$ such that $f(t, \cdot) \in \Gamma_0(\mathrm{H})$ for all $t \geq 0$.

Let $f \in \Gamma_0(\cdot, \mathrm{H})$. For the ease of notations, we denote by

$$\partial f(t, x) := \partial(f(t, \cdot))(x)$$

for all $(t, x) \in \mathbb{R}_+ \times \mathrm{H}$. Then $\partial f : \mathbb{R}_+ \times \mathrm{H} \rightrightarrows \mathrm{H}$ is a parameterized set-valued map. Moreover, we introduce the notation

$$f^{-1}(\cdot, \mathbb{R}) := \{x \in \mathrm{H} \mid \forall t \geq 0, \ f(t, x) \in \mathbb{R}\}.$$

If f is t-independent, note that $f \in \Gamma_0(\mathrm{H})$ and that $f^{-1}(\cdot, \mathbb{R})$ coincides with the classical notation $f^{-1}(\mathbb{R})$.

3.1   Setting and main objective of the paper

In this paper we focus on the sensitivity analysis, with respect to the parameter $t \geq 0$, of the general nonlinear variational inequality of second kind given by

$$(\mathrm{VI}(A(t, \cdot), f(t, \cdot), x(t))) \qquad \forall z \in \mathrm{H}, \ \ \langle A(t, y), z - y \rangle + f(t, z) - f(t, y) \geq \langle x(t), z - y \rangle,$$

where $(A, f) \in \mathcal{A}(\cdot, \mathrm{H}) \times \Gamma_0(\cdot, \mathrm{H})$ and where $x : \mathbb{R}_+ \to \mathrm{H}$ is a given function. From Section 2.4, the above variational inequality admits for all $t \geq 0$ a unique solution $y(t) \in \mathrm{H}$ given by

$$y(t) = \mathrm{prox}_{A(t, \cdot), f(t, \cdot)}(x(t)).$$

Our main objective in this paper is to derive sufficient conditions on A, f and x under which $y : \mathbb{R}_+ \to \mathrm{H}$ is differentiable at $t = 0$ and to provide an explicit formula for $y'(0)$.

In the literature numerous generalized notions of differentiability have been introduced in order to deal with nonsmooth functions. In particular, notions of *semi-differentiability*, *twice epi-differentiability* and *proto-differentiability* have been introduced respectively in [16], [22] and [23]. These wonderful tools turned out to be sufficient in order to deal with the case where A and f are t-independent. We refer to [24, 26] for the finite-dimensional case and to [3] for the infinite-dimensional case.[1]

In view of dealing with the t-dependence of A and f, we need to extend the notions of semi-differentiability, twice epi-differentiability and proto-differentiability above mentioned to the t-dependent framework. This is the aim of Sections 3.2, 3.3 and 3.4.

Remark 15. We mention here that these generalizations are not trivial and are not sufficient in order to fully adapt the strategy developed in [3, 24, 26] to the t-dependent framework. Indeed, we will provide a simple example (see Example 31) showing that the situation is more complicated in the t-dependent setting. As a consequence, we will introduce in Section 4.1 a new concept called *convergent supporting hyperplane* (see Definition 32) that allows us to conclude. Actually, this condition turns out to be necessary and sufficient in a sense that can be made precise (see Proposition 39). We refer to Section 4.1 for a detailed discussion about this issue. Similarly, we will also prove that, in contrary to the t-independent framework, the generalized twice epi-differentiability of a function $f \in \Gamma_0(\cdot, \mathrm{H})$ cannot be directly related to the twice epi-differentiability of its conjugate function f^*. We refer to Section 6.2 for more details.

3.2 Semi-differentiability of a parameterized single-valued map

In this section our aim is to generalize the classical notion of semi-differentiability to the case where the single-valued map considered depends on the parameter $t \geq 0$.

Definition 16 (Semi-differentiability). Let $A : \mathbb{R}_+ \times \mathrm{H} \to \mathrm{H}$ be a parameterized single-valued map and let $x \in \mathrm{H}$. If the limit

$$D_s A(x)(w) := \lim_{\substack{\tau \to 0 \\ w' \to w}} \frac{A(\tau, x + \tau w') - A(0, x)}{\tau}$$

exists in H for all $w \in \mathrm{H}$, we say that A is *semi-differentiable* at x. In that case, $D_s A(x) : \mathrm{H} \to \mathrm{H}$ is a single-valued map called the *semi-derivative* of A at x.

If the single-valued map A is t-independent, Definition 16 recovers the classical notion of semi-differentiability originally introduced in [16].

In the sequel we will denote by $\mathcal{A}_{\mathrm{unif}}(\cdot, \mathrm{H})$ the set of all parameterized single-valued maps $A : \mathbb{R}_+ \times \mathrm{H} \to \mathrm{H}$ such that A is *uniformly Lipschitz continuous*, that is,

$$\exists M \geq 0, \quad \forall t \geq 0, \quad \forall x_1, x_2 \in \mathrm{H}, \quad \|A(t, x_2) - A(t, x_1)\| \leq M \|x_2 - x_1\|,$$

and *uniformly strongly monotone*, that is,

$$\exists \alpha > 0, \quad \forall t \geq 0, \quad \forall x_1, x_2 \in \mathrm{H}, \quad \langle A(t, x_2) - A(t, x_1), x_2 - x_1 \rangle \geq \alpha \|x_2 - x_1\|^2.$$

---

[1]These three references are interested only in the t-independent framework and actually they also only consider the case where $A = \mathrm{Id}$ is the identity operator. Precisely, they proved that if f is twice epi-differentiable, then the classical Moreau's proximity operator $\mathrm{prox}_f = \mathrm{prox}_{\mathrm{Id}, f}$ is proto-differentiable. Even if it is not mentioned in [3, 24, 26], one can easily deduce from the proto-differentiability of prox_f that y is differentiable at $t = 0$ and provide an explicit formula for $y'(0)$. The results presented in the present paper encompass and extend these results.

Proposition 17. *Let $A \in \mathcal{A}_{\mathrm{unif}}(\cdot, \mathrm{H})$. If A is semi-differentiable at $x \in \mathrm{H}$, then $D_s A(x) \in \mathcal{A}(\mathrm{H})$.*

Proof. Let us prove that $D_s A(x)$ is Lipschitz continuous. Let $w_1, w_2 \in \mathrm{H}$. It holds that $\|A(\tau, x + \tau w_2) - A(\tau, x + \tau w_1)\| \leq M\tau \|w_2 - w_1\|$ and thus

$$\left\| \frac{A(\tau, x + \tau w_2) - A(0, x)}{\tau} - \frac{A(\tau, x + \tau w_1) - A(0, x)}{\tau} \right\| \leq M\|w_2 - w_1\|,$$

for all $\tau > 0$. Letting $\tau \to 0$ leads to $\|D_s A(x)(w_2) - D_s A(x)(w_1)\| \leq M\|w_2 - w_1\|$. We conclude that $D_s A(x)$ is Lipschitz continuous. The proof is similar for proving the strong monotonicity of $D_s A(x)$. $\square$

Remark 18. Let $A \in \mathcal{A}_{\mathrm{unif}}(\cdot, \mathrm{H})$ be semi-differentiable at $x \in \mathrm{H}$. If we denote by $M \geq 0$ and $\alpha > 0$ the uniform Lipschitz and strong monotonicity coefficients of A, then $D_s A(x) \in \mathcal{A}(\mathrm{H})$ is M-Lipschitz continuous and α-strongly monotone.

3.3 Proto-differentiability of a parameterized set-valued map

In this section our aim is to generalize the classical notion of proto-differentiability to the case where the set-valued map considered depends on the parameter $t \geq 0$.

Let $A : \mathbb{R}_+ \times \mathrm{H} \rightrightarrows \mathrm{H}$ be a parameterized set-valued map. For all $\tau > 0$, $x \in \mathrm{H}$ and $v \in A(0, x)$, we denote by

$$\begin{aligned}
\Delta_\tau A(x|v) : \quad \mathrm{H} \quad &\rightrightarrows \quad \mathrm{H} \\
w \quad &\mapsto \quad \Delta_\tau A(x|v)(w) := \frac{A(\tau, x + \tau w) - v}{\tau}.
\end{aligned}$$

Definition 19 (Proto-differentiability)**.** Let $A : \mathbb{R}_+ \times \mathrm{H} \rightrightarrows \mathrm{H}$ be a parameterized set-valued map. We say that A is *proto-differentiable* at $x \in \mathrm{H}$ for $v \in A(0, x)$ if $(\Delta_\tau A(x|v))_{\tau > 0}$ G-converges. In that case, we denote by

$$D_p A(x|v) := \text{G-}\lim \Delta_\tau A(x|v)$$

the set-valued map $D_p A(x|v) : \mathrm{H} \rightrightarrows \mathrm{H}$ called the *proto-derivative* of A at x for v.

If the set-valued map A is t-independent, Definition 19 recovers the classical notion of proto-differentiability originally introduced in [23].

Remark 20. Let us point out that if a parameterized single-valued map $A : \mathbb{R}_+ \times \mathrm{H} \to \mathrm{H}$ is semi-differentiable at $x \in \mathrm{H}$, then A is proto-differentiable at x for $A(0, x)$ with $D_p A(x|A(0, x)) = D_s A(x)$.

One can easily prove the two following results. In the t-independent case, we recover [26, p.331-333].

Proposition 21. *Let $A : \mathbb{R}_+ \times \mathrm{H} \to \mathrm{H}$ be a parameterized single-valued map and $B : \mathbb{R}_+ \times \mathrm{H} \rightrightarrows \mathrm{H}$ be a parameterized set-valued map. Let $x \in \mathrm{H}$ and $v \in A(0, x) + B(0, x)$. If A is semi-differentiable at x, then $A + B$ is proto-differentiable at x for v if and only if B is proto-differentiable at x for $v - A(0, x)$. In that case it holds that*

$$D_p(A + B)(x|v) = D_s A(x) + D_p B(x|v - A(0, x)).$$

Proposition 22. *Let $A : \mathbb{R}_+ \times \mathrm{H} \rightrightarrows \mathrm{H}$ be a parameterized set-valued map, $x \in \mathrm{H}$ and $v \in A(0, x)$. Then, A is proto-differentiable at x for v if and only if A^{-1} is proto-differentiable at v for x. In that case, it holds that*

$$D_p(A^{-1})(v|x) := (D_p A(x|v))^{-1}.$$

3.4 Twice epi-differentiability of parameterized extended-real-valued functions

In this section our aim is to generalize the classical notion of twice epi-differentiability to the case where the extended-real-valued function considered depends on the parameter $t \geq 0$.

Let $f \in \Gamma_0(\cdot, H)$. For all $\tau > 0$, $x \in f^{-1}(\cdot, \mathbb{R})$ and $v \in \partial f(0, x)$, we denote by

$$\begin{aligned}
\Delta_\tau^2 f(x|v) : \quad H \quad &\to \quad \mathbb{R} \cup \{+\infty\} \\
w \quad &\mapsto \quad \Delta_\tau^2 f(x|v)(w) := \frac{f(\tau, x + \tau w) - f(\tau, x) - \tau \langle v, w \rangle}{\tau^2}.
\end{aligned} \tag{1}$$

Definition 23 (Twice epi-differentiability)**.** Let $f \in \Gamma_0(\cdot, H)$. We say that f is *twice epi-differentiable* at $x \in f^{-1}(\cdot, \mathbb{R})$ for $v \in \partial f(0, x)$ if $(\Delta_\tau^2 f(x|v))_{\tau > 0}$ M-converges. In that case, we denote by

$$d_e^2 f(x|v) := \text{E-lim} \, \Delta_\tau^2 f(x|v)$$

the extended-real-valued function $d_e^2 f(x|v) : H \to \overline{\mathbb{R}}$ called the *second epi-derivative* of f at x for v.

If the extended-real-valued function f is t-independent, Definition 23 recovers the classical notion of twice epi-differentiability originally introduced in [22] (up to the multiplicative constant $\frac{1}{2}$).

Remark 24. Let $f \in \Gamma_0(\cdot, H)$ be twice epi-differentiable at $x \in f^{-1}(\cdot, \mathbb{R})$ for $v \in \partial f(0, x)$. Even if $\Delta_\tau^2 f(x|v)$ is with values in $\mathbb{R} \cup \{+\infty\}$ for all $\tau > 0$, it may be possible that there exists $w \in H$ such that $d_e^2 f(x|v)(w) = -\infty$. We refer to Example 31. This is an important difference with the t-independent framework. We refer to Section 4.1 for a detailed discussion about this issue.

Remark 25. Let $f \in \Gamma_0(\cdot, H)$ be twice epi-differentiable at $x \in f^{-1}(\cdot, \mathbb{R})$ for $v \in \partial f(0, x)$. It might be possible that $d_e^2 f(x|v)$ is not positively homogeneous of degree two. We refer to Example 38. Indeed, note that $\Delta_\tau^2 f(x|v)(\lambda w) \neq \lambda^2 \Delta_{\lambda\tau}^2 f(x|v)(w)$ in general for $\lambda > 0$ and $w \in H$. This is a second important difference with the t-independent framework.

Remark 26. Let $f : \mathbb{R}_+ \times H \to \mathbb{R}$ be a function of class C^2 on $\mathbb{R}_+ \times H$. A Taylor expansion of f around $x \in H$ with $v = \nabla_x f(0, x)$ implies that $(\Delta_\tau^2 f(x|v)(w))_{\tau > 0}$ converges pointwise to

$$\frac{1}{2} \langle \nabla_{xx}^2 f(0, x)(w), w \rangle + \langle \nabla_{tx}^2 f(0, x), w \rangle$$

for all $w \in H$. This remark is coherent with Remark 25 since the above expression is not positively homogeneous of degree two with respect to the variable w.

Remark 27. Note that Section 6.1 will be devoted to a discussion about the pointwise convergence mentioned in Remark 26 and about the choice of Formula (1).

4 Main results

The whole paper is based on Attouch's theorems (see Theorems 12 and 13) and on Proposition 28 below. Its proof is simple and similar to the t-independent case (see, e.g., [25, Proposition 2.7]).

Proposition 28. *Let* $f \in \Gamma_0(\cdot, H)$, $x \in f^{-1}(\cdot, \mathbb{R})$ *and* $v \in \partial f(0, x)$. *Then,* $\Delta_\tau^2 f(x|v)$ *belongs to* $\Gamma_0(H)$ *with*

$$\partial \left(\Delta_\tau^2 f(x|v) \right) = \Delta_\tau(\partial f)(x|v),$$

for all $\tau > 0$.

One can easily deduce that the twice epi-differentiability of a function $f \in \Gamma_0(\cdot, \mathrm{H})$ is strongly related to the proto-differentiability of its subdifferential operator ∂f.

This relation in the t-independent framework is simple and recalled in Section 4.1 (see Proposition 29). However, a simple example (see Example 31) shows that the situation is more complicated in the t-dependent setting. This is the reason why we introduce a new concept called *convergent supporting hyperplane* (see Definition 32) that allows us to state and prove the counterpart of Proposition 29 in the t-dependent framework (see Theorem 34). This condition actually turns out to be necessary and sufficient in a sense that can be made precise (see Proposition 39).

Finally Section 4.2 is devoted to the proto-differentiability of the generalized proximity operator (see Theorem 41) and to our initial motivation, that is, to the sensitivity analysis of general nonlinear variational inequalities of second kind (see Theorem 42).

4.1   Convergent supporting hyperplane

In the t-independent framework, the following proposition relating the twice epi-differentiability of a function $f \in \Gamma_0(\mathrm{H})$ and the proto-differentiability of its subdifferential operator ∂f is a well-known result. We refer to [24, 26] for the finite-dimensional case and to [3] for the infinite-dimensional one.

Proposition 29. *Let $f \in \Gamma_0(\mathrm{H})$ (that is t-independent), $x \in f^{-1}(\mathbb{R})$ and $v \in \partial f(x)$. The following assertions are equivalent:*

  (i) *f is twice epi-differentiable at x for v;*

  (ii) *∂f is proto-differentiable at x for v and $D_p(\partial f)(x|v)$ is a maximal monotone operator.*

In that case $d_e^2 f(x|v)$ belongs to $\Gamma_0(\mathrm{H})$ with $d_e^2 f(x|v)(0) = 0$ and

$$D_p(\partial f)(x|v) = \partial(d_e^2 f(x|v)).$$

This classical result is based on Attouch's theorems (see Theorems 12 and 13) and on Proposition 28. However it is worth to note here that Proposition 29 is also based on the following lemma.

Lemma 30. *Let $f \in \Gamma_0(\mathrm{H})$ (that is t-independent), $x \in f^{-1}(\mathbb{R})$ and $v \in \partial f(x)$. Then, $\Delta_\tau^2 f(x|v)(0) = 0$ and $\Delta_\tau^2 f(x|v)(w) \geq 0$ for all $w \in \mathrm{H}$ and all $\tau > 0$. In particular, it holds that $0 \in \partial(\Delta_\tau^2 f(x|v))(0)$ for all $\tau > 0$.*

Let us recall the proof of Proposition 29 in order to highlight the crucial role of Lemma 30.

Proof of Proposition 29. Let us assume that f is twice epi-differentiable at x for v. Our first aim is to prove that $d_e^2 f(x|v) \in \Gamma_0(\mathrm{H})$. Firstly, let us assume by contradiction that there exists $w \in \mathrm{H}$ such that $d_e^2 f(x|v)(w) = -\infty$. Since $(w, -1) \in \mathrm{Epi}(d_e^2 f(x|v)) = \limsup \mathrm{Epi}(\Delta_\tau^2 f(x|v))$, there exist $(t_n) \to 0$ and $(w_n, \lambda_n) \to (w, -1)$ such that $(w_n, \lambda_n) \in \mathrm{Epi}(\Delta_{t_n}^2 f(x|v))$, that is, $\Delta_{t_n}^2 f(x|v)(w_n) \leq \lambda_n$ for all $n \in \mathbb{N}$. For sufficiently large n, it holds that $\Delta_{t_n}^2 f(x|v)(w_n) \leq \lambda_n < 0$ which raises a contradiction with Lemma 30. Secondly, since $\mathrm{Epi}(d_e^2 f(x|v)) = \liminf \mathrm{Epi}(\Delta_\tau^2 f(x|v))$, we easily conclude that $d_e^2 f(x|v)$ is lower semi-continuous and convex (see Remarks 1 and 2 and Proposition 28). Thirdly, since $(0, 0) \in \mathrm{Epi}(\Delta_\tau^2 f(x|v))$ for all $\tau > 0$, we deduce that $(0, 0) \in \mathrm{Epi}(d_e^2 f(x|v))$

14

and thus $d_e^2 f(x|v)(0) \leq 0$. This concludes that $d_e^2 f(x|v)$ is proper. Our second aim is to prove that $d_e^2 f(x|v)(0) = 0$. By contradiction, let us assume that $d_e^2 f(x|v)(0) < 0$. Then, there exists $\varepsilon > 0$ such that $(0, -\varepsilon) \in \mathrm{Epi}(d_e^2 f(x|v)) = \limsup \mathrm{Epi}(\Delta_\tau^2 f(x|v))$. Thus, there exist $(t_n) \to 0$ and $(w_n, \varepsilon_n) \to (0, \varepsilon)$ such that $(w_n, -\varepsilon_n) \in \mathrm{Epi}(\Delta_{t_n}^2 f(x|v))$, that is, $\Delta_{t_n}^2 f(x|v)(w_n) \leq -\varepsilon_n$ for all $n \in \mathbb{N}$. For sufficiently large n, it holds that $\Delta_{t_n}^2 f(x|v)(w_n) \leq -\varepsilon_n < 0$ which constitutes a contradiction with Lemma 30. Our last aim is to prove that ∂f is proto-differentiable at x for v with $D_p(\partial f)(x|v) = \partial(d_e^2 f(x|v))$. This result directly follows from Proposition 28 and Theorem 12.

Now let us assume that ∂f is proto-differentiable at x for v and $D_p(\partial f)(x|v)$ is a maximal monotone operator. From Proposition 28, we get that $(\partial(\Delta_\tau^2 f(x|v)))_{\tau>0} = (\Delta_\tau(\partial f)(x|v))_{\tau>0}$ G-converges to the maximal monotone operator $D_p(\partial f)(x|v)$. We deduce from Theorem 13 that $D_p(\partial f)(x|v) = \partial\varphi$ for some $\varphi \in \Gamma_0(\mathrm{H})$. Our aim is now to apply Theorem 12 in order to conclude that $(\Delta_\tau^2 f(x|v))_{\tau>0}$ M-converges. Let us consider $z_\tau := 0$ and $\xi_\tau := 0$ for all $\tau > 0$. In particular, we have $(z_\tau, \xi_\tau) \to (0, 0)$ and since $0 \in \partial(\Delta_\tau^2 f(x|v))(0)$ (see Lemma 30), we deduce that $(z_\tau, \xi_\tau) \in \mathrm{Gr}(\partial(\Delta_\tau^2 f(x|v)))$ for all $\tau > 0$. From Remark 8, we deduce that $(0, 0) \in \mathrm{Gr}(\partial\varphi)$ and thus $0 \in \mathrm{dom}(\varphi)$. Considering $\psi := \varphi - \varphi(0) \in \Gamma_0(\mathrm{H})$, we have $\psi(0) = 0$ and $\partial\psi = \partial\varphi$. Since $\Delta_\tau^2 f(x|v)(\xi_\tau) = 0 \to \psi(0)$, the proof is complete from Theorem 12. □

However it should be noted that Lemma 30 is not true in the general t-dependent setting. As a consequence the above proof cannot be fully adapted to the t-dependent framework. Actually, in the t-dependent setting, we can even prove that the above proposition is not true with a simple counter-example provided below.

Example 31. Let $\mathrm{H} = \mathbb{R}$ and $f(t, x) := |x - t|$ for all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}$. Let us consider $x = 0$ and $v = 0 \in \partial f(0, x)$. One can easily compute that

$$\Delta_\tau^2 f(x|v)(w) = \frac{|w - 1| - 1}{\tau},$$

for all $\tau > 0$ and all $w \in \mathbb{R}$. One can easily deduce that f is twice epi-differentiable at x for v with

$$d_e^2 f(x|v)(w) = \begin{cases} -\infty & \text{if } w \in [0, 2], \\ +\infty & \text{if } w \notin [0, 2]. \end{cases}$$

In particular note that $d_e^2 f(x|v) \notin \Gamma_0(\mathrm{H})$ and that $d_e^2 f(x|v)(0) \neq 0$.

We deduce from Example 31 that Proposition 29 does not admit an exact counterpart in the t-dependent framework. Thus our aim is now to introduce a new concept that allows to recover Proposition 29 in the t-dependent framework. We refer to Section 2.4 for the notion of *supporting hyperplane*.

Definition 32 (Convergent supporting hyperplane). Let $(f_\tau)_{\tau>0}$ be a parameterized family of functions in $\Gamma_0(\mathrm{H})$. We say that $(f_\tau)_{\tau>0}$ admits a *convergent supporting hyperplane* if there exists $(z_\tau, \xi_\tau, \beta_\tau)_{\tau>0} \to (z, \xi, \beta)$ such that $(z_\tau, \xi_\tau, \beta_\tau)$ is a supporting hyperplane of f_τ for all $\tau > 0$.

Remark 33. Let us point out that the existence of a convergent supporting hyperplane to a parameterized family $(f_\tau)_{\tau>0}$ of functions in $\Gamma_0(\mathrm{H})$ is equivalent to the existence of $(z_\tau, \xi_\tau)_{\tau>0} \to (z, \xi)$ such that $(z_\tau, \xi_\tau) \in \mathrm{Gr}(\partial f_\tau)$ for all $\tau > 0$ and such that $f_\tau(z_\tau) \to \gamma$ for some $\gamma \in \mathbb{R}$.

We are now in position in order to state and prove the counterpart of Proposition 29 in the t-dependent framework, under the assumption of the existence of a convergent supporting hyperplane. The following theorem is in this sense.

Theorem 34. *Let $f \in \Gamma_0(\cdot, \mathrm{H})$, $x \in f^{-1}(\cdot, \mathbb{R})$ and $v \in \partial f(0, x)$. Let us assume that $(\Delta_\tau^2 f(x|v))_{\tau>0}$ admits a convergent supporting hyperplane $(z_\tau, \xi_\tau, \beta_\tau)_{\tau>0} \to (z, \xi, \beta)$. Then the following assertions are equivalent:*

  (i) *f is twice epi-differentiable at x for v;*

  (ii) *∂f is proto-differentiable at x for v and $D_p(\partial f)(x|v)$ is a maximal monotone operator.*

In that case $d_e^2 f(x|v)$ belongs to $\Gamma_0(\mathrm{H})$ with $\beta \leq d_e^2 f(x|v)(0) \leq 0$ and

$$D_p(\partial f)(x|v) = \partial(d_e^2 f(x|v)).$$

Proof. Let us assume that f is twice epi-differentiable at x for v. Our first aim is to prove that $d_e^2 f(x|v) \in \Gamma_0(\mathrm{H})$. Firstly, let us assume by contradiction that there exists $w \in \mathrm{H}$ such that $d_e^2 f(x|v)(w) = -\infty$. Let $\lambda \in \mathbb{R}$ such that $\lambda < \langle \xi, w \rangle + \beta$. Since $(w, \lambda) \in \mathrm{Epi}(d_e^2 f(x|v)) = \limsup \mathrm{Epi}(\Delta_\tau^2 f(x|v))$, we know that there exist $(t_n) \to 0$ and $(w_n, \lambda_n) \to (w, \lambda)$ such that $(w_n, \lambda_n) \in \mathrm{Epi}(\Delta_{t_n}^2 f(x|v))$, that is, $\Delta_{t_n}^2 f(x|v)(w_n) \leq \lambda_n$ for all $n \in \mathbb{N}$. From Definition 32, we get that $\langle \xi_{t_n}, w_n \rangle + \beta_{t_n} \leq \lambda_n$ for all $n \in \mathbb{N}$. Letting $n \to +\infty$ raises a contradiction. We prove that $d_e^2 f(x|v)$ is lower semi-continuous, convex and that $d_e^2 f(x|v)(0) \leq 0$ as in the proof of Proposition 29. Our second aim is to prove that $d_e^2 f(x|v)(0) \geq \beta$. By contradiction, let us assume that $d_e^2 f(x|v)(0) < \beta$. Then, there exists $\varepsilon > 0$ such that $(0, \beta - \varepsilon) \in \mathrm{Epi}(d_e^2 f(x|v)) = \limsup \mathrm{Epi}(\Delta_\tau^2 f(x|v))$. Thus, there exist $(t_n) \to 0$ and $(w_n, \lambda_n) \to (0, \beta - \varepsilon)$ such that $(w_n, \lambda_n) \in \mathrm{Epi}(\Delta_{t_n}^2 f(x|v))$, that is, $\Delta_{t_n}^2 f(x|v)(w_n) \leq \lambda_n$ for all $n \in \mathbb{N}$. From Definition 32, we get that $\langle \xi_{t_n}, w_n \rangle + \beta_{t_n} \leq \lambda_n$ for all $n \in \mathbb{N}$. Letting $n \to +\infty$ constitutes a contradiction. Our last aim is to prove that ∂f is proto-differentiable at x for v with $D_p(\partial f)(x|v) = \partial(d_e^2 f(x|v))$. This result directly follows from Proposition 28 and Theorem 12.

Now let us assume that ∂f is proto-differentiable at x for v and $D_p(\partial f)(x|v)$ is a maximal monotone operator. From Proposition 28, we get that $(\partial(\Delta_\tau^2 f(x|v)))_{\tau>0} = (\Delta_\tau(\partial f)(x|v))_{\tau>0}$ G-converges to the maximal monotone operator $D_p(\partial f)(x|v)$. We deduce from Theorem 13 that $D_p(\partial f)(x|v) = \partial \varphi$ for some $\varphi \in \Gamma_0(\mathrm{H})$. Our aim is now to apply Theorem 12 in order to conclude that $(\Delta_\tau^2 f(x|v))_{\tau>0}$ M-converges. From Definition 32, we have $(z_\tau, \xi_\tau) \to (z, \xi)$ and since $\xi_\tau \in \partial(\Delta_\tau^2 f(x|v))(z_\tau)$, we deduce that $(z_\tau, \xi_\tau) \in \mathrm{Gr}(\partial(\Delta_\tau^2 f(x|v)))$ for all $\tau > 0$. From Remark 8, we deduce that $(z, \xi) \in \mathrm{Gr}(\partial \varphi)$ and thus $z \in \mathrm{dom}(\varphi)$. Considering $\psi := \varphi - \varphi(z) + \langle \xi, z \rangle + \beta \in \Gamma_0(\mathrm{H})$, we have $\partial \psi = \partial \varphi$ and $\Delta_\tau^2 f(x|v)(z_\tau) = \langle \xi_\tau, z_\tau \rangle + \beta_\tau \to \psi(z)$ since $(z_\tau, \xi_\tau, \beta_\tau)_{\tau>0}$ is a convergent supporting hyperplane of $(\Delta_\tau^2 f(x|v))_{\tau>0}$. The proof is thereby completed using Theorem 12. $\square$

Remark 35. In the t-independent framework, Theorem 34 exactly coincides with Proposition 29. Indeed, in that case, Lemma 30 ensures that $(\Delta_\tau^2 f(x|v))_{\tau>0}$ admits a convergent supporting hyperplane with $(z_\tau, \xi_\tau, \beta_\tau) = (0, 0, 0)$ for all $\tau > 0$.

Remark 36. In Example 31, one can easily see that $(\Delta_\tau^2 f(x|v))_{\tau>0}$ does not admit a convergent supporting hyperplane.

Remark 37. In contrary to Proposition 29, it might be possible that $d_e^2 f(x|v)(0) \neq 0$ in Theorem 34. We refer to Example 38 below.

Example 38. Let $\mathrm{H} = \mathbb{R}$ and $f(t, x) := |x - t^2|$ for all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}$. Let us consider $x = 0$ and $v = 0 \in \partial f(0, x)$. One can easily compute that

$$\Delta_\tau^2 f(x|v)(w) = \frac{|w - \tau| - \tau}{\tau},$$

for all $\tau > 0$ and all $w \in \mathbb{R}$. One can easily deduce that f is twice epi-differentiable at x for v with

$$d_e^2 f(x|v)(w) = \begin{cases} -1 & \text{if } w = 0, \\ +\infty & \text{if } w \neq 0. \end{cases}$$

In particular note that $d_e^2 f(x|v)(0) \neq 0$ and thus $d_e^2 f(x|v)$ is not positively homogeneous of degree two.

We conclude this section by proving that the concept of *convergent supporting hyperplane* is a necessary and sufficient condition in a sense given by the following proposition.

Proposition 39. *Let $f \in \Gamma_0(\cdot, \mathrm{H})$, $x \in f^{-1}(\cdot, \mathbb{R})$ and $v \in \partial f(0, x)$. If f is twice epi-differentiable at x for v, then the following assertions are equivalent:*

(i) $d_e^2 f(x|v) \in \Gamma_0(\mathrm{H})$;

(ii) $(\Delta_\tau^2 f(x|v))_{\tau > 0}$ *admits a convergent supporting hyperplane.*

Proof. From Theorem 34, we only need to prove the implication (i)$\Rightarrow$(ii). Let us assume that $d_e^2 f(x|v) \in \Gamma_0(\mathrm{H})$. From Theorem 12, we deduce that there exists $(z_\tau, \xi_\tau)_{\tau > 0} \to (z, \xi)$ such that $(z_\tau, \xi_\tau) \in \mathrm{Gr}(\partial(\Delta_\tau^2 f(x|v)))$ for all $\tau > 0$ and $\Delta_\tau^2 f(x|v)(z_\tau) \to d_e^2 f(x|v)(z)$. From Remark 14, we know that $d_e^2 f(x|v)(z) \in \mathbb{R}$. This concludes the proof from Remark 33. $\qquad \square$

4.2 Proto-differentiability of the proximity operator and sensitivity analysis

Let $(A, f) \in \mathcal{A}(\cdot, \mathrm{H}) \times \Gamma_0(\cdot, \mathrm{H})$. For the simplicity of notations we introduce

$$\begin{aligned} \mathrm{Prox}_{A,f} : \quad \mathbb{R}_+ \times \mathrm{H} \quad &\to \quad \mathrm{H} \\ (t, x) \quad &\mapsto \quad \mathrm{Prox}_{A,f}(t, x) := \mathrm{prox}_{A(t, \cdot), f(t, \cdot)}(x). \end{aligned}$$

Using the notations introduced at the beginning of Section 3, note that $\mathrm{Prox}_{A,f} = (A + \partial f)^{-1}$. From Propositions 21 and 22, one can easily conclude the following lemma.

Lemma 40. *Let $(A, f) \in \mathcal{A}(\cdot, \mathrm{H}) \times \Gamma_0(\cdot, \mathrm{H})$ and $x \in \mathrm{H}$. Let us assume that A is semi-differentiable at $v := \mathrm{Prox}_{A,f}(0, x)$. Then, $\mathrm{Prox}_{A,f}$ is proto-differentiable at x for v if and only if ∂f is proto-differentiable at v for $v_0 := x - A(0, v) \in \partial f(0, v)$. In that case, it holds that*

$$D_p(\mathrm{Prox}_{A,f})(x|v) = \Big(D_s A(v) + D_p(\partial f)(v|v_0)\Big)^{-1}.$$

From Theorem 34, Proposition 39 and Lemma 40, we deduce the following theorem.

Theorem 41. *Let $(A, f) \in \mathcal{A}(\cdot, \mathrm{H}) \times \Gamma_0(\cdot, \mathrm{H})$ and $x \in \mathrm{H}$. If the following assertions are satisfied:*

(i) $A \in \mathcal{A}_{\mathrm{unif}}(\cdot, \mathrm{H})$;

(ii) A *is semi-differentiable at* $v := \mathrm{Prox}_{A,f}(0, x)$;

(iii) f *is twice epi-differentiable at v for $v_0 := x - A(0, v) \in \partial f(0, v)$;*

(iv) $d_e^2 f(v|v_0) \in \Gamma_0(\mathrm{H})$;

17

then $\mathrm{Prox}_{A,f}$ *is proto-differentiable at* x *for* v *with*

$$D_p(\mathrm{Prox}_{A,f})(x|v) = \mathrm{prox}_{D_s A(v),\, d_e^2 f(v|v_0)}.$$

Proof. From Proposition 17, we know that $D_s A(v) \in \mathcal{A}(\mathrm{H})$. Hence, the mapping $\mathrm{prox}_{D_s A(v),\, d_e^2 f(v|v_0)}$: $\mathrm{H} \to \mathrm{H}$ is well-defined (see Section 2.4). The proof of Theorem 41 easily follows from Theorem 34, Proposition 39 and Lemma 40. $\qquad\square$

Now we return to the initial motivation of the present paper, that is, the sensitivity analysis, with respect to the parameter $t \geq 0$, of the general nonlinear variational inequality of second kind given by

$$(\mathrm{VI}(A(t,\cdot), f(t,\cdot), x(t))) \qquad\qquad \forall z \in \mathrm{H}, \quad \langle A(t,y), z - y \rangle + f(t,z) - f(t,y) \geq \langle x(t), z - y \rangle,$$

where $(A, f) \in \mathcal{A}(\cdot, \mathrm{H}) \times \Gamma_0(\cdot, \mathrm{H})$ and where $x : \mathbb{R}_+ \to \mathrm{H}$ is a given function. From Section 2.4, the above variational inequality admits for all $t \geq 0$ a unique solution $y(t) \in \mathrm{H}$ given by

$$y(t) = \mathrm{prox}_{A(t,\cdot), f(t,\cdot)}(x(t)) = \mathrm{Prox}_{A,f}(t, x(t)).$$

The next theorem is the major result of the present paper. It provides sufficient conditions on A, f and x under which $y : \mathbb{R}_+ \to \mathrm{H}$ is differentiable at $t = 0$ and provides an explicit formula for $y'(0)$.

Theorem 42. *Let* $(A, f) \in \mathcal{A}(\cdot, \mathrm{H}) \times \Gamma_0(\cdot, \mathrm{H})$ *and let* $x : \mathbb{R}_+ \to \mathrm{H}$ *be a function. We consider the function* $y : \mathbb{R}_+ \to \mathrm{H}$ *defined by*

$$y(t) := \mathrm{prox}_{A(t,\cdot), f(t,\cdot)}(x(t)),$$

for all $t \geq 0$. *If the following assertions are satisfied:*

  (i) x *is differentiable at* $t = 0$;

  (ii) $A \in \mathcal{A}_{\mathrm{unif}}(\cdot, \mathrm{H})$;

  (iii) A *is semi-differentiable at* $y(0)$;

  (iv) f *is twice epi-differentiable at* $y(0)$ *for* $v_0 := x(0) - A(0, y(0)) \in \partial f(0, y(0))$;

  (v) $d_e^2 f(y(0)|v_0) \in \Gamma_0(\mathrm{H})$;

then $y : \mathbb{R}_+ \to \mathrm{H}$ *is differentiable at* $t = 0$ *with*

$$y'(0) = \mathrm{prox}_{D_s A(y(0)),\, d_e^2 f(y(0)|v_0)}(x'(0)).$$

Proof. For the ease of notations, we denote by $\mathrm{prox}_\chi := \mathrm{prox}_{D_s A(y(0)),\, d_e^2 f(y(0)|v_0)}$. Recall that prox_χ is $\frac{1}{\alpha}$-Lipschitz continuous where $\alpha > 0$ denotes the uniform strong monotonicity coefficient of $A \in \mathcal{A}_{\mathrm{unif}}(\cdot, \mathrm{H})$ (see Section 2.4 and Remark 18). From Theorem 41, we deduce that $\mathrm{Prox}_{A,f}$ is proto-differentiable at $x(0)$ for $y(0)$ with

$$D_p(\mathrm{Prox}_{A,f})(x(0)|y(0)) = \mathrm{prox}_\chi.$$

By contradiction, let us assume that there exist $\varepsilon > 0$ and $(t_n) \to 0$ such that

$$\varepsilon \leq \left\| \frac{y(t_n) - y(0)}{t_n} - \mathrm{prox}_\chi(x'(0)) \right\|,$$

18

for all $n \in \mathbb{N}$. Since $\mathrm{prox}_\chi(x'(0)) = D_p(\mathrm{Prox}_{A,f})(x(0)|y(0))(x'(0))$, we deduce that there exist $(w_n, z_n)_{n \in \mathbb{N}} \to (x'(0), \mathrm{prox}_\chi(x'(0)))$ and $N \in \mathbb{N}$ such that

$$(w_n, z_n) \in \mathrm{Gr}\Big(\Delta_{t_n}(\mathrm{Prox}_{A,f})(x(0)|y(0))\Big)$$

for all $n \geq N$. We deduce that

$$\mathrm{prox}_{A(t_n,\cdot),f(t_n,\cdot)}(x(0) + t_n w_n) = y(0) + t_n z_n,$$

for all $n \geq N$. Finally, we obtain that

$$\varepsilon \leq \left\| \frac{\mathrm{prox}_{A(t_n,\cdot),f(t_n,\cdot)}(x(t_n)) - \mathrm{prox}_{A(t_n,\cdot),f(t_n,\cdot)}(x(0) + t_n w_n)}{t_n} \right\|$$
$$+ \left\| \frac{\mathrm{prox}_{A(t_n,\cdot),f(t_n,\cdot)}(x(0) + t_n w_n) - y(0)}{t_n} - z_n \right\| + \|z_n - \mathrm{prox}_\chi(x'(0))\|,$$

for all $n \in \mathbb{N}$. The second term is equal to zero for all $n \geq N$. Using the $\frac{1}{\alpha}$-Lipschitz continuity of $\mathrm{prox}_{A(t_n,\cdot),f(t_n,\cdot)}$, we obtain that $\varepsilon \leq \frac{1}{\alpha}\|\frac{x(t_n) - x(0)}{t_n} - w_n\| + \|z_n - \mathrm{prox}_\chi(x'(0))\|$ for all $n \geq N$. This raises a contradiction since $w_n \to x'(0)$ and $z_n \to \mathrm{prox}_\chi(x'(0))$. The proof is complete. $\qquad\square$

Remark 43. Let us assume that all assumptions of Theorem 42 are satisfied. Then its conclusion can be rewritten as follows. If $y(t)$ is the unique solution of the variational inequality $(\mathrm{VI}(A(t,\cdot), f(t,\cdot), x(t)))$ for all $t \geq 0$, then $y : \mathbb{R}_+ \to \mathrm{H}$ is differentiable at $t = 0$ and $y'(0)$ is the unique solution of $(\mathrm{VI}(D_s A(y(0)), d_e^2 f(y(0)|v_0), x'(0)))$.

5 Applications to parameterized convex minimization problems

In this section our aim is to apply our main result (Theorem 42) to the sensitivity analysis of parameterized convex minimization problems. We start with a general result (see Proposition 44 in Section 5.1), then we give a simple illustration in the one-dimensional setting (see Section 5.2). We conclude by studying in details the case of parameterized smooth convex minimization problems with inequality constraints in a finite-dimensional setting (see Proposition 47 in Section 5.3).

5.1 A general result

In this section we will prove from Theorem 42 that the derivative of the solution of a parameterized convex minimization problem is still, under some appropriate assumptions, the solution of a convex minimization problem. The following proposition is in this sense.

Proposition 44. Let $f \in \Gamma_0(\cdot, \mathrm{H})$ and $\ell : \mathbb{R}_+ \to \mathrm{H}$ be given functions. Let $g : \mathbb{R}_+ \times \mathrm{H} \to \mathbb{R}$ be such that $g(t,\cdot)$ is differentiable on H with $\nabla_x g(t,\cdot) \in \mathcal{A}(\mathrm{H})$ for all $t \geq 0$. Then, for all $t \geq 0$, the parameterized convex minimization problem

$$(\mathcal{M}_1(t)) \qquad \underset{x \in \mathrm{H}}{\mathrm{argmin}} \left[f(t,x) + g(t,x) - \langle \ell(t), x \rangle \right]$$

admits a unique solution denoted by $y(t)$. If moreover the following assumptions are satisfied:

(i) ℓ is differentiable at $t = 0$;

(ii) $\nabla_x g \in \mathcal{A}_{\mathrm{unif}}(\cdot, \mathrm{H})$;

(iii) $\nabla_x g$ is of class C^1 on $\mathbb{R}_+ \times \mathrm{H}$;

(iv) f is twice epi-differentiable at $y(0)$ for $v_0 := \ell(0) - \nabla_x g(0, y(0)) \in \partial f(0, y(0))$;

(v) $d_e^2 f(y(0), v_0) \in \Gamma_0(\mathrm{H})$;

then $y : \mathbb{R}_+ \to \mathrm{H}$ is differentiable at $t = 0$ and $y'(0)$ is the unique solution of the convex minimization problem given by

$$(\mathcal{M}_1'(0)) \qquad \operatorname*{argmin}_{x \in \mathrm{H}} \left[d_e^2 f(y(0), v_0)(x) + \frac{1}{2} \langle \nabla_{xx}^2 g(0, y(0))(x), x \rangle \right.$$
$$\left. + \langle \nabla_{tx}^2 g(0, y(0)) - \ell'(0), x \rangle \right].$$

Proof. Since $\nabla_x g(t, \cdot)$ is monotone, we deduce that $g(t, \cdot)$ is convex on H for all $t \geq 0$. As a consequence, for all $t \geq 0$, $y(t)$ is a solution of the convex minimization problem $(\mathcal{M}_1(t))$ if and only if

$$0 \in \partial \Big(f(t, \cdot) + g(t, \cdot) - \langle \ell(t), \cdot \rangle \Big)(y(t)) = \partial f(t, y(t)) + \partial g(t, y(t)) - \ell(t).$$

In the above equation, since $g(t, \cdot)$ is differentiable and $\langle \ell(t), \cdot \rangle$ is linear continuous, the subdifferential of the sum is equal to the sum of the subdifferentials. Denoting by $A(t, \cdot) := \partial g(t, \cdot) = \nabla_x g(t, \cdot) \in \mathcal{A}(\mathrm{H})$, one can easily conclude that the convex minimization problem $(\mathcal{M}_1(t))$ admits a unique solution $y(t)$ given by

$$y(t) = \operatorname{prox}_{A(t, \cdot), f(t, \cdot)}(\ell(t)).$$

Let us prove the second part of Proposition 44. Since $\nabla_x g$ is of class C^1 on $\mathbb{R}_+ \times \mathrm{H}$, one can easily prove that A is semi-differentiable on H with

$$D_s A(x)(w) = \nabla_{xx}^2 g(0, x)(w) + \nabla_{tx}^2 g(0, x),$$

for all x, $w \in \mathrm{H}$. From Theorem 42, we deduce that $y : \mathbb{R}_+ \to \mathrm{H}$ is differentiable at $t = 0$ with

$$y'(0) = \operatorname{prox}_{D_s A(y(0)), \, d_e^2 f(y(0)|v_0)}(\ell'(0)).$$

We deduce that

$$\ell'(0) \in D_s A(y(0))(y'(0)) + \partial \Big(d_e^2 f(y(0)|v_0) \Big)(y'(0)).$$

Since the function $x \in \mathrm{H} \mapsto \frac{1}{2} \langle \nabla_{xx}^2 g(0, y(0))(x), x \rangle \in \mathbb{R}$ is convex and differentiable on H, we obtain that

$$0 \in \partial \Big(d_e^2 f(y(0)|v_0) + \frac{1}{2} \langle \nabla_{xx}^2 g(0, y(0))(\cdot), \cdot \rangle + \langle \nabla_{tx}^2 g(0, y(0)) - \ell'(0), \cdot \rangle \Big)(y'(0)),$$

which concludes the proof. $\qquad \square$

5.2 Illustration with a one-dimensional example

In this section we denote by $x^+ := \max(0, x)$ and by $x^- := \min(0, x)$ for all $x \in \mathbb{R}$.

In order to illustrate Proposition 44, we study in this section a one-dimensional example. Precisely we consider the parameterized one-dimensional convex minimization problem given by

$$(\mathcal{M}_2(t)) \qquad \underset{x \in \mathbb{R}}{\operatorname{argmin}} \left(a(t)|x - b(t)| + \frac{c(t)}{2}x^2 - d(t)x \right),$$

where

- $a : \mathbb{R}_+ \to \mathbb{R}$ is differentiable at $t = 0$ and $a(t) > 0$ for all $t \geq 0$;

- $b : \mathbb{R}_+ \to \mathbb{R}$ is twice differentiable at $t = 0$ and $b'(0) = 0$;

- $c : \mathbb{R}_+ \to \mathbb{R}$ is of class C^1 on $\mathbb{R}_+$ and $\alpha \leq c(t) \leq \beta$ for all $t \geq 0$ for some $0 < \alpha \leq \beta$;

- $d : \mathbb{R}_+ \to \mathbb{R}$ is differentiable at $t = 0$.

In order to apply Proposition 44, one has to consider $\mathrm{H} = \mathbb{R}$ and

$$f(t, x) := a(t)|x - b(t)|, \quad g(t, x) := \frac{c(t)}{2}x^2 \quad \text{and} \quad \ell(t) := d(t),$$

for all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}$. Almost all hypotheses of Proposition 44 are easily checkable. Actually the only difficult part is to check that f is twice epi-differentiable and that its second epi-derivative belongs to $\Gamma_0(\mathrm{H})$. This follows from the following lemma.

Lemma 45. *Let $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ be defined by $f(t, x) := a(t)|x - b(t)|$ for all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}$, where*

- $a : \mathbb{R}_+ \to \mathbb{R}$ *is differentiable at $t = 0$ and $a(t) > 0$ for all $t \geq 0$;*

- $b : \mathbb{R}_+ \to \mathbb{R}$ *is twice differentiable at $t = 0$ and $b'(0) = 0$.*

Then f is twice epi-differentiable at every $x \in \mathbb{R}$ for every $v \in \partial f(0, x)$ and we obtain five different cases:

(i) *If $x > b(0)$, then $v = a(0)$ and we obtain that $d_e^2 f(x|v)(w) = a'(0)w$ for all $w \in \mathbb{R}$;*

(ii) *If $x = b(0)$, then $v \in [-a(0), a(0)]$.*

    (a) *If $v = a(0)$, we obtain that*

$$d_e^2 f(x|v)(w) = \begin{cases} a'(0)w - a(0)(b''(0))^+ & \text{if} \quad w \geq 0, \\ +\infty & \text{if} \quad w < 0, \end{cases}$$

    for all $w \in \mathbb{R}$.

    (b) *If $v \in (-a(0), a(0))$, we obtain that*

$$d_e^2 f(x|v)(w) = \begin{cases} \left(\dfrac{a(0) - v}{2} \right) b''(0) - a(0)(b''(0))^+ & \text{if} \quad w = 0, \\ +\infty & \text{if} \quad w \neq 0, \end{cases}$$

    for all $w \in \mathbb{R}$.

21

(c) If $v = -a(0)$, we obtain that

$$d_e^2 f(x|v)(w) = \begin{cases} -a'(0)w - a(0)(-b''(0))^+ & \text{if} \quad w \le 0, \\ +\infty & \text{if} \quad w > 0, \end{cases}$$

for all $w \in \mathbb{R}$.

(iii) If $x < b(0)$, then $v = -a(0)$ and we obtain that $d_e^2 f(x|v)(w) = -a'(0)w$ for all $w \in \mathbb{R}$.

In each above case, it holds that $d_e^2 f(x|v) \in \Gamma_0(\mathrm{H})$.

Proof. One can easily compute that

$$\Delta_\tau^2 f(x|v)(w) = \begin{cases} \left(\dfrac{a(\tau) - v}{\tau}\right) w - 2a(\tau) \left(\dfrac{b(\tau) - x}{\tau^2}\right)^+ & \text{if} \quad w \ge \dfrac{b(\tau) - x}{\tau}, \\ -\left(\dfrac{a(\tau) + v}{\tau}\right) w - 2a(\tau) \left(\dfrac{x - b(\tau)}{\tau^2}\right)^+ & \text{if} \quad w \le \dfrac{b(\tau) - x}{\tau}, \end{cases}$$

for all $\tau > 0$, all $(w, x) \in \mathbb{R} \times \mathbb{R}$ and all $v \in \partial f(0, x)$. Lemma 45 follows. $\qquad\square$

Remark 46. Note that Lemma 45 encompasses the result of Example 38. In Example 31, all assumptions of Lemma 45 are satisfied except $b'(0) = 0$.

From Proposition 44 and Lemma 45, we obtain that, for all $t \ge 0$, the convex minimization problem $(\mathcal{M}_2(t))$ admits a unique solution denoted by $y(t)$. Moreover, we know that $y : \mathbb{R}_+ \to \mathrm{H}$ is differentiable at $t = 0$ and that $y'(0)$ is solution of the convex minimization problem

$$(\mathcal{M}_2'(0)) \qquad \underset{x \in \mathbb{R}}{\operatorname{argmin}} \left[d_e^2 f\Big(y(0) | d(0) - c(0)y(0)\Big)(x) + \frac{c(0)}{2}x^2 + (c'(0)y(0) - d'(0))x \right].$$

At this step recall that $d(0) - c(0)y(0) \in \partial f(0, y(0))$. Thus we get three cases:

  (i) if $y(0) > b(0)$, then $d(0) - c(0)y(0) = a(0)$;

  (ii) if $y(0) = b(0)$, then $d(0) - c(0)y(0) \in [-a(0), a(0)]$;

  (iii) if $y(0) < b(0)$, then $d(0) - c(0)y(0) = -a(0)$.

Finally, from Lemma 45 and $(\mathcal{M}_2'(0))$, we obtain that

$$y'(0) = \begin{cases} \left(\dfrac{d-a}{c}\right)'(0) & \text{if} \quad y(0) > b(0), \\[2mm] \left(\left(\dfrac{d-a}{c}\right)'(0)\right)^+ & \text{if} \quad y(0) = b(0) = \dfrac{d(0) - a(0)}{c(0)}, \\[2mm] 0 & \text{if} \quad y(0) = b(0) \in \left(\dfrac{d(0) - a(0)}{c(0)}, \dfrac{d(0) + a(0)}{c(0)}\right), \\[2mm] \left(\left(\dfrac{d+a}{c}\right)'(0)\right)^- & \text{if} \quad y(0) = b(0) = \dfrac{d(0) + a(0)}{c(0)}, \\[2mm] \left(\dfrac{d+a}{c}\right)'(0) & \text{if} \quad y(0) < b(0). \end{cases}$$

The above results are perfectly coherent since it can be proved that the unique solution $y(t)$ of the parameterized convex minimization problem $(\mathcal{M}_2(t))$ is given by

$$
y(t) = \begin{cases}
\dfrac{d(t) - a(t)}{c(t)} & \text{if} \quad b(t) \leq \dfrac{d(t) - a(t)}{c(t)}, \\[3mm]
b(t) & \text{if} \quad \dfrac{d(t) - a(t)}{c(t)} < b(t) < \dfrac{d(t) + a(t)}{c(t)}, \\[3mm]
\dfrac{d(t) + a(t)}{c(t)} & \text{if} \quad \dfrac{d(t) + a(t)}{c(t)} \leq b(t),
\end{cases}
$$

for all $t \geq 0$.

5.3 Applications to parameterized smooth convex minimization problems with inequality constraints

Let m, $d \in \mathbb{N}^*$. In the whole section we denote by $\mathbb{R}_-$ the set of all nonpositive real numbers and by $\mathbb{R}_-^d := \mathbb{R}_- \times \ldots \times \mathbb{R}_-$.

In this section we focus on a general finite-dimensional parameterized convex minimization problem with inequality constraints given by

$$
\underset{\substack{x \in \mathbb{R}^m \\ F(t,x) \in \mathbb{R}_-^d}}{\operatorname{argmin}} \ g(t, x)
$$

where $F := (F_i)_{i=1,\ldots,d} : \mathbb{R}_+ \times \mathbb{R}^m \to \mathbb{R}^d$ and $g : \mathbb{R}_+ \times \mathbb{R}^m \to \mathbb{R}$ are smooth functions.

Precisely, under some appropriate assumptions, we will prove in Proposition 47 that the above minimization problem admits, for all $t \geq 0$, a unique solution denoted by $y(t)$, and that the function $y : \mathbb{R}_+ \to \mathbb{R}^m$ is differentiable at $t = 0$ where $y'(0)$ is the unique solution of a convex minimization problem with linear inequality/equality constraints. We refer to Proposition 47 and Remarks 48 and 49 for details.

To this aim, we need to recall first some classical notions and notations. Let $N(w)$ (resp. $T(w)$) denote the *normal cone* (resp. *tangent cone*) to $\mathbb{R}_-^d$ at some $w \in \mathbb{R}_-^d$. Recall that $N(w)$ and $T(w)$ are both closed convex cones of $\mathbb{R}^d$ containing $0_{\mathbb{R}^d}$, and that they are mutually polar. Let $\delta_Y \in \Gamma_0(\mathbb{R}^d)$ (resp. $\sigma_Y \in \Gamma_0(\mathbb{R}^d)$) denote the *indicator function* (resp. *support function*) of a nonempty closed convex subset $Y \subset \mathbb{R}_-^d$. We refer to standard books like [7, 17, 26] and references therein for details.

If F is of class C^2 on $\mathbb{R}_+ \times \mathbb{R}^m$, we denote by $D^2 F(t, x)$ the classical second-order Fréchet differential of F at $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^m$. We also introduce the following sets:

$$
K(y|v) := \{x \in \mathbb{R}^m \mid \nabla_x F(0, y)x \in T(F(0, y)) \text{ and } \langle x, v \rangle = 0\}
$$

and

$$
Y(y|v) := \{w \in \mathbb{R}^d \mid w \in N(F(0, y)) \text{ and } \nabla_x F(0, y)^\top w = v\},
$$

for all $(y, v) \in \mathbb{R}^m \times \mathbb{R}^m$ such that $F(0, y) \in \mathbb{R}_-^d$.

We are now in position to state and prove the main result of this section.

Proposition 47. *Let $m, d \in \mathbb{N}^*$. Let $F := (F_i)_{i=1,\ldots,d} : \mathbb{R}_+ \times \mathbb{R}^m \to \mathbb{R}^d$ be such that $F_i \in \Gamma_0(\cdot, \mathbb{R}^m)$ for every $i = 1, \ldots, d$. We assume that $C(t) := \{x \in \mathbb{R}^m \mid F(t,x) \in \mathbb{R}^d_-\}$ is not empty for all $t \geq 0$. Let $g : \mathbb{R}_+ \times \mathbb{R}^m \to \mathbb{R}$ be such that $g(t, \cdot)$ is differentiable on $\mathbb{R}^m$ with $\nabla_x g(t, \cdot) \in \mathcal{A}(\mathbb{R}^m)$ for all $t \geq 0$. Then, for all $t \geq 0$, the parameterized convex minimization problem with inequality constraints*

$$(\mathcal{M}_3(t)) \qquad \underset{\substack{x \in \mathbb{R}^m \\ F(t,x) \in \mathbb{R}^d_-}}{\operatorname{argmin}} \ g(t,x)$$

admits a unique solution denoted by $y(t)$. If moreover the following assumptions are satisfied:

(i) $\nabla_x g \in \mathcal{A}_{\mathrm{unif}}(\cdot, \mathbb{R}^m)$;

(ii) $\nabla_x g$ *is of class C^1 on $\mathbb{R}_+ \times \mathbb{R}^m$;*

(iii) F *is of class C^2 on $\mathbb{R}_+ \times \mathbb{R}^m$;*

(iv) $\nabla_t F(0, y(0)) = 0_{\mathbb{R}^d}$;

(v) $y(0) \in C(t)$ *for all $t \geq 0$;*

(vi) $\|\nabla_x F(t, y(0))^\top w\|_{\mathbb{R}^m} \geq \alpha$ *for all $w \in \mathrm{N}(F(t, y(0)))$ with $\|w\|_{\mathbb{R}^d} = 1$ and all $t \geq 0$, for some $\alpha > 0$;*

then $y : \mathbb{R}_+ \to \mathbb{R}^m$ is differentiable at $t = 0$ and $y'(0)$ is the unique solution of the convex minimization problem with linear inequality/equality constraints given by

$$(\mathcal{M}'_3(0)) \qquad \underset{\substack{x \in \mathbb{R}^m \\ x \in \mathrm{K}(y(0)|v_0)}}{\operatorname{argmin}} \left[\frac{1}{2} \langle \nabla^2_{xx} g(0, y(0))(x), x \rangle + \langle \nabla^2_{tx} g(0, y(0)), x \rangle \right.$$
$$\left. + \sigma_{\mathrm{Y}(y(0)|v_0)} \left(\frac{1}{2} D^2 F(0, y(0))(1, x) \right) \right],$$

where $v_0 := -\nabla_x g(0, y(0))$.

Proof. Let us prove the first part of Proposition 47. In order to apply the first part of Proposition 44 with $\mathrm{H} = \mathbb{R}^m$, we introduce $f : \mathbb{R}_+ \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ defined by $f(t,x) := \delta_{C(t)}(x) = \delta_{\mathbb{R}^d_-}(F(t,x))$ for all $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^m$. Thus, the minimization problem $(\mathcal{M}_3(t))$ can be rewritten as

$$\underset{x \in \mathbb{R}^m}{\operatorname{argmin}} \left[f(t,x) + g(t,x) \right].$$

Since $F_i \in \Gamma_0(\cdot, \mathbb{R}^m)$ for every $i = 1, \ldots, d$ and since $C(t)$ is not empty for all $t \geq 0$, one can easily deduce that $f \in \Gamma_0(\cdot, \mathbb{R}^m)$ and thus the first part of Proposition 44 can be applied.

Now let us prove the second part of Proposition 47. In order to apply the second part of Proposition 44 with $\mathrm{H} = \mathbb{R}^m$, we only need to prove that f is twice epi-differentiable at $y(0)$ for v_0 and that $d^2_e f(y(0), v_0) \in \Gamma_0(\mathbb{R}^m)$. To do so, we will essentially adapt the proof of [26, Theorem 13.14 p.594] to the t-dependent framework.

First of all, we need to state some assertions:

- Note that $Y(y(0)|v_0)$ is a closed convex subset of $\mathbb{R}^d$. Moreover, since $v_0 = -\nabla_x g(0, y(0)) \in \partial f(0, y(0)) = \nabla_x F(0, y(0))^\top N(F(0, y(0)))$, we deduce that $Y(y(0)|v_0)$ is nonempty. Finally, since $\|\nabla_x F(0, y(0))^\top x\|_{\mathbb{R}^m} \geq \alpha$ for all $x \in N(F(0, y(0)))$ with $\|x\|_{\mathbb{R}^d} = 1$, one can easily prove that $Y(y(0)|v_0)$ is compact. In particular, it follows that $\sigma_{Y(y(0)|v_0)}(x) = \max_{w \in Y(y(0)|v_0)} \langle w, x \rangle_{\mathbb{R}^d} < +\infty$ for all $x \in \mathbb{R}^d$.

- Note that $K(y(0)|v_0)$ is a closed convex cone of $\mathbb{R}^m$ containing $0_{\mathbb{R}^m}$. Then we can easily deduce that the function $\varphi : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$, defined by

$$\varphi(x) := \delta_{K(y(0)|v_0)}(x) + \sigma_{Y(y(0)|v_0)}\left(\frac{1}{2}D^2 F(0, y(0))(1, x)\right),$$

for all $x \in \mathbb{R}^m$, is such that $\varphi \in \Gamma_0(\mathbb{R}^m)$. In particular we used the fact that each component of some $w \in Y(y(0)|v_0)$ are nonnegative.

- For all $w \in Y(y(0)|v_0)$, $\delta_{\mathbb{R}^d_-}$ is twice epi-differentiable at $F(0, y(0))$ for w and it holds that

$$d_e^2 \delta_{\mathbb{R}^d_-}(F(0, y(0))|w)(\nabla_x F(0, y(0))x) = \delta_{K(y(0)|v_0)}(x),$$

for all $x \in \mathbb{R}^m$. We refer to [26, Exercice 13.17 p.600] for details. In particular, note that the above formula is independent of the choice of $w \in Y(y(0)|v_0)$.

Our aim from now is to prove that f is twice epi-differentiable at $y(0)$ for v_0 with $d_e^2 f(y(0)|v_0) = \varphi$ from the characterization of epi-convergence recalled in Proposition 11 (with $H = \mathbb{R}^m$ is finite-dimensional). Firstly, let $z \in \mathbb{R}^m$ and $(z_\tau)_{\tau>0} \to z$. Adapting the proof of [26, Theorem 13.14 p.594] to the t-dependent framework, one can easily show that

$$\Delta_\tau^2 f(y(0)|v_0)(z_\tau) = \Delta_\tau^2 \delta_{\mathbb{R}^d_-}(F(0, y(0))|w)(\tilde{\Delta}_\tau F) + \left\langle w, \frac{\tau \tilde{\Delta}_\tau F - \tau \nabla_x F(0, y(0))z_\tau}{\tau^2} \right\rangle,$$

for any $w \in Y(y(0)|v_0)$, where $\tilde{\Delta}_\tau F := \frac{F(\tau, y(0)+\tau z_\tau) - F(0, y(0))}{\tau}$. Since $\nabla_t F(0, y(0)) = 0_{\mathbb{R}^d}$, one has $\tilde{\Delta}_\tau F \to \nabla_x F(0, y(0))z$. Moreover, since $\delta_{\mathbb{R}^d_-}$ is twice epi-differentiable at $F(0, y(0))$ for w and since $\frac{\tau \tilde{\Delta}_\tau F - \tau \nabla_x F(0, y(0))z_\tau}{\tau^2} \to \frac{1}{2}D^2 F(0, y(0))(1, z)$, we deduce that

$$\liminf_{\tau \to 0} \Delta_\tau^2 f(y(0)|v_0)(z_\tau) \geq d_e^2 \delta_{\mathbb{R}^d_-}(F(0, y(0))|w)(\nabla_x F(0, y(0))z)$$

$$+ \frac{1}{2}\langle w, D^2 F(0, y(0))(1, z) \rangle.$$

Since $d_e^2 \delta_{\mathbb{R}^d_-}(F(0, y(0))|w)(\nabla_x F(0, y(0))z) = \delta_{K(y(0)|v_0)}(z)$ and since the last inequality is satisfied for any $w \in Y(y(0)|v_0)$ compact, we get that $\liminf_{\tau \to 0} \Delta_\tau^2 f(y(0)|v_0)(z_\tau) \geq \varphi(z)$.

Secondly, our objective is now to exhibit a sequence $(z_\tau)_{\tau>0} \to z$ that satisfies $\limsup_{\tau \to 0} \Delta_\tau^2 f(y(0)|v_0)(z_\tau) \leq \varphi(z)$. If $\varphi(z) = +\infty$, nothing has to be proved. Thus we assume from now that $\varphi(z) < +\infty$, that is, $z \in K(y(0)|v_0)$. It is sufficient to construct a sequence $(z_\tau)_{\tau>0} \to z$ such that

$$\lim_{\tau \to 0} \Delta_\tau^2 f(y(0)|v_0)(z_\tau) = \frac{1}{2}\langle \overline{w}, D^2 F(0, y(0))(1, z) \rangle$$

where $\overline{w} \in \text{argmax}_{w \in Y(y(0)|v_0)} \langle w, D^2 F(0, y(0))(1, z) \rangle$. From [26, Theorem 11.42 and Example 11.43 p.506], we deduce the existence of $\overline{z} \in \mathbb{R}^m$ such that $-\langle v_0, \overline{z} \rangle = \langle \overline{w}, D^2 F(0, y(0))(1, z) \rangle$ and $\eta :=$

$\nabla_x F(0, y(0))\overline{z} + D^2 F(0, y(0))(1, z) \in \mathrm{T}(F(0, y(0)))$. Since $\eta \in \mathrm{T}(F(0, y(0)))$ and $\nabla_x F(0, y(0))z \in \mathrm{T}(F(0, y(0)))$ and from [26, Theorem 13.11 and Proposition 13.12 p.591-592], there exists $\lambda : \mathbb{R}_+ \to \mathbb{R}^d$ such that $\lambda(t) \in \mathbb{R}^d_-$ for all $t \geq 0$ and $\lambda(0) = F(0, y(0))$, $\lambda'(0) = \nabla_x F(0, y(0))z$ and $\lambda''(0) = \eta$. On the other hand, from Assumption (vi) and from the metric regularity property (see [26, Theorem 9.43 and Example 9.44 p.387-388]), we know that

$$\mathrm{d}\left(y(0) + tz + \frac{t^2}{2}\overline{z}, C(t) \right) \leq \frac{1}{\alpha}\mathrm{d}\left(F\left(t, y(0) + tz + \frac{t^2}{2}\overline{z} \right), \mathbb{R}^d_- \right),$$

for sufficiently small $t \geq 0$. Thus we deduce that

$$\mathrm{d}\left(\overline{z}, \frac{C(t) - y(0) - tz}{t^2/2} \right) \leq \frac{1}{\alpha}\mathrm{d}\left(\tilde{\Delta}^2_t F, \frac{\mathbb{R}^d_- - F(0, y(0)) - t\nabla_x F(0, y(0))z}{t^2/2} \right)$$

$$\leq \frac{1}{\alpha}\left\| \tilde{\Delta}^2_t F - \frac{\lambda(t) - \lambda(0) - t\lambda'(0)}{t^2/2} \right\|_{\mathbb{R}^d}$$

for sufficiently small $t \geq 0$, where $\tilde{\Delta}^2_t F := \frac{F(t, y(0) + tz + \frac{t^2}{2}\overline{z}) - F(0, y(0)) - t\nabla_x F(0, y(0))z}{t^2/2}$. Since $\tilde{\Delta}^2_t F \to \eta = \lambda''(0)$, we deduce that there exists $\xi : \mathbb{R}_+ \to \mathbb{R}^m$ such that $\xi(t) \in C(t)$ for all $t \geq 0$ and $\xi(0) = y(0)$, $\xi'(0) = z$ and $\xi''(0) = \overline{z}$. Finally, we define $z_\tau := \frac{\xi(\tau) - \xi(0)}{\tau} \to \xi'(0) = z$. Since $F(t, y(0)) \in \mathbb{R}^d_-$ and $F(t, \xi(t)) \in \mathbb{R}^d_-$ for all $t \geq 0$ and since $\langle z, v_0 \rangle = 0$, we get that

$$\Delta^2_\tau f(y(0)|v_0)(z_\tau) = -\left\langle v_0, \frac{z_\tau - z}{\tau} \right\rangle \longrightarrow -\frac{1}{2}\langle v_0, \overline{z} \rangle = \frac{1}{2}\langle \overline{w}, D^2 F(0, y(0))(1, z) \rangle,$$

which concludes the proof. $\square$

Remark 48. Note that the constraints $x \in \mathrm{K}(y(0)|v_0)$ in the convex minimization problem $(\mathcal{M}'_3(0))$ can be written as linear inequality/equality constraints.

Remark 49. Let us assume that all hypotheses of Proposition 47 are satisfied. In that case, recall that $\mathrm{Y}(y(0)|v_0)$ is a nonempty compact convex subset of $\mathbb{R}^d$ and thus

$$\sigma_{\mathrm{Y}(y(0)|v_0)}\left(\frac{1}{2}D^2 F(0, y(0))(1, x) \right) = \frac{1}{2}\max_{w \in \mathrm{Y}(y(0)|v_0)}\langle w, D^2 F(0, y(0))(1, x) \rangle,$$

for all $x \in \mathbb{R}^m$. In particular, if $D^2 F(0, y(0)) = 0$, then the above term vanishes. In that particular case, the convex minimization problem $(\mathcal{M}'_3(0))$ has a quadratic cost with linear inequality/equality constraints (see Remark 48), while the original parameterized convex minimization problem $(\mathcal{M}_3(t))$ has a nonlinear cost with nonlinear inequality constraints.

6 Additional comments

We conclude this paper with some additional comments about the choice of Formula (1) and about the twice epi-differentiability of the conjugate function.

6.1 Comments on the choice of Formula (1)

In the whole paragraph we consider $f \in \Gamma_0(\cdot, \mathrm{H})$, $x \in f^{-1}(\cdot, \mathbb{R})$ and $v \in \partial f(0, x)$.

As mentioned at the beginning of Section 4, the whole paper is based on Attouch's theorems (see Theorems 12 and 13) and on Proposition 28. One can easily see that these results are totally independent of the definition of $\Delta_\tau^2 f(x|v)$, provided that $\Delta_\tau^2 f(x|v) \in \Gamma_0(\mathrm{H})$ and $\partial(\Delta_\tau^2 f(x|v)) = \Delta_\tau(\partial f)(x|v)$.

As a consequence it is clear that one could easily adapt the whole paper with a different definition of $\Delta_\tau^2 f(x|v)$, as long as $\Delta_\tau^2 f(x|v) \in \Gamma_0(\mathrm{H})$ and $\partial(\Delta_\tau^2 f(x|v)) = \Delta_\tau(\partial f)(x|v)$. For example, instead of Formula (1), one could consider

$$\Delta_\tau^2 f(x|v)(w) := \frac{f(\tau, x + \tau w) - f(0, x) - \tau\langle v, w\rangle}{\tau^2}, \tag{2}$$

or

$$\Delta_\tau^2 f(x|v)(w) := \frac{f(\tau, x + \tau w) - f(\tau, x) - \tau\langle v(\tau), w\rangle}{\tau^2}, \tag{3}$$

where $v(t) \in \partial f(t, x)$ for all $t \geq 0$. This is the reason why we think that it is important to justify our choice of Formula (1). Actually this choice was natural and has prevailed with respect to our initial motivation and followed from the case where f is smooth.

Indeed let us assume that f is smooth on $\mathbb{R}^+ \times \mathrm{H}$ and let $y, z : \mathbb{R}_+ \to \mathrm{H}$ be two given functions differentiable at $t = 0$ that are related by the expression

$$y(t) = \mathrm{prox}_{f(t,\cdot)}(z(t)) = \mathrm{prox}_{\mathrm{Id}, f(t,\cdot)}(z(t)),$$

for all $t \geq 0$. We deduce that $y(t) + \nabla_x f(t, y(t)) = z(t)$ for all $t \geq 0$ and thus $y'(0) + \nabla_{tx}^2 f(0, y(0)) + \nabla_{xx}^2 f(0, y(0))(y'(0)) = z'(0)$. Finally we obtain that $y'(0) = (\mathrm{Id} + \partial\varphi_{y(0)})^{-1}(z'(0)) = \mathrm{prox}_{\varphi_{y(0)}}(z'(0))$ where $\varphi_x : \mathrm{H} \to \mathbb{R}$ is defined by

$$\varphi_x(w) := \frac{1}{2}\left\langle \nabla_{xx}^2 f(0, x)(w), w \right\rangle + \left\langle \nabla_{tx}^2 f(0, x), w \right\rangle,$$

for all $w \in \mathrm{H}$. Hence it was natural to choose a general expression of $\Delta_\tau^2 f(x|v)$ that converges pointwise on H to φ_x (up to an additive constant) whenever f is smooth. Note that Formula (1) does (see Remark 26).

In contrast, if f is smooth and if $\Delta_\tau^2 f(x|v)(w)$ is defined as in (2), then $\Delta_\tau^2 f(x|v)(w)$ does not converge if $\nabla_t f(0, x) \neq 0$. Similarly, if f is smooth and if $\Delta_\tau^2 f(x|v)(w)$ is defined as in (3), then $\Delta_\tau^2 f(x|v)(w)$ converges to $\frac{1}{2}\langle \nabla_{xx}^2 f(0, x)(w), w\rangle$ that is different (even up to an additive constant) from $\varphi_x(w)$ if $\nabla_{tx}^2 f(0, x) \neq 0$.

6.2 Comments on the twice epi-differentiability of the conjugate function

In the t-independent framework, recall that the classical *Fenchel conjugate* $f^* \in \Gamma_0(\mathrm{H})$ of a function $f \in \Gamma_0(\mathrm{H})$ is defined by

$$f^*(v) := \sup_{x \in \mathrm{H}}\{\langle v, x\rangle - f(x)\},$$

for all $v \in \mathrm{H}$. Recall the following result stated by U. Mosco in [15, Theorem 1].

Theorem 50. *Let $(f_\tau)_{\tau > 0}$ be a parameterized family of functions in $\Gamma_0(\mathrm{H})$ and $f \in \Gamma_0(\mathrm{H})$. Then, $f = \mathrm{E\text{-}lim}\, f_\tau$ if and only if $f^* = \mathrm{E\text{-}lim}\, f_\tau^*$.*

It can be easily proved that $f^*(v) + f(x) \geq \langle v, x \rangle$ for all $(x, v) \in \mathrm{H} \times \mathrm{H}$, with equality if and only if $v \in \partial f(x)$. From elementary calculus rules for conjugates, for some $x \in f^{-1}(\mathbb{R})$ and $v \in \partial f(x)$, one can get that the conjugate of $\Delta^2_\tau f(x|v) \in \Gamma_0(\mathrm{H})$ is given by

$$\left(\Delta^2_\tau f(x|v) \right)^* = \Delta^2_\tau f^*(v|x),$$

for all $\tau > 0$. As a consequence, we obtain from Theorem 50 that f is twice epi-differentiable at x for v if and only if f^* is twice epi-differentiable at v for x. In that case, $d^2_e f(x|v)$ and $d^2_e f^*(v|x)$ both belong to $\Gamma_0(\mathrm{H})$ from Proposition 29 and are mutually conjugate, that is, $(d^2_e f(x|v))^* = d^2_e f^*(v|x)$. We refer to [26, Theorem 13.21 p.604] for details.

In the t-dependent framework, for the simplicity of notations, we denote by $f^* : \mathbb{R}_+ \times \mathrm{H} \to \mathbb{R} \cup \{+\infty\}$ the conjugate of any function $f \in \Gamma_0(\cdot, \mathrm{H})$ defined by

$$f^*(t, v) := (f(t, \cdot))^*(v)$$

for all $(t, v) \in \mathbb{R}_+ \times \mathrm{H}$. Our aim in this section is to rely the twice epi-differentiability introduced in this paper (see Definition 23) of f with the twice epi-differentiability of f^* as in the t-independent setting. However, as in Section 4.1, we will see that the situation is more complicated in the t-dependent framework and requires additional assumptions.

Let $f \in \Gamma_0(\cdot, \mathrm{H})$, $x \in f^{-1}(\cdot, \mathbb{R})$ and $v \in \partial f(0, x)$. We introduce the nonnegative function

$$\begin{aligned} \Phi^f_{(x|v)} : \quad \mathbb{R}_+ \quad &\to \quad \mathbb{R}_+ \\ t \quad &\mapsto \quad \Phi^f_{(x|v)}(t) := f^*(t, v) + f(t, x) - \langle v, x \rangle. \end{aligned}$$

Note that $\Phi^f_{(x|v)}(0) = 0$. In the particular case where f is t-independent, it holds that $\Phi^f_{(x|v)}(t) = 0$ for all $t \geq 0$. Imitating the above t-independent framework, it can be proved that

$$\left(\Delta^2_\tau f(x|v) \right)^* = \Delta^2_\tau f^*(v|x) + \frac{\Phi^f_{(x|v)}(\tau)}{\tau^2},$$

for all $\tau > 0$. We deduce the following proposition.

Proposition 51. *Let* $f \in \Gamma_0(\cdot, \mathrm{H})$, $x \in f^{-1}(\cdot, \mathbb{R})$ *and* $v \in \partial f(0, x)$. *If* $\Phi^f_{(x|v)}$ *is twice differentiable at* $t = 0$ *with* $(\Phi^f_{(x|v)})'(0) = 0$, *then* f *is twice epi-differentiable at* x *for* v *with* $d^2_e f(x|v) \in \Gamma_0(\mathrm{H})$ *if and only if* f^* *is twice epi-differentiable at* v *for* x *with* $d^2_e f^*(v|x) \in \Gamma_0(\mathrm{H})$. *In that case*

$$(d^2_e f(x|v))^* = d^2_e f^*(v|x) + \frac{1}{2}(\Phi^f_{(x|v)})''(0).$$

References

[1] H. Attouch. *Variational convergence for functions and operators.* Applicable Mathematics Series. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[2] H. Brezis. Équations et inéquations non linéaires dans les espaces vectoriels en dualité. *Ann. Inst. Fourier (Grenoble)*, 18(1):115–175, 1968.

[3] C. N. Do. Generalized second-order derivatives of convex functions in reflexive Banach spaces. *Trans. Amer. Math. Soc.*, 334(1):281–301, 1992.

[4] S. Fitzpatrick and R. R. Phelps. Differentiability of the metric projection in Hilbert space. *Trans. Amer. Math. Soc.*, 270(2):483–501, 1982.

[5] A. Haraux. How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities. *J. Math. Soc. Japan*, 29(4):615–631, 1977.

[6] J.-B. Hiriart-Urruty. Unsolved problems: at what points is the projection mapping differentiable? *Amer. Math. Monthly*, 89(7):456–458, 1982.

[7] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of convex analysis.* Grundlehren Text Editions. Springer-Verlag, Berlin, 2001.

[8] J.-B. Hiriart-Urruty and A. Seeger. Calculus rules on a new set-valued second order derivative for convex functions. *Nonlinear Anal.*, 13(6):721–738, 1989.

[9] A. J. King and R. T. Rockafellar. Sensitivity analysis for nonsmooth generalized equations. *Math. Programming*, 55(2, Ser. A):193–212, 1992.

[10] A. B. Levy and R. T. Rockafellar. Sensitivity analysis of solutions to generalized equations. *Trans. Amer. Math. Soc.*, 345(2):661–671, 1994.

[11] F. Mignot. Contrôle dans les inéquations variationelles elliptiques. *J. Functional Analysis*, 22(2):130–185, 1976.

[12] B. S. Mordukhovich and R. T. Rockafellar. Second-order subdifferential calculus with applications to tilt stability in optimization. *SIAM J. Optim.*, 22(3):953–986, 2012.

[13] J.-J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France*, 93:273–299, 1965.

[14] U. Mosco. Convergence of convex sets and of solutions of variational inequalities. *Advances in Math.*, 3:510–585, 1969.

[15] U. Mosco. On the continuity of the Young-Fenchel transform. *J. Math. Anal. Appl.*, 35:518–535, 1971.

[16] J.-P. Penot. Differentiability of relations and differential stability of perturbed optimization problems. *SIAM J. Control Optim.*, 22(4):529–551, 1984.

[17] J.-P. Penot. *Calculus without derivatives*, volume 266 of *Graduate Texts in Mathematics*. Springer, New York, 2013.

[18] J.-P. Penot. *Analysis. From concepts to applications.* Universitext. Springer, 2016.

[19] R. Poliquin and T. Rockafellar. Second-order nonsmooth analysis in nonlinear programming. In *Recent advances in nonsmooth optimization*, pages 322–349. World Sci. Publ., River Edge, NJ, 1995.

[20] R. A. Poliquin and R. T. Rockafellar. A calculus of epi-derivatives applicable to optimization. *Canad. J. Math.*, 45(4):879–896, 1993.

[21] R. T. Rockafellar. On the maximal monotonicity of subdifferential mappings. *Pacific J. Math.*, 33:209–216, 1970.

[22] R. T. Rockafellar. Maximal monotone relations and the second derivatives of nonsmooth functions. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 2(3):167–184, 1985.

[23] R. T. Rockafellar. Proto-differentiability of set-valued mappings and its applications in optimization. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 6:449–482, 1989. Analyse non linéaire (Perpignan, 1987).

[24] R. T. Rockafellar. Generalized second derivatives of convex functions and saddle functions. *Trans. Amer. Math. Soc.*, 322(1):51–77, 1990.

[25] R. T. Rockafellar. Second-order convex analysis. J. Nonlinear Convex Anal., 1(1):116, 2000.

[26] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.

[27] A. Shapiro. Directionally nondifferentiable metric projection. J. Optim. Theory Appl., 81(1):203204, 1994.

[28] A. Shapiro. Differentiability properties of metric projections onto convex sets. J. Optim. Theory Appl., 169(3):953964, 2016.

[29] E. H. Zarantonello. Projections on convex sets in Hilbert space and spectral theory. I. Projections on convex sets. pages 237–341, 1971.