UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

A NUMERICAL ALGORITHM FOR MIRROR SWEEPING PROCESSES

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN MATEMÁTICAS APLICADAS

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL MATEMÁTICO

ALDO PATRICIO DANIEL GUTIÉRREZ CONCHA

PROFESOR GUÍA:
EMILIO VILCHES GUTIÉRREZ

PROFESOR CO-GUÍA:
PEDRO PÉREZ AROS

MIEMBROS DE LA COMISIÓN:
SEBASTIÁN DONOSO FUENTES
RAFAEL CORREA FONTECILLA

Este trabajo ha sido parcialmente financiado por CMM ANID BASAL FB210005

SANTIAGO DE CHILE
2023

RESUMEN DE LA TESIS PARA OPTAR
AL GRADO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA,
MENCIÓN MATEMÁTICAS APLICADAS Y
MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL MATEMÁTICO
POR: ALDO PATRICIO DANIEL GUTIÉRREZ CONCHA
FECHA: 2023
PROFESOR GUÍA: EMILIO VILCHES GUTIÉRREZ
PROFESOR CO-GUÍA: PEDRO PERÉZ AROS

Un algoritmo númerico para Procesos de Arrastre de Espejo

Esta tesis explora un caso particular del Proceso de Arraste Degenerado, un problema de evolución descrito por una inclusión diferencial, en un espacio de Hilbert de dimensión infinita. La dinámica involucra un operador representado por el gradiente de una función semicontinua inferior, convexa y propia, inspirada en el mapeo utilizado en el algoritmo de descenso espejo. Nos referimos a esta dinámica como *Mirror Sweeping Process*. Para comenzar introducimos un operador, que llamamos *aproximador*, que generaliza la proyección a conjuntos y mostramos algunas caracterizaciones y propiedades. A continuación mostramos las condiciones necesarias para la suavidad de la preimagen de un conjunto convexo. Finalmente, construimos una familia de soluciones aproximadas al *Mirror Sweeping Process* mediante un algoritmo númerico gobernado por el operador *aproximador*. Demostramos que esta familia converge uniformemente a la única solución del Proceso de Arrastre Degenerado, la cual no requiere ni compacidad de los conjuntos en movimiento ni linealidad del operador.

RESUMEN DE LA TESIS PARA OPTAR
AL GRADO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA,
MENCIÓN MATEMÁTICAS APLICADAS Y
MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL MATEMÁTICO
POR: ALDO PATRICIO DANIEL GUTIÉRREZ CONCHA
FECHA: 2023
PROFESOR GUÍA: EMILIO VILCHES GUTIÉRREZ
PROFESOR CO-GUÍA: PEDRO PERÉZ AROS

A numerical algorithm for mirror sweeping processes

This thesis explores a particular case of the Degenerate Sweeping Process, an evolution problem described by a differential inclusion, in an infinite-dimensional Hilbert space. The dynamic involves an operator represented by the gradient of a lower semicontinuous, convex and proper function, inspired by the mapping used in the mirror descent algorithm. We refer to this dynamic as the *Mirror Sweeping Process*. To begin we introduce an operator, which we call *approximator*, that generalizes the projection to sets and show some characterization and properties. Following that we show necessary conditions for the smoothness of the preimage of a convex set. Finally, we construct a family of approximated solutions to the Mirror Sweeping Process through numerical algorithm governed by the *approximator* operator. We show that this family converges uniformly to the one and only solution of the Degenerate Sweeping Process, which does not require neither compactness of the sets or linearity of the operator.

"The true delight is in the finding out rather than in the knowing"

—Isaac Asimov

Acknowledgments

Comienzo por el comienzo, deseo agradecerle a mi familia, a mi madre por su amor y apoyo incondicional y mostrarme cómo ser mejor persona, a mi padre por darme todo lo que ha podido y por hacerme del colo, a mi hermano Boris por ver películas conmigo cuando era pequeño y ahora de grande verme por quién soy, a mi hermano Bastían por mostrarme que siempre se pueden aprender cosas nuevas y darme una infancia llena de risas, a mi hermana Fernanda por siempre acompañarme en lo que sea, cuando sea y cómo sea. Mi amor por ustedes no cabe en esta tesis.

Agradezco a mis animales, quienes me hacen más humano, Kiara, Sashi, Leia, Woody, Akira, María-Terminator y Anakin.

A Carolina, quien me acompañó en esta tesis, me acompañará en la que vendrá y después también. Gracias por los momentos de paz y de diversión y por hacerme siempre sentir en casa. Gracias por todos los Cauchy-Schwarz y confirmarme que, efectivamente, estaba haciendo bien cada desigualdad triangular. Gracias por todo.

A mis amigos Luquitas y Rodrigo, el primero por acompañarme en cada etapa, por mostrarme el Pokémon ranger y por ser una gran persona en la que puedo confiar, el segundo por todas los "una chelita?", la música compartida y por la informalidad en la severidad, a ambos por su constante y paciente amistad. A Luis y Felipe por compartir el gusto por las matemáticas desde un comienzo. A la Barbi por ser mi primera amiga real de la universidad y por hacer conmigo el Juanjo bot, al Juanjo por las horas de distensión en los primeros años, incluido el Mario Party, y darle el nombre al robot.

A mis amigos del DIM, Gonzalo, el compañero de broma de corta naturaleza, el futbolero lo entiende, gracias por darme la mano y hacerme saber cuando soy chistoso, a Fabían por ser el alma de la fiesta y Benja por las conversaciones nutritivas, a los 3 por sus chistes irreproducibles, Mariano por darme mate y no romperme el cráneo, Manu por toda la cantidad inconmesurable de matraca, Víctor porque su conversación de matraca es un conjunto despreciable y siempre tener buena música en la oficina, a Jurgen por compartir mi aflicción, J por mostrarme memes que no entendía e ir a ver las bolas de los dragones, a Leo por ser una persona sensata y siempre traer grandes juegos, a Pedro por ser una gran persona y siempre escuchar atentamente.

Agradezco a los miembros de la comisión, por aceptar ser parte de ella. Quiero también agradecer a mi profesor guía Emilio, por mostrarme este tema tan entretenido, integrarme a la comunidad del área y ser verdaremente un guía en este proceso. A Pedro Pérez-Aros por sus clases en los seminarios, que me hicieron interesarme en el área, a Sebastían Donoso porque sus clases de teoría de la medida me hicieron volver a creer en mí en esta carrera.

Por último agradecer a Luchito Mena por mostrarme que la perserverancia y el trabajo duro siempre pagan, a Colo-Colo por darme más glorias que penas y Ferrari por darme más penas que glorias y una razón para levantarme los domingos, aunque sea solo para sufrir.

Table of Content

List of Figures

Introduction

The main topic in this thesis concerns the sweeping process, a first-order dynamical system where the normal cone to a moving set is the main component, first introduced in 1971 by Jean Jacques Moreau in a series of papers (see [44, 46–48]) in order to model a variety of elastoplastic mechanical systems (see [41, 45]). The sweeping process is described by the differential inclusion:

$$\begin{cases} \dot{x}(t) \in -N(C(t); x(t)) & \text{a.e. } t \in [0, T]; \\ x(0) = x_0 \in C(0), \end{cases} \tag{$\mathcal{SP}$}$$

which can be seen as a generalized Cauchy problem. The mapping $C\colon [0, T] \rightrightarrows \mathcal{H}$ is a multifunction with nonempty, closed and convex values, and $\mathcal{H}$ is a Hilbert space. The dynamic ($\mathcal{SP}$) describes an element inside C being dragged by the moving set, as it is being swept, in order to stay in the moving set, it must move in an inward normal direction. The reader can consult [35] for an extensive overview. Several variations have been introduced over the years, such as the second order sweeping process, (see e.g., [13, 14, 20, 21]), state-dependent sweeping process, (see e.g., [12, 32, 38]) and degenerate sweeping process (see [33, 39, 40]), the latter is defined by the following differential inclusion:

$$\begin{cases} \dot{x}(t) \in -N(C(t); \mathcal{A}x(t)) + f(t, x(t)) & \text{a.e. } t \in [0, T]; \\ x(0) = x_0 \in \mathcal{A}^{-1}C(0), \end{cases} \tag{$\mathcal{DSP}$}$$

in the same context as ($\mathcal{SP}$), the operator $A\colon \mathcal{H} \to \mathcal{H}$ is usually a Lipschitz, strongly monotone operator and $f\colon [0, T] \times \mathcal{H} \to \mathcal{H}$ is a perturbation function with Lipschitz-like conditions. The case where $\mathcal{A} = I$ and $f \equiv 0$ turns ($\mathcal{DSP}$) into the classical formulation ($\mathcal{SP}$). The term "degenerate" comes from the fact that, even in seemingly easy scenarios, it can fail to admit a solution. For example (see [39]), in the case where

$$\mathcal{H} = \mathbb{R}^2,\ T = 1,\ \mathcal{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } C(t) = [0, 1] \times [t, 1],\ 0 \leq t \leq 1,$$

the operator $\mathcal{A}$ is linear, bounded, self-adjoint and monotone. The moving set is Lipschitzian with respect to the Hausdorff distance, but there is no solution for the starting point $x_0 = (0, 0)$, if this were the case, necessarily it should hold that $Ax(t) = (x_1(t), 0) \in C(t)$ for a.e. $t \in [0, 1]$ which is a contradiction.

The approach to obtain a solution of a sweeping process usually falls between the use of a Moreau-Yosida regularization technique, where one finds a solution of and approximated problem, as in [24, 32, 50, 51], or, a discretization scheme known as the catching-up algorithm.

To introduce this algorithm, for each natural number n, one constructs an approximant of the derivative trough a sequence of discretized points $(x_k^n)_{0\leq k\leq n}$. Each element (x_k^n) must be in $C(t_k^n)$, then, whenever x_k^n is not in the moving set evaluated at the next time-step, that is $x_k^n \notin C(t_{k+1}^n)$, one projects to the moving set in order for the next point to fulfill the condition, namely, $x_{k+1} = \text{proj}_{C(t_{k+1}^n)}(x_k^n)$. The sequence of points will follow the inclusion:

$$\frac{x_{k+1}^n - x_k^n}{t_{k+1}^n - t_k} \in -N_{C(t_{k+1}^n)}(x_{k+1}^n), \quad k = 0, \ldots, n-1,$$

where the gaps are filled with a linear interpolation. Then taking limit in an appropriate manner will yield a solution (see [3, 33, 46]).

In this thesis we will focus on the degenerate sweeping process, with the multifunction C adopting nonempty, closed and convex values and the operator taking the form $\mathcal{A} = \nabla\varphi$, where $\varphi \in \Gamma_0(\mathcal{H})$ is a twice differentiable function such that $\mathcal{A}$ is a strongly monotone, Lipschitz-continuous operator, the mapping φ is based on the *mirror map* used in the mirror descent algorithm. This particular case is referred to as the Mirror Sweeping Process. We will show the existence and uniqueness of a solution for this differential inclusion through a family of discretized solutions, obtained via a catching-up-like algorithm.

The contribution of this work is threefold, we expand on the *approximator operator* introduced in [39] and drop any compactness hypothesis on the moving set, these are usually needed in the context of a degenerate sweeping process where the operator is possibly nonlinear. For instance, in [37, 39, 40] for existence and uniqueness, the set is required to be compact when intersected with a ball, while in [33] the operator must be linear. In [2] the authors achieve an existence and uniqueness results in a finite dimension setting and although compactness is relaxed, uniqueness is lost in the nonlinear case. It is also of great interest the link that can be made to online optimization, this framework enables algorithms and machine learning models to operate in real-time or dynamic environments.

The document is organized as follows, Chapter 1 contains the necessary mathematical background for the tools used in the main result of the thesis. Chapter 2 introduces the problem in hand and provides some preparatory results. Chapter 3 presents our main result, the existence and uniqueness of a solution to the Mirror Sweeping Process ($\mathcal{MSP}$), in the case where the set-valued operator $C(\cdot)$ takes on nonempty, closed and convex values, through a modified catching-up algorithm. We also present an application to online optimization. Finally, some concluding remarks and possible future contributions are presented.

Chapter 1

Preliminaries

In this chapter, we will introduce some necessary mathematical background for the context of this thesis; these include definitions and known results. We also set the notation that will be used from now on.

Unless stated otherwise, we shall work in the context of a Hilbert space. Most of the definitions given in this chapter work in more general settings, but we reduce to the Hilbert case. From now on, let $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ be a real Hilbert space with associated norm $\|x\| = \sqrt{\langle x, x\rangle}$ and distance $d(x, y) = \|x - y\|$. We denote by $\mathbb{B}$ the closed unit ball and $B(x, r)$ (respectively $B[x, r]$) denotes the open (respectively closed) ball of radius $r > 0$ centered at x. The extended real line is defined as $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$. For a function $f\colon \mathcal{H} \to \overline{\mathbb{R}}$, its domain is denoted as $\operatorname{dom} f := \{x \in \mathcal{H} : f(x) < \infty\}$.

For a given interval $I \subset \mathbb{R}$, we denote the space of Lebesgue integrable functions over I, with values in $\mathcal{H}$, by $L^1(I; \mathcal{H})$. In the special case where $\mathcal{H} = \mathbb{R}$, we employ the notation $L^1(I)$. Furthermore, we define $x\colon I \to \mathcal{H}$ to be absolutely continuous if there exists $f \in L^1(I; \mathcal{H})$ and $x_0 \in \mathcal{H}$ and $t_0 \in I$ such that

$$x(t) = x_0 + \int_{t_0}^{t} f(s)ds \text{ for all } t \in I.$$

The set of absolutely continuous functions is denoted by $x \in \mathrm{AC}(I; \mathcal{H})$.

1.1 Elements of analysis

1.1.1 Notions of distance between sets

In the following, we introduce and give some basic properties for a notion of distance between sets. For $x \in \mathcal{H}$ and $C \subset \mathcal{H}$, the distance from x to C is defined as

$$d(C; x) := \inf\{d(x, y) : y \in C\}.$$

We refer to [22, 30, 58] for further details on distance between sets.

Definition 1.1.1 *Let C_1, C_2 be subsets of $\mathcal{H}$. We define the excess of C_1 over C_2 as*

$$e(C_1, C_2) := \sup\{d(C_2; y) : y \in C_1\}.$$

Similarly, the excess of C_2 over C_1 is

$$e(C_2, C_1) := \sup\{d(C_1; y) : y \in C_2\}.$$

Moreover, the Hausdorff distance between C_1 and C_2 is defined by

$$d_H(C_1, C_2) := \max\{e(C_1, C_2), e(C_2, C_1)\}.$$

Remark 1.1.2 It is worth noting that the excess is not always symmetric. Hence, the Hausdorff distance is not necessarily a distance between sets. For example, when one of the sets is unbounded, the distance function can take the value $+\infty$ (See Figure 1.1).

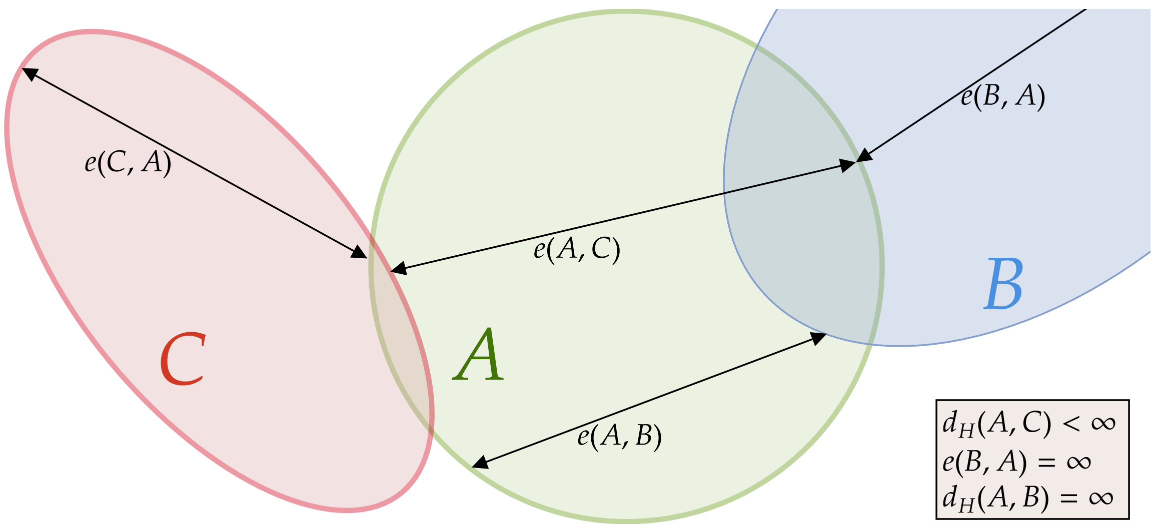

Figure 1.1: Example of the excess and Hausdorff distance where B is an unbounded set.

For a given $\varepsilon > 0$, consider the sets

$$C_1^\varepsilon := \{x \in \mathcal{H} : d(C_1; x) < \varepsilon\} \text{ and } C_2^\varepsilon := \{x \in \mathcal{H} : d(C_2; x) < \varepsilon\}.$$

Then, directly from the definition, we get the following characterization:

$$e(C_1, C_2) = \inf\{\varepsilon > 0 : C_1 \subset C_2^\varepsilon\} \text{ and } e(C_2, C_1) = \inf\{\varepsilon > 0 : \; C_2 \subset C_1^\varepsilon\}.$$

Hence,

$$d_H(C_1, C_2) = \inf\{\varepsilon > 0 : C_1 \subset C_2^\varepsilon, C_2 \subset C_1^\varepsilon\}.$$

Proposition 1.1.3 *Let C_1, C_2, C_3 be subsets of $\mathcal{H}$. Then the following assertions hold:*

(i) $e(C_1, C_2) = 0 \iff C_1 \subset \overline{C}_2$ *and* $d_H(C_1, C_2) = 0 \iff \overline{C}_1 = \overline{C}_2$,

(ii) $e(C_1, C_3) \leq e(C_1, C_2) + e(C_2, C_3)$ *and* $d_H(C_1, C_3) \leq d_H(C_1, C_2) + d_H(C_2, C_3)$.

1.1.2 Convergences

In this subsection, we recall the notion of convergence, mainly of the weak convergence and some related properties, see, e.g., [8, 16, 25].

Definition 1.1.4 *A sequence $(x_n)_n \subset \mathcal{H}$ is called Cauchy if for every $\varepsilon > 0$ there exists a positive natural number N such that*

$$\|x_n - x_n\| < \varepsilon \quad \text{for all } n, m \geq N.$$

In a Hilbert space, every Cauchy sequence is convergent, and the reciprocal statement is always true for an inner product space. It is also known that every Cauchy sequence is bounded, that is,

$$\sup_{n\in\mathbb{N}} \|x_n\| < +\infty.$$

Definition 1.1.5 *A sequence $(x_n)_n \subset \mathcal{H}$ is said to weakly converge to $x \in \mathcal{H}$ if*

$$\langle x_n, y\rangle \to \langle x, y\rangle \quad \text{for all } y \in \mathcal{H}.$$

The weak convergence of $(x_n)_n$ to x is denoted by $x_n \rightharpoonup x$.

The following result summarizes elementary properties of the weak convergence.

Proposition 1.1.6 *Consider a sequence $(x_n)_n \subset \mathcal{H}$ and $x \in \mathcal{H}$. Then*

(i) $x_n \to x \implies x_n \rightharpoonup x$.

(ii) $x_n \to x \iff x_n \rightharpoonup x$ *and* $\|x_n\| \to \|x\|$.

(iii) $x_n \rightharpoonup x \implies (x_n)_n$ *is bounded and* $\|x\| \leq \liminf \|x_n\|$.

(iv) Every bounded sequence in $\mathcal{H}$ has a weakly convergent subsequence.

The following lemma is the celebrated Mazur's lemma (see, e.g., [8, Corollary 3.36]).

Lemma 1.1.7 (Mazur's lemma) *Let $(x_n)_n \subset \mathcal{H}$ be a sequence that converges weakly to $x \in \mathcal{H}$. Then there exists a sequence $(y_n)_n$ of convex combinations of the elements of $(x_n)_n$ which converges strongly to x. Specifically,*

$$y_n = \sum_{k=0}^{n} \alpha_{k,n} x_k \to x,$$

where for all $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} \alpha_{k,n} = 1 \text{ and } (\alpha_{k,n})_{0\leqslant k\leqslant n} \in [0,1]^{n+1} \ \forall n \in \mathbb{N}.$$

Now, we introduce the Dunford-Pettis theorem, a significant result that characterizes weak compactness on $L^1(\mathcal{H})$. In order to enunciate the theorem, first, we need the concept of uniform integrability.

Definition 1.1.8 *We say that $K \subset L^1(\mathcal{H})$, is uniformly integrable if*

(i) K is bounded on $L^1(\mathcal{H})$.

(ii) For every $\varepsilon > 0$, there exists $\delta_\varepsilon > 0$ such that for every measurable set A,

$$\mu(A) < \delta_\varepsilon \implies \sup_{v \in K} \int_A |v(x)| d\mu(x) < \varepsilon.$$

See, e.g., [5, Theorem 2.4.5] for the following version of the Dunford-Pettis Theorem.

Theorem 1.1.9 (Dunford-Pettis) *Let $K \subset L^1(\mathcal{H})$. Then, the following assertions are equivalent:*

(i) K is relatively compact for the weak topology $\sigma(L^1(\mathcal{H}), L^\infty(\mathcal{H}))$.

(ii) K is uniformly integrable.

(iii) For every sequence $(x_n)_n \subset K$, it is possible to extract a convergent subsequence for the topology $\sigma(L^1(\mathcal{H}), L^\infty(\mathcal{H}))$.

1.1.3 Operators

Consider X, Y to be two nonempty sets. A set-valued operator (or mapping) $T \colon X \rightrightarrows 2^Y$ is an object that maps elements from X to subsets of Y. The domain of an operator is the set of elements in $x \in X$ such that Tx is a nonempty set. If for every $x \in \operatorname{dom} T$, the set Tx is a singleton, then T is said to be at most single-valued. Each operator is uniquely characterized by its graph, defined as

$$\operatorname{gph} T := \{(x, y) \in X \times Y : y \in Tx\}.$$

The inverse of an operator T, denoted by T^{-1} is defined by

$$T^{-1}y := \{x \in X : (x, y) \in \operatorname{gph} T\} \quad \forall y \in Y.$$

Now, we recall the motion of monotone operator.

Definition 1.1.10 *Let α be a positive real number. A set-valued operator $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is said to be:*

(i) monotone if

$$\langle y_2 - y_1, x_2 - x_1 \rangle \geq 0, \quad \text{for all } (x_1, y_1), (x_2, y_2) \in \operatorname{gph} T;$$

(ii) maximally monotone if it is monotone and there does not exist a monotone operator $\widetilde{T}$ such that $\operatorname{gph} \widetilde{T}$ properly contains $\operatorname{gph} T$;

(iii) $\alpha-$*strongly monotone if*

$$\langle y_2 - y_1, x_2 - x_1 \rangle \geq \alpha \|x_2 - x_1\|^2, \quad \textit{for all } (x_1, y_1), (x_2, y_2) \in \operatorname{gph} T;$$

(iv) $\alpha-$*cocoercive if*

$$\langle y_2 - y_1, x_2 - x_1 \rangle \geq \alpha \|y_2 - y_1\|^2, \quad \textit{for all } (x_1, y_1), (x_2, y_2) \in \operatorname{gph} T.$$

We introduce the notion of the adjoint operator. Firstly, we recall the set of linear continuous mappings, defined for two Hilbert spaces $\mathcal{H}, \mathcal{H}'$, as

$$\mathcal{L}(\mathcal{H}, \mathcal{H}') := \{T \colon \mathcal{H} \to \mathcal{H}' \text{ such that } T \text{ is linear and bounded}\}.$$

When $\mathcal{H}' = \mathcal{H}$, we denote it as $\mathcal{L}(\mathcal{H})$.

Definition 1.1.11 *Let $T \in \mathcal{L}(\mathcal{H})$, the adjoint operator, $T^* \colon \mathcal{H} \to \mathcal{H}$, is the operator characterized by the following equality*

$$\langle T^* y, x \rangle = \langle y, Tx \rangle \quad \textit{for all } x, y \in \mathcal{H}.$$

Whenever $T \equiv T^*$, the operator T is said to be self-adjoint.

1.1.4 Differentiability

The goal of this subsection is to recall the notion of Gâteaux derivative, mainly in order to ensure that the second derivative of the operator involved in the Degenerate Sweeping Process in study is self-adjoint. This affirmation simplifies operations in later chapters.

Definition 1.1.12 *Let $C \subset \mathcal{H}$ and $T \colon C \to \mathbb{R}$. We say that T is Gâteaux differentiable at $x \in C$ if there exists $DT(x) \in \mathcal{L}(\mathcal{H}, \mathbb{R})$ such that, for every $y \in \mathcal{H}$*

$$DT(x)y = \lim_{t \to 0^+} \frac{T(x + ty) - T(x)}{t}.$$

The map $DT(x)$ is the Gâteaux derivative of T at x.

Higher order derivatives are defined in the same way. It is of our concern only up to the second derivative, defined as the operator $DT^2(x) \in \mathcal{L}(\mathcal{H}, \mathcal{L}(\mathcal{H}, \mathbb{R}))$ characterized by

$$DT^2(x)y = \lim_{t \to 0^+} \frac{DT(x + ty) - DT(x)}{t}.$$

Remark 1.1.13 For a Gâteaux differentiable function $f \colon C \to \mathbb{R}$, the gradient operator of f at $x \in C$ is characterized by the following equality

$$Df(x)y = \langle \nabla f(x), y \rangle \quad \forall y \in \mathcal{H}.$$

Furthermore, if f is twice Gâteaux differentiable, then $D^2 f(x)$ is identified with the operator $\nabla^2 f(x)$ defined by the equality:

$$(D^2 f(x)y)z = \langle z, \nabla^2 f(x) y \rangle.$$

When there is no confusion, we will use either notation indistinctively.

In this thesis, when we refer to a differentiable function, we will mean a differentiable function in the sense of Fréchet. It is a well-known fact that if a function is Fréchet differentiable at a point x in its domain, then it is also Gâteaux differentiable at x, and the two derivatives coincide (see, e.g., [8, Lemma 2.61]).

Proposition 1.1.14 *Let $C \subset \mathcal{H}$ be a nonempty set and $f\colon C \to \mathbb{R}$ a function twice Fréchet differentiable at $x \in C$. Then it holds that*

$$(D^2 f(x)y)z = (D^2 f(x)z)y \quad \forall y, z \in \mathcal{H}.$$

As a consequence of the above equality, $\nabla^2 f(x)$ is self-adjoint.

The next notion of differentiability will be needed in the next chapter.

Definition 1.1.15 *A map $f\colon \mathcal{H} \to \mathcal{H}$ is said to be* strictly differentiable *at $\overline{x} \in \operatorname{dom} f$ it here exists a mapping $l \in \mathcal{L}(\mathcal{H})$ such that, for every $x, y \in \operatorname{dom} f$ with $x \neq y$, one has*

$$\lim_{x,y \to \overline{x}} \frac{\|f(x) - f(y) - l(x-y)\|}{\|x-y\|} = 0.$$

Finally, whenever f is continuously differentiable at a point, then it is strictly differentiable at this point (see [55, Proposition 2.56]).

1.2 Elements of Convex Analysis

This section introduces the necessary results and definitions from Convex Analysis. We refer to [8, 11, 28, 43, 55, 61] for further details.

We recall the notion of convexity, along with some variants and basic properties. Consider $x, y \in \mathcal{H}$, we denote by $[x, y]$ the line segment between x and y, defined formally as

$$[x, y] := \{\lambda x + (1-\lambda)y : 0 \leq \lambda \leq 1\}.$$

A set C contained in $\mathcal{H}$ is said to be *convex* if, for every pair of elements, the line segment lies in C, i.e., $[x, y] \subset C$. In turn, a function $f\colon C \to \overline{\mathbb{R}}$ defined over a convex set C is said to be convex if, for every pair of elements $x, y \in C$, one has

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \quad \forall \lambda \in [0, 1].$$

Moreover, if the above inequality is strict, the function is said to be *strictly convex.*

The *convex hull* of a set $C \subset \mathcal{H}$, denoted as $\operatorname{co}(C)$, is defined as the smallest convex set in $\mathcal{H}$ that contains C. The following characterization helps to visualize $\operatorname{co}(C)$ in an explicit manner.

Proposition 1.2.1 *Let C be a subset of $\mathcal{H}$, then the following equality holds*

$$\operatorname{co}(C) = \left\{ x = \sum_{i=1}^{n} \lambda_i x_i \mid n \in \mathbb{N},\ x_1, \ldots, x_n \in C,\ \lambda_1, \ldots, \lambda_n \in \mathbb{R}_+,\ \sum_{i=1}^{n} \lambda_i = 1 \right\}.$$

To establish a direct connection between set and function convexity, we introduce the following notions.

Definition 1.2.2 *Let $f\colon \mathcal{H} \to \overline{\mathbb{R}}$ be a function, with $x \in \operatorname{dom} f$, its epigraph is defined as*

$$\operatorname{epi} f := \{(x, \lambda) \in \mathcal{H} \times \mathbb{R} : f(x) \leq \lambda\}.$$

For $\lambda \in \mathbb{R}$, the level set is defined as,

$$[f \leq \lambda] := \{x \in \mathcal{H} : f(x) \leq \lambda\}.$$

Proposition 1.2.3 *Let $f\colon \mathcal{H} \to \overline{\mathbb{R}}$ be a function, then*

(i) f is convex if and only if $\operatorname{epi} f$ is convex.

(ii) If f is a convex function, then its lower-level sets are convex.

The following concepts are related to minimization problems.

Definition 1.2.4 *Consider a function $f\colon \mathcal{H} \to \overline{\mathbb{R}}$. Then f is said to be*

(i) coercive if

$$\lim_{\|x\| \to +\infty} f(x) = +\infty,$$

(ii) supercoercive if

$$\lim_{\|x\| \to +\infty} \frac{f(x)}{\|x\|} = +\infty,$$

(iii) proper, if its domain is nonempty and it never takes the value $-\infty$.

From the definition, it is straightforward to prove that supercoerciveness implies coerciveness.

Proposition 1.2.5 *Let $f\colon \mathcal{H} \to \overline{\mathbb{R}}$ be a proper function. Suppose that f is differentiable on $\operatorname{dom} f$ and its domain is an open and convex set, then the following are equivalent:*

(i) f is convex.

(ii) For all $x, y \in \operatorname{dom} f$

$$f(x) \geq f(y) + \langle x - y, \nabla f(y)\rangle.$$

(iii) For all $x, y \in \operatorname{dom} f$

$$\langle \nabla f(x) - \nabla f(y), x - y\rangle \geq 0,$$

i.e., the gradient is monotone.

The projection of a point $x \in H$ onto a set $C \subset \mathcal{H}$ is defined as the set of points in C that attain the distance of x to C. Formally, the projection (or best approximation) of $x \in \mathcal{H}$ onto a set C is

$$\operatorname{Proj}_C(x) := \{y \in C : \|y - x\| = d_C(x)\}. \tag{1.1}$$

When the above object is a singleton, we denote it by $\operatorname{proj}_C(x)$.

The following theorem involves the projection onto a nonempty, closed, and convex set. It provides existence, uniqueness, and a characterization of the notion of projection. See, for example, [8, Theorem 3.16] or [27].

Theorem 1.2.6 *Let $C \subset \mathcal{H}$ be a nonempty, closed, and convex set and an element $x \in \mathcal{H}$. Then the projection of x onto C is a singleton. Furthermore, p is the projection of x onto C if and only if*

$$\langle x - p, y - p \rangle \leq 0 \quad \forall y \in \mathcal{H}.$$

As a consequence of the above inequality, we obtain that the projection onto a closed and convex set is Lipschitz continuous with constant 1, i.e.,

$$\|\mathrm{proj}_C(x_1) - \mathrm{proj}_C(x_2)\| \leq \|x_1 - x_2\| \quad \forall x_1, x_2 \in \mathcal{H}. \tag{1.2}$$

Now, we introduce a stronger version of convexity.

Definition 1.2.7 *Let $f \colon \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a proper function, then f is said to be strongly convex with constant $L > 0$ if the function $f - \frac{L}{2}\|\cdot\|^2$ is convex.*

Clearly, if f is strongly convex, then it is strictly convex. The following result improves Proposition 1.2.5 for a strongly convex function.

Proposition 1.2.8 *Let $f \colon \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a proper function and L a positive real number. Then, if f is differentiable on $\mathrm{dom}\, f$ and its domain is an open and convex set, the following are equivalent:*

(i) f *is strongly convex with constant* L.

(ii) $\forall x, y \in \mathrm{dom}\, f : f(x) \geq f(y) + \langle x - y, \nabla f(y) \rangle + \frac{L}{2}\|x - y\|^2$.

(iii) $\forall x, y \in \mathrm{dom}\, f : \langle \nabla f(x) - \nabla f(y), x - y \rangle \geq L\|x - y\|^2$, *i.e., the gradient is* $L-$*strongly monotone.*

The following notion comes in hand in relaxing continuity hypotheses when dealing with minimization problems

Definition 1.2.9 *Let $f \colon \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a function and x an element of $\mathcal{H}$. Then f is lower semicontinuous (lsc) at x if for every sequence $(x_n)_n$ converging to x,*

$$f(x) \leq \liminf_{n \to \infty} f(x_n),$$

or equivalently,

$$\forall \varepsilon > 0, \exists \delta > 0 : y \in B(x, \delta) \implies f(x) - \varepsilon \leq f(y).$$

Proposition 1.2.10 *Let $f \colon \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$, then the following are equivalent*

(i) f *is lsc.*

(ii) $\mathrm{epi}\, f$ *is closed.*

(iii) For every $\lambda \in \mathbb{R}$, *the lower level sets* $[f \leq \lambda]$ *are closed.*

Given a Hilbert space $\mathcal{H}$, we denote by $\Gamma(\mathcal{H})$ the set of convex and lower semicontinuous (lsc) functions with values in the extended real line. It is immediate to see that the functions $f_1 \equiv +\infty$ and $f_2 \equiv -\infty$ lie in $\Gamma(\mathcal{H})$, which is not always desirable. In order to fix this problem, one defines $\Gamma_0(\mathcal{H}) := \Gamma(\mathcal{H}) \setminus \{f_1, f_2\}$, which is equivalent to the definition

$$\Gamma_0(\mathcal{H}) := \{f \colon \mathcal{H} \to \overline{\mathbb{R}} : f \text{ is convex, lsc and proper}\}.$$

We introduce the notion of the Fenchel conjugate, an object widely used in convex analysis.

Definition 1.2.11 *Let $f \colon \mathcal{H} \to \overline{\mathbb{R}}$, the Fenchel conjugate of f is defined as the mapping $f^* \colon \mathcal{H} \to \overline{\mathbb{R}}$, where*

$$f^*(x) := \sup_{y \in \mathcal{H}} \{\langle x, y\rangle - f(y)\}.$$

It is clear from the definition that if there is an element $x \in \mathcal{H}$ such that $f(x) = -\infty$, then $f^* \equiv \infty$. It is also straightforward that f^* is convex and lsc.

The following result gives the main properties of the Fenchel conjugate.

Proposition 1.2.12 *Consider a proper function $f \colon \mathcal{H} \to \overline{\mathbb{R}}$ and $x \in \operatorname{dom} f$. Then the following hold*

(i) $f \in \Gamma(\mathcal{H}) \iff (f^*)^* = f$,

(ii) $f \in \Gamma_0(\mathcal{H}) \iff f^* \in \Gamma_0(\mathcal{H})$.

1.2.1 Subdifferential and Normal Cone

Definition 1.2.13 *Let $f \colon \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a convex function. The (convex) subdifferential of f is the set-valued mapping $\partial f \colon \mathcal{H} \rightrightarrows \mathcal{H}$ defined, for every $x \in \operatorname{dom} f$, as*

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

Whenever $x \notin \operatorname{dom} f$, we set $\partial f(x) = \emptyset$.

The next result is a main resul in convex analysis. It will be useful when dealing with constrained minimization problems (see, e.g., [8]).

Lemma 1.2.14 *Let f, g be functions in $\Gamma_0(\mathcal{H})$ such that one of the following holds:*

(i) $\operatorname{dom} f \cap \operatorname{int} \operatorname{dom} g \neq \emptyset$.

(ii) $\operatorname{dom} g = \mathcal{H}$.

Then, for $x \in \operatorname{dom} f \cap \operatorname{int} \operatorname{dom} g$, one has

$$\partial(f + g)(x) = \partial f(x) + \partial g(x).$$

Now, we present the concept of the normal cone to a convex set.

Definition 1.2.15 *For a nonempty and convex set $C \subset \mathcal{H}$, the normal cone of C at $\overline{x} \in C$ is defined as*

$$N(C;\overline{x}) := \{x \in \mathcal{H} : \langle x, y - \overline{x}\rangle \le 0 \quad \text{for all } y \in C\}.$$

Whenever $\overline{x} \notin C$, we set $N(C;\overline{x}) = \emptyset$.

It follows from the definition that if $\overline{x} \in \operatorname{int} C$, then $N(C;\overline{x}) = \{0\}$. The normal cone appears naturally in a wide range of topics in optimization. For a set $C \subset \mathcal{H}$, the indicator function of C is defined by

$$\delta_C(x) := \begin{cases} 0 & \text{if } x \in C, \\ +\infty, & \text{if } x \notin C. \end{cases}$$

Moreover, the following characterization holds $N(C;\overline{x}) = \partial\delta_C(\overline{x})$.

Proposition 1.2.16 *Let $f \colon \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a convex function. Then, the subdifferential is a monotone operator, i.e.,*

$$\langle u - v, x - y\rangle \ge 0 \quad \forall (x,u),(y,v) \in \operatorname{gph} \partial f.$$

As a consequence of the above inequality, for a nonempty, convex set $C \subset \mathcal{H}$, its normal cone is a monotone operator, i.e., it holds that

$$\langle u - v, x - y\rangle \ge 0, \tag{1.3}$$

for every $u \in N(C;x)$ and $v \in N(C;y)$.

Remark 1.2.17 In view of these results, the inequality in Theorem 1.2.6 turns into

$$p = \operatorname{proj}_C(x) \iff x - p \in N(C;p).$$

1.2.2 Existence of Minimizers

Consider a function $f \colon \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ and let C be a subset of $\mathcal{H}$. In this section, we provide conditions for the existence and optimality of the following problem:

$$\inf_{x \in C} f(x). \tag{$\mathcal{P}$}$$

The following proposition is a generalization of the Fermat rule for an unconstrained minimization problem.

Proposition 1.2.18 *Let $f \colon \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a function with $\overline{x} \in \operatorname{dom} f$. Then $\overline{x}$ is a local minimizer of f if and only if $0 \in \partial f(\overline{x})$.*

Proposition 1.2.19 *Let f be a function in $\Gamma_0(\mathcal{H})$, and $C \subset \mathcal{H}$ a closed convex subset of $\mathcal{H}$ with $C \cap \operatorname{dom} f \neq \emptyset$. Then, if f is coercive or C is bounded, then the set*

$$\operatorname{argmin}_{x \in C} f(x) := \{x^* \in C : \inf_{x \in C} f(x) = f(x^*)\}$$

is nonempty. Consequently, ($\mathcal{P}$) has at least one solution. Moreover, if f is strictly convex, the above set is a singleton, which ensures the uniqueness of the solution of ($\mathcal{P}$).

The following result establishes an optimality condition for the problem $(\mathcal{P})$ (see, e.g., [8, Proposition 27.8]).

Proposition 1.2.20 *Let $f \in \Gamma_0(\mathcal{H})$ and $C \subset \mathcal{H}$ be a closed and convex set. Then if $\overline{x}$ is a minimizer for $(\mathcal{P})$ the following inclusion holds:*

$$0 \in \nabla f(\overline{x}) + N(C; \overline{x}).$$

1.2.3 Legendre functions

In this subsection, we will consider the concept of Legendre function. This concept will play a main role in the study of the Mirror Sweeping Process, primarily due to the duality established between the gradient of a Legendre function and its conjugate. The main reference for this subsection is [10, Chapter 7].

Definition 1.2.21 *A function $f \in \Gamma_0(\mathcal{H})$, is said to be*

(i) essentially smooth if its subdifferential is locally bounded and single-valued on its domain.

(ii) essentially strictly convex if the inverse of its subdifferential is locally bounded on its domain and f is strictly convex on every convex subset of the domain of the subdifferential.

(iii) Legendre if it is essentially smooth and essentially strictly convex.

In finite dimension, the definition is stated as follows. See also [57] for the finite dimension setting.

Definition 1.2.22 *A function $f \in \Gamma_0(\mathbb{R}^n)$, is said to be*

(i) essentially smooth in the classical sense if it is differentiable on $\operatorname{int}\operatorname{dom} f \neq \emptyset$ and $\|\nabla f(x_n)\| \to \infty$ as x_n converges to a point in the boundary of $\operatorname{dom} f$.

(ii) essentially strictly convex in the classical sense if it is strictly convex on every convex subset of the domain of the subdifferential.

(iii) Legendre in the classical sense if it is essentially smooth and essentially strictly convex in the classical sense.

We will give some necessary results for our context. The first lemma allows us to check if a function is essentially strictly convex in an easier way.

Lemma 1.2.23 *Suppose that the function f belongs to $\Gamma_0(\mathcal{H})$. It holds that if f is strictly convex, then it is essentially strictly convex.*

The following theorem relates properties of f and its conjugate.

Theorem 1.2.24 *Let $f \in \Gamma_0(\mathcal{H})$. Then f is essentially smooth if and only if its conjugate f^* essentially strictly convex.*

Corollary 1.2.25 *Consider $f\colon \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ in $\Gamma_0(\mathcal{H})$. Then f is Legendre if and only if f^* is Legendre.*

Remark 1.2.26 Thanks to Proposition 1.2.12, the functions f^* belong to $\Gamma_0(\mathcal{H})$, then it makes sense to talk about f^* being of Legendre type.

The theorem below plays a major role in our setting, allowing us to know in terms of the conjugate, the inverse of the operator ∇f.

Theorem 1.2.27 *Let $f\colon \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a function in $\Gamma_0(\mathcal{H})$. Then, if f is Legendre, the map*

$$\nabla f\colon \operatorname{int}\operatorname{dom} f \to \operatorname{int}\operatorname{dom} f^*$$

is bijective. Moreover, the following equality holds:

$$(\nabla f)^{-1} = \nabla f^*.$$

Using the previous results, we can introduce the concept of the Bregman divergence (see, e.g., [7]).

Definition 1.2.28 *Let $f \in \Gamma_0(\mathcal{H})$ be a differentiable function such that the interior of the domain is nonempty. The Bregman divergence associated to f is defined as:*

$$D_f(x,y) := \begin{cases} f(x) - f(y) - \langle x - y, \nabla f(y)\rangle & \text{if } y \in \operatorname{int}\operatorname{dom} f, \\ +\infty & \text{otherwise.} \end{cases}$$

From the definition, it follows, because of convexity and Proposition 1.2.5 (ii), that the Bregman divergence is always greater or equal than 0.

Figure 1.2 illustrates the notion of Bregman distance.

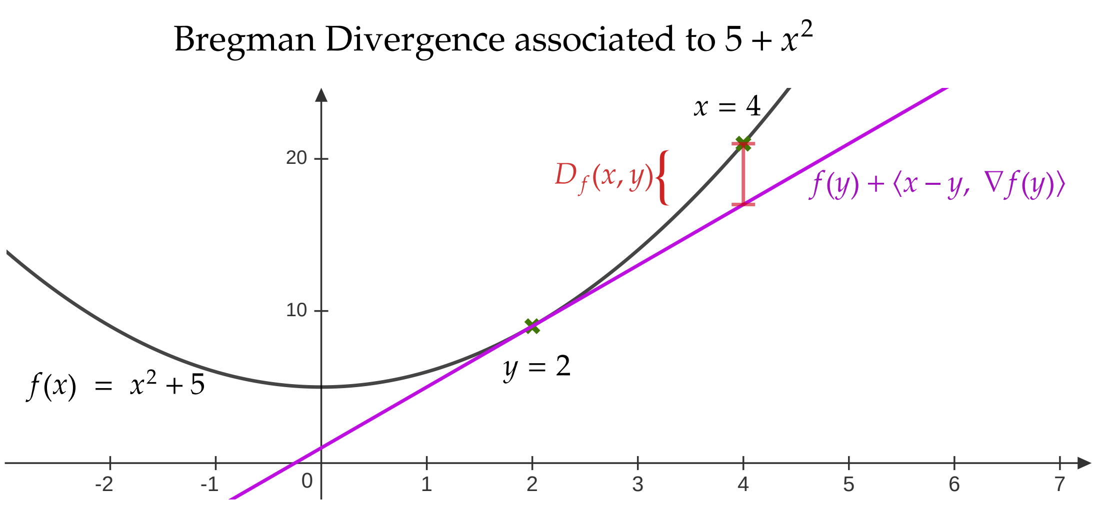

Figure 1.2: Bregman divergence associated to a parabola.

Remark 1.2.29 In the specific case where $f = \frac{1}{2}\|\cdot\|^2$, one has

$$D_f(x,y) = D_f(y,x) = \frac{1}{2}\|x-y\|^2 \quad \forall x,y \in \mathcal{H}.$$

Some elementary properties are listed below.

Proposition 1.2.30 *Under the assumptions of Definition 1.2.28 and $x \in \mathcal{H},\ y \in \operatorname{int}\operatorname{dom} f$.*

(i) $D_f(x,y) \geq 0, \quad \forall x,y \in \mathcal{H}$.

(ii) $D(\cdot,y) \in \Gamma_0(\mathcal{H})$ *and* $\operatorname{dom} D(\cdot,y) = \operatorname{dom} f$.

(iii) If f is essentially strictly convex at y, then $D(x,y) = 0$ if and only if $x = y$.

The Bregman divergence is not necessarily symmetric. For example, if $\mathcal{H} = \mathbb{R}^2$ and $f(x,y) := x^2 + \frac{1}{2}y^3$, then

$$D_f\left(\begin{pmatrix}1\\2\end{pmatrix},\begin{pmatrix}2\\1\end{pmatrix}\right) = 3, \quad D_f\left(\begin{pmatrix}2\\1\end{pmatrix},\begin{pmatrix}1\\2\end{pmatrix}\right) = 3.5.$$

With this in consideration, the following definition is introduced.

Definition 1.2.31 *Under the same assumptions from Definition 1.2.28. We define the $f-$Symmetrized Bregman divergence associated to f as:*

$$S_f(x,y) := D_f(x,y) + D_f(y,x).$$

From the above equality, one gets

$$S_f(x,y) = \langle \nabla f(x) - \nabla f(y), x - y\rangle.$$

It is also possible to define a projection associated to the Bregman divergence of a function.

Definition 1.2.32 *Let C be a closed convex set in $\mathcal{H}$. For $y \in \operatorname{int}\operatorname{dom} f$, the Bregman projection of y onto C is defined as:*

$$\operatorname{Proj}_C^f(y) := \operatorname{argmin}_{z\in C} D_f(z,y).$$

Whenever the set $\operatorname{Proj}_C^f(y)$ *is a singleton, we denote it as* $\operatorname{proj}_C^f(y)$.

The following result ensures the existence of the Bregman projection to a convex set. See [10, Theorem 7.6.8, Corollary 7.6.9].

Theorem 1.2.33 *Let $f \in \Gamma_0(\mathcal{H})$ be a Legendre function. Suppose that $C \subset \mathcal{H}$ is closed convex set such that $C \cap \operatorname{dom} f \neq \emptyset$, and $y \in \operatorname{int}\operatorname{dom} f$. Then the projection Proj_C^f is a singleton and belongs to the set* $\operatorname{int}\operatorname{dom} f$.

1.3 Elements of Nonsmooth Analysis

In this section, we recall some concepts of Variational Analysis. We refer to [11, 24, 26, 42] for further details.

1.3.1 Subdifferential and Normal Cone

Definition 1.3.1 *Let $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a lsc function and $x \in \operatorname{dom} f$. Then, an element $\zeta \in \mathcal{H}$ is said to be*

(i) A Fréchet subgradient of f at x if for every positive real number $\varepsilon > 0$, there exists U, a neighbourhood of x, such that

$$\langle \zeta, y - x \rangle \leq f(y) - f(x) + \varepsilon \|y - x\| \quad \textit{for all } y \in U.$$

The Fréchet subdifferential of f at x, denoted as $\partial^F(x)$, is the set of all Fréchet subgradients of f at x.

(ii) A limiting subgradient of f at x if there exists a sequence of points $((x_n, f(x_n)))_{n\in\mathbb{N}}$ converging to $(x, f(x))$ and a sequence $(\zeta_n)_{n\in\mathbb{N}}$, that converges weakly to ζ, where $\zeta_n \in \partial^F f(x_n)$. The Mordukhovich limiting subdifferential of f at x, denoted as $\partial^L(x)$, is the set of all limiting subgradients of f at x.

(iii) A proximal subgradient of f at x if there exists a real number $\sigma \geq 0$ and U, a neighbourhood of x, such that

$$\langle \zeta, y - x \rangle \leq f(y) - f(x) + \sigma \|y - x\|^2 \quad \textit{for all } y \in U.$$

The proximal subdifferential of f at x, denoted as $\partial^P(x)$, is the set of all proximal subgradients of f at x.

For all three cases, whenever $x \notin \operatorname{dom} f$, its subdifferential is defined as $\emptyset$.

Definition 1.3.2 *Let C be a nonempty subset of $\mathcal{H}$, the normal proximal cone of C at $\overline{x}$, denoted by $N^P(C; \overline{x})$ is defined as the set of points $x = t(y - \overline{x})$, where $\overline{x} \in \operatorname{Proj}_C(y), t$ is a nonnegative real number and $y \notin C$.*

Now, we introduce the Clarke normal cone and the Clarke subdifferential.

Definition 1.3.3 *Let C be a nonempty subset of $\mathcal{H}$, $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ a lsc function and $x \in \operatorname{dom} f$. The Clarke normal cone and the Clarke subdifferential are defined, respectively, as:*

$$N^{\mathrm{Cl}}(C; x) := \overline{\mathrm{co}}(N^L(C; x)) \textit{ and } \partial^{\mathrm{Cl}} f(x) := \overline{\mathrm{co}}(\partial^L(C; x)).$$

Given the nature of these definitions, for a nonempty set $C \subset \mathcal{H}$, the following inclusions hold

$$N^P(C; x) \subset N^F(C; x) \subset N^L(C; x) \subset N^{\mathrm{Cl}}(C; x).$$

1.3.2 Prox-regularity

In this section, we introduce the concept of prox-regular set, which is widely used in the context of sweeping processes because it generalizes convexity. We refer to [26, 29] for more details.

Definition 1.3.4 *Consider a closed set $C \subset \mathcal{H}$ and a positive real number ρ. The set C is said to be ρ-uniformly prox-regular if for every $\zeta \in N^P(C;x) \cap \mathbb{B}$ and any $0 < t < \rho$, one has*

$$x = \text{proj}_C(x + t\zeta).$$

The following theorem provides a characterization of prox-regularity (see [26, Theorem 16]).

Theorem 1.3.5 *Let C be a closed subset of $\mathcal{H}$ and $\rho \in (0, +\infty]$. Then, the following statements are equivalent:*

(i) *C is $\rho-$uniformly prox-regular.*

(ii) *For every $x, y \in C$ and $\zeta \in N^P(C;x)$ one has*

$$\langle \zeta, y - x \rangle \leq \frac{1}{2\rho} \|\zeta\| \|y - x\|^2.$$

(iii) *For any $x \in C$, $\zeta \in N^P(C;x) \cap \mathbb{B}$, and any nonnegative real number $t < \rho$, it holds that $x = \text{Proj}_C(x + t\zeta)$.*

(iv) *For any $x_i \in C, \zeta_i \in N^P(C;x_i) \cap \mathbb{B}$, one has*

$$\langle \zeta_1 - \zeta_2, x_1 - x_2 \rangle \geq -\frac{1}{\rho} \|x_1 - x_2\|^2.$$

(v) *For any positive constant $\gamma < 1$, the projection exists and is unique for every $x \in U_\rho^\gamma(C) := \{y \in \mathcal{H} : d_C(y) < \gamma\rho\}$. Furthermore, the mapping $\text{proj}_C \colon U_\rho(C) \to C$ is Lipschitz with constant $(1 - \gamma)^{-1}$ in $U_\rho^\gamma(C)$.*

As convex sets, prox-regular sets are normally regular.

Proposition 1.3.6 *Let $C \subset \mathcal{H}$ be a closed and $\rho-$uniformly prox-regular set. Then,*

$$N^P(C;x) = N^F(C;x) = N^L(C;x) = N^{\text{Cl}}(C;x) \quad \text{for all } x \in C.$$

Whenever there is no ambiguity, we write $N(C;x)$.

1.3.3 Smoothness via inverse image

The following definition and result will be useful for an intermediary result involving the preimage of a convex set. We refer to [31] for more details.

Definition 1.3.7 *Let $h \colon \mathcal{H} \to \mathcal{H}'$ be a differentiable mapping defined between two Hilbert spaces and $C \subset \mathcal{H}'$. We say that the set $h^{-1}(C)$ has the* normal cone inverse image property *at $\overline{x} \in h^{-1}(C)$ with respect to the Clarke normal cone if there exists a constant $\eta > 0$ and a neighbourhood U of $\overline{x}$ such that for every $x \in U \cap h^{-1}(C)$ the following inclusion holds:*

$$N^{\mathrm{Cl}}(h^{-1}(C); x) \cap \mathbb{B}_{\mathcal{H}} \subset Dh(x)^*(N^{\mathrm{Cl}}(C; h(x)) \cap \eta \mathbb{B}_{\mathcal{H}'}), \tag{1.4}$$

where $Dh(x)^*$ denotes the adjoint operator of $Dh(x)$.

Moreover, we say that the set $h^{-1}(C)$ has the uniform normal cone inverse image property *(UNCIIP) with η, if there exists some $\eta > 0$ such that at (1.4) holds for all $x \in h^{-1}(C)$.*

The next proposition, based on [31, Proposition 3.13], provides a necessary condition of the UNCIIP property.

Proposition 1.3.8 *Let $h \colon \mathcal{H} \to \mathcal{H}$ be a differentiable function and $C \subset \mathcal{H}$ be a closed convex set. If there exist $\eta > 0$ such that for all $x \in \mathcal{H}$ it holds:*

$$d(h^{-1}(C); x) \leq \eta d(C; h(x)). \tag{1.5}$$

Then, the set $h^{-1}(C)$ satisfies the UNCIIP with constant η.

When condition (1.5) holds, we say that the set-valued inverse mapping of h is metrically calm relatively to C with constant $\eta > 0$. We refer the reader to [26, 42, 55] for further details.

Now, we introduce a composition rule for the limiting subdifferential, it will be useful for an auxiliary result used in the proof of the main result. See [42, Theorem 3.41, Corollary 3.42] for more details.

Proposition 1.3.9 *Let $C \subset \mathcal{H}$ be a closed set and $g \colon \mathcal{H} \to \mathcal{H}$ be a Lipschitz continuous and differentiable function around $\bar{x}$ with $g(\bar{x}) \in C$. If g is metrically calm relatively to C, around $\overline{x}$, then*

$$N(g^{-1}(C); \bar{x}) \subset Dg(\bar{x})^* N(C; g(\bar{x})),$$

where equality holds if g is strictly differentiable at $\overline{x}$.

Chapter 2

Preparatory Results

2.1 Mirror Sweeping Process

Let $I := [0, T]$ be a real interval. We consider the following problem:

$$\begin{cases} \dot{x}(t) \in -N(C(t); \nabla\varphi(x(t))) + f(t, x(t)) \quad \text{a.e. } t \in I, \\ x(T_0) = x_0, \end{cases} \tag{$\mathcal{MSP}$}$$

where $\nabla\varphi(x_0) \in C(0)$ and $N(C(t); \cdot)$ denotes the normal cone and $\varphi \in \Gamma_0(\mathcal{H})$. In what follows, the above differential inclusion will be referred as *Mirror Sweeping Process*.

The following assumptions will be needed for the well-posedness of the Mirror Sweeping Processes and will be our standing hypotheses.

Hypotheses on $\varphi\colon \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$: The function φ belongs to $\Gamma_0(\mathcal{H})$ and is twice differentiable over an open convex set $\mathcal{D}$ contained in the domain of $\nabla\varphi$. Moreover, the following conditions will be used in the paper.

$(\mathcal{H}_\varphi^1)$ There exists $m > 0$ such that

$$\langle \nabla\varphi(x) - \nabla\varphi(y), x - y \rangle \geq m\|x - y\|^2 \text{ for all } x, y \in \mathcal{D}.$$

$(\mathcal{H}_\varphi^2)$ There exists $M > 0$ such that

$$\|\nabla\varphi(x) - \nabla\varphi(y)\| \leq M\|x - y\| \text{ for all } x, y \in \mathcal{D}.$$

$(\mathcal{H}_\varphi^3)$ There exists $\vartheta > 0$ such that

$$\|D^2\varphi(x) - D^2\varphi(y)\| \leq \vartheta\|x - y\| \text{ for all } x, y \in \mathcal{D}.$$

Hypotheses on the moving sets: The set-valued map $C\colon I \rightrightarrows \mathcal{H}$ has closed, nonempty and convex values. Moreover, the following condition will be used in the paper.

($\mathcal{H}_C$) C is $\kappa-$Lipschitz with respect to the excess, i.e.,

$$e(C(t), C(s)) \leq \kappa|t-s| \quad \forall t, s \in I.$$

Hypotheses on the perturbation: The perturbation term $f\colon I \times \mathcal{H} \to \mathcal{H}$ satisfies the following assumptions:

($\mathcal{H}_1^f$) For every $x \in \mathcal{H}$, the map $t \mapsto f(t,x)$ is measurable;

($\mathcal{H}_2^f$) For all $r > 0$, there exists $\mu_r \geq 0$ such that for a.e. $t \in I$, the map $x \mapsto f(t,x)$ is Lipschitz of constant μ_r on $r\mathbb{B}$;

($\mathcal{H}_3^f$) There exist nonnegative constants c, d such that

$$\|f(t,x)\| \leq c\|x\| + d \text{ for all } x \in \mathcal{H}.$$

We observe that ($\mathcal{H}_C$) includes the classical case where the multifunction $C(\cdot)$ is $\kappa-$Lipschitz with respect to the Hausdorff distance, that is,

$$d_H(C(t), C(s)) = \sup_{z\in\mathcal{H}} |d(C(t), z) - d(C(s), z)| \leq \kappa|t-s| \text{ for all } t, s \in I.$$

2.2 Preparatory Results

In this section, we introduce some results that will be of use later in the thesis.

We start with two different versions of Grönwall's Lemma. For the continuous version see, e.g., [54], and for the discrete version, see [36, Lemma 6].

Lemma 2.2.1 (Gronwall's lemma) *Let $u, \alpha, \beta\colon [t_0, t_1] \to \mathbb{R}$ be integrable functions. Assume that u is absolutely continuous on $[t_0, t_1]$ and*

$$\dot{u}(t) \leq \alpha(t) + \beta(t)u(t) \text{ a.e. } t \in [t_0, t_1].$$

Then, for all $t \in [t_0, t_1]$

$$u(t) \leq u(t_0)\exp\left(\int_{t_0}^{t} \beta(\tau)d\tau\right) + \int_{t_0}^{t} \alpha(s)\exp\left(\int_{s}^{t} \beta(\tau)d\tau\right) ds.$$

Lemma 2.2.2 (Discrete Gronwall's lemma) *Let $(\alpha_i), (\beta_i), (\gamma_i)$ and (u_i) be sequences of non-negative real numbers such that*

$$u_{i+1} \leq \alpha_i + \beta_i(\sum_{l=0}^{i-1} u_{l-1}) + (1+\gamma_i)u_i \quad \text{for } i > 0.$$

Then, for all $j > 0$,

$$u_j \leq \left(u_0 + \sum_{k=0}^{j-1} \alpha_k\right) \cdot \exp\left(\sum_{k=0}^{j-1} k\beta_k + \gamma_k\right).$$

2.2.1 Approximator Operator

Now, we introduce the operator P_C^{φ}, that can be seen as a generalization of the projection operator; in fact, whenever $\varphi = \frac{1}{2}\|\cdot\|^2$, the operator P_C^{φ} is just the projection onto the convex set C (see formula (1.1)). In what follows, we provide a direct proof that this operator is well defined and Lipschitz continuous.

Lemma 2.2.3 *Let $\varphi \in \Gamma_0(\mathcal{H})$ be a function satisfying $(\mathcal{H}_\varphi^1)$, $(\mathcal{H}_\varphi^2)$. Suppose that φ is differentiable over an open convex set $\mathcal{D}$. Then, the map $x \to \nabla\varphi$ is bijective with $(\nabla\varphi)^{-1} = \nabla\varphi^*$. Moreover, $\nabla\varphi^*$ is Lipschitz continuous of constant $\frac{1}{m}$ and $\frac{1}{M}-$strongly monotone.*

Proof. For the purpose of proving that $\nabla\varphi$ is bijective, we first prove that φ is a Legendre function (see Definition 1.2.21).

First, by virtue of the differentiability of φ and assumption $(\mathcal{H}_\varphi^2)$, the function φ is essentially smooth. Second, due to hypothesis $(\mathcal{H}_\varphi^1)$ and Proposition 1.2.8, the function φ is strongly convex of constant m, in particular φ is a strictly convex function. Thus, because of Lemma 1.2.23, φ is essentially strictly convex which means that φ is a Legendre function. As a result, by Theorem 1.2.27,

$$x \to \nabla\varphi \quad \text{is bijective with} \quad (\nabla\varphi)^{-1} = \nabla\varphi^*. \tag{2.1}$$

Moving forward, by the enhanced Baillon-Haddad theorem (see [56, Theorem 3.1]), we obtain that for every pair of elements x, y in the domain of φ,

$$\langle \nabla\varphi(x) - \nabla\varphi(y), x - y\rangle \geq \frac{1}{M}\|\nabla\varphi(x) - \nabla\varphi(y)\|^2, \tag{2.2}$$

that is, $\nabla\varphi$ is $\frac{1}{M}-$cocoercive (see Definition 1.1.10 (iv)). Set $w = \nabla\varphi(x), z = \nabla\varphi(y)$, from (2.1), we derive that $x = \nabla\varphi^*(w)$ and $y = \nabla\varphi^*(z)$. Therefore, (2.2) turns into

$$\langle w - z, \nabla\varphi^*(w) - \nabla\varphi^*(z)\rangle \geq \frac{1}{M}\|w - z\|^2,$$

which concludes that $\nabla\varphi^*$ is $\frac{1}{M}-$strongly monotone (see Definition 1.1.10 (iii)). Using the same transformation along with hypothesis $(\mathcal{H}_\varphi^1)$, we obtain that

$$\langle w - z, \nabla\varphi^*(w) - \nabla\varphi^*(z)\rangle \geq m\|\nabla\varphi^*(w) - \nabla\varphi^*(z)\|^2.$$

This inequality, when used in conjunction with the Cauchy-Schwarz inequality, results in $\nabla\varphi^*$ being Lipschitz continuous with constant $\frac{1}{m}$. □

We are now in position to introduce the main component of the algorithm that will be defined in the next section. This operator was introduced in [39] as the *approximation* operator associated to a strongly and maximal monotone operator.

It was proved that for a nonempty, closed, convex and bounded set, the approximator operator is well defined. Thanks to the Lipschitz continuity hypothesis and the duality between $\nabla\varphi$ and $\nabla\varphi^*$ we are able to remove the boundedness that was required. The proof in [39, Lemma 1] relies on maximal monotone operator theory, while ours makes use of a constrained optimization problem involving φ^*.

Definition 2.2.4 *Under the assumptions of Lemma 2.2.3. Consider $C \subset \mathcal{H}$ to be nonempty, closed and convex such that $C \subset \mathcal{D}$. Let us define the operators $P_C^\varphi : \mathcal{H} \to \mathcal{H}$ and $Q_C^\varphi : \mathcal{H} \to \mathcal{H}$, respectively as,*

$$P_C^\varphi(x) := \{y \in \mathcal{H} \colon x \in y + N(C; \nabla\varphi(y))\} \quad \textit{and} \quad Q_C^\varphi(x) := \left(\nabla\varphi^*(\cdot) + N(C;\cdot)\right)^{-1}(x).$$

We will refer to P_C^φ as the *approximator operator*. To assert that the recently defined operators are well-defined we need the following technical lemma, that takes advantage of the duality provided in Lemma 2.2.3. The minimizer of the function defined in this lemma, constrained to a nonempty, closed and convex set C, will ultimately play the role of the operator Q_C^φ.

Lemma 2.2.5 *Consider $\varphi \in \Gamma_0(\mathcal{H})$ to be a function satisfying $(\mathcal{H}_\varphi^2)$ and a fixed $x \in \mathcal{H}$. Then the function $f \colon \mathcal{H} \to \mathcal{H}$ defined by*

$$f(z) := \varphi^*(z) - \langle x, z\rangle$$

is supercoercive and strongly convex with constant $\frac{1}{M}$.

Proof. By Lemma 2.2.3 we know that $\nabla\varphi^*$ is $\frac{1}{M}$ strongly monotone. Hence, using this fact and Proposition 1.2.8, we get that φ^* is strongly convex of constant $\frac{1}{M}$. By the same result, one has, for every $z, y \in \operatorname{dom}\varphi^*$,

$$\varphi^*(z) \geq \varphi^*(y) + \langle z - y, \nabla\varphi^*(y)\rangle + \frac{1}{M}\|z-y\|^2.$$

Fix $y \in \operatorname{dom}\varphi^*$. By means of the Cauchy-Schwarz inequality, the following inequality holds for every $z \in \mathcal{H} \setminus \{0\}$:

$$\frac{f(z)}{\|z\|} \geq \frac{\varphi^*(y) - \|z-y\|\|\nabla\varphi^*(y)\| - \|x\|\|z\|}{\|z\|} + \frac{1}{M}\frac{\|z-y\|^2}{\|z\|}.$$

Then, as y and x are fixed, the right-hand side approaches infinity as $\|z\| \to +\infty$, which gives us supercoercivity of f by definition. ☐

With this construction, we are able to show that the operators P_C^φ and Q_C^φ are well-defined. We can also show Lipschitz continuity of the operators and the relationship between them.

Proposition 2.2.6 *Let C be a nonempty, closed and convex set in $\mathcal{H}$ and let $\varphi \in \Gamma_0(\mathcal{H})$ be a differentiable function over an open convex set $\mathcal{D}$ such that $C \subset \mathcal{D}$ satisfying $(\mathcal{H}_\varphi^1)$, $(\mathcal{H}_\varphi^2)$. Under the conditions stated above, the operator Q_C^φ is well-defined over $\mathcal{H}$ and is Lipschitz continuous of constant $1/m$. Moreover, the following formula holds*

$$P_C^\varphi(x) = \nabla\varphi^*\left(Q_C^\varphi(x)\right). \tag{2.3}$$

Consequently, the operator P_C^φ is well-defined on $\mathcal{H}$ and is Lipschitz continuous of constant $1/m^2$.

Proof. We proceed to show that Q_C^φ is well-defined. Fix $x \in \mathcal{H}$ and define the function $f(z) := \varphi^*(z) - \langle x, z\rangle$. We consider the optimization problem

$$\min_{z \in C} f(z). \tag{$\mathcal{P}_f$}$$

First, we can see that, by virtue of Lemma 2.2.5, f is a strongly convex and supercoercive function. In turn, this implies that f is strictly convex and supercoercive. Second, due to Proposition 1.2.19, the problem $(\mathcal{P}_f)$ admits one and only one solution $z^* \in C$. Using the necessary optimality condition given in Proposition 1.2.20 we must also have

$$0 \in \partial(f(\cdot) + \delta_C(\cdot))(z^*).$$

Therefore, since $C \subset \mathcal{D}$, using the sum rule from Proposition 1.2.14, it follows that

$$x \in \nabla\varphi^*(z^*) + N(C; z^*).$$

Hence, by definition of Q_C^φ and uniqueness of the minimizers for $(\mathcal{P}_f)$, we obtain that $z^* = Q_C^\varphi(x)$. Finally, for $x_i \in \mathcal{H}$, where $i = 1, 2$, we set $z_i^* = Q_C^\varphi(x_i)$. By the monotonicity of the normal cone (Proposition 1.3), it follows that

$$\langle x_1 - \nabla\varphi^*(z_1^*) - (x_2 - \nabla\varphi^*(z_2^*)), z_1^* - z_2^* \rangle \geq 0.$$

Thus, due to $(\mathcal{H}_\varphi^1)$, we obtain that

$$\begin{aligned} \langle x_1 - x_2, z_1^* - z_2^* \rangle &\geq \langle \nabla\varphi^*(z_1^*) - \nabla\varphi^*(z_2^*), z_1^* - z_2^* \rangle \\ &\geq m\|z_1^* - z_2^*\|^2, \end{aligned}$$

which implies the Lipischitzianity of Q_C^φ. We now turn to the approximator operator, P_C^φ. To prove (2.3), we proceed as follows. Consider $z = Q_C^\varphi(x)$. It follows from the definition, that

$$x \in \nabla\varphi^*(z) + N(C; z).$$

Setting $y = \nabla\varphi^*(z)$, recalling that $\nabla\varphi^*$ is bijective, we obtain that $z = \nabla\varphi(y)$. Therefore, it holds that

$$x \in y + N(C; \nabla\varphi(y)).$$

By definition of the operator, the above inclusion shows that $y = \nabla\varphi^*(z) \in P_C^\varphi(x)$. Conversely, considering $y \in P_C^\varphi(x)$ and setting $z = \nabla\varphi(y)$ completes the proof of the equality.

Since $\nabla\varphi^*$ is a bijection and (2.3) holds, the operator is well-defined. To achieve Lipschitzianity, let $x_1, x_2 \in \mathcal{H}$. Using Lemma 2.2.3 and the fact that Q_C^φ is Lipschitz of constant $\frac{1}{m}$, we get that

$$\begin{aligned} \|P_C^\varphi(x_1) - P_C^\varphi(x_2)\| &= \|\nabla\varphi^*(Q_C^\varphi(x_1)) - \nabla\varphi^*(Q_C^\varphi(x_2))| \\ &\leq \frac{1}{m}\|Q_C^\varphi(x_1) - Q_C^\varphi(x_2)\| \\ &\leq \frac{1}{m^2}\|x_1 - x_2\|, \end{aligned}$$

which concludes the proof. □

Remark 2.2.7 This definition and result gain in interest if we realize that in the case where $\varphi = \frac{1}{2}\|\cdot\|^2$, Definition 2.2.4 yields

$$P_C^\varphi(x) = \{y \in \mathcal{H} : x - y \in N(C; y)\}.$$

The above equality and Remark 1.2.17 imply that the approximator and the projection coincide. Furthermore, in this context, $\frac{1}{2}\|\cdot\|^2$ satisfies $(\mathcal{H}_\varphi^1)$ with $m = 1$, and $(\mathcal{H}_\varphi^2)$ with $M = 1$. Then, Proposition 2.2.6 yields that P_C^φ is Lipschitz with constant 1, the same constant as in (1.2).

Remark 2.2.8 Under the assumptions of Proposition 2.2.6, we obtain that

$$x \in \nabla\varphi^*(Q_C^\varphi(x)) + N(C; Q_C^\varphi(x)) \quad \forall\, x \in \mathcal{H}.$$

By the previous proof, $Q_C^\varphi(x)$ takes values on C. Hence, the operation $N_C(Q_C^\varphi(x))$ is well defined. Moreover, the following equality holds for all $x \in \mathcal{H}$

$$\nabla\varphi(P_C^\varphi(x)) = Q_C^\varphi(x).$$

The forthcoming results will prove helpful in laying the groundwork for the introduction of the discrete approximation of the Mirror Sweeping Process. These preliminary findings are poised to play a pivotal role in shaping the algorithm. For a similar result, using the Hausdorff distance for a nonempty, closed, convex and bounded set, we refer to [39, Lemma 2].

Lemma 2.2.9 *Let $\varphi \in \Gamma_0(\mathcal{H})$ be a differentiable function satisfying $(\mathcal{H}_\varphi^1)$, $(\mathcal{H}_\varphi^2)$. Consider $C_1, C_2 \in \mathcal{H}$ be nonempty, closed and convex sets. If $x \in \mathcal{D}$ with $\nabla\varphi(x) \cap C_1 \neq \emptyset$, then*

$$\|x - P_{C_2}^\varphi(x)\| \leq \frac{1}{m} e(C_1, C_2).$$

Proof. Let $y \in P_{C_2}^\varphi(x)$, by definition of the operator, we have that $x - y \in N(C_2; \nabla\varphi(y))$, which entails, by definition of the normal cone in the convex case (Definition 1.2.15),

$$\langle x - y, z - \nabla\varphi(y))\rangle \leq 0 \quad \forall z \in C_2.$$

Then, by using the above inequality and $(\mathcal{H}_\varphi^1)$, we get for all $z \in C_2$

$$\begin{aligned} m\|x-y\|^2 &\leq \langle \nabla\varphi(x) - \nabla\varphi(y), x - y\rangle \\ &= \langle z - \nabla\varphi(y), x - y\rangle + \langle \nabla\varphi(x) - z, x - y\rangle \\ &\leq \|x - y\| \cdot (\|\nabla\varphi(x) - z\|). \end{aligned}$$

Then, taking $\inf_{z \in C_2}$ and using the fact that $\nabla\varphi(y) \in C_2$ we arrive to

$$m\|x-y\|^2 \leq \|x-y\| \cdot d(C_2; \nabla\varphi(x)).$$

Finally, by the hypothesis $\nabla\varphi(x) \cap C_1 \neq \emptyset$,

$$\|x - y\| \leq \frac{1}{m} d(C_2; \nabla\varphi(x)) \leq \frac{1}{m} e(C_1, C_2).$$

□

Lemma 2.2.10 *Consider the same assumptions from the above lemma. Assume, in addition to $(\mathcal{H}_\varphi^1)$ and $(\mathcal{H}_\varphi^2)$, that $(\mathcal{H}_1^f)$ and $(\mathcal{H}_3^f)$ hold. Then, for all $x \in \mathcal{D}$ with $\nabla\varphi(x) \cap C_1 \neq \emptyset$, and $t_0, t_1 \in [0, T]$, it holds that*

$$\left\| x - P_{C_2}^\varphi\Big(x + \int_{t_0}^{t_1} f(s, x)ds\Big) \right\| \leq \frac{1}{m} e(C_1, C_2) + \frac{|t_1 - t_0|}{m^2}(c\|x\| + d).$$

Proof. According to Proposition 2.2.6, the operator P^{φ}_{C} is Lipschitz continuous with constant $\frac{1}{m^2}$, then, by virtue of Lemma 2.2.9 and $(\mathcal{H}^f_3)$ the following hold:

$$\begin{aligned}\left\|x - P^{\varphi}_{C_2}(x + \int_{t_0}^{t_1} f(s,x)ds)\right\| &\leq \|x - P^{\varphi}_{C_2}(x)\| + \|P^{\varphi}_{C_2}(x) - P^{\varphi}_{C_2}(x + \int_{t_0}^{t_1} f(s,x)ds)\| \\ &\leq \frac{1}{m}e(C_1, C_2) + \frac{1}{m^2}\int_{t_0}^{t_1} \|f(s,x)\|ds \\ &\leq \frac{1}{m}e(C_1, C_2) + \frac{|t_1 - t_0|}{m^2}(c\|x\| + d),\end{aligned}$$

which conclude the proof. □

2.2.2 Smoothness via inverse image

Finally, we show a that the preimage of a convex set through the mapping $\nabla\varphi$ is prox-regular and that this same mapping is metrically calm with respect to the set-valued map $C(\cdot)$.

The next result shows the prox-regularity of the preimage of a convex set C. It will be useful to make use of the normal cone of a prox-regular set. The preimage of a convex set through a continuous mapping is not always convex, for example, if we consider $f : \mathbb{R} \to \mathbb{R}$ as $f(x) = x^2$, the preimage of the interval $[1,2]$ is $[-\sqrt{2}, -1] \cup [1, \sqrt{2}]$ (see Figure 2.1). In

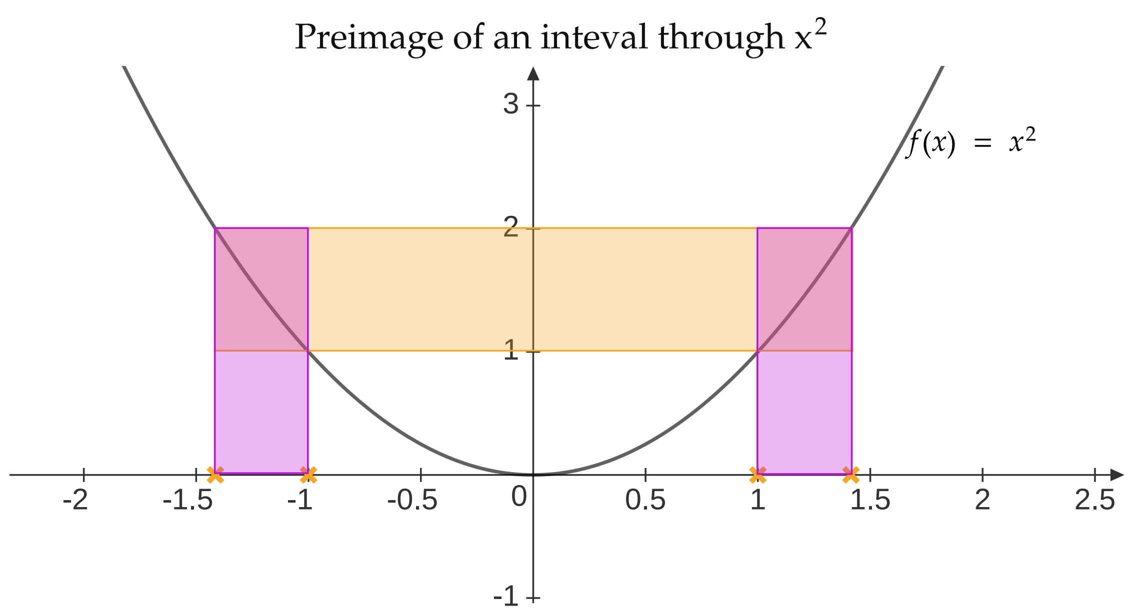

Figure 2.1: Nonconvex preimage of the convex set $[1,2]$ via x^2.

our setting $\nabla\varphi$ will take the role of the mapping h.

Proposition 2.2.11 *Let $h\colon \mathcal{H} \to \mathcal{H}$ be a Lipschitz continuous mapping with constant $M > 0$, such that its derivative Dh is Lipschitz continuous with constant $\vartheta > 0$. If the family $(C(t))_{t\in[0,T]} \subset \mathcal{H}$ is convex and has the UNCIIP for every $t \in [0,T]$ with the same constant*

$\eta > 0$. Then the family $(h^{-1}(C(t)))_{t\in[0,T]}$ is $\widetilde{\rho}-$uniformly prox-regular, where the constant $\widetilde{\rho}$ is defined as

$$\widetilde{\rho} := \frac{1}{2\eta\vartheta}.$$

Proof. Fix $t \in [0,T]$ and two elements u_1, u_2 in the preimage set $h^{-1}(C(t))$, we also consider

$$x_i^* \in N(h^{-1}(C(t)); u_i) \cap \mathbb{B} \quad \text{for } i = 1, 2.$$

By Theorem 1.3.5, it is enough to prove that

$$\langle x_1^* - x_2^*, u_1 - u_2\rangle \geq -\frac{1}{\widetilde{\rho}}\|u_1 - u_2\|^2. \tag{2.4}$$

First, by the UNCIIP (see Definition 1.3.7), there exists y_i^* such that

$$y_i^* \in N(C(t); h(u_i)) \cap \mathbb{B}, \text{ where } x_i^* = \eta Dh(u_i)^* y_i^* \quad \text{for } i = 1, 2. \tag{2.5}$$

Second, thanks to convexity of $C(t)$, by Proposition 1.3 one has

$$\langle y_1^* - y_2^*, h(u_1) - h(u_2)\rangle \geq 0. \tag{2.6}$$

Next, by the Lipschitzianity of Dh, for all $s \in [0,1]$, we have that

$$\|Dh(u_2 + s(u_1 - u_2)) - Dh(u_1)\| \leq \vartheta\|s(u_2 - u_1)\|. \tag{2.7}$$

Then, since h is Lipschitz continuous, we have that

$$\begin{aligned}
\langle \eta y_1^*, h(u_1) - h(u_2)\rangle &= \langle \eta y_1^*, \int_0^1 Dh(u_2 + s(u_1 - u_2))(u_1 - u_2)ds\rangle \\
&= \langle \eta y_1^*, Dh(u_1)(u_1 - u_2)\rangle \\
&\quad + \langle \eta y_1^*, \int_0^1 \left(Dh(u_2 + s(u_1 - u_2)) - Dh(u_1)\right)(u_1 - u_2)ds\rangle.
\end{aligned}$$

Hence, due to (2.7) and the fact that $\|y_1^*\| \leq 1$, we obtain that

$$\begin{aligned}
\langle \eta y_1^*, h(u_1) - h(u_2)\rangle &\leq \langle \eta y_1^*, Dh(u_1)(u_1 - u_2)\rangle + \eta\|y_1^*\| \cdot \vartheta\|u_1 - u_2\|^2 \\
&\leq \langle \eta Dh(u_1)^* y_1^*, u_1 - u_2\rangle + \eta\vartheta\|u_1 - u_2\|^2 \\
&= \langle x_1^*, u_1 - u_2\rangle + \eta\vartheta\|u_1 - u_2\|^2,
\end{aligned}$$

where the last equality is due to (2.5). Similarly, we get that

$$\langle \eta y_2^*, h(u_2) - h(u_1)\rangle \leq \langle x_2^*, u_2 - u_1\rangle + \eta\vartheta\|u_1 - u_2\|^2.$$

Adding both inequalities, one gets

$$\eta\langle y_1^* - y_2^*, h(u_1) - h(u_2)\rangle \leq \langle x_1^* - x_2^*, u_1 - u_2\rangle + 2\eta\vartheta\|u_1 - u_2\|^2.$$

Then, by (2.6) we obtain

$$0 \leq \langle x_1^* - x_2^*, u_1 - u_2\rangle + 2\eta\vartheta\|u_1 - u_2\|^2,$$

or, equivalently,

$$-2\eta\vartheta\|u_1 - u_2\|^2 \leq \langle x_1^* - x_2^*, u_1 - u_2\rangle,$$

which shows (2.4) with $\widetilde{\rho} := \frac{1}{2\eta\vartheta}$. Hence, the set $C(t)$ is $\widetilde{\rho}-$uniformly prox-regular for all $t \in [0,T]$. ∎

Lemma 2.2.12 *Let $\varphi \in \Gamma_0(\mathcal{H})$ be a function satisfying $(\mathcal{H}^1_\varphi)$, $(\mathcal{H}^2_\varphi)$ and $(\mathcal{H}^3_\varphi)$. Assume that $C\colon I \rightrightarrows \mathcal{H}$ has nonempty, closed, and convex values and $C(t) \subset \mathcal{D}$ for all $t \in I$. If φ is differentiable over an open convex set $\mathcal{D}$, then the mapping $\nabla\varphi$ is metrically calm relatively to C with constant $\frac{1}{m}$. Consequently for all $t \in I$, the set $\nabla\varphi^*(C(t))$ satisfies the UNCIIP with constant $\frac{1}{m}$.*

Proof. Fix $t \in I$ and consider an arbitrary element $x \in \mathcal{H}$. If $x \in \nabla\varphi^{-1}(C(t))$ the conclusion is immediate. Suppose $x \in H \setminus \nabla\varphi^{-1}(C(t))$, by definition one has

$$\begin{aligned} d(\nabla\varphi^{-1}(C(t)); x) &:= \inf_{w \in \nabla\varphi^{-1}(C(t))} \|x - w\| \\ &= \inf_{\nabla\varphi^{-1}(z): z \in C(t)} \|x - \nabla\varphi^{-1}(z)\|. \end{aligned}$$

By virtue of Lemma 2.2.3, we know that $\nabla\varphi$ is bijective with $\nabla\varphi^{-1} = \nabla\varphi^*$ and $\nabla\varphi^*$ is Lipschitz with constant $\frac{1}{m}$. Therefore, we obtain that

$$\begin{aligned} d(\nabla\varphi^{-1}(C(t)); x) &= \inf_{\nabla\varphi^*(z): z \in C(t)} \|\nabla\varphi^*(\nabla\varphi(x)) - \nabla\varphi^*(z)\| \\ &\leq \frac{1}{m} \inf_{z \in C(t)} \|\nabla\varphi(x) - z\| \\ &= \frac{1}{m} d(C(t); \nabla\varphi(x)), \end{aligned}$$

which implies that $\nabla\varphi$ is metrically calm relatively to $C(t)$ with constant $\frac{1}{m}$.

Finally, by virtue of Proposition 1.3.8 and Theorem 1.2.27, the set $\nabla\varphi^*(C(t))$ satisfies the UNCIIP with constant $\frac{1}{m}$. □

Corollary 2.2.13 *Under the assumptions of the above lemma, $\nabla\varphi^*(C(t))$ is $\frac{1}{2\eta\vartheta}-$uniformly prox-regular, for every $t \in [0, T]$.*

Proof. Thanks to the previous lemma, $\nabla\varphi^*(C)$ satisfies the UNCIIP with constant $\frac{1}{m}$. We conclude by virtue of Proposition 2.2.11. □

Chapter 3

Existence of solutions for the Mirror Sweeping Process

3.1 Vanilla Catching-up Algorithm

There are mainly two ways to solve a sweeping process. In this subsection, we will mention some results involving a discretization scheme. See [1] for a quick review of some variants of Moreau's Sweeping Process. The first instance of the catching-up algorithm is introduced in [46], to solve the classical sweeping process:

$$\begin{cases} \dot{x}(t) \in -N(C(t); x(t)) & \text{a.e. } t \in [0, T]; \\ x(0) = x_0 \in C(0), \end{cases} \tag{3.1}$$

where $C(\cdot)$ takes nonempty, closed and convex values. This dynamic can be interpreted as the point $x(t)$ being "swept" by the moving set $C(t)$. As the moving set evolves over time, if the initial point is not touched by the boundary of the moving set, it must lie within the interior. In this case, it remains unaffected, as the normal cone in the interior is $\{0\}$. Whenever $x(t)$ is reached by its boundary, or in other words, "caught up" by it, it is forced to move in an inward normal direction, in order to stay in the moving set.

The above interpretation is the foundation of the catching-up algorithm. Consider a discretization of n points for the interval $[0, T]$:

$$0 = t_0^n < t_1^n < t_2^n < \cdots < t_{n-1}^n < t_n^n = T.$$

Given a starting point $x_0 \in C(0)$, define the sequence of points

$$x_0^n = x_0, \quad x_{i+1}^n = \text{proj}_{C(t_{i+1}^n)}(x_i^n), \quad i = 0, \ldots n-1. \tag{3.2}$$

In this scheme, for each step $i = 0, \ldots, n-1$ one projects the previous point, x_i^n, to the set evaluated at the next time in the partition, that is, $C(t_{i+1}^n)$. As in the previous interpretation, if $x_i^n \in C(t_{i+1}^n)$ the projection does nothing. Conversely, if $x_i^n \notin C(t_{i+1}^n)$, the next point in the sequence is the projection to the closed convex set $C(t^n_{i+1})$. By Remark 1.2.17, the sequence defined in (3.2) fulfills the following inclusion:

$$\frac{x^n_{i+1} - x^n_i}{t^n_{i+1} - t_i} \in -N_{C(t^n_{i+1})}(x^n_{i+1}), \quad i = 0, \ldots, n-1. \tag{3.3}$$

The left hand side will act as an approximant of the derivative $\dot{x}(\cdot)$. To fill in the gaps, a linear interpolation is used to defined an approximated solution $x_n(\cdot)$ over the whole interval I, then a suitable limit $x(\cdot)$ is found, which will solve (3.1). See figure 3.1 for a visual example.

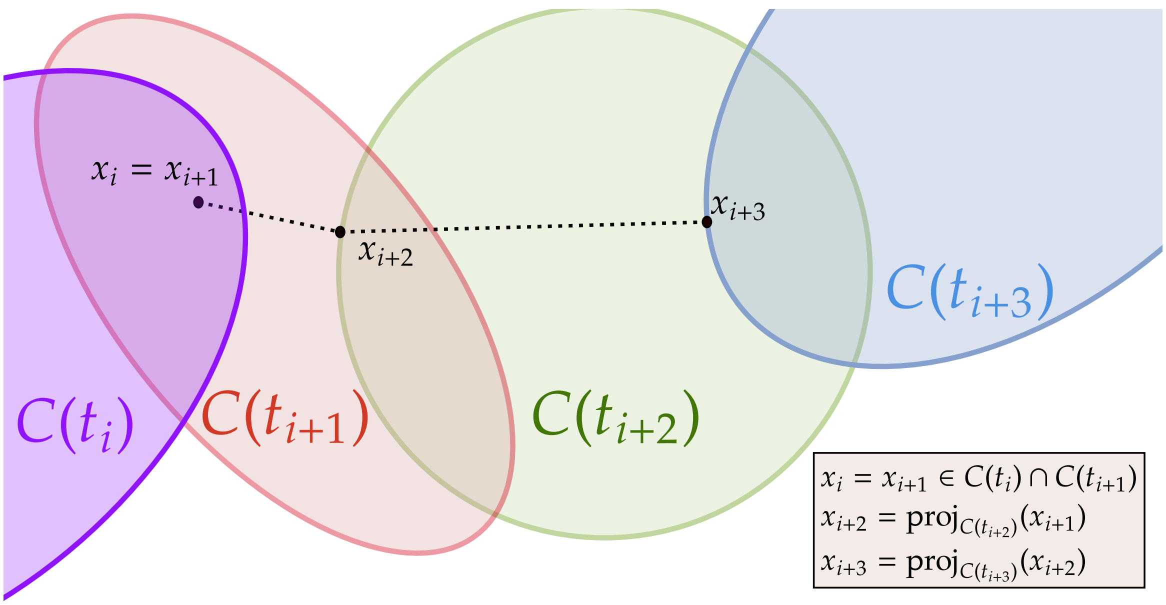

Figure 3.1: Steps of the catching-up algorithm.

In the case of the Mirror Sweeping Process, the projecting step is the problematic one, here the point being "caught up" is $\nabla\varphi(x(\cdot))$. Consider the same interval discretization and $x_0 \in \nabla\varphi(C(0))$. If one were to project as in the previous scheme, we obtain

$$x^n_0 = x_0, \quad x^n_{i+1} = \text{proj}_{C(t^n_{i+1})}(\nabla\varphi(x^n_i)), \quad i = 0, \ldots n-1.$$

The previous scheme implies:

$$\nabla\varphi(x^n_i) - x^n_{i+1} \in N_{C(t^n_{i+1})}(x^n_{i+1}), \quad i = 0, \ldots, n-1,$$

which is not helpful for either approximating the derivative nor having the normal cone of the set evaluated on the operator. The requirements for finding a suitable scheme for $(\mathcal{MSP})$, like in (3.3) is approximating the derivative on the left hand side, and the normal cone of the set $C(\cdot)$ at $\nabla\varphi(\cdot)$ on the right hand side. This reinterpretation is what makes the operator P^φ_C so valuable in this work. It will be the most important aspect for the algorithm defined in the next section.

3.2 Mirror Catching-up Algorithm

In this chapter, we will introduce a catching-up type algorithm for the Mirror Sweeping Process.

To this end, let $I := [0,T]$, for every $n \in \mathbb{N}$ we consider, for $0 \le i \le n$, the points $t_i^n := \frac{iT}{n}$ that define a partition $\mathcal{P}_n := \{t_0^n = 0, t_1^n, \dots, t_i^n, \dots, t_{n-1}^n, t_n^n = T\}$ of the set I. For the sake of convenience we also define $I_i^n = (t_i^n, t_{i+1}^n]$ and $\mu_n := \frac{T}{n}$ for every natural number n and $0 \le i \le n$. For every $n \in \mathbb{N}$ we consider the following scheme:

$$\begin{cases} x_0^n := x_0 \in (\nabla\varphi)^{-1}(C(0)), \\ x_{i+1}^n := P^{\varphi}_{C(t_{i+1}^n)}\left(x_i^n + \displaystyle\int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds\right), \quad \text{for } 0 \le i \le n-1. \end{cases} \tag{$\mathcal{DMCU}$}$$

From this discrete definition, we extend to an approximated solution over the whole interval I, defined by a polygonal of the points in ($\mathcal{DMCU}$):

$$\begin{cases} x_n(0) := x_0 \in (\nabla\varphi)^{-1}(C(0)), \\ x_n(t) := x_i^n + \dfrac{t - t_i^n}{\mu_n}\left(x_{i+1}^n - x_i^n - \displaystyle\int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds\right) + \displaystyle\int_{t_i^n}^{t} f(s, x_i^n)ds, \quad \text{for } 0 \le i \le n-1, t \in I_i^n, \end{cases} \tag{$\mathcal{MCU}$}$$

where, for both schemes we assume that all the hypotheses from section 2.1 hold.

We will prove that the sequence defined by the approximate solutions in ($\mathcal{MCU}$) satisfies a Cauchy criterion and thus converges. It will be shown that the limit of this sequence is a solution of ($\mathcal{MSP}$). In order the achieve this, we will need some bounds on the elements described in the discrete scheme ($\mathcal{DMCU}$). The consecutive points in the discrete setting will allow us to make use of the normal cone of the set-valued mapping $C(\cdot)$, this proves useful when dealing with the continuous scheme.

The first lemma in this section ensures that the discrete scheme is well-defined and provides a useful characterization involving the normal cone. We reiterate that we assume that all assumptions from Section 2.1 hold.

Lemma 3.2.1 *Let $C(t)$ be a nonempty, closed and convex set for all t in the interval $I = [0,T]$. Then the scheme described in ($\mathcal{DMCU}$) is well-defined. Furthermore, it holds that*

$$x_{i+1}^n - x_i^n \in -N_{C(t_{i+1}^n)}(\nabla\varphi(x_{i+1}^n)) + \int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds, i \in \{0, \dots, n-1\}. \tag{3.4}$$

and

$$\nabla\varphi(x_{i+1}^n) = Q^{\varphi}_{C(t_{i+1}^n)}(x_i^n + \int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds) \in C(t_i^n) \quad \textit{for } i \in \{0, \dots, n-1\}. \tag{3.5}$$

Proof. Set $n \in \mathbb{N}$ and $0 \le i \le n$. We know that for any $z \in \mathcal{H}$, the operator $P^{\varphi}_{C(t_{i+1}^n)}(z)$ exists because of Proposition 2.2.6. Furthermore, due to $(\mathcal{H}_1^f)$ we can integrate the perturbation function. Then, $(\mathcal{DMCU})$ is well defined and by definition of the approximator operator, it holds that

$$x_i^n + \int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds \in x_{i+1}^n + N_{C(t_{i+1}^n)}(\nabla\varphi(x_{i+1}^n)), \quad \text{for } i \in \{0, \dots, n-1\},$$

which implies

$$x_{i+1}^n - x_i^n \in -N_{C(t_{i+1}^n)}(\nabla\varphi(x_{i+1}^n)) + \int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds, i \in \{0, \dots, n-1\}.$$

In addition, by virtue of Remark 2.2.8,

$$\begin{aligned}
\nabla\varphi(x_{i+1}^n) &= \nabla\varphi(P^{\varphi}_{C(t_{i+1}^n)}(x_i^n + \int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds)) \\
&= Q^{\varphi}_{C(t_{i+1}^n)}((x_i^n + \int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds)) \in C(t_i^n), \quad \text{for } i \in \{0, \dots, n-1\}.
\end{aligned}$$

□

The following remark is useful to ensure that the scheme is well-defined.

Remark 3.2.2 By virtue of Remark 2.2.8 and the definition of the discrete scheme, it is easily seen that $\nabla\varphi(x_{i+1}^n) \cap C(t_{i+1}^n) \neq \emptyset$ for every $i \in \{0, \dots, n-1\}$. Thus, the normal cone $N_{C(t_{i+1}^n)}(\nabla\varphi(x_{i+1}^n))$ is nonempty for $i \in \{0, \dots, n-1\}$.

The following proposition shows bounds of the discrete sequence defined by $(\mathcal{DMCU})$. They will be promptly used to show bounds on the continuous approximated solution defined by $(\mathcal{MCU})$.

Proposition 3.2.3 *Let $C(t)$ be a nonempty, closed and convex set for all t in the interval $I = [0, T]$. Then the following hold*

$$\|x_i^n - x_{i+1}^n\| \leq \kappa\mu_n + \frac{\mu_n}{m^2}(c\|x_i^n\| + d) \quad \textit{for } i \in \{0, \dots, n-1\}, \tag{3.6}$$

and

$$\|x_i^n\| \leq \left(\|x_0^n\| + \frac{T}{m}(\frac{d}{m} + \kappa)\right) \exp\left(\frac{cT}{m}\right) \quad \textit{for } i \in \{1, \dots, n\}. \tag{3.7}$$

Proof. Let us first prove (3.6). By virtue of Lemma 2.2.10 and the definition of $(\mathcal{DMCU})$, one gets

$$\begin{aligned}
\|x_i^n - x_{i+1}^n\| &= \|x_i^n - P^{\varphi}_{C(t_{i+1}^n)}(x_i^n + \int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds)\| \\
&\leq \frac{1}{m}e(C(t_i^n), C(t_{i+1}^n)) + \frac{|t_{i+1}^n - t_i^n|}{m^2}(c\|x_i^n\| + d).
\end{aligned}$$

Using the fact that

$$e(C(t_i^n), C(t_{i+1}^n)) \leq \kappa(t_{i+1}^n - t_i^n) = \kappa\mu_n \qquad \text{for } i \in \{0, \dots, n-1\},$$

we obtain that

$$\|x_i^n - x_{i+1}^n\| \leq \frac{\kappa\mu_n}{m} + \frac{\mu_n}{m^2}(c\|x_i^n\| + d),$$

which proves the desired inequality. In order to deduce (3.7) we make use of (3.6), this equation implies that

$$\|x_{i+1}^n\| \leq \|x_i^n\|(1 + \frac{c\mu_n}{m^2}) + \frac{\mu_n}{m}(\frac{d}{m} + \kappa).$$

Set, for a fixed n, the constants $\alpha_i = \frac{\mu_n}{m}(\frac{d}{m} + \kappa)$, $\beta_i = 0$, $\gamma_i = \frac{c\mu_n}{m}$ and $u_i = \|x_i^n\|$ for all $0 \leq i \leq n$. By virtue of Lemma 2.2.2, we get that for any $0 \leq j \leq n$,

$$\begin{aligned}
\|x_j^n\| &\leq \left(\|x_0^n\| + \sum_{k=0}^{j-1} \frac{\mu_n}{m}(\frac{d}{m} + \kappa)\right) \exp\left(\sum_{k=0}^{j-1} \frac{c\mu_n}{m}\right) \\
&= \left(\|x_0^n\| + j\frac{\mu_n}{m}(\frac{d}{m} + \kappa)\right) \exp\left(j\frac{c\mu_n}{m}\right).
\end{aligned}$$

Finally, by definition of μ_n, we know that, $\frac{j}{n} \leq 1$ for every $0 \leq j \leq n$, this fact implies the following formula:

$$\|x_j^n\| \leq \left(\|x_0\| + \frac{T}{m}(\frac{d}{m} + \kappa)\right) \exp\left(\frac{cT}{m}\right),$$

which completes the proof. □

Considering the continuous setting, the following results are crucial to arrive at the desired solution; these results build from the discrete setting. We begin by asserting that $(\mathcal{MCU})$ is well defined, getting an explicit formula and an upper bound for the derivative of the approximated continuous solution, which will help us to get a normal cone characterization for this case.

From now on, for each $n \in \mathbb{N}$ we consider the functions $\theta_n, \delta_n \colon I \to I$ defined, respectively, as

$$\theta_n(t) := t_{i+1}^n, \delta_n(t) := t_i^n,$$

for each $t \in I_i^n$ and $i \in \{0, \dots, n\}$. It is clear that, as n tends to infinity, both θ_n and δ_n tend to the identity function in I.

Proposition 3.2.4 *Consider $n \in \mathbb{N}$ and $x_n(\cdot) : I \to H$ as defined in $(\mathcal{MCU})$, then one gets for $i \in \{0, \dots, n-1\}$, $t_1, t_2 \in I_i^n$ the following formulas for a.e. $t \in I_i^n$:*

$$\dot{x}_n(t) = \frac{1}{\mu_n}\left(x_{i+1}^n - x_i^n - \int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds\right) + f(t, x_n(\delta_n(t))). \tag{3.8}$$

and

$$\|\dot{x}_n(t)\| \leq (cL + d)\left(2 + \frac{1}{m^2}\right) + \kappa, \tag{3.9}$$

where $L := \left(\|x_0\| + \frac{T}{m}(\frac{d}{m} + \kappa)\right) \exp\left(\frac{cT}{m}\right)$.

Proof. We consider $t_1, t_2 \in I_i^n$, where, without loss of generality, $t_1 < t_2$. If t_2 is sufficiently close to t_1, explicitly $|t_1 - t_2| < \min\{|t_1 - \delta_n(t_1)|,\ |\theta_n(t_1) - t_1|\}$, one has $\delta_n(t_1) = \delta_n(t_2)$. Then, by simple calculations

$$\begin{aligned} x_n(t_2) - x_n(t_1) &= \frac{t_2 - t_1}{\mu_n}\left(x_{i+1}^n - x_i^n - \int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds\right) \\ &\quad + \int_{\delta_n(t_2)}^{t_2} f(s, x_n(\delta_n(t_1))) - \int_{\delta_n(t_1)}^{t_1} f(s, x_n(\delta_n(t_1)))ds \\ &= \frac{t_2 - t_1}{\mu_n}\left(x_{i+1}^n - x_i^n - \int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds\right) + \int_{t_1}^{t_2} f(s, x_n(\delta_n(t_1)))ds. \end{aligned}$$

Therefore, $x_n(\cdot)$ is absolutely continuous, it also holds

$$\dot{x}_n(t) = \frac{1}{\mu_n}\left(x_{i+1}^n - x_i^n - \int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds\right) + f(t, x_n(\delta_n(t))).$$

which is the desired equation (3.8).

To prove (3.9), we use (3.8), $(\mathcal{H}_3^f)$ and equation (3.6) from Proposition 3.2.3 to obtain

$$\|\dot{x}_n(t)\| \leq \frac{1}{\mu_n}(\kappa\mu_n + \frac{\mu_n}{m^2}(c\|x_i^n\| + d)) + \frac{|t_i^n - t_i^{n+1}|}{\mu_n}(c\|x_i^n\| + d) + \|f(t, x_n(\delta_n(t)))\|,$$

by simple calculations and equation (3.7) from Proposition 3.2.3 we obtain that

$$\|\dot{x}_n(t)\| \leq \kappa + \frac{1}{m^2}(cL + d) + cL + d + cL + d = (cL + d)\left(2 + \frac{1}{m^2}\right) + \kappa.$$

This finishes the proof. □

Given the recently proved nature of the derivative of $x_n(\cdot)$, one has, almost by definition, the following inclusion involving the normal cone of the right endpoint of each interval in the partition $\mathcal{P}_n$.

Corollary 3.2.5 *Consider $n \in \mathbb{N}$, $x_n(\cdot) : I \to \mathcal{H}$ as defined in $(\mathcal{MCU})$ and the functions $\theta_n(t) := t_{i+1}^n$ and $\delta_n(t) := t_i^n$. Then, it holds that*

$$\dot{x}_n(t) \in -N(C(\theta_n(t)); \nabla\varphi(x_n(\theta_n(t)))) + f(t, x_n(\delta_n(t))) \quad \text{for a.e. } t \in I. \tag{3.10}$$

Proof. By definition, it is clear that and $x_{i+1}^n = x_n(\theta_n(t))$. Hence, as a consequence of (3.8) and (3.4), the inclusion

$$\dot{x}_n(t) - f(t, x_n(\delta_n(t))) \in -N(C(\theta_n(t)); \nabla\varphi(x_n(\theta_n(t))))$$

holds, which concludes the proof. □

The remainder of this chapter is devoted to prove the main result of the thesis, finding a solution to the Mirror Sweeping Process. Firstly we provide a Cauchy criterion for $(x_n)_{n\in\mathbb{N}}$.

Proposition 3.2.6 *There exists a constant $\mathcal{C}_{n,k} > 0$ such that for every pair of natural numbers n, k it holds that*

$$\sup_{t\in[0,T]} \|x_n(t) - x_k(t)\| \leq \mathcal{C}_{n,k},$$

where $x_n(\cdot)$ and $x_k(\cdot)$ are defined by $(\mathcal{MCU})$ and the constant $\mathcal{C}_{n,k}$ depends on μ_n, μ_k and goes to 0 as n, k tend to $+\infty$.

To prove this proposition, we need several results. The first one helps us in estimating the distance between two approximated solutions given by $(\mathcal{MCU})$, associated to two different natural numbers n, k. The second one is a technical lemma involving the normal cone of the preimage through $\nabla\varphi$ of the moving set, and the second derivative of φ. These elements appear when proving a Cauchy criterion for the sequence of a approximated solutions with the help of the $\varphi-$Symmetrized Bregman divergence (see Definition 1.2.31).

Lemma 3.2.7 *Let $n, k \in \mathbb{N}$ and $t, t_1, t_2 \in I_i^n \cap I_j^k$ for some $0 \leq i \leq n,\ 0 \leq j \leq n$ such that $t_1 < t_2$. Then, the following hold for the sequence of functions defined by $(\mathcal{MCU})$:*

(i)

$$\|x_n(t_1) - x_n(t_2)\| \leq \mu_n(\gamma + \kappa),$$

(ii)

$$\|x_n(\theta_n(t)) - x_k(\theta_k(t))\| \leq \frac{\kappa(\mu_n + \mu_k)}{m},$$

(iii)

$$\|x_n(t_1) - x_k(t_1)\| \leq (\mu_n + \mu_k)\left(\kappa\left(1 + \frac{1}{m}\right) + \gamma\right),$$

where $\gamma := (cL + d)\left(2 + \frac{1}{m^2}\right)$.

Proof. As in the proof of Proposition 3.2.4, we have

$$\|x_n(t_1) - x_n(t_2)\| = \frac{t_1 - t_2}{\mu_n}\left(x_{i+1}^n - x_i^n - \int_{t_i^n}^{t_{i+1}^n} f(s, x_i^n)ds\right) + \int_{t_2}^{t_1} f(s, x_n(\delta_n(t_1)))ds.$$

Hence, because of $(\mathcal{H}_3^f)$ and Lemma 3.2.3,

$$\begin{aligned}\|x_n(t_1) - x_n(t_2)\| &\leq (\kappa\mu_n + \frac{\mu_n}{m^2}(cL + d)) + |t_i^n - t_i^{n+1}|(cL + d) + |t_1 - t_2|(cL + d)\\ &\leq \kappa\mu_n + \mu_n(cL + d)(2 + \frac{1}{m^2}),\end{aligned}$$

which proves (i).

To establish (ii), first we note that, thanks to Remark 3.2.2, $\nabla\varphi(x_\iota(\theta_\iota(t))) \in C(\theta_\iota(t))$, with $\iota = n, k$. Then, by virtue of $(\mathcal{H}_C)$ and (ii) in Proposition 1.1.3, one gets

$$\begin{aligned}e(C(\theta_n(t)), \{\nabla\varphi(x_k(\theta_k(t)))\}) &\leq e(C(\theta_n(t), C(\theta_k(t))) + e(C(\theta_k(t)), \{\nabla\varphi(x_k(\theta_k(t)))\})\\ &\leq \kappa|\theta_n(t) - \theta_k(t)|.\end{aligned}$$

The previous inequality and $(\mathcal{H}^1_\varphi)$ yield:

$$\begin{aligned}\|x_n(\theta_n)(t) - x_k(\theta_k)(t)\| &\leq \frac{1}{m}\|\nabla\varphi(x_n(\theta_n(t))) - \nabla\varphi(x_k(\theta_k(t)))\| \\ &\leq \frac{\kappa}{m}|\theta_n(t) - \theta_k(t)|,\end{aligned}$$

we can see that, for $t \in I^n_i \cap I^k_j$, we have $|\theta_n(t) - \theta_k(t)| \leq |\theta_n(t) - t| + |t - \theta_k(t)| \leq \mu_n + \mu_k$, which concludes this part of the proof. Finally, one gets

$$\|x_n(t_1) - x_k(t_1)\| \leq \|x_n(t_1) - x_n(\theta_n(t))\| + \|x_n(\theta_n(t)) - x_k(\theta_k(t))\| + \|x_k(\theta_k(t)) - x_k(t)\|.$$

Then by virtue of (i),

$$\|x_n(t_1) - x_n(\theta_n(t))\| \leq \kappa\mu_n + (cL + d)\left(2\mu_n + \frac{\mu_n}{m^2}\right),$$

and

$$\|x_k(\theta_k(t)) - x_k(t)\| \leq \kappa\mu_k + (cL + d)\left(2\mu_k + \frac{\mu_k}{m^2}\right).$$

Thus, because of (ii),

$$\|x_n(\theta_n(t)) - x_k(\theta_k(t))\| \leq \frac{\kappa(\mu_n + \mu_k)}{m},$$

which concludes the proof. □

Lemma 3.2.8 *The following inclusions hold for every $n, k \in \mathbb{N}$, for a.e. $t \in I$,*

$$v_n(t) := -\frac{1}{M(\widetilde{\gamma} + \kappa)}D^2\varphi(x_n(\theta_n(t)))(\dot{x}_n(t) - f(t, x_n(\delta_n(t)))) \in N(\nabla\varphi^*(C(t)); x_n(\theta_n(t))) \cap \mathbb{B},$$

$$v_k(t) := -\frac{1}{M(\widetilde{\gamma} + \kappa)}D^2\varphi(x_k(\theta_k(t)))(\dot{x}_k(t) - f(t, x_k(\delta_k(t)))) \in N(\nabla\varphi^*(C(t)); x_k(\theta_k(t))) \cap \mathbb{B}.$$

Consequently, for a.e. $t \in I$

$$\langle v_n(t) - v_k(t), x_n(\theta_n(t)) - x_k(\theta_k(t))\rangle \geq -\frac{1}{\tilde{\rho}}\|x_n(\theta_n(t)) - x_k(\theta_k(t))\|^2.$$

The next remark will be necessary for the proof.

Remark 3.2.9 By $(\mathcal{H}^2_\varphi)$ and $(\mathcal{H}^3_\varphi)$, one gets $\|D^2\varphi(u)\| \leq M$ for every u where $\nabla\varphi$ is differentiable.

Proof. Firstly, by virtue of Corollary 2.2.13, $\nabla\varphi^*(C(t))$ is $\frac{1}{2\eta\vartheta}-$uniformly prox-regular, for every $t \in [0, T]$. Secondly, we check for the hypotheses for the composition rule in Proposition 1.3.9. Consider the mapping $g = \nabla\varphi$. By $(\mathcal{H}^2_\varphi)$, g is Lipschitz continuous. Furthermore, Hypothesis $(\mathcal{H}^3_\varphi)$ gives us strict differentiability.

Finally, thanks to Lemma 2.2.12, g is metrically calm relatively to C. Thus, by virtue of the composition rule, for a.e. $t \in I$ and for any $\overline{x} \in \nabla\varphi^*(C(t))$, one gets the set equality:

$$N(\nabla\varphi^*(C(t)); \overline{x}) = D\nabla\varphi(\overline{x})^* N(C(t); \nabla\varphi(\overline{x})).$$

Therefore,

$$N(\nabla\varphi^*(C(t));\overline{x}) = D^2\varphi(\overline{x})N(C(t);\nabla\varphi(\overline{x})), \tag{3.11}$$

where we used that $D^2\varphi$ is self adjoint (Proposition 1.1.14). Now we consider $n,k\in\mathbb{N}$. By virtue of Corollary 3.2.5, we obtain that

$$\begin{aligned}
-\dot{x}_n(t)+f(t,x_n(\delta_n(t))) &\in N(C(\theta_n(t));\nabla\varphi(x_n(\theta_n(t)))),\\
-\dot{x}_k(t)+f(t,x_k(\delta_k(t))) &\in N(C(\theta_k(t));\nabla\varphi(x_k(\theta_k(t)))).
\end{aligned}$$

Then, it is clear that

$$\begin{aligned}
-D^2\varphi(x_n(\theta_n(t)))(\dot{x}_n(t)-f(t,x_n(\delta_n(t)))) &\in D^2\varphi(x_n(\theta_n(t))N(C(\theta_n(t));\nabla\varphi(x_n(\theta_n(t)))),\\
-D^2\varphi(x_k(\theta_k(t)))(\dot{x}_k(t)-f(t,x_k(\delta_k(t)))) &\in D^2\varphi(x_k(\theta_k(t)))N(C(\theta_k(t));\nabla\varphi(x_k(\theta_k(t)))),
\end{aligned}$$

which, by (3.11), imply that

$$\begin{aligned}
-D^2\varphi(x_n(\theta_n(t)))(\dot{x}_n(t)-f(t,x_n(\delta_n(t)))) &\in N(\nabla\varphi^*(C(\theta_n(t));x_n(\theta_n(t)))),\\
-D^2\varphi(x_k(\theta_k(t)))(\dot{x}_k(t)-f(t,x_k(\delta_k(t)))) &\in N(\nabla\varphi^*(C(\theta_k(t));x_k(\theta_k(t)))).
\end{aligned}$$

Next, from the proof of Proposition 3.2.4, for $\iota\in\{n,k\}$ one gets that

$$\|\dot{x}_\iota(t)-f(t,x_\iota(\delta_\iota(t)))\|\leq\widetilde{\gamma}+\kappa, \tag{3.12}$$

where $\widetilde{\gamma}:=\gamma-(cL+d)$. Thus, as a consequence of the upper bound in Remark 3.2.9 and (3.12) the proof of the normal cone inclusion is complete.

Finally, the prox-regularity of $\nabla\varphi^*(C(t))$ and assertion (ii) of Theorem 1.3.5 end the proof. □

We now have all the necessary equipment to prove the Cauchy criterion in Proposition 3.2.6.

Proof of Proposition 3.2.6. We shall prove the convergence of (x_n) to a solution of the degenerate sweeping process $(\mathcal{MSP})$. To this end, we consider the $\varphi-$Symmetrized Bregman divergence between $x_n(t)$ and $x_k(t)$:

$$\Upsilon(t):=S_\varphi(x_n(t),x_k(t))=\langle\nabla\varphi(x_n(t))-\nabla\varphi(x_k(t)),x_n(t)-x_k(t)\rangle.$$

According to $(\mathcal{H}_\varphi^2)$, we observe that Υ is absolutely continuous with derivative

$$\begin{aligned}
\dot{\Upsilon}(t) &= \langle D^2\varphi(x_n(t))\dot{x}_n(t)-D^2\varphi(y_n(t))\dot{x}_k(t),x_n(t)-x_k(t)\rangle\\
&\quad+\langle\nabla\varphi(x_n(t))-\nabla\varphi(x_k(t)),\dot{x}_n(t)-\dot{x}_k(t)\rangle.
\end{aligned}$$

For the sake of readability, let us write,

$$\dot{\Upsilon}(t)=I_1(t)+I_2(t),$$

where

$$\begin{aligned}
I_1(t) &:= \langle D^2\varphi(x_n(t))\dot{x}_n(t)-D^2\varphi(x_k(t))\dot{x}_k(t),x_n(t)-x_k(t)\rangle,\\
I_2(t) &:= \langle\nabla\varphi(x_n(t))-\nabla\varphi(x_k(t)),\dot{x}_n(t)-\dot{x}_k(t)\rangle.
\end{aligned}$$

In order to arrive at the necessary condition we will find and upper bound for both I_1 and I_2.

Estimation of $I_1(t)$: We observe that $I_1(t) = I_{11}(t) + I_{12}(t) + I_{13}(t)$, where

$$\begin{array}{rcl} I_{11}(t) & := & \langle D^2\varphi(x_n(t))\dot{x}_n(t) - D^2\varphi(x_n(\theta_n(t)))\dot{x}_n(t), x_n(t) - x_k(t)\rangle, \\ I_{12}(t) & := & \langle D^2\varphi(x_n(\theta_n(t)))\dot{x}_n(t) - D^2\varphi(x_k(\theta_k(t)))\dot{x}_k(t), x_n(t) - x_k(t)\rangle, \\ I_{13}(t) & := & \langle D^2\varphi(x_k(t))\dot{x}_k(t) - D^2\varphi(x_k(\theta_k(t)))\dot{x}_n(t), x_n(t) - x_k(t)\rangle. \end{array}$$

We proceed in finding the required upper bounds. By virtue $(\mathcal{H}^2_\varphi)$, (i) and (iii) from Lemma 3.2.7, and (3.9), we obtain that for a.e. $t \in I_i^n$

$$\begin{array}{rcl} I_{11}(t) & \leq & \vartheta\|\dot{x}_n(t)\|\|x_n(t) - x_n(\theta_n(t))\|\|x_n(t) - x_k(t)\| \\ & \leq & \vartheta(\gamma+\kappa)\mu_n(\gamma+\kappa)(\mu_n+\mu_k)(\kappa(1+\dfrac{1}{m})+\gamma) \\ & = & \vartheta(\gamma+\kappa)^2\mu_n(\mu_n+\mu_k)(\kappa(1+\dfrac{1}{m})+\gamma), \end{array}$$

analogously, one gets

$$I_{13}(t) \leq \vartheta(\gamma+\kappa)^2\mu_k(\mu_n+\mu_k)(\kappa(1+\frac{1}{m})+\gamma).$$

The estimation of I_{12} needs a little more work. We note that $I_{12}(t) = J_1(t) + J_1(t) + J_1(t)$, where

$$\begin{array}{rcl} J_1(t) & = & \langle D^2\varphi(x_n(\theta_n(t)))\dot{x}_n(t) - D^2\varphi(x_k(\theta_k(t)))\dot{x}_k(t), x_n(t) - x_n(\theta_n(t))\rangle, \\ J_2(t) & = & \langle D^2\varphi(x_n(\theta_n(t)))\dot{x}_n(t) - D^2\varphi(x_k(\theta_k(t)))\dot{x}_k(t), x_n(\theta_n(t)) - x_k(\theta_k(t))\rangle, \\ J_3(t) & = & \langle D^2\varphi(x_n(\theta_n(t)))\dot{x}_n(t) - D^2\varphi(x_k(\theta_k(t)))\dot{x}_k(t), x_k(\theta_k(t)) - x_k(t)\rangle. \end{array}$$

Using Remark 3.2.9, (i) from Lemma 3.2.7 and (3.9), one gets

$$\begin{array}{rcl} J_1(t) & \leq & (M\|\dot{x}_n(t)\| + M\|\dot{x}_k(t)\|)\,\|x_n(t) - x_n(\theta_n(t))\| \\ & \leq & \dfrac{2M(\gamma+\kappa)}{m}\mu_n(\gamma+\kappa) \\ & = & \dfrac{2M\mu_n(\gamma+\kappa)^2}{m}, \end{array}$$

and similarly,

$$J_3(t) \leq \frac{2M\mu_k(\gamma+\kappa)^2}{m}.$$

For $J_2(t)$, set $z_n(t) = \dot{x}_n(t) - f(t, x_n(\delta_n(t)))$, $z_k(t) = \dot{x}_k(t) - f(t, x_k(\delta_k(t)))$. We note that $J_2(t) = J_{21}(t) + J_{22}(t) + J_{23}(t)$, where

$$\begin{array}{rcl} J_{21}(t) & := & D^2\varphi(x_n(\theta_n(t)))z_n(t) - D^2\varphi(x_k(\theta_k(t)))\dot{x}_k(t)z_k(t), x_n(\theta_n(t)) - x_k(\theta_k(t))\rangle, \\ J_{22}(t) & := & [D^2\varphi(x_n(\theta_n(t))) - D^2\varphi(x_k(\theta_k(t)))]f(t, x_n(\delta_n(t))), x_n(\theta_n(t)) - x_k(\theta_k(t))\rangle, \\ J_{23}(t) & := & D^2\varphi(x_k(\theta_k(t)))(f(t, x_n(\delta_n(t))) - f(t, x_k(\delta_k(t)))), x_n(\theta_n(t)) - x_k(\theta_k(t))\rangle. \end{array}$$

By virtue of Lemma 3.2.8 and (iii) from Lemma 3.2.7, it holds that

$$\begin{aligned} J_{21}(t) &\leq \frac{M(\widetilde{\gamma}+\kappa)}{\widetilde{\rho}}\|x_n(\theta_n(t)) - x_k(\theta_k(t))\|^2 \\ &\leq \frac{M(\widetilde{\gamma}+\kappa)}{\widetilde{\rho}}[(\mu_n+\mu_k)(\kappa(1+\frac{1}{m})+\gamma)]^2 \text{ for a.e. } t \in I. \end{aligned}$$

Using $(\mathcal{H}_\varphi^3)$, $(\mathcal{H}_3^f)$, Proposition 3.2.3 and (ii) from Lemma 3.2.7, we obtain that

$$\begin{aligned} J_{22}(t) &\leq \vartheta\|x_n(\theta_n(t)) - x_k(\theta_k(t))\| \cdot \|f(t, x_n(\delta_n(t)))\| \\ &\leq \vartheta\kappa(cL+d)\frac{\mu_n+\mu_k}{m}. \end{aligned}$$

Lastly, using Remark 3.2.9, $(\mathcal{H}_2^f)$, Proposition 3.2.3 and (iii) from Lemma 3.2.7, it holds that

$$\begin{aligned} J_{23}(t) &\leq M\mu_{\|x_0\|+L}\|x_n(\theta_n(t)) - x_k(\theta_k(t))\|^2 \\ &\leq M\mu_{\|x_0\|+L}(\mu_n+\mu_k)(\kappa(1+\frac{1}{m})+\gamma). \end{aligned}$$

Estimation of $I_2(t)$: We note that $I_2(t) = I_{21}(t) + I_{22}(t) + I_{23}(t)$, where

$$\begin{aligned} I_{21}(t) &:= \langle \nabla\varphi(x_n(\theta_n(t))) - \nabla\varphi(x_k(\theta_k(t))), \dot{x}_n(t) - \dot{x}_k(t)\rangle, \\ I_{22}(t) &:= \langle \nabla\varphi(x_k(\theta_k(t))) - \nabla\varphi(x_k(t)), \dot{x}_n(t) - \dot{x}_k(t)\rangle, \\ I_{23}(t) &:= \langle \nabla\varphi(x_n(t)) - \nabla\varphi(x_n(\theta_n(t))), \dot{x}_n(t) - \dot{x}_k(t)\rangle. \end{aligned}$$

To estimate I_{21}, we note that $I_{21}(t) = G_1(t) + G_2(t)$, where

$$\begin{aligned} G_1(t) &= \langle \nabla\varphi(x_n(\theta_n(t))) - \nabla\varphi(x_k(\theta_k(t))), \dot{x}_n(t) - f(t, x_n(\delta_n(t))) - (\dot{x}_k(t) - f(t, x_k(\delta_k(t))))\rangle \\ G_2(t) &= \langle \nabla\varphi(x_n(\theta_n(t))) - \nabla\varphi(x_k(\theta_k(t))), f(t, x_n(\delta_n(t))) - f(t, x_k(\delta_k(t)))\rangle \end{aligned}$$

On the one hand (3.10) in Corollary 3.2.5 and the monotonicity of the normal cone imply that

$$G_1(t) \leq 0 \text{ for a.e. } t \in I.$$

On the other hand, $(\mathcal{H}_\varphi^2)$, $(\mathcal{H}_2^f)$, Proposition 3.2.3 and (ii) and (iii) from Lemma 3.2.7 imply that

$$\begin{aligned} G_2(t) &\leq M\|x_n(\theta_n(t)) - x_k(\theta_k(t))\| \cdot \mu_{\|x_0\|+L}\|x_n(\delta_n(t)) - x_k(\delta_k(t))\| \\ &\leq M\mu_{\|x_0\|+L}\frac{\kappa}{m}(\mu_n+\mu_k)(\mu_n+\mu_k)(\kappa(1+\frac{1}{m})+\gamma) \\ &= \frac{M\kappa\mu_{\|x_0\|+L}}{m}(\kappa(1+\frac{1}{m})+\gamma)(\mu_n+\mu_k)^2 \\ &= \frac{M\kappa\mu_{\|x_0\|+L}(\kappa(1+m)+\gamma m)}{m^2}(\mu_n+\mu_k)^2. \end{aligned}$$

By virtue of $(\mathcal{H}_\varphi^2)$, (3.9) in Proposition 3.2.4 and (i) from Lemma 3.2.7, we arrive at the following inequalities

$$I_{22}(t) \leq 2M\mu_k(\gamma+\kappa)^2 \quad \text{and} \quad I_{23}(t) \leq 2M\mu_n(\gamma+\kappa)^2.$$

Finally, putting everything together, we arrive at the following inequality:

$$\dot{\Upsilon}(t) \le C_{n,k} \quad \text{for a.e. } t \in I,$$

where $C_{n,k} \to 0$ as $n, k \to \infty$ as it only involves the terms μ_n, μ_k that tend to 0, and some positive finite constants.

This, and the fact that $\Upsilon(0) = 0$ imply that, for all $t \in I$,

$$\Upsilon(t) \le TC_{n,k}.$$

Now, we define the auxiliary function $H(t) := \int_0^t \|x_n(\tau) - x_k(\tau)\|^2 d\tau$. It is clear that H is absolutely continuous on I and $\dot{H}(t) = \|x_n(t) - x_k(t)\|^2$, then hypothesis $(\mathcal{H}_\varphi^1)$ implies that

$$m\dot{H}(t) \le \Upsilon(t) \le TC_{n,k} \quad \text{for all } t \in I.$$

Finally, integrating over $[0, T]$, it results that there exists a constant $\mathcal{C}_{n,k} > 0$ that satisfies the desired conclusion. □

The next result proves the convergence of $(\mathcal{MCU})$ to a solution of the Mirror Sweeping Process. The proof follows known techniques when approximating a solution of a sweeping process (see [15, 19, 46, 49, 59]).

Proposition 3.2.10 *The family $(x_n)_{n\in\mathbb{N}}$ converges uniformly to a solution of the Mirror Sweeping Process,*

$$\begin{cases} \dot{x}(t) \in -N(C(t); \nabla\varphi(x(t))) + f(t, x(t)) & a.e.\ t \in I, \\ x(T_0) = x_0. \end{cases} \tag{$\mathcal{MSP}$}$$

This solution is absolutely continuous, it holds that $\|\dot{x}(t)\| \le \gamma + \kappa$ and $\|\dot{x}(t) - f(t, x(t))\| \le \widetilde{\gamma} + \kappa$ for a.e. $t \in I$.

Proof. First, we will show that there is, in fact, a limit. Indeed, by Proposition 3.2.6, we have a Cauchy criterion for the family of approximated solutions defined by the Mirror Catching-up Algorithm as n approaches infinity. Hence, when $n \to \infty$, $x_n(\cdot)$ converges uniformly to a limit $x(\cdot)$ in the space of continuous functions defined over the interval I, namely $x(\cdot) \in C(I; \mathcal{H})$.

Secondly, due to (3.9) in Proposition 3.2.4, we know that

$$\|\dot{x}_n(t)\| \le \gamma + \kappa \quad \text{for all } n \in \mathbb{N}. \tag{3.13}$$

Therefore, the set $K := \{\dot{x}_n : n \in \mathbb{N}\} \subset L^1(I; \mathcal{H})$ is uniformly integrable (Definition 1.1.8). Then, by the Dunford-Pettis theorem (Theorem 1.1.9), K is relatively compact in the weak topology. As a consequence, there exists a subsequence $\dot{x}_{n_j}$ that converges weakly to some z. This limit satisfies

$$\|z(t)\| \le \gamma + \kappa, \text{ a.e. } t \in I,$$

thanks to the fact that the set $\{w \in L^1(I; \mathcal{H}) : \|w(t)\| \le \gamma + \kappa, \text{ a.e. } t \in I\}$ is convex and closed.

In order to make use of the structure of the Hilbert space $\mathcal{H}$, fix $h \in \mathcal{H}$, we have that

$$\langle h, \int_0^t \dot{x}_{n_j}(\tau)d\tau\rangle = \int_0^t \langle h\mathbb{1}_{[0,t]}(\tau), \dot{x}_{n_j}(\tau)\rangle d\tau.$$

This implies the following weak convergence in $\mathcal{H}$,

$$\int_0^t \dot{x}_{n_j}(\tau)d\tau \rightharpoonup \int_0^t z(\tau)d\tau. \tag{3.14}$$

To take advantage of the above, we note that, by absolute continuity of $x_n(\cdot)$, we have the equality

$$x_{n_j}(t) = x_0 + \int_0^t \dot{x}_{n_j}(\tau)d\tau.$$

Hence, as x_n converges strongly, (3.14) implies that

$$x_{n_j}(t) \to x(t) = x_0 + \int_0^t z(\tau)d\tau. \tag{3.15}$$

By definition, this means that the limit $x(\cdot)$ is absolutely continuous with $\dot{x}(t) = z(t)$ for a.e. $t \in I$. Thus $\dot{x}_{n_j}(\cdot)$ weakly converges to $\dot{x}(\cdot)$.

Now, we show that the limit fulfills the condition $\nabla\varphi(x(t)) \in C(t)$ for all $t \in I$. By Remark 3.2.2, we know that $\nabla\varphi(x_n(\theta_n(t))) \in C(\theta_n(t))$, it is also true that $\nabla\varphi(x_n(\theta_n(t))) \to \nabla\varphi(x_n(t))$. Then, by virtue of $(\mathcal{H}_C)$ for every $t \in I$

$$d(C(t); \nabla\varphi(x_n(t))) \leq e(C(t), C(\theta_n(t))) \leq \kappa|\theta_n(t) - t|,$$

as $n \to \infty$, one gets $\theta_n(t) - t \to 0$, which, by closedness of $C(t)$, means that $\nabla\varphi(x(t)) \in C(t)$.

As in (3.12), for all $n \in \mathbb{N}$, it holds that

$$\|\dot{x}_n(t) - f(t, x_n(\delta_n(t)))\| \leq \widetilde{\gamma} + \kappa. \tag{3.16}$$

The following inclusion holds due to Corollary 3.2.5,

$$\dot{x}_n(t) \in -N(C(\theta_n(t)); \nabla\varphi(x_n(\theta_n(t)))) + f(t, x_n(\delta_n(t))) \text{ for all } n \in \mathbb{N}. \tag{3.17}$$

Applying Mazur's lemma (Lemma 1.1.7), there exists a sequence z_{n_j} of convex combinations of the elements of x_{n_j} such that $\sum_{j=0}^{l} \alpha_{j,l} x_{n_j}$ converges strongly to $\dot{\widetilde{x}}(t) - f(t, \widetilde{x}(t))$ in $L^1(I, \mathcal{H})$, where for all $l \in \mathbb{N}$,

$$\sum_{j=0}^{l} \alpha_{j,l} = 1, \quad (\alpha_{j,l})_{0 \leqslant j \leqslant l} \in [0,1]^{l+1} \text{ for all } l \in \mathbb{N}. \tag{3.18}$$

By considering the extraction of a subsequence, it is possible to assume that, for a negligible set $J \subset I$, the derivatives $\dot{x}(t)$ and $\dot{x}_{n_j}(t)$ remain well-defined for all $t \in I \setminus J$. We now fix $t \in I \setminus J$, firstly, by definition of the normal cone in the convex case and (3.17), for any $w \in C(T)$ the following holds:

$$
\begin{aligned}
&\sum_{j=0}^{l}\alpha_{j,l}\left\langle -[\dot{x}_{n_j}(t) - f(t, x_{n_j}(\delta_{n_j}(t)))], w - \nabla\varphi(x_{n_j}(\theta_{n_j}(t)))\right\rangle \\
&\quad = \sum_{j=0}^{l}\alpha_{j,l}\left\langle -[\dot{x}_{n_j}(t) - f(t, x_{n_j}(\delta_{n_j}(t)))], \nabla\varphi(x(t)) - \nabla\varphi(x_{n_j}(\theta_{n_j}(t)))\right\rangle \\
&\quad\quad + \sum_{j=0}^{l}\alpha_{j,l}\left\langle -[\dot{x}_{n_j}(t) - f(t, x_{n_j}(\delta_{n_j}(t)))], w - \nabla\varphi(x(t))\right\rangle \\
&\quad \le \sum_{j=0}^{l}\alpha_{j,l}\left\langle -[\dot{x}_{n_j}(t) - f(t, x_{n_j}(\delta_{n_j}(t)))], \nabla\varphi(x(t)) - \nabla\varphi(x_{n_j}(\theta_{n_j}(t)))\right\rangle. \qquad (3.19)
\end{aligned}
$$

Secondly, by (3.9), we know that

$$
\begin{aligned}
&\langle -[\dot{y}_{n_j}(t) - f(t, x_{n_j}(\delta_{n_j}(t)))], \nabla\varphi(x(t)) - \nabla\varphi(x_{n_j}(\theta_{n_j}(t)))\rangle \\
&\le (\gamma + \kappa)\|\nabla\varphi(x(t)) - \nabla\varphi(x_{n_j}(\theta_{n_j}(t)))\|,
\end{aligned}
$$

which, by the continuity hypothesis of f, along with all the previous results, imply that as $j \to \infty$, it holds:

$$
-\dot{x}(t) + f(t, x(t)) \in N(C(t); \nabla\varphi(x(t))) \quad \forall t \in I \setminus J.
$$

Henceforth, it follows that

$$
\dot{x}(t) \in -N(C(t); \nabla\varphi(x(t))) + f(t, x(t)) \quad \text{a.e. } t \in I.
$$

Finally, by (3.16) and and the continuity hypothesis of f, it holds that $\|\dot{x}(t) - f(t, x(t))\| \le \widetilde{\gamma} + \kappa$. ☐

Lemma 3.2.11 *The solution to* $(\mathcal{MSP})$ *is unique.*

Proof. We assume that there exists two solutions $x_1(\cdot), x_2(\cdot)$, then

$$
\begin{cases}
\dot{x}_1(t) \in -N(C(t); \nabla\varphi(x_1(t))) + f(t, x_1(t)) & \text{a.e. } t \in I,\\
\dot{x}_2(t) \in -N(C(t); \nabla\varphi(x_2(t))) + f(t, x_2(t)) & \text{a.e. } t \in I,\\
x_1(0) = x_0 = x_2(0).
\end{cases}
$$

The proof is similar to Proposition 3.2.6. As in Proposition 3.2.6, we consider the φ-Symmetrized Bregman divergence between $x_1(t)$ and $x_2(t)$:

$$
\Upsilon(t) := \langle \nabla\varphi(x_1(t)) - \nabla\varphi(x_2(t)), x_1(t) - x_2(t)\rangle,
$$

with derivative

$$
\begin{aligned}
\dot{\Upsilon}(t) &= \langle D^2\varphi(x_1(t))\dot{x}_1(t) - D^2\varphi(x_2(t))\dot{x}_2(t), x_1(t) - x_2(t)\rangle \\
&\quad + \langle \nabla\varphi(x_1(t)) - \nabla\varphi(x_2(t)), \dot{x}_1(t) - \dot{x}_2(t)\rangle.
\end{aligned}
$$

Analogously to the previous proof, we write:

$$\dot{\Upsilon}(t) = I_1(t) + I_2(t),$$

where

$$\begin{aligned} I_1(t) &:= \langle D^2\varphi(x_1(t))\dot{x}_1(t) - D^2\varphi(x_2(t))\dot{x}_2(t), x_1(t) - x_2(t)\rangle, \\ I_2(t) &:= \langle \nabla\varphi(x_1(t)) - \nabla\varphi(x_2(t)), \dot{x}_1(t) - \dot{x}_2(t)\rangle. \end{aligned}$$

Estimation of $I_2(t)$: Let us first find and upper bound for I_2. By the definition of the solutions, it follows that, for a.e. $t \in I$:

$$\begin{aligned} \dot{x}_1(t) - f(t, x_1(t)) &\in -N(C(t); \nabla\varphi(x_1(t))), && (3.20) \\ \dot{x}_2(t) - f(t, x_2(t)) &\in -N(C(t); \nabla\varphi(x_2(t))), && (3.21) \\ x_1(0) = x_2(0) &= x_0. && (3.22) \end{aligned}$$

We break up I_2 into two terms:

$$I_2(t) = I_{21}(t) + I_{22}(t),$$

where

$$\begin{aligned} I_{21}(t) &:= \langle \nabla\varphi(x_1(t)) - \nabla\varphi(x_2(t)), \dot{x}_1(t) - f(t, x_1(t)) - (\dot{x}_2(t) - f(t, x_2(t)))\rangle, \\ I_{22}(t) &:= \langle \nabla\varphi(x_1(t)) - \nabla\varphi(x_2(t)), f(t, x_1(t)) - f(t, x_2(t))\rangle. \end{aligned}$$

On the one hand, due to (3.20), (3.21) and the monotonicity of the normal cone, it holds that $I_{21}(t) \leq 0$ for a.e. $t \in I$. On the other hand, we use that the interval $I = [0, T]$ is finite, the fact that both $\nabla\varphi(x_1(t))$ and $\nabla\varphi(x_2(t))$ must belong to $C(t)$, (3.22) and $(\mathcal{H}_C)$ to obtain that

$$\nabla\varphi(x_1(t)) \in \mathbb{B}\left(\nabla\varphi(x_0), \kappa T\right) \text{ and } \nabla\varphi(x_2(t)) \in \mathbb{B}\left(\nabla\varphi(x_0), \kappa T\right) \text{ for a.e. } t \in I. \tag{3.23}$$

Using (3.23) and Lemma 2.2.3, for $i = 1, 2$, one gets

$$\begin{aligned} \|x_i(t) - x_0\| &= \|\nabla\varphi^*(\nabla\varphi((x_i(t)))) - \nabla\varphi^*(\nabla\varphi(x_0))| \\ &\leq \frac{1}{m}\|\nabla\varphi(x_i(t)) - \nabla\varphi(x_0)\| \\ &\leq \frac{\kappa T}{m}, \end{aligned}$$

which implies that

$$x_1(t) \in \mathbb{B}\left(x_0, \frac{\kappa T}{m}\right) \text{ and } x_2(t) \in \mathbb{B}\left(x_0, \frac{\kappa T}{m}\right) \text{ for a.e. } t \in I. \tag{3.24}$$

Then, by using (3.24), $(\mathcal{H}_\varphi^2)$ and $(\mathcal{H}_2^f)$, for a.e. $t \in I$, it holds that

$$I_{22}(t) \leq M\mu_{\|x_0\| + \frac{\kappa T}{m}}\|x_1(t) - x_2(t)\|^2.$$

Estimation of $I_1(t)$: For the case of I_1, we are in need of the following claim, related to Lemma 3.2.8.

Claim 1: The following differential inclusions hold for a.e. $t \in I$:

$$\begin{aligned} \frac{1}{\widetilde{\gamma}+\kappa} D^2\varphi(x_1(t))(\dot{x}_1(t) - f(t,x_1(t))) \quad &\in \quad -N(\nabla\varphi^*(C(t)); x_1(t)) \cap \mathbb{B}, \\ \frac{1}{\widetilde{\gamma}+\kappa} D^2\varphi(x_2(t))(\dot{x}_2(t) - f(t,x_2(t))) \quad &\in \quad -N(\nabla\varphi^*(C(t)); x_2(t)) \cap \mathbb{B}. \end{aligned}$$

Proof of Claim 1. Due to (3.11) in the proof of Lemma 3.2.8, we have, for a.e. $t \in I$, for every $\overline{x} \in \nabla\varphi^*(C(t))$, the following equality

$$N(\nabla\varphi^*(C(t)); \overline{x}) = D^2\varphi(\overline{x})N(C(t); \nabla\varphi(\overline{x})). \tag{3.25}$$

By assumption, $x_1(\cdot)$ and $x_2(\cdot)$ are both solutions, then it holds that $\nabla\varphi(x_i(t)) \in C(t)$ for a.e. $t \in I$, with $i = 1, 2$. As a consequence, $x_1(t),\ x_2(t) \in \nabla\varphi^*(C(t))$ for a.e. $t \in I$, then (3.25) holds for both $x_1(\cdot)$ and $x_2(\cdot)$.

The inclusions (3.20), (3.21) and (3.25) imply that, for a.e. $t \in I$,

$$\begin{aligned} -D^2\varphi(x_1(t))(\dot{x}_1(t) - f(t,x_1(t))) \in D^2\varphi(x_1(t))N(C(t); \nabla\varphi(x_1(t))), \\ -D^2\varphi(x_2(t))(\dot{x}_2(t) - f(t,x_2(t))) \in D^2\varphi(x_2(t))N(C(t); \nabla\varphi(x_2(t))), \end{aligned}$$

which, by (3.25) and the upper bound in Proposition 3.2.10, imply that

$$\begin{aligned} \frac{1}{\widetilde{\gamma}+\kappa} D^2\varphi(x_1(t))(\dot{x}_1(t) - f(t,x_1(t))) \in -N(\nabla\varphi^*(C(t)); x_1(t))) \cap \mathbb{B}, \\ \frac{1}{\widetilde{\gamma}+\kappa} D^2\varphi(x_2(t))(\dot{x}_2(t) - f(t,x_2(t))) \in -N(\nabla\varphi^*(C(t)); x_2(t))) \cap \mathbb{B}. \end{aligned}$$

which is our claim. $\square$

With this claim in hand, we observe that, for $t \in I$,

$$I_1(t) = I_{11}(t) + I_{12}(t),$$

where

$$\begin{aligned} I_{11}(t) &:= \langle D^2\varphi(x_1(t))(\dot{x}_1(t) - f(t,x_1(t))) - D^2\varphi(x_2(t))(\dot{x}_2(t) - f(t,x_2(t))), x_1(t) - x_2(t)\rangle, \\ I_{12}(t) &:= \langle D^2\varphi(x_1(t))f(t,x_1(t)) - D^2\varphi(x_2(t))f(t,x_2(t)), x_1(t) - x_2(t)\rangle. \end{aligned}$$

Estimation of $I_{11}(t)$: Due to Claim 1 and prox-regularity of $\nabla\varphi^*(C(t))$ for $t \in I$ (Corollary 2.2.13), it holds that

$$I_{11}(t) \le \frac{\widetilde{\gamma}+\kappa}{\widetilde{\rho}}\|x_1(t) - x_2(t)\|^2 \quad \text{for a.e. } t \in I.$$

Estimation of $I_{12}(t)$: We observe that

$$I_{12}(t) = J_1(t) + J_2(t),$$

where

$$\begin{aligned} J_1(t) &:= \langle D^2\varphi(x_1(t))f(t,x_1(t))) - D^2\varphi(x_2(t))f(t,x_1(t))), x_1(t) - x_2(t)\rangle, \\ J_2(t) &:= \langle D^2\varphi(x_2(t))f(t,x_1(t)) - D^2\varphi(x_2(t))f(t,x_2(t)), x_1(t) - x_2(t)\rangle. \end{aligned}$$

Estimation of $J_1(t)$: Set $\sigma := \max\{\|f(t,x_1(t))\|, \|f(t,x_2(t))\|\}$, by virtue of $(\mathcal{H}_3^f)$ and (3.24), we know that $\sigma < \infty$, this and $(\mathcal{H}_\varphi^3)$ imply that:

$$\begin{aligned} J_1(t) &\leq \|D^2\varphi(x_1(t))f(t,x_1(t))) - D^2\varphi(x_2(t))f(t,x_1(t)))\| \cdot \|x_1(t) - x_2(t)\| \\ &\leq \|f(t,x_1(t))\| \cdot \|D^2\varphi(x_1(t)) - D^2\varphi(x_2(t))\| \cdot \|x_1(t) - x_2(t)\| \\ &\leq \sigma\vartheta\|x_1(t) - x_2(t)\|^2 \quad \text{for a.e. } t \in I. \end{aligned}$$

Estimation of $J_2(t)$: From By using Remark 3.2.9, $(\mathcal{H}_2^f)$ and (3.24), we get for a.e. $t \in I$,

$$\begin{aligned} J_2(t) &\leq \|D^2\varphi(x_2(t))f(t,x_1(t)) - D^2\varphi(x_2(t))f(t,x_2(t))\| \cdot \|x_1(t) - x_2(t)\| \\ &\leq \|D^2\varphi(x_2(t))\| \cdot \|f(t,x_1(t)) - f(t,x_2(t))\| \cdot \|x_1(t) - x_2(t)\| \\ &\leq M\mu_{\|x_0\|+\frac{\kappa T}{m}}\|x_1(t) - x_2(t)\|^2. \end{aligned}$$

In summary, we have arrived at the following inequality:

$$\dot{\Upsilon}(t) \leq C\|x_1(t) - x_2(t)\|^2 \quad \text{for a.e. } t \in I, \tag{3.26}$$

where C is a positive and finite constant. By a straightforward calculation,

$$\Upsilon(t) - \Upsilon(0) = \Upsilon(t) \leq C\int_0^t \|x_1(\tau) - x_2(\tau)\|^2 d\tau,$$

where we used (3.22) to conclude that $\Upsilon(0) = 0$.

Finally, we define the auxiliary function

$$H(t) := \int_0^t \|x_1(\tau) - x_2(\tau)\|^2 d\tau.$$

It follows immediately that H is absolutely continuous on I and

$$\dot{H}(t) = \|x_1(t) - x_2(t)\|^2,$$

from hypothesis $(\mathcal{H}_\varphi^1)$ and (3.26), it may be concluded that

$$m\dot{H}(t) \leq \Upsilon(t) \leq CH(t) \quad \forall t \in I,$$

Using Gronwall's Lemma on $H(\cdot)$, we obtain

$$H(t) \leq H(0)\exp\left(\int_{t_0}^t C d\tau\right) \quad \forall t \in I,$$

from (3.22), $H(0) = 0$, therefore we obtain that

$$H(t) = 0 \quad \forall t \in I,$$

or, equivalently,

$$\|x_1(t) - x_2(t)\|^2 = 0,$$

which implies the desired conclusion. □

Consequently, due to existence and bounds proved in Proposition 3.2.10, and the uniqueness in Lemma 3.2.11, we are able the state the main result of the thesis.

Theorem 3.2.12 *Assume that* $(\mathcal{H}_\varphi^1)$, $(\mathcal{H}_\varphi^2)$, $(\mathcal{H}_\varphi^3)$, $(\mathcal{H}_C)$, $(\mathcal{H}_1^f)$, $(\mathcal{H}_1^f)$ *and* $(\mathcal{H}_1^f)$ *hold. Then, the Mirror Sweeping Process:*

$$\begin{cases} \dot{x}(t) \in -N(C(t); \nabla\varphi(x(t))) + f(t, x(t)) & a.e.\ t \in I, \\ x(T_0) = x_0, \end{cases} \tag{$\mathcal{MSP}$}$$

admits a unique and Lipschitz continuous solution. This solutions satisfies $\|\dot{x}(t)\| \leq \gamma + \kappa$ *and* $\|\dot{x}(t) - f(t, x(t))\| \leq \widetilde{\gamma} + \kappa$ *for a.e.* $t \in I$.

We reiterate that no compactness hypothesis was necessary on the moving set, the operator is possibly nonlinear and the space is not necessary separable. Thus expanding the literature on the Degenerate Sweeping Process on infinite dimension.

3.3 Application to online optimization

In this thesis, we consider $(\mathcal{MSP})$ to link sweeping processes and online optimization. In [18], a dynamic using continuous-time mirror descent was used to propose a competitive randomized algorithm for the $k-$server problem, while in [52] this dynamic was linked to a degenerate sweeping process. We proceed to give an idea of the dynamic.

The mirror descent algorithm is a first order method of optimization for a problem of the type $(\mathcal{P}_f)$. It is related to the projected gradient descent, which, given a sufficiently smooth function $f\colon \mathcal{H} \to \mathbb{R}$ and a starting point $y_0 = x_0$, performs a gradient step:

$$y_{k+1} = x_k - t_k \nabla f(x_k), \quad t_k > 0, \tag{3.27}$$

and then projects back to the constraint set C :

$$x_{k+1} = \mathrm{Proj}_C(y_{k+1}).$$

Hypotheses such as Lipschitzianity of the gradient, and control of the term t_k, are enough to imply that projected gradient descent converges to a solution of the optimization problem.

On spaces where the norm is not induced by an inner product and therefore we can not use the Riesz representation theorem. Thus, step (3.27) is not possible, so an alternative is required. This alternative serves as a motivation for mirror descent.

The notion of a mirror map is employed to perform a gradient step in mirror descent, usually defined in finite dimension. Let $\mathcal{D} \subset \mathbb{R}^n$ such that the constraint set C satisfies $C \subset \overline{\mathcal{D}}$. The definition in [17] says that $\phi : \mathcal{D} \to \mathbb{R}$ is *mirror map* if the following are satisfied:

(i) ϕ is strictly convex and differentiable.

(ii) The gradient of ϕ is surjective.

(iii) The gradient of ϕ diverges in the border of $\mathcal{D}$, i.e.,

$$\lim_{x \to \mathrm{bdry}\, \mathcal{D}} ||\nabla\phi(x)|| = +\infty.$$

We observe that (i) and (iii) imply that ϕ is Legendre in the classical sense (see Definition 1.2.22). Statement (ii) is not always required. For example, in [9] ϕ must satisfy (i) but with strong convexity, and (iii), the work in [6] also drops hypothesis (ii). In this context, given a starting point x_0, mirror descent performs a gradient step in the "dual space":

$$\nabla\phi(y_{k+1}) = \nabla\phi(x_k) - t_k g_k, \quad t_k > 0,\ g_k \in \partial f(t_k, x_k), \tag{3.28}$$

where the objective function f depends on the time and state and is subdifferentiable. Mirror descent then projects back to the constraint set C :

$$x_{k+1} = \text{Proj}_C^{\varphi}(y_{k+1}), \tag{3.29}$$

here, $\text{Proj}_C^{\varphi}(y_{k+1})$ is the Bregman projection of y_{k+1} onto C. Whenever C is a nonempty, closed and convex set, this projection is well defined thanks to Theorem 1.2.33. See Figure 3.2 for a visual example of the stpes in the mirror descent algorithm. We refer the reader to

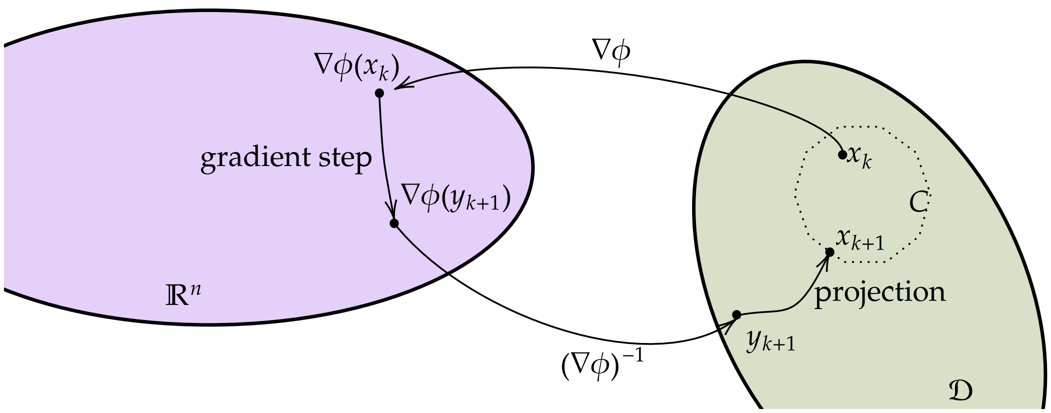

Figure 3.2: Example of Mirror descent. See [17, p. 299]

[4, 6, 9, 23, 34] for a deeper look into the mirror descent algorithm and its uses.

In an attempt to describe a continuous dynamic for mirror descent (see [9, 53, 60]), for a sufficiently small τ, we can propose the following steps instead of (3.28) and (3.29):

$$\begin{aligned} \nabla\phi(y(t+\tau)) &= \nabla\phi(x(t)) - \tau g(t, x(t)), \quad t_k > 0,\ g(t, x(t)) \in \partial f(t, x(t)) \\ x(t+\tau) &= \text{Proj}_C^{\phi}(y(t+\tau)). \end{aligned}$$

By definition of the Bregman projection and Theorem 1.2.27 we have the following set of equalities:

$$\begin{aligned} x(t+\tau) &= \operatorname{argmin}_{x\in C} D_\phi(x(t), y(t+\tau)) \\ &= \operatorname{argmin}_{x\in C} D_\phi\left(x(t), \nabla\phi^*(\nabla\phi(x(t)) - \tau g(t, x(t)))\right). \end{aligned}$$

Then, by the optimality condition in Proposition 1.2.20, one has

$$\nabla_x D_\phi\left(x(t+\eta), \nabla\phi^*(\nabla\phi(x_t) - \tau g(t, x(t)))\right) \in -N_C(x(t+\tau)). \tag{3.30}$$

Clearly, with these hypotheses D_φ is differentiable with respect to x and

$$\nabla_x D_\phi\left(z_1, z_2\right) = \nabla\varphi(z_1) - \nabla\phi(z_2).$$

Therefore, (3.30) turns into,

$$\nabla\phi(x(t+\tau)) - \nabla\phi[\nabla\phi^*(\nabla\phi(x_t) - \tau g(t, x(t)))] \in -N_C(x(t+\tau)),$$

which is equivalent to

$$(\nabla\phi(x(t+\tau)) - \nabla\phi(x(t))) - \tau g(t, x(t)) \in -N_C(x(t+\tau)).$$

Consequently, broadly speaking, dividing by τ and taking the limit $\tau \to 0$, one gets

$$\nabla^2\phi(x(t))\dot{x}(t) \in -g(t, x(t)) - N_C(x(t)), \tag{3.31}$$

which is the proposed dynamic in [18], used to obtain a $O((\log k)^2)$-competitive randomized algorithm for the k-server problem on hierarchically separated trees. Moreover, (3.31) is equivalent to

$$\dot{z}(t) \in -N_C(\nabla\varphi(z(t))) + h(t, z(t)) \text{ a.e. } t \in [0, T], \tag{3.32}$$

where we use the transformation $z(t) = \nabla\phi(x(t)),\ h(t, z(t)) = g(t, \nabla\phi^*(z(t))$ and take $\varphi = \phi^*$. Formulation (3.32) was used in [52] and motivates us to define the Mirror Sweeping Process ($\mathcal{MSP}$) and extend it to more general settings.

Theorem 3.2.12 guarantees the existence and uniqueness of a solution to (3.31) and (3.32).

Proposition 3.3.1 *Let C be a nonempty, closed and convex set satisfying $(\mathcal{H}_C)$. Let $\phi : \mathcal{H} \to \mathbb{R}$ satisfy $(\mathcal{H}^1_\varphi)$, $(\mathcal{H}^2_\varphi)$ and $(\mathcal{H}^3_\varphi)$. Assume that $g : [0, T] \times \mathcal{H} \to \mathcal{H}$ satisfies $(\mathcal{H}^f_1)$, $(\mathcal{H}^f_2)$ and $(\mathcal{H}^f_3)$. Then, for any initial point $x_0 \in C$, the dynamic* (3.32) *admits a unique, Lipschitz continuous solution.*

This result enhances the findings presented in [18, Theorem 2.2], as no compactness on the set is required. It is also possible to extend it to a setting with moving sets that are Lipschitz with respect to the excess.

Furthermore, we give an algorithm to find said solution, this will prove extremely useful when computing an explicit solution for the dynamic. Just as in the proof of the well-definedness of P^φ_C, the approximator operator can be obtained through the resolution of a constrained optimization problem.

For a possible choice for the *mirror map* (see [18]), consider $\delta > 0,\ n \in \mathbb{N},\ 1 \leq k \leq n$ and a set of positive weights $\{w_1, \ldots, w_n\}$. Set

$$\begin{aligned} \varphi \colon P_\delta &\to \mathbb{R} \\ x &\to \sum_{i=1}^n w_i x_i \log(x_i), \end{aligned}$$

where the domain is defined as

$$P_\delta := \left\{ x \in [\delta, 1]^n \ :\ \sum_{i=1}^n x_i = n - k \right\}.$$

Then, φ satisfies all the hypotheses of our setting. The previous function was used to find a competitive algorithm for the fractional weighted paging problem whenever $\delta = \frac{1}{2k}$. We refer to [18] for more details.

Conclusions

In this thesis, we have have applied techniques of convex and nonsmooth analysis to conduct an in-depth study of a Degenerate Sweeping Process in an infinite dimension Hilbert space.

In Chapter 2, we have rekindled the approximator operator introduced in [39], dropping the boundedness required for the set and providing a different approach for the proof. This approach allows for an explicit computation of the approximator operator. We provide several properties and characterizations of the operator. In addition, we provide conditions to where the preimage of a convex set, through a sufficiently smooth mapping, is prox-regular, and we supply explicitly the constant of prox-regularity.

In Chapter 3, via a modified catching-up algorithm, by means of the approximator operator, we proved the existence of a Lipschitz solution to a particular case of the degenerate sweeping process. This particular case is of great interest, as it provides a setting where the dynamic admits a unique, Lipschitz continuous solution. This result, as far as we know, has not been achieved in the literature in the case where the operator is possibly nonlinear and the moving set does not meet a compactness assumption.

With regards to the recently introduced Mirror Sweeping Process, there are multiple possible avenues to extend the work. For example, one can further investigate the hypotheses on the moving set, such as changing the convexity for prox-regularity, for a more general setting. It would also be interesting to study the case where the moving set is Lipschitz with respect to the *truncated excess* (or *truncated Hausdorff distance*), as seen in [51] both the excess and Hausdorff distance can be restrictive in some practical applications.

Finally, this path uncovers a bridge between the rich frameworks of differential inclusions and online optimization, which we hope to exploit in the future with real world applications.

Bibliography

[1] S. Adly. *A variational approach to nonsmooth dynamics*. SpringerBriefs in Mathematics. Springer, Cham, 2017. Applications in unilateral mechanics and electronics, With a foreword by J.-B. Hiriart-Urruty.

[2] S. Adly and T. Haddad. Well-posedness of nonconvex degenerate sweeping process via unconstrained evolution problems. *Nonlinear Anal. Hybrid Syst.*, 36:100832, 13, 2020.

[3] S. Adly and B. Le. Unbounded second-order state-dependent Moreau's sweeping processes in Hilbert spaces. *J. Optim. Theory Appl.*, 169(2):407–423, 2016.

[4] Z. Allen-Zhu and L. Orecchia. Linear coupling: an ultimate unification of gradient and mirror descent. In *8th Innovations in Theoretical Computer Science Conference*, volume 67 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 3, 22. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2017.

[5] H. Attouch, G. Buttazzo, and G. Michaille. *Variational analysis in Sobolev and BV spaces*, volume 6 of *MPS/SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA, 2006. Applications to PDEs and optimization.

[6] J.Y. Audibert, S. Bubeck, and G. Lugosi. Regret in online combinatorial optimization. *Math. Oper. Res.*, 39(1):31–45, 2014.

[7] H.-H. Bauschke, J. Borwein, and P. Combettes. Bregman monotone optimization algorithms. *SIAM J. Control Optim.*, 42(2):596–636, 2003.

[8] H.-H. Bauschke and P. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, Cham, second edition, 2017. With a foreword by Hédy Attouch.

[9] A. Beck and M. Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Oper. Res. Lett.*, 31(3):167–175, 2003.

[10] J. Borwein and J. Vanderwerff. *Convex functions: constructions, characterizations and counterexamples*, volume 109 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2010.

[11] J. Borwein and Q. Zhu. *Techniques of variational analysis*, volume 20 of *CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC*. Springer-Verlag, New York, 2005.

[12] C. Bounkhel, M.and Castaing. State dependent sweeping process in p-uniformly smooth and q-uniformly convex Banach spaces. *Set-Valued Var. Anal.*, 20(2):187–201, 2012.

[13] M. Bounkhel and R. Al-Yusof. First and second order convex sweeping processes in reflexive smooth Banach spaces. *Set-Valued Var. Anal.*, 18(2):151–182, 2010.

[14] M. Bounkhel and D.-L. Azzam. Existence results on the second-order nonconvex sweeping processes with perturbations. *Set-Valued Anal.*, 12(3):291–318, 2004.

[15] M. Bounkhel and L. Thibault. Nonconvex sweeping process and prox-regularity in Hilbert space. *J. Nonlinear Convex Anal.*, 6(2):359–374, 2005.

[16] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[17] S. Bubeck. Convex optimization: Algorithms and complexity. *Found. Trends Mach. Learn.*, 8(3–4):231–357, nov 2015.

[18] S. Bubeck, M. Cohen, Y. Lee, J. Lee, and A. Mądry. k-server via multiscale entropic regularization. In *STOC'18—Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 3–16. ACM, New York, 2018.

[19] C. Castaing. Équation différentielle multivoque avec contrainte sur l'état dans les espaces de Banach. *Travaux Sém. Anal. Convexe*, 8(2):exp. no. 13, 20, 1978.

[20] C. Castaing. Quelques problèmes d'évolution du second ordre. In *Seminaire d'Analyse Convexe, Vol. 18 (Montpellier, 1988)*, pages Exp. No. 5, 18. Univ. Sci. Tech. Languedoc, Montpellier, 1988.

[21] C. Castaing, Trúóng Xuân Dúc Hā, and M. Valadier. Evolution equations governed by the sweeping process. *Set-Valued Anal.*, 1(2):109–139, 1993.

[22] C. Castaing and M. Valadier. *Convex analysis and measurable multifunctions*. Lecture Notes in Mathematics, Vol. 580. Springer-Verlag, Berlin-New York, 1977.

[23] Y. Censor and S. A. Zenios. Proximal minimization algorithm with D-functions. *J. Optim. Theory Appl.*, 73(3):451–464, 1992.

[24] F. Clarke. Nonsmooth analysis in systems and control theory. In *Mathematics of complexity and dynamical systems. Vols. 1–3*, pages 1137–1151. Springer, New York, 2012.

[25] F. Clarke. *Functional analysis, calculus of variations and optimal control*, volume 264 of *Graduate Texts in Mathematics*. Springer, London, 2013.

[26] G. Colombo and L. Thibault. Prox-regular sets and applications. In *Handbook of nonconvex analysis and applications*, pages 99–182. Int. Press, Somerville, MA, 2010.

[27] F. Deutsch. *Best approximation in inner product spaces*, volume 7 of *CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC*. Springer-Verlag, New York, 2001.

[28] I. Ekeland and R. Temam. *Convex analysis and variational problems*. Studies in Mathematics and its Applications, Vol. 1. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1976. Translated from the French.

[29] D. Gao and D. Motreanu, editors. *Handbook of nonconvex analysis and applications*. International Press, Somerville, MA, 2010.

[30] S. Hu and N. Papageorgiou. *Handbook of multivalued analysis. Vol. I*, volume 419 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 1997. Theory.

[31] A. Jourani and E. Vilches. Positively α-far sets and existence results for generalized perturbed sweeping processes. *J. Convex Anal.*, 23(3):775–821, 2016.

[32] A. Jourani and E. Vilches. Moreau-Yosida regularization of state-dependent sweeping processes with nonregular sets. *J. Optim. Theory Appl.*, 173(1):91–116, 2017.

[33] M. Kecies, T. Haddad, and M. Sene. Degenerate sweeping process with a Lipschitz perturbation. *Appl. Anal.*, 100(14):2927–2949, 2021.

[34] K. Kiwiel. Proximal minimization methods with generalized Bregman functions. *SIAM J. Control Optim.*, 35(4):1142–1168, 1997.

[35] M. Kunze and M Monteiro-Marques. An introduction to Moreau's sweeping process. In *Impacts in mechanical systems (Grenoble, 1999)*, volume 551 of *Lecture Notes in Phys.*, pages 1–60. Springer, Berlin, 2000.

[36] M. Kunze and M. D. P. Monteiro Marques. BV solutions to evolution problems with time-dependent domains. *Set-Valued Anal.*, 5(1):57–72, 1997.

[37] M. Kunze and M. D. P.. Monteiro Marques. Existence of solutions for degenerate sweeping processes. *J. Convex Anal.*, 4(1):165–176, 1997.

[38] M. Kunze and M. D. P. Monteiro Marques. On parabolic quasi-variational inequalities and state-dependent sweeping processes. *Topol. Methods Nonlinear Anal.*, 12(1):179–191, 1998.

[39] M. Kunze and M. D. P. Monteiro Marques. On the discretization of degenerate sweeping processes. *Portugal. Math.*, 55(2):219–232, 1998.

[40] M. Kunze and M. D. P. Monteiro Marques. Degenerate sweeping processes. In *Variations of domain and free-boundary problems in solid mechanics (Paris, 1997)*, volume 66 of *Solid Mech. Appl.*, pages 301–307. Kluwer Acad. Publ., Dordrecht, 1999.

[41] M. D. P. Monteiro Marques. *Differential inclusions in nonsmooth mechanical problems*, volume 9 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Verlag, Basel, 1993. Shocks and dry friction.

[42] B.S. Mordukhovich. *Variational analysis and generalized differentiation. I*, volume 330 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006. Basic theory.

[43] B.S. Mordukhovich and N. Nam. *Convex analysis and beyond. Vol. 1. Basic theory*. Springer Series in Operations Research and Financial Engineering. Springer, Cham, 2022.

[44] J. J. Moreau. Rafle par un convexe variable. II. In *Travaux du Séminaire d'Analyse Convexe, Vol. II*, Secrétariat des Math., Publ. No. 122, pages Exp. No. 3, 36. U.E.R. de Math., Univ. Sci. Tech. Languedoc, Montpellier, 1972.

[45] J. J. Moreau. On unilateral constraints, friction and plasticity. In *New variational techniques in mathematical physics (Centro Internaz. Mat. Estivo (C.I.M.E.), II Ciclo, Bressanone, 1973)*, Centro Internazionale Matematico Estivo (C.I.M.E.), pages 171–322. Ed. Cremonese, Rome, 1974.

[46] J. J. Moreau. Evolution problem associated with a moving convex set in a Hilbert space. *J. Differential Equations*, 26(3):347–374, 1977.

[47] J.J. Moreau. Rafle par un convexe variable. I. In *Travaux du Séminaire d'Analyse Convexe, Vol. I, Exp. No. 15*, pages 1–43. U.É.R. de Math., Univ. Sci. Tech. Languedoc, Montpellier, 1971.

[48] J.J. Moreau. Numerical aspects of the sweeping process. *Comput. Methods Appl. Mech. Engrg.*, 177(3-4):329–349, 1999.

[49] F. Nacry. Truncated nonconvex state-dependent sweeping process: implicit and semi-implicit adapted Moreau's catching-up algorithms. *J. Fixed Point Theory Appl.*, 20(3):Paper No. 121, 32, 2018.

[50] F. Nacry and L. Thibault. Regularization of sweeping process: old and new. *Pure Appl. Funct. Anal.*, 4(1):59–117, 2019.

[51] D. Narváez and E. Vilches. Moreau-Yosida regularization of degenerate state-dependent sweeping processes. *J. Optim. Theory Appl.*, 193(1-3):910–930, 2022.

[52] D. Narváez and E. Vilches. Moreau-Yosida regularization of degenerate state-dependent sweeping processes. *(Accepted in) J. Convex Anal*, 2023.

[53] A. S. Nemirovsky and D. B. Yudin. *Problem complexity and method efficiency in optimization*. Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons, Inc., New York, 1983. Translated from the Russian and with a preface by E. R. Dawson, A Wiley-Interscience Publication.

[54] B. G. Pachpatte. *Inequalities for differential and integral equations*, volume 197 of *Mathematics in Science and Engineering*. Academic Press, Inc., San Diego, CA, 1998.

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

[56] P. Pérez-Aros and E. Vilches. An enhanced Baillon-Haddad theorem for convex functions defined on convex sets. *Appl. Math. Optim.*, 83(3):2241–2252, 2021.

[57] R. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, NJ, 1970.

[58] E. Schechter. *Handbook of analysis and its foundations*. Academic Press, Inc., San Diego, CA, 1997.

[59] M. Sene and L. Thibault. Regularization of dynamical systems associated with prox-regular moving sets. *J. Nonlinear Convex Anal.*, 15(4):647–663, 2014.

[60] S. Sorin. No-regret algorithms in on-line learning, games and convex optimization. *Mathematical Programming*, 03 2023.

[61] C. Zălinescu. *Convex analysis in general vector spaces*. World Scientific Publishing Co., Inc., River Edge, NJ, 2002.