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

MIRROR SWEEPING PROCESSES DRIVEN BY BOUNDED VARIATION MOVING SETS

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

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL MATEMÁTICA

CAROLINA ISABEL CHIU LARA

PROFESOR GUÍA:
EMILIO VILCHES GUTIÉRREZ

PROFESOR CO-GUÍA:
PEDRO PÉREZ AROS

MIEMBROS DE LA COMISIÓN:
RAFAEL CORREA FONTECILLA
SEBASTIÁN DONOSO FUENTES

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
INGENIERA CIVIL MATEMÁTICA
POR: CAROLINA ISABEL CHIU LARA
FECHA: 2023
PROFESOR GUÍA: EMILIO VILCHES GUTIÉRREZ
PROFESOR CO-GUÍA: PEDRO PERÉZ AROS

Procesos de Arrastre de Espejo impulsados por conjuntos en movimiento de variación acotada

La presente tesis se ocupa del estudio del llamado *Mirror Sweeping Process*, un proceso de arrastre degenerado que utiliza un *mirror map* como su operador asociado. Primero se muestra que cuando el conjunto móvil es convexo y Lipschitz continuo, la inclusión diferencial puede ser regularizada mediante una familia de ecuaciones parciales ordinarias, las cuales poseen una única solución. Se prueba que esta familia de soluciones converge a una solución del proceso de arrastre. Luego, suponemos que el conjunto se mueve con retracción acotada con respecto al exceso, donde se utiliza una técnica de factorización para parametrizar el conjunto en función de la longitud de arco, y se rellenan los saltos del conjunto con una familia de geodésicas.

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
INGENIERA CIVIL MATEMÁTICA
POR: CAROLINA ISABEL CHIU LARA
FECHA: 2023
PROFESOR GUÍA: EMILIO VILCHES GUTIÉRREZ
PROFESOR CO-GUÍA: PEDRO PERÉZ AROS

Mirror Sweeping Processes Driven by Bounded Variation Moving Sets

This thesis is concerned with the study of the so called Mirror Sweeping Process, a degenerate sweeping process that uses a mirror map as its associated operator. First, it is shown that when the moving set is convex and Lipschitz continuous, the differential inclussion can be regularized by a family of well-posed ordinary differential equations. This family of solutions is proven to converge uniformly to a solution of the sweeping process. Then, we suppose the moving set is convex and of bounded retraction with respect to the excess, where we use a factorization technique to parametrize the moving set by means of the arc length, and filling in the jumps with a suitable family of geodesics.

"One can never be overdressed or overeducated."
—Oscar Wilde.

Acknowledgments

Primero y siempre quiero agradecir a mi familia, porque de ustedes aprendo las cosas realmente importantes. A mis padres, por apoyarme en todo lo que hago, por creer en mi y darme siempre todo lo que necesito. A mis hermanas, Andrea y Agus, porque con ustedes siempre puedo ser yo misma. Gracias por su amor incondicional.

Aldo, mi compañero de vida, mi mayor colaborador y mi mejor amigo. Sin ti esta tesis literalmente no estaría terminada. Gracias por hacer el catching-up en el caso degenerado. Y gracias por estar conmigo en todas, por motivarme a ser mejor y recordarme que debo poner los pies en la tierra.

A todas las personas que conocí en la facultad y que hicieron mi paso por pregrado un poco menos difícil y definitivamente más ameno. Mis primeros amigos de plan común que sin querer se convirtieron también en mis amigos de especialidad: Benja, porque siempre puedo contar contigo para todo. Javier, porque de a poquito me dejaste entrar en tu vida y por siempre compartirme tus pautas. Jota, por enseñarme perserverancia a través de tus acciones. Estoy muy agradecida de que todas las estrellas se alinearan para llegar a este momento juntos. A Bastián, porque nuestra amistad fue más fuerte que nuestro robot de Intro 2.

A mis amigos que conocí en el DIM: Mariano, porque aunque apenas nos conocíamos viviste conmigo unos de los momentos que más marcaron mi carrera y desde entonces no nos separamos. Pedro, porque un pequeño acto de bondad de tu parte unió nuestros caminos. Gonzalo, porque nada es más validante que cuando me celebras los chistes. Leo, por educarme en jueguitos de mesa y ser (casi) siempre la voz de la razón. Fabián, por todas las veces que quisiste cantar conmigo canciones de Disney en el karaoke. Manu, por las incontables horas de estudio y colaboración. A mis compañeros de oficina, por la música, las risas, los dramas y los chismes. A Víctor, gracias por aguantar tantas horas de Taylor Swift.

También me gustaría agradecer a mi profesor guía, Emilio Vilches, por explicarme mil veces las cosas y apoyarme en cada paso del camino. Por exponerme al mundo de la academia y a la hermosa comunidad que constituye nuestra área. No podría haber pedido un mejor guía en este proceso. A Pedro Pérez Aros, porque sus clases me convencieron de especializarme en optimización. A Sebastián Donoso, porque su curso de Teoría de la Medida me devolvió el amor por las matemáticas y me hizo creer en mi misma. A Rafael Correa, porque su educación trasciende la sala de clases. Gracias a los tres por acceder a ser parte de mi comisión. A Jocelyn Dunstan, por ser la primera en creer en mí.

Finalmente, a mi mascota, Sila, por ser mi compañera más fiel. A mis bandas favoritas: One Direction, por enseñarme que el cambio siempre es bueno; y 5 Seconds of Summer, por escribir el soundtrack de mi vida.

Table of Content

Introduction

This thesis is centered around the existence of solutions for Degenerate Sweeping Processes driven by convex sets with bounded variation. The study of sweeping processes dates back to the seventies, where J.J. Moreau introduced a series of papers [24, 25, 27, 28] modelling an elasto-plastic mechanical system by using a first-order differential inclusion of the form

$$\begin{cases} \dot{x}(t) \in -N_{C(t)}(x(t)) \text{ a.e. } t \in [0,T], \\ x(0) = x_0 \in \mathcal{H}. \end{cases}$$

Here, $C\colon [0,T] \rightrightarrows \mathcal{H}$ is a time-dependent moving set with nonempty, closed and convex values on a Hilbert space $\mathcal{H}$, and $N_{C(t)}(\cdot)$ denotes the (outward) normal cone of $C(t)$, in the sense of convex analysis. The interpretation of this formulation is understood as a point $x(t)$ which is "swept" by a moving set $C(t)$. Whenever the point is on the interior of $C(t)$, it does not move, since the normal cone is reduced to zero, but once the points is "caught up" by the boundary of $C(t)$, the point moves inwards the set for almost every $t \in [0,T]$. The solution of the sweeping process maps the trajectory of the particle over the interval of time $[0,T]$. As mentioned before, this problem first attempted to model an elasto-plastic system, but since then sweeping processes have many other applications, including nonsmooth mechanics [2, 9], crowd motion [22], switched electrical circuits [1], among many more.

Given the development of different techniques to solve differential inclusions, we have a vastly studied multitude of variations to Moreau's sweeping process. Namely, some of the most relevant examples are the state-dependent sweeping process [29] and the second order sweeping process [7]. We will focus in the degenerate sweeping process, i.e., the differential inclusion of the form

$$\begin{cases} \dot{x}(t) \in -N_{C(t)}(Ax(t)), \\ x(0) = x_0, \end{cases} \tag{$\mathcal{DSP}$}$$

where once again C is closed and convex-valued over a Hilbert space, and it is a Lipschitz map with respect to the Hausdorff distance. Moreover, $A : \mathcal{H} \to \mathcal{H}$ is a strongly monotone (and possibly nonlinear) operator, i.e.,

$$\langle Ax - Ay, x - y \rangle \geq m\|x - y\|^2 \text{ for all } x, y \in \mathcal{H}, \tag{1}$$

where m is a positive constant. The system ($\mathcal{DSP}$) was first introduced by Monteiro-Marques in [19, 21] as a model of contact mechanics, with A a multi-valued map. A notion of a solution for this case is the pair (u,v) such that $u(t) \in \operatorname{dom}(A)$, $v(t) \in A(u(t)) \cap C(t)$, and $\dot{u}(t) \in -N_{C(t)}(v(t))$ a.e. for $t \in [0,T]$.

This sweeping process is called degenerate as it, depending on the operator, may not posses a solution or it may not be unique. For example, let us consider $\mathcal{H} = \mathbb{R}$ and a monotone operator A with $\mathrm{dom}(A) = \{0\}$, such that $A(0) = \mathbb{R}$. Let us also consider the time interval $[0, 1]$ and the initial data $x_0 = 0$. Hence, the solution must be $x(t) \equiv 0$, but $A(x(t)) = \mathbb{R}$, which means $v(t)$ could be any selection of $C(\cdot)$. For more examples on degenerate sweeping processes with no solution, refer to [21].

The present work addresses the so called Mirror Sweeping Process:

$$\begin{cases} \dot{x}(t) \in -N_{C(t)}\left(\nabla\varphi(x(t))\right) + f(t, x(t)) \quad \text{a.e. } t \in I := [T_0, T], \\ x(T_0) = x_0, \end{cases} \tag{$\mathcal{MSP}$}$$

where $\nabla\varphi(x_0) \in C(T_0)$, and f is a continuous function. In this case A is the gradient of a *mirror map* φ, which is an operator mostly used to perform the Mirror Gradient Descent in the context of online optimization [11]. We prove the well-posedness of solutions of ($\mathcal{MSP}$) in two settings: first, we assume that C moves in a Lipschitzian way with respect to the Hausdorff distance, whereas in the second case we suppose it moves with bounded variation. To deal with the Lipschitz setting we will use the Moreau-Yosida regularizaton technique seen in [18, 29, 40]. We extend the aforementioned work by proving the existence and uniqueness of solutions without the necessity of the separability of $\mathcal{H}$, as seen in [39]. We also drop the compactness of C, the latter being a strongly used hypotheses in other works that also deal with the degenerate sweeping process in an absolutely continuous setting (see for instance [19–21]).

As for the bounded variation setting, we will consider the retraction of C with respect to the excess, a quantity that measures the separation of two closed sets. The technique to find a solution for this case relies on reducing this problem to the case where the moving set is Lipschitz with respect to the excess, using a suitable parametrization of C by means of its arc length ℓ_C of the form $C = \widetilde{C} \circ \ell_C$, where $\widetilde{C}$ is the unique Lipschitz that holds the equality. Seeing that the domain of $\widetilde{C}$ is the image of the arc length, in order to extend the domain to the whole interval where C is defined, we use a family of geodesics to appropriately fill in the jumps. Once again, this result is an extension of the work seen in [36], where the author presents a bounded retraction sweeping process with $A \equiv Id_{\mathcal{H}}$. It was proven in [16] that the Mirror Sweeping Process has a unique solution when it is driven by a set that is Lispchitz with respect to the excess, by means of a catching-up-like algorithm.

This thesis is organized as follows: Chapter 1 contains all the preliminary notions and results that will be used throughout this work. Chapter 2 deals with the existence and uniqueness of solutions for the Mirror Sweeping Process in a Lipschitz setting using the Moreau-Yosida regularization. Finally, in Chapter 3 we prove the existence of solutions for the bounded retraction case using the factorization technique previously explained. We conclude this document by pointing out the main conclusions and future work regarding this particular type of sweeping process.

Chapter 1

Preliminaries

In this chapter we will give preliminary background about definitions, relevant known results, and the notation that will be used in the following chapters.

1.1 Functions of bounded variation

In this subsection we will mainly draw on the work of [27]. In what follows, (X, d) denotes an extended complete metric space, i.e. X is a set and $d\colon X \times X \to [0, \infty]$ satisfies the axioms of a distance, but it can also take on the value of ∞. For a subset $Y \subset X$, the topological notions of interior, closure and boundary will be denoted by $\mathrm{int}(Y), \mathrm{cl}(Y)$ and $\mathrm{bdry}(Y)$, respectively. The closed ball centered at x with radius r is denoted by $\mathbb{B}(x, r)$ and we refer to the closed unit ball as $\mathbb{B}$.

As usual, we define the distance between a subset $Y \subset X$ and an element $x \in X$, as $d(x, Y) := \inf_{y \in Y} d(x, y)$. Moreover, if (Y, d_Y) is a metric space, the continuity set of a function $f\colon Y \to X$ will be denoted as $\mathrm{Cont}(f)$ and naturally, $\mathrm{Discont}(f) := Y \setminus \mathrm{Cont}(f)$.

For a subset $S \subset Y$, $f\colon Y \to X$ is said to be Lipschitz continuous if

$$\mathrm{Lip}(f) := \sup_{t \neq s} \frac{d(f(s), f(t))}{d_Y(t, s)} < +\infty.$$

It is clear that if f is Lipschitz, then $d(f(s), f(t)) \leq \mathrm{Lip}(f) d_Y(t, s)$. The set of Lipschitz functions is denoted by $\mathrm{Lip}(Y, X) := \{f\colon Y \to X\colon\ \mathrm{Lip}(f) < +\infty\}$.

Let $I \subset \mathbb{R}$ be an interval, and the function $f\colon I \to X$. Given $J = [s, t] \subset I$, and considering all finite sequences of the form $s = \tau_0 \leq \cdots \leq \tau_n = t$, one defines the variation of f over J, denoted $\mathrm{V}(f, J)$, as the supremum of $\sum_{i=1}^{n} d(f(\tau_{i-1}), f(\tau_i))$.

If $\mathrm{V}(f, J) < \infty$ then we say that f is of *bounded variation*. As usual, we denote

$$\mathrm{BV}(I; X) := \{f\colon I \to X\colon\ \mathrm{V}(f, I) < \infty\}.$$

Let us notice that the completeness of X allows for $f \in \mathrm{BV}(I;X)$ to have one-sided limits $f(t^-)$ and $f(t^+)$ at every point t, with $f(\inf I^-) := f(\inf I)$ if $\inf I \in I$ and similarly, $f(\sup I^+) := f(\sup I)$ if $\sup I \in I$. We also deduce by the definition of the functions of bounded variation that $\mathrm{Discont}(f)$ is at most countable.

1.2 Convex sets in Hilbert spaces

This section covers the basics of convex analysis that will be used in the following chapters. For more details, we refer to [3, 5, 6, 37].

Unless otherwise stated, we assume that $\mathcal{H}$ is a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $||x|| := \langle x, x\rangle^{1/2}$. The metric is naturally defined as $d(x,y) := ||x-y||$, for all $x, y \in \mathcal{H}$.

We recall a subset $C \subset \mathcal{H}$ is convex if $\lambda x + (1-\lambda)y \in C$ whenever $x, y \in C$ and $\lambda \in [0,1]$. In particular, $\mathcal{H}$ and $\emptyset$ are convex.

Let us set $D_\rho := \{x \in \mathcal{H} \colon ||x|| \leq \rho\}$, $\rho > 0$, and

$$\mathcal{C}_{\mathcal{H}} := \{K \subset \mathcal{H} \colon K \text{ is nonempty, closed and convex}\}.$$

The *lower level sets* of a function $f \colon \mathcal{H} \to [-\infty, +\infty]$ are the sets $\Gamma_\alpha(f) := \{x \in \mathcal{H} \colon f(x) \leq \alpha\}$, where $\alpha \in \mathbb{R}$. The *epigraph* of f is defined by

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

Furthermore, a function $f : \mathcal{H} \to [-\infty, +\infty]$ is convex if its epigraph is a convex subset of $\mathcal{H} \times \mathbb{R}$. Moreover, if f is convex, then its lower level sets are convex.

Lastly, we say f is *lower-semicontinuous* (lsc) if for any $x_0 \in \mathcal{H}$ one has

$$\liminf_{x \to x_0} f(x) \geq f(x_0).$$

As usual, we will denote $\Gamma_0(\mathcal{H})$ the set of functions that are proper, convex and lower-semicontinuous. We say a function is proper if $\operatorname{dom} f \neq \emptyset$ and $\inf f(x) > -\infty$.

In the remaining part of this section, we will introduce the concept of the metric projection, which is fundamental for the understanding of this work.

Definition 1.2.1 *Let C be a nonempty subset of $\mathcal{H}$ and $x \in \mathcal{H}$. Then, $p \in C$ is the projection on C if $d(x, C) = ||x - p||$. We denote $\mathrm{Proj}_C(x)$ as the (possibly empty) set of points which attain this infimum. If $\mathrm{Proj}_C(x)$ is a singleton, we write $\mathrm{proj}_C(x)$.*

If every point in $\mathcal{H}$ has at least one projection onto C, then C is *proximinal*. If every point in $\mathcal{H}$ has exactly one projection onto C, then C is a *Chebyshev set*. The next theorem is known as the *Projection theorem* and its proof can be found in [3].

Theorem 1.2.2 (Projection Theorem) *Let C be a nonempty closed convex subset of $\mathcal{H}$. Then C is a Chebyshev set and, for every $x \in \mathcal{H}$,*

$$p = \mathrm{proj}_C(x) \quad \textit{and} \quad (\forall y \in C)\ \langle y - p, x - p\rangle \leq 0.$$

A direct consequence of this theorem is the fact that the projection is $1-$Lipschitz continuous, i.e.,

$$\|\operatorname{proj}_C(x) - \operatorname{proj}_C(y)\| \leq \|x - y\|. \tag{1.1}$$

Next, we define the normal cone in the sense of convex analysis.

Definition 1.2.3 *The normal cone to a convex set $C \subset \mathcal{H}$ at a point $\bar{x} \in C$ is defined by*

$$N_C(\bar{x}) := \{\xi \in \mathcal{H} \colon \langle \xi, y - \bar{x} \rangle \leq 0 \,\forall y \in \mathcal{H}\}.$$

Remark 1.2.4 By means of Theorem 1.2.2 and Definition 1.2.3, one has:

$$p = \operatorname{proj}_C(x) \Longleftrightarrow x - p \in N_C(p).$$

Finally, we introduce the Fenchel conjugate of a function $f \colon \mathcal{H} \to [-\infty, +\infty]$ as the proper, convex function $f^* \colon \mathcal{H} \to [-\infty, +\infty]$ defined by

$$f^*(\phi) = \sup_{x \in \mathcal{H}} \{\langle \phi, x \rangle - f(x)\}. \tag{1.2}$$

1.3 Retraction with respect to the excess

In this section we introduce some notions of rating for the displacements of sets, as well as the definition of retraction of sets, which will be used mostly in Chapter 3.

Definition 1.3.1 (Hausdorff distance) *In a metric space (X, d), the Hausdorff distance between two non-empty subsets A and B is defined by*

$$d_H(A, B) := \max\{e(A, B), e(B, A)\},$$

where e denotes the excess or separation of A from B given by

$$e(A, B) := \sup_{a \in A} d(a, B) = \sup_{a \in A} \inf_{b \in B} d(a, b)$$

Clearly, $e(A, B) = 0$ if and only if A is contained in $\operatorname{cl} B$. Moreover,

$$e(A, B) = \inf\{\rho > 0 : A \subset B + D_\rho\}, \quad \forall A, B \subset X,$$

with $D_\rho = \{x \in X \ : \ ||x|| \leq \rho\}$, $\rho > 0$.

Let us remark that the Hausdorff distance, while satisfying the axioms of a metric, can take the value $+\infty$ in the case where the sets are unbounded.

The retraction $\mathrm{R}(f; J)$ of f over the interval $J = [s, t]$ is defined by means of the excess as the supremum of $\sum_{i=1}^{n} e(f(\tau_{i-1}), f(\tau_i))$, where we consider all finite sequences $s = \tau_0 \leq \cdots \leq \tau_n = t$. If $\mathrm{R}(f; J)$ is finite, we call f of bounded retraction, and we denote the sets of functions with these properties as $\mathrm{BR}(I; X)$.

From now on, we will consider the bounded retraction of sets on $\mathcal{C}_\mathcal{H}$ over an interval $I = [0, T]$, and we will use the following notation:

(i) $\mathrm{BR}^r(I;\mathcal{C}_\mathcal{H}) := \{C \in \mathrm{BR}(I;\mathcal{C}_\mathcal{H}) : e(C(t), C(t+)) = 0, \forall t \in I\}$.

(ii) $\mathrm{BR}^l(I;\mathcal{C}_\mathcal{H}) := \{C \in \mathrm{BR}(I;\mathcal{C}_\mathcal{H}) : e(C(t-), C(t)) = 0, \forall t \in I\}$.

For an interval $I = [0,T]$ and $C \in \mathrm{BR}(I,\mathcal{C}_\mathcal{H})$, we define the arc length ℓ_C with respect to the excess e, as

$$\ell_C(t) := \mathrm{R}(C;[0,t]),\ 0 \leq t \leq T. \tag{1.3}$$

The arc length is an increasing function such that $\ell_C(0) = 0$ and $\ell_C(T) = \mathrm{R}(C;[0,T])$. If $C \in \mathrm{BR}(I,\mathcal{C}_\mathcal{H})$, then for each $t \in I$ we have the lateral limits:

$$\begin{aligned} C(t+) &:= \liminf_{s\to t+} C(s) = \{x \in \mathcal{H} : \lim_{s\to t+} d(x, C(s)) = 0\}; \\ C(t-) &:= \liminf_{s\to t-} C(s) = \{x \in \mathcal{H} : \lim_{s\to t-} d(x, C(s)) = 0\}. \end{aligned}$$

Which implies that $C(t+)$ and $C(t-)$ are nonempty, closed and convex sets. This gives us the following results:

$$\begin{aligned} e(C(t), C(t+)) &= \lim_{s\to t+} e(C(t), C(s)) = \ell_C(t+) - \ell_C(t); \\ e(C(t-), C(t)) &= \lim_{s\to t-} e(C(s), C(t)) = \ell_C(t) - \ell_C(t-). \end{aligned}$$

Finally, for $C \in \mathrm{BR}(I;\mathcal{C}_\mathcal{H})$, we set

$$\mathrm{Cont}(C) := \{t \in I : e(C(t), C(t+)) = e(C(t-), C(t)) = 0\}$$

and $\mathrm{Discont}(C) := I \setminus \mathrm{Cont}(C)$.

1.3.1 Bregman Distance

In this section we will introduce the main properties of the Bregman distance, which is an asymmetric proximity measure between a point and a reference point. This object is mainly used in optimization problems as a mean to exploit the nonlinear geometry of constraints. For more details see [6, 8] and the references therein.

1.3.2 Legendre functions

Firstly, let us recall the concept of Legendre functions (for more details see [6, Chapter 7]).

Definition 1.3.2 *We will say $f \in \Gamma_0(\mathcal{H})$ is:*

(i) essentially smooth if ∂f is both locally bounded and single-valued on its domain;

(ii) essentially strictly convex, if $(\partial f)^{-1}$ is locally bounded on its domain and f is strictly convex on every convex subset of $\mathrm{dom}\,\partial f$*;*

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

Moreover, f is Legendre if and only if f^* is. This follows from the fact that f is essentially smooth if and only if f^* is essentially strictly convex (see [6, Theorem 7.3.2.]).

Theorem 1.3.3 *Let $f\colon \mathcal{H} \to [-\infty, \infty]$ be function in $\Gamma_0(\mathcal{H})$. If f is of Legendre type, then the map*

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

is bijective. Moreover, the following equality holds:

$$(\nabla f)^{-1} = \nabla f^*. \tag{1.4}$$

1.3.3 Bregman distance

Definition 1.3.4 (Bregman distance) *Let $f \in \Gamma_0(\mathcal{H})$ be a differentiable function of Legendre type. The Bregman distance corresponding to f is defined by*

$$D_f\colon X \times \operatorname{int}\operatorname{dom} f \to [0, \infty]\colon (x, y) \mapsto f(x) - f(y) + \nabla f(y) \tag{1.5}$$

The Bregman distance is really a *divergence* measuring how far away a second point is from a reference point, and it is asymmetrical most of the time. It also does not satisfy the triangular inequality, and therefore it can not be considered a distance, even though we will call it so.

We now present some properties of the Bregman distance.

Lemma 1.3.5 *Let $f \in \Gamma_0(\mathcal{H})$ with a nonempty domain. Suppose $x \in \mathcal{H}$ and $y \in \operatorname{int}\operatorname{dom} f$. Then:*

(i) $D_f(x, y) = f(x) - f(y) + \max\langle \partial f(y), y - x\rangle$.

(ii) $D_f(\cdot, y)$ *is convex, lsc, proper with* $\operatorname{dom} D(\cdot, y) = \operatorname{dom} f$.

(iii) If f is differentiable on $\operatorname{int}\operatorname{dom} f$ *and essentially strictly convex, and* $x \in \operatorname{int}\operatorname{dom} f$, *then* $D_f(x, y) = D_{f^*}(\nabla f(y), \nabla f(x))$.

(iv) If (y_n) is sequence in $\operatorname{int}\operatorname{dom} f$ *converging to y, then* $D(y, y_n) \to 0$.

1.4 Elements of nonsmooth analysis

The main purpose of this section is to review relevant notions of nonsmooth analysis that will be used throughtout this work. For more details on this subject we refer to [12, 13].

1.4.1 Cones and subdifferentials

We start by defining the *subdifferential* of a proper function $f\colon \mathcal{H} \to]-\infty,+\infty]$ as the set-valued operator

$$\partial f\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto \{u \in \mathcal{H}\colon (\forall y \in \mathcal{H})\ \langle y-x, u\rangle + f(x) \leq f(y)\}.$$

Let $x \in \mathcal{H}$, then f is subdifferentiable at x if $\partial f(x) \neq \emptyset$. The elements of $\partial f(x)$ are the subgradients of f at x. With this concept in mind, we can redefine the normal cone from Definition 1.2.3 as $N_C(\bar{x}) := \partial \delta_C(\bar{x})$, where $\delta_C(x)$ is the indicator function of C defined as 0 if $x \in C$ and $+\infty$ otherwise.

A vector $v \in \mathcal{H}$ is called a *Fréchet subgradient* of f at x if for each $\varepsilon > 0$ there exists some neighborhood U of x such that

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

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

Let us also define the *limiting subgradient* of f at x. If there exists a sequence $((x_n, f(x_n)))_n$ converging to $(x, f(x))$ and a sequence $(v_n)_n$ converging weakly to v such that $v_n \in \partial^F f(x_n)$, then x is a limiting subgradient. We denote $\partial^L f(x)$ the set of all limiting subgradients of f at x.

As usual, $\partial^F f(x) = \partial^L f(x) = \emptyset$ whenever $x \notin \operatorname{dom} f$, and for $C \in \mathcal{H}$ we define the Fréchet normal cone and limiting normal cone, respectively, as $N_C^F(x) := \partial^F \delta_C(x)$ and $N_C^L(x) := \partial^L \delta_C(x)$.

Similarly, a *proximal subgradient* of f at x is a vector $v \in \mathcal{H}$ such that there exists some real number $\sigma > 0$ and some neighborhood U of x that satifies

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

The set $\partial^P f(x)$ of all proximal subgradients of f at x is the proximal subdifferential and once again we set $\partial^P f(x) = \emptyset$ whenever $f(x)$ is not finite. Just as before, the *normal proximal cone* $N_C^P(x)$ is the set of all proximal normal vectors $v \in \mathcal{H}$, that is, the set of all vectors $v \in \mathcal{H}$ that satisfy that there exists some $\sigma > 0$ and a neighborhood U of x such that

$$\langle v, y-x\rangle \leq \sigma\|y-x\|^2 \text{ for all } y \in U \cap C.$$

It is clear by definition of each cone, that

$$N_C^P(x) \subset N_C^F(x) \subset N_C^L(x).$$

Finally, let us define the Clarke normal cone, which is used to describe properties of prox-regular sets, as seen in the next subsection.

Definition 1.4.1 *Let $\mathcal{H}$ be a Hilbert space, then the Clarke normal cone can be defined as*

$$N_C^{Cl}(x) = \operatorname{cl}\left(\operatorname{co}(N_C^L(x))\right),$$

where $\overline{\operatorname{co}}$ denotes the closed convex hull and $N_C^L(x)$ is the limiting normal cone. Moreover, the Clarke subdifferential of f at x is defined by

$$\partial^{Cl} f(x) := \{v \in \mathcal{H}\colon (v, -1) \in N^{Cl}(\operatorname{epi} f, (x, f(x)))\}.$$

1.4.2 Prox-regularity

We will now introduce the notion of prox-regularity, a relevant concept used, among other areas, in the theory of differential inclusions. For more details on prox-regularity, see for instance [13].

Definition 1.4.2 *A closed set $C \subset \mathcal{H}$ and a continuous function $\rho\colon C \to]0,\infty]$ is said to be $\rho(\cdot)-$prox-regular if for every $\zeta \in N_C^P(x) \cap \mathbb{B}$ and any $0 < t < \rho(x)$, one has*

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

Moreover, we say C is uniformly-prox regular whenever $\rho(x) = r$ for all $x \in C$, where $r \in]0, +\infty]$.

Theorem 1.4.3 *Let C be a closed subset of $\mathcal{H}$ and $r \in]0, +\infty]$. Then the following assertions are equivalent.*

(i) The set C is $r-$prox-regular.

(ii) For any $x, y \in C$ and $v \in N_C^P(x)$ one has

$$\langle v, y - x \rangle \leq \frac{1}{2r}\|v\| \cdot \|y - x\|^2.$$

(iii) For any $x_i \in C$, $v_i \in N_C^P(x_i) \cap \mathbb{B}$ with $i = 1, 2$ one has

$$\langle v_1, v_2, x_1 - x_2 \rangle \geq -\frac{1}{r}\|x_1 - x_2\|^2.$$

(iv) For any positive $\gamma < 1$ the mapping $\operatorname{proj}_C(\cdot)$ is well-defined on $U_r^\gamma(C)$ with $(1-\gamma)^{-1}$ as a Lipschitz constant and $U_r^\gamma(C) := \{y \in \mathcal{H}\colon d_C(y) < \gamma\rho\}$. Moreover, the projection is unique for every $x \in U_r^\gamma(C)$.

Finally, for both convex and prox-regular sets we have the following property.

Proposition 1.4.4 *Let $C \subset \mathcal{H}$ be a closed and $\rho-$uniformly prox-regular set, with $\rho > 0$. Then,*

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

1.5 Smoothness of the inverse image

We will now present a property known as the *uniform normal cone inverse image property* (UNCIIP) [17], which will be useful to prove that the inverse image of a convex set through a differentiable mapping is prox-regular.

Definition 1.5.1 *Consider two Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$ and a differentiable mapping $h : \mathcal{H}_1 \to \mathcal{H}_2$. Let $C \subset \mathcal{H}_1, D \subset \mathcal{H}_2$ be closed sets, we say that the inverse image set $h^{-1}(D) \cap C$ satisfies the normal cone inverse image property at $\overline{x} \in h^{-1}(D) \cap C$ with respect to the Clarke normal cone if there exists a constant $k > 0$ and a neighbourhood U of $\overline{x}$ such that for every $x \in U \cap (h^{-1}(D) \cap C)$ the following inclusion holds*

$$N^{\mathrm{Cl}}(h^{-1}(D) \cap C; x) \cap \mathbb{B}_{\mathcal{H}_1} \subset Dh(x)^*(N^{\mathrm{Cl}}(D; h(x)) \cap k\mathbb{B}_{\mathcal{H}_2}) + N(C; x). \tag{1.6}$$

Therefore, we say the set $h^{-1}(D) \cap C$ has the UNCIIP if there exists some $k > 0$ such that (1.6) holds for all $x \in h^{-1}(D) \cap C$. Finally, in order to make this property useful to our ends, we will present a proposition whose proof can be found in [17].

Proposition 1.5.2 *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces. Assume $C \subset \mathcal{H}_1$ and $D \subset \mathcal{H}_2$ are two closed convex sets and consider $h\colon \mathcal{H}_1 \to \mathcal{H}_2$ a differentiable function. If there exists $k > 0$ such that for all $w \in \mathcal{H}_2$, and all $x \in C$:*

$$d(x; h^{-1}(D - w) \cap C) \le kd(h(x) + w; D), \tag{1.7}$$

Then the inverse image $h^{-1}(D - w) \cap C$ satisfies the UNCIIP with constant $k > 0$.

Definition 1.5.3 (Metrically calm) *Let $F\colon \mathcal{H}_1 \to \mathcal{H}_2$ be a mapping between two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. Consider $D \subset \mathcal{H}_2$ and $\bar{x} \in F^{-1}(D)$. We say F is metrically calm at $\bar{x}$ relatively to the set D if there exist a constant $\gamma > 0$ and a neighborhood U of $\bar{x}$ such that*

$$d(x; F^{-1}(D)) \le \gamma d(F(x); D).$$

It is easy to see that F is metrically calm at $\bar{x}$ relatively to the set D if condition (1.7) is fulfilled with $h \equiv F$, $C = \mathcal{H}$ and $w = 0$.

1.6 Measure Theory

In this section, we review some tools from measure theory and bounded variation functions. We refer to [32] for more details.

Definition 1.6.1 *The Borel $\sigma-$algebra of a topological space (X, τ) is the $\sigma-$algebra generated by the family τ of open sets. Members of the Borel $\sigma-$algebra are Borel sets. The Borel $\sigma-$algebra is denoted $\mathcal{B}$.*

We recall that a $\mathcal{H}-$valued measure on I is a map $\mu\colon \mathcal{B}(I) \to \mathcal{H}$ such that for a family of disjoint sets $(A_n)_{n\in\mathbb{N}} \subset \mathcal{H}$, it holds that

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

We will now present the connection between functions of bounded variation and Borel vector measures on the real line. In what follows, we assume I is a non-degenerate interval.

First, let us recall the notion of *total variation* of a measure. Let $\mu\colon \mathcal{B}(I) \to \mathcal{H}$ be a vector measure, then the total variation of μ denoted by $|\mu|\colon \mathcal{B}(I) \to [0, \infty]$ is defined by

$$|\mu|(B) := \sup \left\{ \sum_{n=1}^{\infty} \|\mu(B_n)\|_{\mathcal{H}} \colon B = \bigcup_{n=1}^{\infty} B_n, B_n \in \mathcal{B}(I), B_h \cap B_k = \emptyset \text{ if } h \neq k \right\}.$$

A vector measure μ is called with bounded variation whenever $|\mu|(I) < \infty$. In this case, $\|\mu\| := |\mu|(I)$ defines a norm on the space of measures with bounded variation.

We will now present a result whose proof can be found in [14].

Theorem 1.6.2 *If $f \in \mathcal{H}^I$ is with bounded variation, then there exists a unique vector measure $\mu_f\colon \mathcal{B}(I) \to \mathcal{H}$ such that for every $c, d \in I$ with $c < d$ we have*

$$\begin{aligned} \mu_f(]c,d[) &= f(d-) - f(c+), \qquad & \mu_f([c,d]) &= f(d+) - f(c-), \\ \mu_f([c,d[) &= f(d-) - f(c-), \qquad & \mu_f(]c,d]) &= f(d+) - f(c+). \end{aligned} \tag{1.8}$$

Moreover, if μ_f is with bounded variation and if $f^- := f(t^-)$, then $\mu_f = \mu_{f^-}$. On the other hand, if $\mu : \mathcal{B}(I) \to \mathcal{H}$ is a vector measure with bounded variation, which implies that if the map $f_\mu\colon I \to \mathcal{H}$ is defined by $f_\mu(t) := \mu([\inf I, t[\cap I)$, then $f_\mu \in \mathrm{BV}(I; \mathcal{H})$ and $\mu_{f_\mu} = \mu$.

Usually, μ_f is known as the *differential measure* or *Lebesgue-Stieltjes measure associated with* f. From this definition, it is possible to derive the following characterization of the total variation.

Proposition 1.6.3 *Let $f\colon I \to \mathcal{H}$ be a function with bounded variation and let $f^-\colon I \to \mathcal{H}$ be defined by $f^- := f(t^-)$. Define $\mathrm{V}_f\colon I \to]0, \infty]$ as*

$$V_f(t) := \mathrm{V}(f^-; [\inf I, t] \cap I).$$

Then, $|\mu| = \mu_{V_f} = \mathrm{V}(f^-, I)$.

Let us now show its connection to the distributional derivative. Let $f \in \mathrm{BV}(I; \mathcal{H})$ and consider $\bar{f}\colon \mathbb{R} \to \mathcal{H}$ defined by:

$$\bar{f}(t) := \begin{cases} f(t) & \text{if } t \in I, \\ f(\inf I) & \text{if } \inf I \in \mathbb{R}, t \notin I, t \leq \inf I, \\ f(\sup f) & \text{if } \sup I \in \mathbb{R}, t \notin I, t \geq \sup I. \end{cases}$$

Then, according to [32, Proposition 2.9], we have that $\mu_f(B) = D\bar{f}(B)$ for every $B \in \mathcal{B}(\mathbb{R})$, where $D\bar{f}$ is such that

$$-\int_{\mathbb{R}} h'(t)\bar{f}(t)dt = \int_{\mathbb{R}} h dD\bar{f} \quad \forall h \in C_c^1(\mathbb{R}; \mathbb{R}).$$

Here, $C_c^1(\mathbb{R}; \mathbb{R})$ stands for the space of real continuously differentiable functions with compact support. We call $D\bar{f}$ the distributional derivative of $\bar{f}$, and we notice that it is concentrated on I, in the sense that $D\bar{f}(B) = \mu_f(B \cap I)$ for each $B \in \mathcal{B}$. Hence, for $f \in \mathrm{BV}(I;\mathcal{H})$ we will write

$$Df := D\bar{f} = \mu_f.$$

Furthermore, by Proposition 1.6.3 follows that

$$\|Df\| = |Df|(I) = \|\mu_f\| = V(f, I) \quad \forall f \in \mathrm{BV}^r(I;\mathcal{H}).$$

The following result provides some calculus rules for bounded variation functions. See, e.g., Lemma A.5. and Theorem A.7 of [32].

Proposition 1.6.4 *Assume that $I, J \subset \mathbb{R}$ are intervals and that $h\colon I \to J$ is nondecreasing.*

(i) *$Dh\,(h^{-1}(B)) = \mathcal{L}^1(B)$ for every $B \in \mathcal{B}(h(\mathrm{Cont}(h)))$, where $\mathcal{L}^1$ is the Lebesgue measure.*

(ii) *If $f \in \mathrm{Lip}(J;\mathcal{H})$ and $g : I \to \mathcal{H}$ is defined by*

$$g(t) := \begin{cases} f'(h(t)) & \text{if } t \in \mathrm{Cont}(h), \\ \dfrac{f(h(t+)) - f(h(t-))}{h(t+) - h(t-)} & \text{if } t \in \mathrm{Discont}(h). \end{cases}$$

then $f \circ h \in BV(I;\mathcal{H})$ and $D(f \circ h) = gDh$. This result holds even if f' is replaced by any of its $\mathcal{L}^1$-representatives.

Chapter 2

Mirror Sweeping Process on a Lipschitz setting

This chapter is divided in three parts. First, we will introduce the motivation behind the formulation of the Mirror Sweeping Process. Then, we will present the main hypotheses on the mirror map and the moving set. Lastly, we will show the well-posedness of the Mirror Sweeping Process in the case where C moves in a Lipschitzian way.

2.1 Mirror Sweeping Process

As previously stated we are interested in a degenerate sweeping process where the associated operator is a mirror map, which is mainly used to perform the Mirror Gradient Descent method [30]. The principal significance of this method is that it allows us to solve optimization problems in spaces whose norm does not derive from an inner product, e.g., a Banach space B. In this case, it does not make sense to compute the term $x - \eta \nabla f(x)$, since the gradient is an element of the dual space, and differently from when we work in a Hilbert space, we do not count on Riesz representation theorem that makes the primal and dual space isometric. Therefore, the advantage of using the Mirror Gradient Descent lies in the fact that it first maps the point $x \in B$ into the dual space B^*, to then perform the usual gradient update in said space and project the result back to the primal space. In the case that this projection lies outside the constraint set $X \subset B$, we might have to project one more time.

In order to go from the primal to the dual space, we use the so called *mirror map* [10, Section 4.1], which we proceed to define as follows. Let $D \subset \mathbb{R}^n$ be a convex open set such that $X \subset \operatorname{cl} D$. We say that $\varphi \colon D \to \mathbb{R}$ is a mirror map if it satisfies the following properties:

(i) φ is strictly convex and differentiable.

(ii) The gradient of φ takes all possible values, that is $\nabla \varphi(D) = \mathbb{R}^n$.

(iii) The gradient of φ diverges on the boundary of D, that is

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

Let us remark that this definition can be extended to infinite dimension spaces.

This way, we map $x \in X \cap D$ through $\nabla\varphi$ to then update the gradient by computing $\nabla\varphi(x) - \eta\nabla f(x)$. And by (ii) from the definition of mirror map, there exists $y \in D$ such that $\nabla\varphi(y) = \nabla\varphi(x) - \eta\nabla f(x)$. As mentioned before, it is possible that y does not belong to the set of constraints and therefore it might need to be projected back onto X. To this end, one uses the Bregman distance associated to φ, which we recall from Definition 1.3.4.

In this context, we introduce the Mirror Sweeping Process as the following system:

$$\begin{cases} \dot{x}(t) \in -N_{C(t)}\left(\nabla\varphi(x(t))\right) + f(t, x(t)) & \text{a.e. } t \in I := [T_0, T], \\ x(T_0) = x_0, \end{cases}$$

An application of this dynamic is the continuous Mirror Gradient method, as seen in [4]. This dynamic is used in [11] to obtain a competitive randomized algorithm for the $k-$server problem on a tree structure.

We will show the well-posedness of this sweeping process in a convex setting using the Moreau-Yosida regularization method seen in [18, 29, 40]. Let $\Phi \in \Gamma_0(\mathcal{H})$, then the Moreau-Yosida regularization Φ_λ of Φ is a $C^{1,1}$ function defined by

$$\Phi_\lambda(x) := \inf_{y \in \mathcal{H}} \left\{ \Phi(y) + \frac{1}{2\lambda}\|y - x\|^2 \right\}.$$

It is clear that $\Phi_\lambda \to \Phi$ as $\lambda \to 0^+$.

We will use the Moreau-Yosida regularization as a way to approximate the solution of $(\mathcal{P}_L)$ over the interval $[T_0, T]$. This process provides a Lipschitzian solution over each interval of the partition $\{[T_0, T_0 + \sigma], [T_0 + \sigma, T_0 + 2\sigma], \ldots, [T_0 + n\sigma, T]\}$, where σ is a positive constant yet to be determined, and $n \in \mathbb{N}$ is such that the partition is well-defined. Next, we take this approximation and show a Cauchy criterion for the convergence to a Lipschitz solution for the Mirror Sweeping Process.

This regularization method has been used before for degenerate sweeping processes in [21] and for state-dependent sweeping processes in [18, 29, 40]. Differently from this works, we present a proof of well-posedness similar to what it is done in [39] for a nondegenerate sweeping process, where the authors do not use any compactness hypotheses over the moving set. Besides adding the mirror map, we also consider a perturbation function in the dynamic.

2.2 Main Hypotheses

In this section, we list the main hypotheses used throughout this chapter.

Hypotheses on $\varphi\colon \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$: The function φ belongs to $\Gamma_0(\mathcal{H})$ and is twice differentiable. Moreover, the following conditions will be assumed in the rest of the chapter.

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

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

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

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

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

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

Hypotheses on the moving sets: The set-valued map $C\colon I \rightrightarrows \mathcal{H}$ has closed, nonempty and convex values. Moreover, we will assume the following property is satisfied:

($\mathcal{H}_C$) There exists $\kappa \geq 0$ such that for $s, t \in I$

$$\sup_{z\in\mathcal{H}} |d(z, C(t)) - d(z, C(s))| \leq \kappa|t - s|.$$

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 on $r\mathbb{B}$;

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

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

2.3 Preparatory Lemmas

We will now give preliminary results that will be useful in the remainder of this chapter regarding properties of the distance function and some results on set-valued maps. We start with Gronwall's Lemma, and we refer to [31] for its proof.

Lemma 2.3.1 (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 $t \in [t_0, t_1]$

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

We now present Mazur's Lemma, whose proof can be found in [15, Chapter 1].

Lemma 2.3.2 (Mazur's Lemma) *Let $(u_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ be a sequence converging weakly to $\bar{u}$. Then there is a sequence of convex combinations $(v_n)_{n\in\mathbb{N}}$ such that*

$$v_n = \sum_{k=n}^{N} \lambda_k u_k \quad \text{where } \sum_{k=n}^{N} \lambda_k = 1 \text{ and } \lambda_k \geq 0, n \leq k \leq N,$$

which converges to $\bar{u}$ in norm.

The next lemma will allow us to find an upper bound for the distance between the moving set and the trajectories of an approximated problem of the Mirror Sweeping Process.

Lemma 2.3.3 *Assume that $(\mathcal{H}^2_\varphi)$ and $(\mathcal{H}_C)$ hold. Let $x\colon [t_0, t_1] \to \mathcal{H}$ be an absolutely continuous function on $[t_0, t_1]$. If $\theta(t) := d_{C(t)}(\nabla\varphi(x(t)))$ for all $t \in [t_0, t_1]$, then θ is absolutely continuous on $[t_0, t_1]$ and*

$$\theta(t)\dot{\theta}(t) \leq \kappa\theta(t) + \langle \nabla\varphi(x(t)) - \operatorname{proj}_{C(t)}(\nabla\varphi(x(t))), D^2\varphi(x(t))\dot{x}(t)\rangle \ \text{a.e. } t \in [t_0, t_1].$$

Proof. First, let us prove that θ is absolutely continuous. This comes from the fact that, thanks to $(\mathcal{H}^2_\varphi)$ and $(\mathcal{H}_C)$, for any $s, t \in [t_0, t_1]$, one has

$$\begin{aligned}
|\theta(t) - \theta(s)| &= |d_{C(t)}(\nabla\varphi(x(t))) - d_{C(s)}(\nabla\varphi(x(s)))| \\
&\leq |d(\nabla\varphi(x(t)), \nabla\varphi(x(s))) + d_{C(t)}(\nabla\varphi(x(s))) - d_{C(s)}(\nabla\varphi(x(s)))| \\
&\leq |d(\nabla\varphi(x(t)), \nabla\varphi(x(s)))| + |d_{C(t)}(\nabla\varphi(x(s))) - d_{C(s)}(\nabla\varphi(x(s)))| \\
&\leq \|\nabla\varphi(x(t)) - \nabla\varphi(x(s))\| + \kappa|t - s| \\
&\leq M\|x(t) - x(s)\| + \kappa|t - s|.
\end{aligned}$$

Now, to prove the inequality let us set

$$\psi(t, x) := \frac{1}{2} d^2_{C(t)}(\nabla\varphi(x)) \text{ for } t \in [t_0, t_1] \text{ and } x \in \mathcal{H}.$$

Then, $\psi(t, \cdot)$ is continuously differentiable around $x(t)$ (see for instance [12, Theorem 5.1]) with

$$\nabla_2\psi(t, x(t)) = x(t) - \operatorname{proj}_{C(t)}(x(t)), \tag{2.1}$$

where $\nabla_2\psi$ stands for the derivative of ψ with respect to the second variable.

Let $t \in]t_0, t_1[$ such that $\dot{\theta}(t)$ and $\dot{x}(t)$ exist. Then, for all $s > 0$ small enough,

$$\begin{aligned}
\frac{1}{2s}[\theta^2(t+s) - \theta^2(t)] &= \frac{1}{s}[\psi(t+s, x(t+s)) - \psi(t, x(t+s))] \\
&\quad + \frac{1}{s}[\psi(t, x(t+s)) - \psi(t, x(t))] \\
&\leq \frac{\kappa}{2}[d_{C(t+s)}(x(t+s)) + d_{C(t)}(x(t))] \\
&\quad + \frac{1}{s}[\psi(t, x(t+s)) - \psi(t, x(t))].
\end{aligned} \tag{2.2}$$

Here, the last inequality holds since, by definition of ψ, we have

$$\begin{aligned}
\frac{1}{s}[\psi(t+s,x(t+s))-\psi(t,x(t+s))] &= \frac{1}{s}\frac{1}{2}[d^2_{C(t+s)}(\nabla\varphi(x(t+s)))-d^2_{C(t)}(\nabla\varphi(x(t+s)))]\\
&= \frac{1}{2s}\Big[\big(d_{C(t+s)}(\nabla\varphi(x(t+s)))+d_{C(t)}(\nabla\varphi(x(t+s)))\big)\\
&\quad\cdot\big(d_{C(t+s)}(\nabla\varphi(x(t+s)))-d_{C(t)}(\nabla\varphi(x(t+s)))\big)\Big]\\
&\leq \frac{\kappa}{2}\Big[d_{C(t+s)}(\nabla\varphi(x(t+s)))+d_{C(t)}\big(\nabla\varphi(x(t+s))\big)\Big].
\end{aligned} \tag{2.3}$$

Moreover, since $\dot{x}$ is well-defined in t, there exists $\varepsilon>0$ such that $\varepsilon\to 0$ in $\mathcal{H}$ whenever $s\to 0$, and therefore we can write

$$x(t+s)=x(t)+s\dot{x}(t)+s\varepsilon.$$

Let us define $x_t=x(t)+s\dot{x}(t)+s\varepsilon, y_t=x(t)+s\dot{x}(t)$, then

$$\begin{aligned}
\frac{1}{s}\psi(t,x(t+s)) &= \frac{1}{2s}d^2_{C(t)}(\nabla\varphi(x_t))\\
&\leq \frac{1}{2s}\left(d^2(\nabla\varphi(x_t),\nabla\varphi(y_t))+d^2_{C(t)}(\nabla\varphi(y_t))\right)\\
&= \frac{1}{2s}\|\nabla\varphi(x_t)-\nabla\varphi(y_t)\|^2+\frac{1}{s}\psi(t,y_t)
\end{aligned} \tag{2.4}$$

Using $(\mathcal{H}^2_\varphi)$ and 2.4, it yields, for $w\in C(t)$,

$$\begin{aligned}
\frac{1}{s}\psi(t,x(t+s)) &\leq \frac{1}{2s}M^2\|x_t-y_t\|^2+\frac{1}{s}\psi(t,y_t)\\
&= \frac{M^2}{2s}\|s\varepsilon\|\cdot\|x_t-y_t\|+\frac{1}{s}\psi(t,y_t)\\
&\leq \frac{M^2}{2}\|\varepsilon\|\cdot\|(x_t-w)-(y_t-w)\|+\frac{1}{s}\psi(t,y_t)\\
&\leq \frac{M^2}{2}\|\varepsilon\|\cdot(\|x(t)+s\dot{x}(t)+s\varepsilon-w\|+\|x(t)+s\dot{x}(t)-w\|)+\frac{1}{s}\psi(t,y_t)
\end{aligned}$$

Since $w\in C(t)$ is arbitrary, we can take infimum over $C(t)$ to obtain

$$\frac{1}{s}\psi(t,x(t+s)) \leq \frac{M^2}{2}\|\varepsilon\|\cdot\big(d_{C(t)}(x_t)+d_{C(t)}(y_t)\big)+\frac{1}{s}\psi(t,y_t). \tag{2.5}$$

Finally, we notice that if we use inequalities (2.3) and (2.5) in (2.2), we obtain

$$\begin{aligned}
\frac{1}{2s}[\theta(t+s)^2-\theta(t)^2] &\leq \frac{\kappa}{2}[d_{C(t+s)}(\nabla\varphi(x(t+s)))+d_{C(t)}(\nabla\varphi(x(t+s)))]\\
&\quad+\frac{1}{s}[\varphi(t,\nabla\varphi(x(t+s))))-\psi(t,x(t))]\\
&\leq \frac{\kappa}{2}\left(d_{C(t+s)}(x(t+s))+d_{C(t)}(x(t+s))\right)\\
&\quad+\frac{M^2}{2}\|\varepsilon\|\cdot\big(d_{C(t)}(x_t)+d_{C(t)}(y_t)\big)+\frac{1}{s}\psi(t,y_t)-\frac{1}{s}\psi(t,x(t)).
\end{aligned}$$

Setting $\delta(s) := \frac{M^2}{2}\|\varepsilon\| \cdot \left(d_{C(t)}(x_t) + d_{C(t)}(y_t)\right)$, we have that

$$
\begin{aligned}
\frac{1}{2s}[\theta(t+s)^2 - \theta(t)^2] &\leq \frac{\kappa}{2}\left(d_{C(t+s)}(\nabla\varphi(x(t+s))) + d_{C(t)}(\nabla\varphi(x(t+s)))\right) \\
&\quad +\delta(s) + \frac{1}{s}\psi(t, y_t) - \frac{1}{s}\psi(t, x(t)) \\
&= \frac{\kappa}{2}[d_{C(t+s)}(\nabla\varphi(x(t+s))) + d_{C(t)}(\nabla\varphi(x(t+s)))] \\
&\quad +\frac{1}{s}\left(\psi(t, \nabla\varphi(x(t) + s\dot{x}(t))) - \psi(t, \nabla\varphi(x(t)))\right) + \delta(s),
\end{aligned}
$$

Hence, taking $s \searrow 0$, we conclude. □

2.4 Well-posedness of Mirror Sweeping Process

In this section we prove the existence and uniqueness of solutions for the perturbed Mirror Sweeping Process as discussed in the previous sections. Let us recall the Mirror Sweeping Process in a Lipschitz setting:

$$
\begin{cases}
\dot{x}(t) \in -N_{C(t)}\left(\nabla\varphi(x(t))\right) + f(t, x(t)) & \text{a.e. } t \in I := [T_0, T], \\
x(T_0) = x_0,
\end{cases}
\tag{$\mathcal{P}_L$}
$$

where $\nabla\varphi(x_0) \in C(T_0)$.

Let $\lambda > 0$. We will find $x_\lambda \colon I \to \mathcal{H}$ with $x_\lambda(0) = x_0$ such that for a.e. $t \in I$,

$$
\begin{cases}
\dot{x}_\lambda(t) = -\dfrac{1}{2\lambda}\nabla d^2_{C(t)}(\nabla\varphi(x_\lambda(t))) + f(t, x_\lambda(t)), \\
\nabla\varphi(x_0) \in C(T_0).
\end{cases}
\tag{$\mathcal{P}_\lambda$}
$$

Remark 2.4.1 Without loss of generality, we will assume $\eta > 0$ such that

$$
|T - T_0| = \sigma < \frac{m}{\kappa + (M+m)\beta}\eta,
$$

where, for all $x \in \mathbb{B}(x_0, \eta)$, we will have

$$
\|f(t,x)\| \leq \beta := c\eta + c\|x_0\| + d. \tag{2.6}
$$

Proposition 2.4.2 *Assume that $(\mathcal{H}^2_\varphi)$, $(\mathcal{H}_C)$, $(\mathcal{H}^f_1)$, $(\mathcal{H}^f_2)$ and $(\mathcal{H}^f_3)$. Then, for every $\lambda > 0$ there exists a unique solution $x_\lambda \in \mathbb{B}(x_0, \eta)$ of $(\mathcal{P}_L)$ defined on its maximal interval of existence $[T_0, T_\lambda[\subset [T_0, T]$, where η is given in Remark 2.4.1.*

Proof. The proof falls naturally into two claims.

Claim 1: For all $x \in B(x_0, \eta)$ we have that $d_{C(t)}(\nabla\varphi(x(t))) < +\infty$ for all $t \in [T_0, T]$.

Proof of Claim 1. Let $x \in B(x_0, \eta)$. Let us notice that, thanks to $(\mathcal{H}_C)$ and $(\mathcal{H}^2_\varphi)$, we have that

$$\begin{aligned} d_{C(t)}(\nabla\varphi(x)) &\le d(\nabla\varphi(x), \nabla\varphi(x_0)) + d_{C(t)}(\nabla\varphi(x_0)) \\ &= d(\nabla\varphi(x), \nabla\varphi(x_0)) + d_{C(t)}(\nabla\varphi(x_0)) - d_{C(T_0)}(\nabla\varphi(x_0)) \\ &\le M\|x(t) - x_0\| + \kappa|t - T_0| \\ &< M\eta + \kappa\sigma \\ &< +\infty, \end{aligned}$$

which completes the proof. □

Claim 2: The map

$$g(t, x) := -\frac{1}{\lambda}(\nabla\varphi(x) - \mathrm{proj}_{C(t)}(\nabla\varphi(x))) + f(t, x(t)) \quad \text{for } (t, x) \in [T_0, T] \times \mathbb{B}(x_0, \eta),$$

has the following properties:

(i) For all $x \in \mathbb{B}(x_0, \eta)$, the map $t \mapsto g(t, x)$ is measurable;

(ii) For all $t \in [T_0, T]$, the map $x \mapsto g(t, x)$ is ℓ^η_λ−Lipschitz on $\mathbb{B}(x_0, \eta)$ with

$$\ell^\eta_\lambda := \frac{2M}{\lambda} + \mu_{\|x_0\|+\eta}.$$

Here, $\mu_{\|x_0\|+\eta}$ is the constant given by $(\mathcal{H}^f_2)$ for $r = \|x_0\| + \eta$.

Proof of Claim 2. The first assertion follows directly from $(\mathcal{H}_C)$ and $(\mathcal{H}^f_1)$. To prove *(ii)*, we observe that according to (1.1), for all $t \in [T_0, T]$, the map $u \mapsto \mathrm{proj}_{C(t)}(u)$ is 1-Lipschitz in $\mathcal{H}$. Therefore, *(ii)* is a consequence of $(\mathcal{H}^2_\varphi)$ and $(\mathcal{H}^f_2)$. □

Hence, Claim 2 and the Cauchy-Lipschitz Theorem [38, p. 819], guarantee a unique solution $x_\lambda(\cdot) \in \mathbb{B}(x_0, \eta)$, for all $\lambda > 0$, of $(\mathcal{P}_L)$ on its maximal interval of existence $[T_0, T_\lambda[\subset [T_0, T]$. □

Remark 2.4.3 Claim 1 guarantees that the projection of $\nabla\varphi(x(t))$ over $C(t)$ is well-defined for all $x \in \mathbb{B}(x_0, \eta)$. In more general settings, such as the case where $C(t)$ is ρ-prox-regular (see, e.g.,[39]), one can define η in a more convenient way, such that $d_{C(t)}(\nabla\varphi(x(t)))$ belongs to $U^\gamma_\rho(C(t))$, and therefore, the projection exists. In our case, however, since we are working with a Chebyshev set, it suffices us to consider $\eta > 0$ a finite constant.

Now we show that the trajectories of $(\mathcal{P}_L)$ stay uniformly close to the moving sets (with respect to λ).

Lemma 2.4.4 *Suppose that $(\mathcal{H}^1_\varphi)$, $(\mathcal{H}^2_\varphi)$ and $(\mathcal{H}_C)$ hold. Let $\theta_\lambda(t) := d_{C(t)}(\nabla\varphi(x_\lambda(t)))$ for all $t \in [T_0, T)$. Then, θ_λ is absolutely continuous on $[T_0, T_\lambda[$, and a.e. for $t \in [T_0, T_\lambda[$,*

$$\dot\theta_\lambda(t) \le (\kappa + M\beta) - \frac{m}{\lambda}\theta_\lambda(t), \tag{2.7}$$

where β is given by (2.6).

Consequently,

$$\theta_\lambda(t) = d_{C(t)}(\nabla\varphi(x_\lambda(t))) \leq \frac{\kappa + M\beta}{m}\lambda \quad \textit{for all } t \in [T_0, T_\lambda[. \tag{2.8}$$

Proof. From Lemma 2.3.3, we deduce that θ_λ is absolutely continuous. To prove the inequality, let $t \in \Omega_\lambda := \{t \in [T_0, T_\lambda[\colon \theta_\lambda(t) \neq 0\}$ where $\dot{x}_\lambda(t)$ exists and set

$$\Phi(t, x) := \nabla\varphi(x) - \mathrm{proj}_{C(t)}(\nabla\varphi(x)).$$

Once again using Lemma 2.3.3, we have that

$$\begin{aligned}
\theta_\lambda(t)\dot{\theta}_\lambda(t) &\leq \kappa\theta_\lambda(t) + \langle \Phi(t, x_\lambda(t)), D^2\varphi(x_\lambda(t))\dot{x}_\lambda(t)\rangle \\
&= \kappa\theta_\lambda(t) - \frac{1}{\lambda}\langle \Phi(t, x_\lambda(t)), D^2\varphi(x_\lambda(t))\Phi(t, x_\lambda(t))\rangle \\
&\quad + \langle \nabla\Phi(t, x_\lambda(t)), D^2\varphi(x_\lambda(t))f(t, x_\lambda(t))\rangle \\
&\leq \kappa\theta_\lambda(t) - \frac{m}{\lambda}\theta_\lambda^2(t) + M\beta\theta_\lambda(t),
\end{aligned}$$

where we have used hypotheses $(\mathcal{H}_\varphi^1)$ and $(\mathcal{H}_\varphi^2)$, the fact that φ is twice differentiable, and that, by defintion, $\|\Phi(t, x_\lambda(t))\| = \theta_\lambda(t)$. Hence, we have shown that (2.7) holds for $t \in \Omega_\lambda$. Now, let $t \notin \Omega_\lambda$, i.e., $\theta_\lambda(t) = 0$. Then, θ_λ attains a minimum at this point and therefore $\dot{\theta}_\lambda(t) = 0$ for a.e. $t \notin \Omega_\lambda$ and (2.7) also holds in this case.

Thus, the inequality (2.7) holds in any case. The rest of the proof follows from Gronwall's Lemma (lemma 2.3.1) with $\alpha(t) \equiv \kappa + M\beta$ and $\beta(t) \equiv -\frac{m}{\lambda}$, obtaining for a.e. $t \in [T_0, T_\lambda[$,

$$\theta_\lambda(t) \leq \theta_\lambda(T_0)\exp\left(-\int_{T_0}^t \frac{m}{\lambda}d\tau\right) + \int_{T_0}^t (\kappa + M\beta)\exp\left(-\int_s^t \frac{m}{\lambda}d\tau\right)ds$$

And given that $\theta_\lambda(t) = d_{C(T_0)}(x_0) = 0$, we have that

$$\theta_\lambda(t) \leq \exp\left(-\frac{m}{\lambda}t\right)\int_{T_0^t}(\kappa + M\beta)\exp\left(\frac{m}{\lambda}s\right)ds.$$

Therefore, we conclude that

$$\theta_\lambda(t) \leq \frac{\kappa + M\beta}{m}\lambda,$$

which completes the proof. ☐

Remark 2.4.5 Given Lemma 2.4.4, we obtain that x_λ is a Lipschitz continuous map on $[T_0, T_\lambda[$ for any $\lambda > 0$.

Proposition 2.4.6 *For almost every $t \in [T_0, T_\lambda[$, one has*

$$||\dot{x}_\lambda(t) - f(t, x(t))|| \leq \gamma := \frac{\kappa + M\beta}{m}.$$

Consequently, for a.e. $t \in [T_0, T_\lambda[$

$$\|\dot{x}_\lambda(t)\| \leq \omega := \frac{\kappa + (M + m)\beta}{m}.$$

Proof. Let us recall that, according to $(\mathcal{P}_\lambda)$,

$$\dot{x}_\lambda(t) = -\frac{1}{\lambda}(\nabla\varphi(x_\lambda(t)) - \mathrm{proj}_{C(t)}(\nabla\varphi(x_\lambda(t)))) + f(t, x_\lambda(t)).$$

Therefore,

$$\|\dot{x}_\lambda(t) - f(t, x_\lambda(t))\| = \frac{1}{\lambda}\theta_\lambda(t).$$

Thanks to Lemma 2.4.4, we have that

$$\|\dot{x}_\lambda(t) - f(t, x_\lambda(t))\| \leq \frac{\kappa + M\beta}{m}.$$

To conclude, we notice that

$$\|\dot{x}_\lambda(t)\| \leq \|\dot{x}_\lambda(t) - f(t, x_\lambda(t))\| + \|f(t, x_\lambda(t))\|.$$

□

We conclude from Proposition 2.4.6 that $x_\lambda(t)$ is Lipschitz solution of $(\mathcal{P}_L)$ over the interval $[T_0, T_\lambda[$, where $x_\lambda(T_\lambda) \subset \mathbb{B}(x_0, \eta)$. Since T_λ is finite, we have that $x(T_\lambda) := \lim_{t \nearrow T_\lambda} x_\lambda(t)$ is well defined on $\mathcal{H}$. Therefore, we can extend the Lipschitz solution to $[T_0, T_\lambda]$, and we have that

$$||x(T_\lambda) - x_0|| \leq ||x(T_\lambda) - x(T_0)|| \leq \omega(T_\lambda - T_0) < \eta.$$

This yields that necessarily, $T_\lambda = T_0 + \sigma$, otherwise the previous inequality would allow us to extend x_λ on the right of T_λ in a solution of $(\mathcal{P}_L)$ with the extension of x_λ included in $\mathbb{B}(x_0, \eta)$, which contradicts the maximality of the interval $[T_0, T_\lambda[$.

Therefore, we have shown that for any $\lambda > 0$, $(\mathcal{P}_L)$ has a unique Lipschitz solution x_λ on the whole of the interval $[T_0, T]$ with $x_\lambda([T_0, T_\lambda]) \subset \mathbb{B}(x_0, \eta)$.

We will now prove that $(x_\lambda)_{\lambda>0}$ is a Cauchy sequence when $\lambda \searrow 0$. Firstly, let us show a preliminary results that will be used in the proof.

Remark 2.4.7 Let us notice that thanks to $(\mathcal{H}^2_\varphi)$ and (1.4) in Theorem 1.3.3, it holds that $\nabla\varphi^*$ is $\frac{1}{m}-$Lipschitz.

Lemma 2.4.8 *Consider $\lambda, \mu > 0$, and let us denote $w_\iota(t) := \mathrm{proj}_{C(t)}(\nabla\varphi(x_\iota(t)))$ for $\iota \in \{\lambda, \mu\}$. There exists $\widetilde{\rho}$ such that for all $t \in [T_0, T]$, the set $[\nabla\varphi]^{-1}(C(t))$ is $\widetilde{\rho}-$uniformly prox-regular. Furthermore, for a.e. $t \in [T_0, T]$*

$$v_\lambda(t) := -\frac{1}{\gamma M} D^2(\nabla\varphi^*(w_\lambda(t)))z_\lambda(t) \in N_{[\nabla\varphi]^{-1}(C(t))}(\nabla\varphi^*(w_\lambda(t))) \cap \mathbb{B},$$
$$v_\mu(t) := -\frac{1}{\gamma M} D^2(\nabla\varphi^*(w_\mu(t)))z_\mu(t) \in N_{[\nabla\varphi]^{-1}(C(t))}(\nabla\varphi^*(w_\mu(t))) \cap \mathbb{B},$$

where $z_\lambda(t) := \dot{x}_\lambda(t) - f(t, x_\lambda(t))$ and $z_\mu(t) := \dot{x}_\mu(t) - f(t, x_\mu(t))$.

Consequently, for a.e. $t \in [T_0, T]$

$$\langle v_\lambda(t) - v_\mu(t), \nabla\varphi^*(w_\lambda(t))) - \nabla\varphi^*(w_\mu(t)))\rangle \geq -\frac{1}{\widetilde{\rho}}||\nabla\varphi^*(w_\lambda(t))) - \nabla\varphi^*(w_\mu(t)))||^2.$$

Proof. Let us begin by proving that $\nabla\varphi$ is metrically calm. Let $t \in [T_0, T]$ and $x \in \mathcal{H}$. If $x \in \nabla\varphi^{-1}(C(t))$, then the result is immediate. Therefore, we will assume $x \in \mathcal{H}\backslash\nabla\varphi^{-1}(C(t))$. By (1.4), we know that $\nabla\varphi^{-1} = \nabla\varphi^*$. Then, one has

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

Therefore, $\nabla\varphi$ is $\frac{1}{m}-$metrically calm relatively to $C(t)$. This implies by Proposition 1.5.2 that $\nabla\varphi^*(C(t))$ satisfies the UNCIIP with constant $\frac{1}{m}$. By [13, Theorem 31] and [17, Proposition 3.14], $\nabla\varphi^*(C(t))$ is $\widetilde{\rho}-$prox-regular for some $\widetilde{\rho} > 0$.

Now, thanks to $(\mathcal{H}_\varphi^2)$ and $(\mathcal{H}_\varphi^3)$ we have that $\nabla\varphi$ is strictly differentiable, and therefore we can apply the chain rule for nonsmooth functions [23, Theorem 3.41] and the fact that $D^2\varphi$ is self-adjoint [3, Fact 2.66] to obtain the following equality for $\bar{x} \in \nabla\varphi^*(C(t))$:

$$N_{\nabla\varphi^*(C(t))}(\bar{x}) = D^2\varphi(\bar{x})N_{C(t)}(\nabla\varphi(\bar{x})). \tag{2.9}$$

Next, for $\iota \in \{\lambda, \mu\}$ we have that $z_\iota(t) := \dot{x}_\iota(t) - f(t, x_\iota(t)) \in -N_{C(t)}(w_\iota(t))$. Then, recalling the bounds for $\dot{x}_\iota$ in Proposition 2.4.6, and the fact that thanks to $(\mathcal{H}_\varphi^2)$ and $(\mathcal{H}_\varphi^3)$ we have that $\|D^2\varphi(u)\| \leq M$ for all $u \in \mathcal{H}$, it is clear that

$$-\frac{1}{\gamma M}D^2\varphi(\nabla\varphi^*(w_\iota(t)))z_\iota(t) \in D^2\varphi(\nabla\varphi^*(w_\iota(t)))N_{C(t)}(w_\iota(t)) \cap \mathbb{B}.$$

And using (2.9) with $\bar{x} = \nabla\varphi^*(w_\iota(t))$, it yields

$$-\frac{1}{\gamma M}D^2\varphi(\nabla\varphi^*(w_\iota(t)))z_\iota(t) \in N_{\nabla\varphi^*(C(t))}(\nabla\varphi^*(w_\iota(t))) \cap \mathbb{B}.$$

Finally, the inequality holds by *(iii)* from Theorem 1.4.3. □

Proposition 2.4.9 *There exists a constant $\mathcal{C} > 0$ such that for all positive numbers λ, μ, the following inequality holds:*

$$\sup_{t\in[T_0,T]} ||x_\lambda(t) - x_\mu(t)|| \leq \mathcal{C}(\lambda + \mu).$$

Proof. For the sake of readability, let us consider the following notation:

$$\begin{aligned} &w_\lambda(t) := \mathrm{proj}_{C(t)}(\nabla\varphi(x_\lambda(t))); && w_\mu(t) := \mathrm{proj}_{C(t)}(\nabla\varphi(x_\mu(t))), \\ &z_\lambda(t) := \dot{x}_\lambda(t) - f(t, x_\lambda(t)); && z_\mu(t) := \dot{x}_\mu(t) - f(t, x_\mu(t)), \\ &u_\lambda(t) := \nabla\varphi(x_\lambda(t)) - \mathrm{proj}_{C(t)}(\nabla\varphi(x_\lambda(t)); && u_\mu(t) := \nabla\varphi(x_\mu(t)) - \mathrm{proj}_{C(t)}(\nabla\varphi(x_\mu(t)). \end{aligned}$$

Then, using Proposition 2.4.6 for $\lambda, \mu > 0$, one has

$$\max\{||\dot{x}_\lambda||, ||\dot{x}_\mu||\} \leq \gamma \text{ and } \max\{\|\dot{x}_\lambda(t)\|, \|\dot{x}_\mu(t)\|\} \leq \omega \text{ for a.e. } t \in [T_0, T].$$

We now consider the $\varphi-$Symmetrized Bregman Divergence between $x_\lambda(t)$ and $x_\mu(t)$:

$$\Upsilon(t) := \langle \nabla\varphi(x_\lambda(t)) - \nabla\varphi(x_\mu(t)), x_\lambda(t) - x_\mu(t)\rangle \text{ for all } t \in [T_0, T].$$

Thanks to $(\mathcal{H}_\varphi^2)$, it is clear that Υ is absolutely continuous with

$$\dot{\Upsilon}(t) = I_1(t) + I_2(t) \text{ for a.e. } t \in [T_0, T],$$

where

$$\begin{aligned} I_1(t) &:= \langle D^2\varphi(x_\lambda(t))\dot{x}_\lambda(t) - D^2\varphi(x_\mu(t))\dot{x}_\mu(t), x_\lambda(t) - x_\mu(t)\rangle, \\ I_2(t) &:= \langle \nabla\varphi(x_\lambda(t)) - \nabla\varphi(x_\mu(t)), \dot{x}_\lambda(t) - \dot{x}_\mu(t)\rangle. \end{aligned}$$

We will now find suitable bounds for these quantities.

Estimation of $I_1(t)$: Let us write I_1 for $t \in [T_0, T]$ as

$$I_1(t) = I_{11}(t) + I_{12}(t) + I_{13}(t),$$

where

$$\begin{aligned} I_{11}(t) &:= \langle D^2\varphi(x_\lambda(t))\dot{x}_\lambda(t) - D^2\varphi(\nabla\varphi^*(w_\lambda(t)))\dot{x}_\lambda(t), x_\lambda(t) - x_\mu(t)\rangle, \\ I_{12}(t) &:= \langle D^2\varphi(\nabla\varphi^*(w_\lambda(t)))\dot{x}_\lambda(t) - D^2\varphi(\nabla\varphi^*(w_\mu(t)))\dot{x}_\mu(t), x_\lambda(t) - x_\mu(t)\rangle, \\ I_{13}(t) &:= \langle D^2\varphi(\nabla\varphi^*(w_\mu(t)))\dot{x}_\mu(t) - D^2\varphi(x_\mu(t))\dot{x}_\mu(t), x_\lambda(t) - x_\mu(t)\rangle. \end{aligned}$$

We will first estimate $I_{11}(t)$ and $I_{13}(t)$. To do this, let us recall that by Proposition 2.4.2, we have that $x_\lambda, x_\mu \in \mathbb{B}(x, \eta)$. Therefore, by using Lemma 2.4.4, Proposition 2.4.6, and the hypotheses $(\mathcal{H}_\varphi^1)$ and $(\mathcal{H}_\varphi^3)$, we obtain for a.e. $t \in [T_0, T]$,

$$\begin{aligned} I_{11}(t) &\le \vartheta\|\dot{x}_\lambda(t)\| \cdot \|x_\lambda(t) - \nabla\varphi^*(w_\lambda(t))\| \cdot \|x_\lambda(t) - x_\mu(t)\| \\ &\le \vartheta \cdot \omega \cdot \|\nabla\varphi^*(\nabla\varphi(x_\lambda(t))) - \nabla\varphi^*(w_\lambda(t))\| \cdot \|x_\lambda(t) - x_\mu(t)\| \\ &\le \vartheta \cdot \omega \cdot \frac{1}{m}\|\nabla\varphi(x_\lambda(t)) - w_\lambda(t)\| \cdot \eta \\ &= \frac{\vartheta\omega\eta}{m} \cdot \theta_\lambda(t) \\ &\le \frac{\vartheta\omega\eta\gamma}{m}\lambda. \end{aligned}$$

Similarly, we obtain that, for a.e. $t \in [T_0, T]$,

$$I_{13}(t) \le \frac{\vartheta\omega\eta\gamma}{m}\mu.$$

Moving onto the estimation of I_{12}, we will write for $t \in [T_0, T]$

$$I_{12}(t) = J_1(t) + J_2(t) + J_3(t),$$

where

$$\begin{aligned} J_1(t) &:= \langle D^2\varphi(\nabla\varphi^*(w_\lambda(t)))\dot{x}_\lambda(t) - D^2\varphi(\nabla\varphi^*(w_\mu(t)))\dot{x}_\mu(t), x_\lambda(t) - \nabla\varphi^*(w_\lambda(t))\rangle, \\ J_2(t) &:= \langle D^2\varphi(\nabla\varphi^*(w_\lambda(t)))\dot{x}_\lambda(t) - D^2\varphi(\nabla\varphi^*(w_\mu(t)))\dot{x}_\mu(t), \nabla\varphi^*(w_\lambda(t)) - \nabla\varphi^*(w_\mu(t))\rangle, \\ J_3(t) &:= \langle D^2\varphi(\nabla\varphi^*(w_\lambda(t)))\dot{x}_\lambda(t) - D^2\varphi(\nabla\varphi^*(w_\mu(t)))\dot{x}_\mu(t), \nabla\varphi^*(w_\mu(t)) - x_\mu(t)\rangle. \end{aligned}$$

In the same way as before, we will first give the estimation of $J_1(t)$ and $J_3(t)$, which are obtained in a similar way.

First we notice that according to $(\mathcal{H}_\varphi^2)$, we have that

$$\|D^2\varphi(u)\| \leq M \text{ for all } u \in \mathcal{H}. \tag{2.10}$$

Secondly, thanks to $(\mathcal{H}_\varphi^1)$ and Lemma 2.4.4, we know that for a.e. $t \in [T_0, T]$,

$$\begin{aligned}
J_1(t) &\leq 2M \cdot \omega \cdot \|x_\lambda(t) - \nabla\varphi^*(w_\lambda(t))\| \\
&\leq 2M \cdot \omega \cdot \frac{1}{m}\theta_\lambda(t) \\
&\leq 2\frac{M\omega\gamma}{m}\lambda
\end{aligned}$$

Similarly, for a.e. $t \in [T_0, T]$

$$J_3(t) \leq 2\frac{M\omega\gamma}{m}\mu.$$

Now, to estimate $J_2(t)$, let us notice that

$$J_2(t) = J_{21}(t) + J_{22}(t) + J_{23}(t) \text{ for a.e. } t \in [T_0, T],$$

where

$$\begin{aligned}
J_{21}(t) &:= \langle D^2\varphi(\nabla\varphi^*(w_\lambda(t)))z_\lambda(t) - D^2\varphi(\nabla\varphi^*(w_\mu(t)))z_\mu(t), \nabla\varphi^*(w_\lambda(t)) - \nabla\varphi^*(w_\mu(t))\rangle, \\
J_{22}(t) &:= \langle\left[D^2\varphi(\nabla\varphi^*(w_\lambda(t))) - D^2\varphi(\nabla\varphi^*(w_\mu(t)))\right] f(t, x_\lambda(t)), \nabla\varphi^*(w_\lambda(t)) - \nabla\varphi^*(w_\mu(t))\rangle, \\
J_{23}(t) &:= \langle D^2\varphi(\nabla\varphi^*(w_\mu(t)))\left[f(t, x_\lambda(t)) - f(t, x_\mu(t))\right], \nabla\varphi^*(w_\lambda(t)) - \nabla\varphi^*(w_\mu(t))\rangle.
\end{aligned}$$

To estimate J_{21} we resort to Lemma 2.4.8.

$$\begin{aligned}
J_{21}(t) &= -M\gamma\langle v_\lambda(t) - v_\mu(t), \nabla\varphi^*(w_\lambda(t)) - \nabla\varphi^*(w_\mu(t))\rangle \\
&\leq \frac{M\gamma}{\tilde{\rho}}\|\nabla\varphi^*(w_\lambda(t)) - \nabla\varphi^*(w_\mu(t))\|^2 \\
&\leq \frac{M\gamma}{\tilde{\rho}m^2}\|w_\lambda(t) - w_\mu(t)\|^2 \\
&\leq \frac{M^3\gamma}{\tilde{\rho}m^2}\|x_\lambda(t) - x_\mu(t)\|^2,
\end{aligned}$$

where we used $(\mathcal{H}_\varphi^1)$, $(\mathcal{H}_\varphi^2)$ and (1.1).

For J_{22} we use hypotheses $(\mathcal{H}_\varphi^1)$,$(\mathcal{H}_\varphi^2)$,$(\mathcal{H}_\varphi^3)$ and (1.1).

$$\begin{aligned}
J_{22}(t) &\leq \vartheta\|f(t, x_\lambda(t))\| \cdot \|\nabla\varphi^*(w_\lambda(t)) - \nabla\varphi^*(w_\mu(t))\|^2 \\
&\leq \vartheta \cdot \beta \cdot \frac{1}{m^2}\|w_\lambda(t) - w_\mu(t)\|^2 \\
&\leq \frac{\vartheta\beta}{m^2}\|\nabla\varphi(x_\lambda(t)) - \nabla\varphi(x_\mu(t))\|^2 \\
&\leq \frac{\vartheta\beta M^2}{m^2}\|x_\lambda(t) - x_\mu(t)\|^2.
\end{aligned}$$

Finally, to estimate J_{23} we use $(\mathcal{H}_\varphi^1)$, $(\mathcal{H}_2^f)$ and (2.10).

$$\begin{aligned}
J_{23}(t) &\leq M\|f(t,x_\lambda(t)) - f(t,x_\mu(t))\| \cdot \|\nabla\varphi^*(w_\lambda(t)) - \nabla\varphi^*(w_\mu(t))\| \\
&\leq M \cdot \mu_{\|x_0\|+\eta}\|x_\lambda(t) - x_\mu(t)\| \cdot \frac{1}{m}\|w_\lambda(t) - w_\mu(t)\| \\
&\leq M \cdot \mu_{\|x_0\|+\eta}\|x_\lambda(t) - x_\mu(t)\| \cdot \frac{1}{m}\|x_\lambda(t) - x_\mu(t)\| \\
&= \frac{M}{m}\mu_{\|x_0\|+\eta}\|x_\lambda(t) - x_\mu(t)\|^2,
\end{aligned}$$

Estimation of $I_2(t)$: let us notice that

$$I_2(t) = I_{21}(t) + I_{22}(t) \text{ for a.e. } t \in [T_0, T],$$

where

$$\begin{aligned}
I_{21}(t) &:= \langle u_\lambda(t) - u_\mu(t), \dot{x}_\lambda(t) - \dot{x}_\mu(t)\rangle, \\
I_{22}(t) &:= \langle w_\lambda(t) - w_\mu(t), \dot{x}_\lambda(t) - \dot{x}_\mu(t)\rangle.
\end{aligned}$$

To estimate $I_{21}(t)$, we use Proposition 2.4.6 to obtain for a.e. $t \in [T_0, T]$,

$$\begin{aligned}
I_{21}(t) &= \langle u_\lambda(t) - u_\mu(t), \dot{x}_\lambda(t) - \dot{x}_\mu(t)\rangle \\
&= \langle -\lambda z_\lambda(t) + \mu z_\mu(t), \dot{x}_\lambda(t) - \dot{x}_\mu(t)\rangle \\
&\leq (\lambda\|z_\lambda(t)\| + \mu\|z_\mu(t)\|)(\|\dot{x}_\lambda(t)\| + \|\dot{x}_\mu(t)\|) \\
&\leq 2\gamma\omega(\lambda + \mu).
\end{aligned}$$

For $I_{22}(t)$ we consider hypotheses $(\mathcal{H}_\varphi^2)$, $(\mathcal{H}_\varphi^2)$ and (1.1), and we notice that, according to $(\mathcal{P}_\lambda)$ and Remark 1.2.4, $\dot{x}_\lambda(t) - f(t, x_\lambda(t)) \in -N_{C(t)}(\text{proj}_{C(t)}(\nabla\varphi(x_\lambda)))$ (analogously for $\dot{x}_\mu$). Therefore,

$$\begin{aligned}
I_{22}(t) &= \langle w_\lambda(t) - w_\mu(t), \dot{x}_\lambda(t) - f(t, x_\lambda(t)) - (\dot{x}_\mu(t) - f(t, x_\mu(t)))\rangle \\
&\quad + \langle w_\lambda(t) - w_\mu(t), f(t, x_\lambda(t)) - f_\mu(t, x_\lambda(t))\rangle \\
&\leq \|w_\lambda(t) - w_\mu(t)\| \cdot \|f(t, x_\lambda(t)) - f(t, x_\mu(t)) \\
&\leq M\mu_{\|x_0\|+\eta}\|x_\lambda(t) - x_\mu(t)\|^2.
\end{aligned}$$

Final steps: considering all the previous estimates, we obtain that there exists $c_1, c_2 > 0$ such that

$$\dot{\Upsilon}(t) \leq c_1(\lambda + \mu) + c_2\|x_\lambda(t) - x_\mu(t)\|^2 \quad \text{for a.e. } t \in [T_0, T].$$

Therefore, for all $t \in [T_0, T]$,

$$\Upsilon(t) \leq c_1(\lambda + \mu)(t - T_0) + c_2\int_{T_0}^t \|x_\lambda(s) - x_\mu(s)\|^2 ds.$$

It follows by virtue of $(\mathcal{H}_\varphi^1)$ that for all $t \in [T_0, T]$,

$$m\|x_\lambda(t) - x_\mu(t)\|^2 \leq \Upsilon(t) \leq c_1(\lambda + \mu)(t - T_0) + c_2\int_{T_0}^t \|x_\lambda(s) - x_\mu(s)\|^2 ds.$$

Hence, we have found positive constants $\mathcal{C}_1, \mathcal{C}_2$, such that for all $t \in [T_0, T]$,

$$m\|x_\lambda(t) - x_\mu(t)\|^2 \leq \mathcal{C}_1(\lambda + \mu) + \mathcal{C}_2 \int_{T_0}^{t} \|x_\lambda(s) - x_\mu(s)\|^2 ds.$$

We conclude by using Gronwall's Lemma, obtaining for $t \in [T_0, T]$,

$$\|x_\lambda(t) - x_\mu(t)\|^2 \leq \mathcal{C}_1 \exp(\mathcal{C}_2(t - T_0))(\lambda + \mu).$$

$\square$

In the remaining of this chapter we will prove that the family $(x_\lambda)_{\lambda>0}$ converges uniformly to a unique solution of $(\mathcal{P}_L)$.

Lemma 2.4.10 *The family $(x_\lambda)_{\lambda>0}$ converges uniformly to a $\frac{\kappa+(M+m)\beta}{m}-$Lipschitz solution of $(\mathcal{P}_L)$.*

Proof. The convergence of $x_\lambda(\cdot)$ is guaranteed over the interval $[T_0, T]$ thanks to the Cauchy criteria given by Proposition 2.4.9. Therefore, $(x_\lambda(\cdot))_\lambda$ uniformly converges to $x(\cdot) \in \mathcal{C}([T_0, T], \mathcal{H})$ when $\lambda \searrow 0$. Hence, it suffices to prove that the limit of x_λ satisfies $(\mathcal{P}_L)$.

On the one hand, according to Proposition 2.4.6, we have that

$$\theta_\lambda(t) = d_{C(t)}(\nabla\varphi(x_\lambda(t))) \leq \frac{\kappa + M\beta}{m}\lambda,$$

and therefore $\nabla\varphi(x(t)) \in C(t)$ when $\lambda \searrow 0$ for all $t \in [T_0, T]$.

Next, let us recall from Proposition 2.4.6 that for a.e. $t \in [T_0, T]$

$$\|\dot{x}(t)\| \leq \frac{\kappa + (M + m)\beta}{m}, \tag{2.11}$$

and thus, there exists a sequence $(\lambda_n)_n$ such that $\dot{x}_{\lambda_n}(\cdot)$ weakly converges to an element $z(\cdot) \in L^2([T_0, T], \mathcal{H})$ when $\lambda_n \searrow 0$, such that for a.e. $t \in [T_0, T]$,

$$\|z(t)\| \leq \frac{\kappa + (M + m)\beta}{m}.$$

This inequality still holds at the limit thanks to the fact that the set

$$\left\{ v \in L^2([T_0, T], \mathcal{H}) \colon \|v(t)\| \leq \frac{\kappa + (M + m)\beta}{m} \text{ for a.e. } t \in [T_0, T] \right\}$$

is a closed and convex set in $L^2([T_0, T]; \mathcal{H})$.

Next, for $t \in [T_0, T]$ let us fix $h \in \mathcal{H}$. Then, if we write

$$\left\langle h, \int_{T_0}^{t} \dot{x}_{\lambda_n}(s) ds \right\rangle = \int_{T_0}^{t} \langle h\mathbb{1}_{[T_0,t]}(s), \dot{x}_{\lambda_n}(s)\rangle ds,$$

we see that

$$\int_{T_0}^{t} \dot{x}_{\lambda_n}(s) \rightharpoonup \int_{T_0}^{t} z(s) ds.$$

Moreover, since $(x_\lambda(t))_\lambda$ converges to $x(t)$ strongly in $\mathcal{H}$, we have that

$$x_{\lambda_n}(t) = x_0 + \int_{T_0}^{t} \dot{x}_{\lambda_n}(s)ds$$

converges to $x(t) = x_0 + \int_{T_0}^{t} z(s)ds$. Therefore, $x(\cdot)$ is absolutely continuous with $\dot{x}(t) = z(t)$ for a.e. $t \in [T_0, T]$, and we have that $\dot{x}_{\lambda_n}(\cdot)$ converges weakly to $\dot{x}(\cdot)$ in $L^2([T_0, T], \mathcal{H})$. This entails that

$$\dot{x}_{\lambda_n}(\cdot) - f(\cdot, x_{\lambda_n}(\cdot)) \rightharpoonup \dot{x}(\cdot) - f(\cdot, x(\cdot)) \tag{2.12}$$

in $L^2([T_0, T], \mathcal{H})$, and thanks to Proposition 2.4.6, this limit is bounded for a.e. $t \in [T_0, T]$, and we also have that $\|\dot{x}(t)\| \leq \frac{\kappa+(M+m)\beta}{m}$, which implies that

$$\|x(t) - x_0\| \leq \frac{\kappa + (M+m)\beta}{m}(t - T_0) \leq \eta.$$

Therefore, $x \in \mathbb{B}(x_0, \eta)$, for all $t \in [T_0, T]$.

For the last part of this proof we make use of Mazur's lemma (Lemma 2.3.2) through (2.12), implying that there exist the real numbers $r(n) > n$ for all $n \in \mathbb{N}$ and $s_{k,n} \geq 0$ with $\sum_{k=n}^{r(n)} s_{k,n} = 1$, such that

$$\sum_{k=n}^{r(n)} s_{k,n}(f_{\lambda_k} - \dot{x}_{\lambda_k}) \to (f - \dot{x})$$

strongly in $L^2([T_0, T], \mathcal{H})$, and where we denote $f_{\lambda_k}(t) = f(t, x_{\lambda_k}(t))$. By extracting a subsequence, we can assume $\dot{x}(t)$ and $\dot{x}_{\lambda_n}(t)$ exist for a.e. $t \in [T_0, T]$, and

$$\sum_{k=n}^{r(n)} s_{k,n}(f_{\lambda_k}(t) - \dot{x}_{\lambda_k}(t)) \to (f(t, x(t)) - \dot{x}(t)) \text{ for a.e. } t \in [T_0, T].$$

We suppose that the inequalities in Proposition 2.4.6 also hold a.e. for $t \in [T_0, T]$, and for all λ_n, $n \in \mathbb{N}$. Let us fix $t \in [T_0, T]$ such that $\dot{x}(t)$ and $\dot{x}_{\lambda_n}(t)$ are well defined. Then, we have that

$$\left| \sum_{k=n}^{r(n)} s_{k,n}\langle f_{\lambda_k}(t) - \dot{x}_{\lambda_k}(t), \nabla\varphi(x(t)) - \mathrm{proj}_{C(t)}(\nabla\varphi(x_{\lambda_k}(t)))\rangle \right| \tag{2.13}$$

$$\leq \quad \frac{\kappa + M\beta}{m} \sum_{k=n}^{r(n)} s_{k,n}\|\nabla\varphi(x(t)) - \mathrm{proj}_{C(t)}(\nabla\varphi(x_{\lambda_k}(t)))\|. \tag{2.14}$$

Since both the projection and $\nabla\varphi$ are Lipschitz, and $\nabla\varphi(x(t)) \in C(t)$, we have that, when $n \to +\infty$,

$$\mathrm{proj}_{C(t)}(\nabla\varphi(x_{\lambda_n}(t))) - \nabla\varphi(x(t)) \to 0, \tag{2.15}$$

strongly in $\mathcal{H}$. Therefore, according to (2.13), as $n \to +\infty$, it yields

$$\sum_{k=n}^{r(n)} s_{k,n}\langle f_{\lambda_k}(t) - \dot{x}_{\lambda_k}(t), \nabla\varphi(x(t)) - \mathrm{proj}_{C(t)}(\nabla\varphi(x_{\lambda_k}(t)))\rangle \to 0.$$

Next, let $x' \in \mathcal{H}$, then

$$\begin{aligned}
&\sum_{k=n}^{r(n)} s_{k,n}\langle f_{\lambda_k}(t) - \dot{x}_{\lambda_k}(t), \nabla\varphi(x') - \mathrm{proj}_{C(t)}(\nabla\varphi(x_{\lambda_k}(t)))\rangle \\
= &\left\langle \sum_{k=n}^{r(n)} s_{k,n}(f_{\lambda_k}(t) - \dot{x}_{\lambda_k}(t)), \nabla\varphi(x') - \nabla\varphi(x(t)) \right\rangle \\
&+ \sum_{k=n}^{r(n)} s_{k.n}\langle f_{\lambda_k}(t) - \dot{x}_{\lambda_k}(t), \nabla\varphi(x(t))\rangle
\end{aligned}$$

According to this equality and (2.15), we have that

$$\sum_{k=n}^{r(n)} s_{k,n}\langle f_{\lambda_k}(t) - \dot{x}_{\lambda_k}(t), \nabla\varphi(x') - \mathrm{proj}_{C(t)}(\nabla\varphi(x_{\lambda_k}(t)))\rangle$$

converges to

$$\langle f(t, x(t)) - \dot{x}(t), \nabla\varphi(x') - \nabla\varphi(x_{\lambda_k}(t))\rangle.$$

On the other hand, if we recall Proposition 2.4.6, and the fact that $f_{\lambda_k}(t) - \dot{x}_{\lambda_k}(t) \in N_{C(t)}(\mathrm{proj}_{C(t)}(\nabla\varphi(x_\lambda(t))))$, we have that for all $x' \in C(t)$,

$$\sum_{k=n}^{r(n)} s_{k,n}\langle f_{\lambda_k}(t) - \dot{x}_{\lambda_k}(t), \nabla\varphi(x') - \mathrm{proj}_{C(t)}(\nabla\varphi(x_{\lambda_k}(t)))\rangle \leq 0.$$

Hence, since $\mathrm{proj}_{C(t)}(\nabla\varphi(x_{\lambda_n}(t))) \to \nabla\varphi(x(t))$ as $n \to +\infty$, it follows that

$$\sum_{k=n}^{r(n)} s_{k,n}\langle f_{\lambda_k}(t) - \dot{x}_{\lambda_k}(t), \nabla\varphi(x') - \mathrm{proj}_{C(t)}(\nabla\varphi(x(t)))\rangle \leq 0,$$

i.e.,

$$-\dot{x}(t) + f(t, x(t)) \in N_{C(t)}(\nabla\varphi(x(t))) \quad \text{for a.e. } t \in [T_0, T],$$

which completes the proof. $\square$

Finally, we prove uniqueness of the solution.

Lemma 2.4.11 *The dynamical system $(\mathcal{P}_L)$ has a unique solution.*

Proof. Let x_1 and x_2 be two solutions of $(\mathcal{P}_L)$. This proof is similar to the proof of Proposition 2.4.9. Therefore, we consider the $\varphi-$Symmetrized Bregman Divergence between x_1 and x_2:

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

and its derivative,

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

where

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

Just as the proof in Proposition 2.4.9, we will estimate I_1 and I_2.

Estimation of $I_1(t)$: let us write this expression as

$$I_1(t) := I_{11}(t) + I_2(t),$$

where

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

Here, to find a bound for $I_{11}(t)$ we use the fact that $\nabla\varphi^*(C(t))$ is a prox-regular set with constant $\widetilde{\rho}$, as proven in Lemma 2.4.8. Thanks to the same lemma, we get that for $i \in \{1,2\}$

$$-\frac{1}{\gamma M}D^2\varphi(x_i(t))(\dot{x}_i(t) - f(t,x_i(t))) \in N_{\nabla\varphi^*(C(t))}(\nabla\varphi^*(\mathrm{proj}_{C(t)}(x_i))),$$

and therefore by (iii) in Theorem 1.4.3, we get that, for a.e. $t \in [T_0,T]$,

$$I_{11}(t) \leq \frac{\gamma M}{\widetilde{\rho}}\|x_1(t) - x_2(t)\|^2.$$

To estimate $I_{12}(t)$, let us notice that

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

where

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

For J_1, let us set $\upsilon := \max\{\|f(t,x_1(t))\|, \|f(t,x_2(t))\|\}$. By $(\mathcal{H}_3^f)$, we know that υ is a finite constant. Therefore, using $(\mathcal{H}_\varphi^3)$, we have that

$$\begin{aligned}
J_1(t) &\leq \|D^2\varphi(x_1(t))f(t,x_1(t)) - D^2\varphi(x_2(t))f(t,x_1(t))\| \cdot \|x_1(t) - x_2(t)\| \\
&\leq \|f(t,x_1(t))\| \cdot \|D^2\varphi(x_1(t)) - D^2\varphi(x_2(t))\| \cdot \|x_1(t) - x_2(t)\| \\
&\leq \upsilon\vartheta\|x_1(t) - x_2(t)\|^2
\end{aligned}$$

As for $J_2(t)$, we recall from Proposition 2.4.6 that for a.e. $t \in [T_0,T]$ and $i \in \{1,2\}$, $\|\dot{x}_i(t)\| \leq \omega$, and therefore, $\|x_i(t)\| \leq \omega \cdot \sigma < \eta$, for $i \in \{1,2\}$. Using this fact, $(\mathcal{H}_\varphi^2)$ and $(\mathcal{H}_2^f)$, it follows that

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

Estimation of $I_2(t)$: first, let us write

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

where

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

Notice that since $\dot{x}_i(t) - f(t, x_i(t)) \in -N_{C(t)}(\nabla\varphi(x_i(t)))$, for $i = 1, 2$, then by definition of the normal cone (Definition 1.2.3), we conclude that $I_{21}(t) \leq 0$. As for I_{22}, thanks to $(\mathcal{H}_\varphi^2)$ and $(\mathcal{H}_2^f)$, we have that

$$I_{22}(t) \leq M\mu_{\|x_0\|+\eta}\|x_1(t) - x_2(t)\|^2,$$

Hence, we have arrived to the following bound, for a.e. $t \in [T_0, T]$:

$$\dot{\Upsilon}(t) \leq \mathcal{C}\|x_1(t) - x_2(t)\|^2,$$

with $\mathcal{C}$ a positive finite constant. This inequality implies that

$$\Upsilon(t) \leq \mathcal{C}\int_{T_0}^{t} \|x_1(s) - x_2(s)\|^2 ds.$$

Therefore, defining the absolutely continuous function $F(t) := \int_{T_0}^{t} \|x_1(s) - x_2(s)\|^2 ds$, using $(\mathcal{H}_\varphi^1)$ and the bound for $\dot{\Upsilon}$ to conclude that, for all $t \in [T_0, T]$,

$$m\dot{F}(t) \leq \Upsilon(t) \leq \mathcal{C}F(t),$$

where $\dot{F}(t) = \|x_1(s) - x_2(s)\|^2$ by a straightforward calculation. Hence, by Gronwall's Lemma (lemma 2.3.1), we get that

$$F(t) \leq F(T_0)\exp\left(\int_{T_0}^{t} \mathcal{C}ds\right) \quad \forall t \in [T_0, T].$$

And since x_1 and x_2 are solutions of the sweeping process, they satisfy the initial condition $x_1(T_0) = x_2(T_0) = x_0$, which implies that $F(T_0) = 0$. This means that

$$\|x_1(s) - x_2(s)\|^2 \leq 0,$$

which concludes the proof. $\square$

Chapter 3

Bounded Retraction Sweeping Process

We will now give a solution of the Degenerate Mirror Sweeping Process in the case where the moving set C is of bounded retraction. For the rest of this chapter, let us assume $(\mathcal{H}_\varphi^1)$, $(\mathcal{H}_\varphi^2)$ and $(\mathcal{H}_\varphi^3)$ hold locally over an open set $\mathcal{D}$ contained in the domain of $\nabla\varphi$. For the moving set C, let us assume it has closed, nonempty and convex values. Moreover, C is of bounded retraction over the interval $[0,T]$, i.e., $C \in \mathrm{BR}([0,T];\mathcal{C}_\mathcal{H})$. Finally, we will assume this dynamic is not perturbed, i.e., from now on $f \equiv 0$.

The proof of existence of said solution is inspired by the methods used in [33, 34, 36]; that is, we will prove that for $T > 0$, if $C \in \mathrm{BR}([0,T],\mathcal{C}_\mathcal{H})$, there exists a solution $x \in \mathrm{BV}([0,T];\mathcal{H})$ such that there is a measure $\mu : \mathcal{B}([0,T]) \to [0,\infty]$ and a function $v \in L^1_{loc}(\mu;H)$ that satisfy the following dynamic:

$$\begin{cases} \nabla\varphi(x(t)) \in C(t) \quad \forall t \in [0,T] \\ Dx = v\mu \\ -v(t) \in N_{C(t)}(\nabla\varphi(x(t))) \quad \mu-\text{a.e.}, t \in \mathrm{Cont}(C) \\ \nabla\varphi(x(t)) = \mathrm{Proj}_{C(t)}(\nabla\varphi(x(t^-)), \quad \forall t \in \mathrm{Discont}(C) \setminus \{0\} \\ \nabla\varphi(x(t^+)) = \mathrm{Proj}_{C(t^+)}(\nabla\varphi(x(t))) \quad \forall t \in \mathrm{Discont}(C) \setminus \{0\} \\ \nabla\varphi(x(0)) = \mathrm{Proj}_{C(0)}(\nabla\varphi(x_0)) \\ \nabla\varphi(x(0^+)) = \mathrm{Proj}_{C(0^+)}(\nabla\varphi(x(0)) \end{cases} \tag{$\mathcal{P}_{BR}$}$$

To prove existence, we will show that we can reduce $(\mathcal{P}_{BR})$ to the case where the moving set is 1-Lipschitz continuous by parametrizing the moving set with the arc length $\ell_C(t) = \mathrm{R}(C;[0,t])$, so we can write $C(t) := \widetilde{C}(\ell_C(t))$. This method was first used by Moreau in [26] when reducing the absolutely continuous sweeping process to the Lipschitz case. In our case, this method comes naturally if we notice that whenever the retraction of a moving set C is Lipschitz, $(\mathcal{P}_{BR})$ is reduced to a dynamic much like the one seen in $(\mathcal{P}_L)$, except that the moving set must be Lipschitz with respect to the excess instead of the Hausdorff distance, since, in this case, one lacks of discontinuity points for the moving set. The Lipschitz case with respect to the excess has been studied in [16], where by means of a catching-up-like algorithm, it is proven there exists a unique solution for the Mirror Sweeping Process. From now on, whenever we speak of $(\mathcal{P}_L)$, we will refer to the case in which the moving set is Lipschitz with respect to the excess.

Let us return our attention back to the parametrization of the moving set. The problem with this is that $\widetilde{C}$ is only defined in the image $\ell_C([0,T])$, and therefore we need to connect the jumps of C with a suitable class of geodesics (see for instance [36]), and consequently defining $\widetilde{C}$ on the whole interval $[0, \mathrm{R}(C,[0,T])]$. This way, we can get the solution of the Mirror Sweeping Process driven by $\widetilde{C}$ using the method proposed in [16], and then discard the jumps to get the solution associated with C. It has been shown in [32, 35] that the choice of paths to connect the endpoints of a jump is non-trivial, and in [36] the author presents the suitable class of geodesics for the bounded retraction case that we will be using in this work.

Finally, let us consider the *solution operator* of $(\mathcal{P}_{BR})$,

$$\overline{S} : \mathrm{BR}([0,T];\mathcal{C}_{\mathcal{H}}) \times \mathcal{H} \to \mathrm{BV}([0,T];\mathcal{H}), \tag{3.1}$$

which assigns to each pair (C, x_0) the solution $x \in \mathrm{BV}([0,T];\mathcal{H})$ of $(\mathcal{P}_{BR})$. By virtue of Lemma 2.4.10 , we can define the solution operator associated to $(\mathcal{P}_L)$ in the Lipschitz setting:

$$S\colon \mathrm{Lip}([0,T];\mathcal{C}_{\mathcal{H}}) \to \mathrm{Lip}([0,T];\mathcal{H}).$$

In order to reduce $(\mathcal{P}_{BR})$ to $(\mathcal{P}_L)$, we will rely on the following proposition regarding the rate independence of the sweeping process. For a similar result, we refer to [34].

Proposition 3.0.1 *Let* $S\colon \mathrm{Lip}([0,T];\mathcal{C}_{\mathcal{H}}) \to \mathrm{Lip}([0,T];\mathcal{H})$ *be the solution operator of* $(\mathcal{P}_L)$*. If* $\phi\colon [0,T] \to [0,T]$ *is continuous and non-decreasing, then*

$$S(C \circ \phi, x_0) = S(C, x_0) \circ \phi. \tag{3.2}$$

Proof. Let $x := S(C, x_0)$ be a solution for $(\mathcal{P}_L)$, that is,

$$\dot{x}(t) \in -N_{C(t)}(\nabla\varphi(x(t))) \text{ for a.e. } t \in [0,T]. \tag{3.3}$$

Now, let us consider $z := x \circ \phi = S(C, x_0) \circ \phi$. Since ϕ is non-decreasing,

$$\dot{z}(t) = \dot{x}(\phi(t))\dot{\phi}(t) \text{ for a.e. } t \in [0,T].$$

Then, by using (3.3), we conclude that

$$\dot{z}(t) = \dot{x}(\phi(t))\dot{\phi}(t) \in -N_{C(\phi(t))}(\nabla\varphi(z(t))) \text{ for a.e. } t \in [0,T],$$

where we have used that $\dot{\phi} \geq 0$ almost everywhere. Hence, z is a solution of $(\mathcal{P}_L)$ driven by the sets $t \mapsto (C \circ \phi)(t)$. Therefore, $z = S(C \circ \phi, x_0)$, which concludes the proof. ☐

3.1 Geodesics for the retraction

As stated before, our goal is to redefine $\widetilde{C}$ on the interval $[0,T]$, and therefore we need suitable paths to fill in the jumps of C. To do this, we introduce the class of geodesics with respect to the excess that will allow us reduce $(\mathcal{P}_{BR})$ as discussed above.

Definition 3.1.1 *Assume that $A, B \in \mathcal{C}_{\mathcal{H}}$ and set $\rho := e(A, B)$. We define the curve $\mathcal{F}_{(A,B)} \colon [0, 1] \to \mathcal{C}_{\mathcal{H}}$ by*

$$\mathcal{F}_{(A,B)}(t) := \begin{cases} A & \text{if } t = 0, \\ B + (1-t)D_\rho = B + D_{(1-t)\rho} & \text{if } 0 < t \leq 1. \end{cases} \tag{3.4}$$

Figure 3.1 shows a graphic representation of the geodesic in a discontinuity interval. We see that when $t = 0$, then the geodesic is just the set A, whereas when there is a $t \in [0, 1]$, we see that the geodesic "enhances" the set B with a radius of $(1 - t)\rho$, where ρ is the excess of A over B, represented in the figure as the outer rim. That is, we build our geodesic by modifying the set B just enough so it can sweep the point from A to B at each time of the discontinuity interval. Since we define the geodesic in function of A and B, which in our case are convex sets, then it is direct that $\mathcal{F}_{(A,B)}$ is also convex.

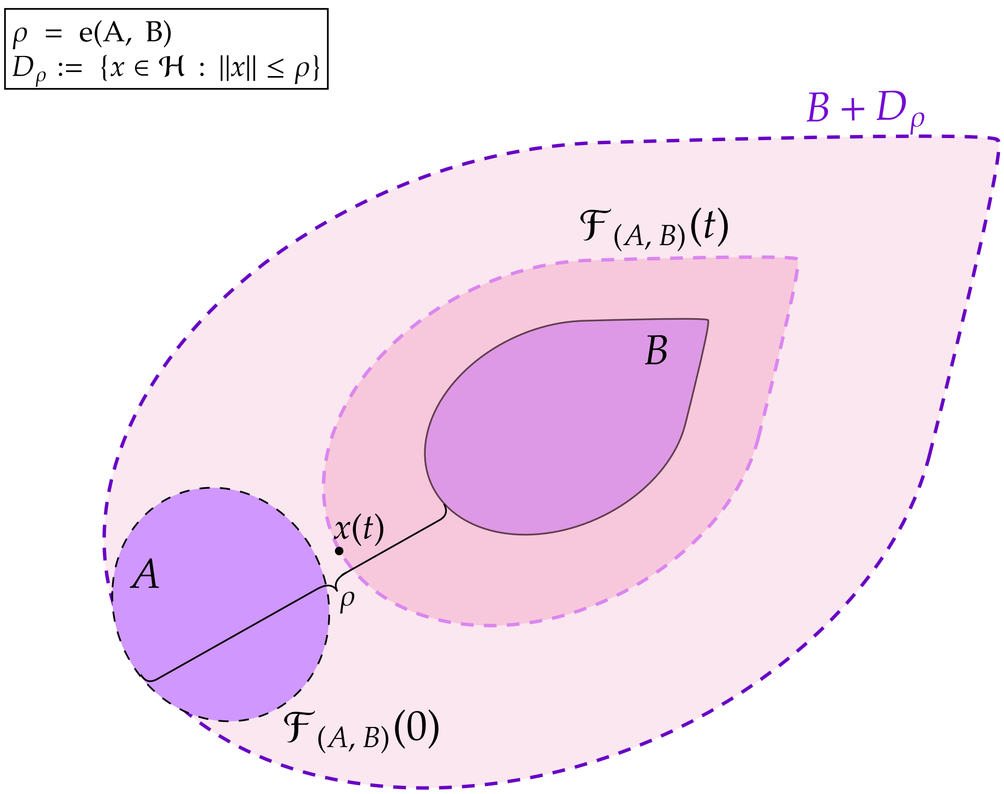

Figure 3.1: Graphic representation of the geodesic $\mathcal{F}_{(A,B)}$.

Proposition 3.1.2 *If $A, B \in \mathcal{C}_{\mathcal{H}}$ and $\mathcal{F}_{(A,B)} := \mathcal{F} \colon [0, 1] \to \mathcal{C}_{\mathcal{H}}$ is defined as in (3.4), we have*

$$e(\mathcal{F}(s), \mathcal{F}(t)) = (t - s)e(A, B) \quad \forall s, t \in [0, 1], s < t. \tag{3.5}$$

In particular, $\operatorname{Lip}(\mathcal{F}) = (t - s)$ *and we can call $\mathcal{F}$ an e-geodesic connecting A to B.*

Proof. We follow the proof of [36, Proposition 6.1].

For every $t > 0$ we have $\mathcal{F}(0) = A \subset B + D_\rho$ and since $D_\rho = D_{(1-t)\rho} + D_{t\rho}$, then $e(\mathcal{F}(0), \mathcal{F}(t)) \leq t\rho$.

If $0 < s \leq t$ we have

$$\begin{aligned}\mathcal{F}(s) &= B + D_{(1-s)\rho} \\ &= B + D_{(1-t)\rho} + D_{(t-s)\rho} \\ &= \mathcal{F}(t) + D_{(t-s)\rho},\end{aligned}$$

And therefore $e(\mathcal{F}(s), \mathcal{F}(t)) \leq (t-s)\rho = (t-s)e(A, B)$. On the other hand we have

$$e(A, B) \leq e(A, \mathcal{F}(s)) + e(\mathcal{F}(s), \mathcal{F}(t)) + e(\mathcal{F}(t), B) \leq se(A, B) + e(\mathcal{F}(s), \mathcal{F}(t)) + (1-t)e(A, B),$$

hence $(t-s)e(A, B) \leq e(\mathcal{F}(s), \mathcal{F}(t))$ and we obtain the desired result. □

The next lemma shows that there exists a unique a solution of the sweeping process driven by a geodesic $\mathcal{F}_{(A,B)}$.

Lemma 3.1.3 *Let $A, B \in \mathcal{C}_{\mathcal{H}}$ such that $\rho = e(A, B) < +\infty$, and let $\mathcal{F}_{(A,B)}\colon [0, 1] \to \mathcal{C}_{\mathcal{H}}$ defined as in (3.4). Then, the system*

$$\begin{cases} -\dot{x}(t) \in N_{\mathcal{F}_{(A,B)}(t)}(\nabla\varphi(x(t))), \quad \mathcal{L}^1 - a.e., t \in [0, 1] \\ \ x(0) = \operatorname{Proj}_{\mathcal{F}_{(A,B)}(0)}(\nabla\varphi(x_0)), \end{cases} \tag{3.6}$$

has a unique absolutely continuous solution.

Proof. It suffices to notice that thanks to Proposition 3.1.2 the moving set is Lipschitz with respect to the excess, and therefore the existence and uniqueness of the solution is guaranteed by [16, Proposition 3.2.10]. □

Remark 3.1.4 In the case where the moving set is of bounded variation, we shall use a geodesic $\mathcal{F}$ that is Lipschitz continuous with respect to the Hausdorff distance (see Definition 1.3.1), in which case we can apply the results shown in Chapter 2 to prove that the sweeping process has a unique solution when driven by $\mathcal{F}$.

3.2 Main Result

In this section we present the main theorem of this thesis, regarding the existence of solutions for $(\mathcal{P}_{BR})$. To prove this, we will use the family of geodesics defined in (3.4) to fill in the jumps of C, which will allow us to extend the solution seen in the previous section.

Theorem 3.2.1 *Let us assume that $C \in \mathrm{BR}([0, T]; \mathcal{C}_{\mathcal{H}})$, $x_0 \in \mathcal{H}$. Then there exists a unique $x \in \mathrm{BV}([0, T]; \mathcal{C}_{\mathcal{H}})$ such that there exists a Borel measure $\mu\colon \mathcal{B}([0, T]) \to [0, +\infty]$ and a function $v \in L^1(\mu; \mathcal{H})$,*

$$
\begin{cases}
\nabla\varphi(x(t)) \in C(t) \quad \forall t \in [0,T] \\
Dx = v\mu \\
-v(t) \in N_{C(t)}(\nabla\varphi(x(t))) \quad \mu - a.e., t \in \mathrm{Cont}(C) \\
\nabla\varphi(x(t)) = \mathrm{Proj}_{C(t)}(\nabla\varphi(x(t^-)), \quad \forall t \in \mathrm{Discont}(C) \setminus \{0\} \\
\nabla\varphi(x(t^+)) = \mathrm{Proj}_{C(t^+)}(\nabla\varphi(x(t))) \quad \forall t \in \mathrm{Discont}(C) \setminus \{0\} \\
\nabla\varphi(x(0)) = \mathrm{Proj}_{C(0)}(\nabla\varphi(x_0)) \\
\nabla\varphi(x(0^+)) = \mathrm{Proj}_{C(0^+)}(\nabla\varphi(x(0))
\end{cases}
\tag{$\mathcal{P}_{BR}$}
$$

Moreover, x is left continuous (respectively: right continuous) if and only if C is left continuous (respectively: right continuous).

The proof we will present in the following section is inspired by [36], where the author finds existence of solution for a nondegenerate sweeping process with bounded retraction. A similar method is used in [32, 35] to solve the nondegenerate dynamic when the moving set is of bounded variation with respect to the Hausdorff distance. In [27], the author solves the nondegenerate sweeping process where the moving set is of locally bounded variation and right continuous. In our case, we use a more general setting and do not require neither left continuity nor right continuity to prove existence of a solution.

3.2.1 Existence of the solution

In this section, we will prove the existence of a solution for the problem $(\mathcal{P}_{BR})$ of the form $\overline{S}(C, y_0) = S(\widetilde{C}) \circ \ell_C$, using the methods previously described. First, let us notice that in the intervals $[s,t]$ where ℓ_C is constant, $\ell_C^{-1}(\tau)$ is a degenerate interval for $\tau \in [s,t]$. Therefore, the following proposition presents the existence of a well-defined function $\widetilde{C}$ that addresses this issue.

Proposition 3.2.2 *Let $C \in \mathrm{BR}([0,T]; \mathcal{C}_{\mathcal{H}})$ and consider the arc length ℓ_C defined in* (1.3). *Then there exists $\widetilde{C} \in \mathrm{Lip}([0,T], \mathcal{C}_{\mathcal{H}})$ such that $C = \widetilde{C} \circ \ell_C$, where*

$$
\widetilde{C}(\tau) = \begin{cases} C(t) & \text{if } \ell_C^{-1}(\tau) \text{ is a singleton and } \ell_C^{-1}(\tau) = \{t\}, \\ C(t^+) & \text{if } \ell_C^{-1}(\tau) \text{ is not a singleton and } \inf \ell_C^{-1}(\tau) = t \end{cases} \tag{3.7}
$$

Figure 3.2 gives us an example of why we extend $\widetilde{C}$ the way that we do. The image in the left represents the arc length, and whenever there is a jump of C, the arc length remains constant, since it is defined by the retraction of the set. On the one hand, t_1 is a continuity point, and if we look at the preimage of $\ell_C(t_1)$ we can clearly see it is a singleton. Hence, for t_1 we choose $\widetilde{C}(t_1) = C(t_1)$. On the other hand, if we take any point $t \in [t_2, t_3]$, we get that the preimage $\ell_C^{-1}(t)$ is no longer a singleton, but a degenerate interval, as shown in the right image. Hence, in this case we choose $\widetilde{C}(t) = C(t_2^+)$, since $\inf \ell_C^{-1}(t) = t_2$.

With this selection, $\widetilde{C}$ is now well defined in the whole interval $[0, \mathrm{R}(C, [0,T])]$.

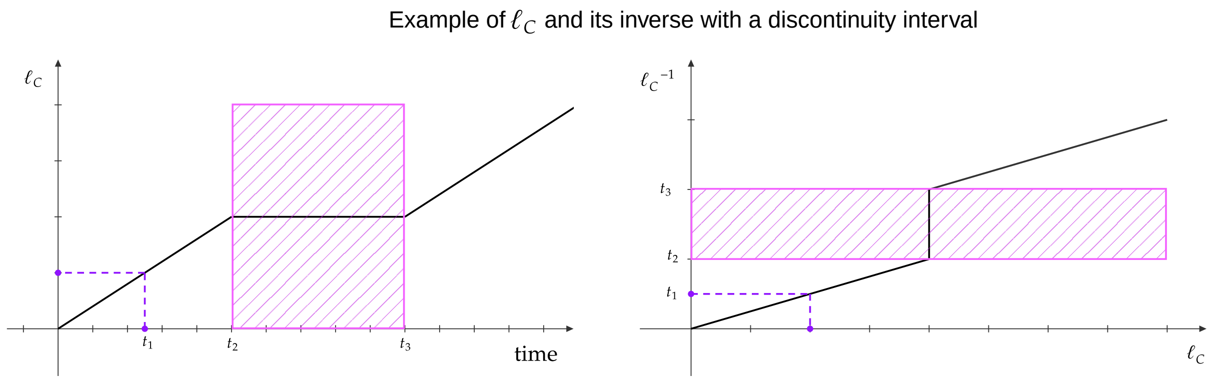

Figure 3.2: Example of ℓ_C and its inverse with a discontinuity inteval.

Proof of Proposition 3.2.2. We will prove $\widetilde{C}$ is $1-$Lipschitz continuous with respect to the excess e, as seen in [36].

Let $\sigma, \tau \in \ell_C([0,T])$ such that $\sigma < \tau$. Let us assume that $s = \inf \ell_C^{-1}(\sigma) \leq \sup \ell_C^{-1}(\sigma) = s^*$, and $t = \inf \ell_C^{-1}(\tau) \leq \sup \ell_C^{-1}(\tau) = t^*$. Then, we have three possible cases:

(i) *Case 1:* $\ell_C^{-1}(\sigma)$ and $\ell_C^{-1}(\tau)$ are both singletons, then

$$\begin{aligned} e(\widetilde{C}(\sigma), \widetilde{C}(\tau)) &= e(C(s), C(t)) \\ &\leq \mathrm{R}([s,t], C) = \tau - \sigma. \end{aligned}$$

(ii) *Case 2:* $\ell_C^{-1}(\sigma)$ is not a singleton but $\ell_C^{-1}(\tau)$ is,

$$\begin{aligned} e(\widetilde{C}(\sigma), \widetilde{C}(\tau)) &= e(C(s^+), C(t)) \\ &= e(C((s+s^*)/2), C(t)) \\ &\leq \mathrm{R}([(s+s^*)/2; t]; C) = \tau - \sigma. \end{aligned}$$

(iii) *Case 3:* $\ell_C^{-1}(\sigma)$ and $\ell_C^{-1}(\tau)$ are not singleton, then

$$\begin{aligned} e(\widetilde{C}(\sigma), \widetilde{C}(\tau)) &= e(C(s^+), C(t^+)) \\ &= e(C((s+s^*)/2), C((t+t^*)/2)) \\ &\leq \mathrm{R}([(s+s^*)/2; (t+t^*)/2]; C) = \tau - \sigma. \end{aligned}$$

And since the case where $\ell_C^{-1}(\tau)$ is not a singleton but $\ell_C^{-1}(\sigma)$ is, is analogous to case 2, we have proven that $\widetilde{C}$ is $1-$Lipschitz with respect to the excess. $\square$

Now that we found a suitable parametrization of the moving set, let us extend the definition of $\widetilde{C}$ over $[0, \mathrm{R}([0,T]; C)]$ as

$$\widetilde{C}(\sigma) := \begin{cases} \mathcal{F}_{(C(t^-),C(t))}\left(\dfrac{\sigma - \ell_C(t^-)}{\ell_C(t) - \ell_C(t^-)}\right) & \text{if } \sigma \in]\ell_C(t^-), \ell_C(t)[, \ell_C(t-) \neq \ell_C(t) \\ \mathcal{F}_{(C(t),C(t^+))}\left(\dfrac{\sigma - \ell_C(t)}{\ell_C(t^+) - \ell_C(t)}\right) & \text{if } \sigma \in]\ell_C(t), \ell_C(t^+)[, \ell_C(t) \neq \ell_C(t^+) \end{cases} \tag{3.8}$$

Moreover, $\widetilde{C}(\ell_C(t^-)) = C(t^-)$ if $\ell_C(t^-) \notin \ell_C([0,T])$ and if $\ell_C(t^-) \neq \ell_C(t)$. We notice that thanks to Proposition 3.1.2, $\widetilde{C}$ remains a 1−Lipschitz continuous function with respect to the excess, given that $\mathcal{F}_{(C(t^-),C(t))}$ and $\mathcal{F}_{(C(t),C(t^+))}$ are geodesics connecting $C(t^-)$ to $C(t)$, and $C(t)$ to $C(t^+)$, respectively. In addition,

$$C(\inf \ell_C^{-1}(\ell_C(t^-))^+) \subset C(t^-) \text{ if } \ell_C(t^-) \in \ell_C([0,T]).$$

Hence, we conclude that $S(\widetilde{C}, x_0) \in \mathrm{Lip}([0,T], \mathcal{H})$.

Now, recalling the rate independence of the system shown in Proposition 3.0.1, the clear candidate for a solution of $(\mathcal{P}_{BR})$ is of the form

$$x := S(\widetilde{C} \circ \ell_C, x_0) = S(\widetilde{C}, x_0) \circ \ell_C, \tag{3.9}$$

where S is the solution operator of $(\mathcal{P}_L)$, and x_0 the initial data.

It only remains to prove that x solves $(\mathcal{P}_{BR})$. The equalities $\nabla\varphi(x(0)) = \mathrm{Proj}_{C(0)}(\nabla\varphi(x_0))$ and $\nabla\varphi(x(0^+)) = \mathrm{Proj}_{C(0^+)}(\nabla\varphi(x(0))$ are easy to check. First, we prove that $\nabla\varphi(x(t)) \in C(t)$ for all $t \in [0,T]$. Let us denote $y := S(\widetilde{C}, x_0)$. First, let us notice that $\nabla\varphi(x(t)) \in C(t)$ whenever $\ell_C^{-1}(\ell_C(t)) = \{t\}$. In the case where $\ell_C^{-1}(\ell_C(t))$ is not a singleton, we have that $\nabla\varphi(x(t)) = \nabla\varphi(y(\ell_C(t))) \in \widetilde{C}(\ell_C(t)) = C(\inf \ell_C^{-1}(\ell_C(t))^+) \subset C(t)$, and therefore we always have that $\nabla\varphi(x(t)) \in C(t)$.

Second, we show that there exists a function $v \in L^1(\mu; \mathcal{H})$ and a measure $\mu\colon \mathcal{B}([0,T]) \to [0,+\infty]$ such that $Dx = v\mu$.

We now turn to the remainder of the equations in the system $(\mathcal{P}_{BR})$. Since we know that y is Lipschitz and ℓ_C is nondecreasing, then $x = y \circ \ell_C \in BV([0,T], \mathcal{H})$. Moreover, x is left continuous (resp. right continuous) if and only if ℓ_C is left continuous (resp. right continuous), and therefore $\mathrm{Discont}(x) = \mathrm{Discont}(\ell_C) = \mathrm{Discont}(C)$.

Next, we consider the function $v\colon [0,T] \to \mathcal{H}$ defined by

$$v(t) := \begin{cases} \dot{y}(\ell_C(t)) & \text{if } t \in \mathrm{Cont}(\ell_C), \\ \dfrac{y(\ell_C(t^+)) - y(\ell_C(t^-))}{\ell_C(t^+) - \ell_C(t^-)} & \text{if } t \in \mathrm{Discont}(\ell_C). \end{cases} \tag{3.10}$$

By virtue of (ii) in Proposition 1.6.4 and once again the fact that ℓ_C is nondecreasing, we have that $y \circ \ell_C = x \in \mathrm{BV}([0,T], \mathcal{H})$ and $Dx = vD\ell_C$. Therefore, the equality $Dx = v\mu$ holds with $\mu = D\ell_C$.

Now set

$$\mathcal{Z} := \{t \in [0,\infty[\colon\, -\dot{y}(t) \notin N_{\widetilde{C}(t)}(\nabla\varphi(y(t)))\}.$$

We will prove that $\mathcal{L}^1(Z) = 0$. Let us notice that by (i) in Proposition 1.6.4,

$$\begin{aligned} & D\ell_C(\{t \in \mathrm{Cont}(\ell_C)\colon\, -(v(t) \notin N_{C(t)}(\nabla\varphi(y(t)))\}) \\ = \;& D\ell_C(\{t \in \mathrm{Cont}(\ell_C)\colon\, -\dot{y}(\ell_C(t)) \notin N_{C(t)}(\nabla\varphi(y(t)))\}) \\ = \;& D\ell_C(\{t \in \mathrm{Cont}(\ell_C)\colon \ell_C(t) \in Z\}) \\ = \;& \mathcal{L}^1(Z) = 0, \end{aligned}$$

where in the last step we have used that, since y is a solution of $(\mathcal{P}_L)$, we have that $\mathcal{L}^1(Z) = 0$. This shows that $-v(t) \in N_{C(t)}(\nabla\varphi(x))$ holds with $\mu = D\ell_C$.

Finally, let us fix $t \in \mathrm{Discont}(\ell_C)$. We observe that by (3.8), if $\sigma \in]\ell_C(t^-), \ell_C(t)[$,

$$\widetilde{C}(\sigma) = \mathcal{F}_{(C(t^-),C(t))}\left(\frac{\sigma - \ell_C(t-)}{\ell_C(t) - \ell_C(t^-)}\right).$$

Then, since $(\mathcal{P}_L)$ has a unique solution, the solution operator has the semigroup property:

$$S(C, y_0)(t) = S(C(\cdot + s), S(C, y_0)(s))(t - s) \qquad \forall t, s, \quad 0 \leq s < t.$$

And therefore, using this property, we have

$$\begin{aligned}
\nabla\varphi(x(t)) &= \nabla\varphi(y(\ell_C(t))) \\
&= \nabla\varphi(S(\widetilde{C}, y_0)(\ell_C(t))) \\
&= \nabla\varphi(S(\widetilde{C}(\cdot + \ell_C(t^-)), S(\widetilde{C}, y_0)(\ell_C(t^-)))(\ell_C(t) - \ell_C(t^-))) \\
&= \nabla\varphi(S(\mathcal{F}_{(C(t^-),C(t))}, y(\ell_C(t^-)))(1)) \\
&= \mathrm{Proj}_{C(t)}(\nabla\varphi(x(t-))).
\end{aligned}$$

On the other hand, if $\sigma \in]\ell_C(t), \ell_C(t+)[$, we have that

$$\widetilde{C}(\sigma) = \mathcal{F}_{(C(t),C(t^+))}\left(\frac{\sigma - \ell_C(t)}{\ell_C(t^+) - \ell_C(t)}\right).$$

Analogously to the previous case,

$$\begin{aligned}
\nabla\varphi(x(t^+)) &= \nabla\varphi(y(\ell_C(t^+))) \\
&= \nabla\varphi(S(\widetilde{C}, y_0)(\ell_C(t))) \\
&= \nabla\varphi(S(\widetilde{C}(\cdot + \ell_C(t)), S(\widetilde{C}, y_0)(\ell_C(t)))(\ell_C(t^+) - \ell_C(t))) \\
&= \nabla\varphi((S(\mathcal{F}_{(C(t),C(t^+))}, y(\ell_C(t)))(1)) \\
&= \mathrm{Proj}_{C(t^+)}(\nabla\varphi(x(t))).
\end{aligned}$$

This concludes that $\nabla\varphi(x(t)) = \mathrm{Proj}_{C(t)}(\nabla\varphi(x(t^-))$ and that, for all $t \in \mathrm{Discont}(C) \setminus \{0\}$, $\nabla\varphi(x(t^+)) = \mathrm{Proj}_{C(t^+)}(\nabla\varphi(x(t)))$, which completes the proof.

Conclusion

This thesis has dealt with the study of a degenerate sweeping process in two different settings, using convex analysis and nonsmooth optimization techniques.

In Chapter 2, we study the so called Mirror Sweeping Process in the case where the moving set is convex and Lipschitz continuous. By means of the Moreau-Yosida regularization method, we find approximated solutions over sub-intervals, that are then proven to converge uniformly to a solution of the Mirror Sweeping Process. This solution is absolutely continuous and, moreover, is proven to be unique. To prove convergence, we drop the usual hypotheses of separability of the Hilbert space and compactness of the moving set, which, to the best of our knowledge, is new in degenerate sweeping processes with nonlinear operators.

In Chapter 3 we find a solution for the Mirror Sweeping Process in the case where the moving set is convex and with bounded retraction. In this case we parametrize the set by means of the arc length, and fill in the discontinuity jumps with a family of convex geodesics that move in a Lipschitzian way with respect to the excess, reducing our problem to one that we already know it has a solution, as seen in [16] via a catching-up-like algorithm.

There are still many open questions regarding the Mirror Sweeping Process, the first one being the extension of this work onto a moving set with a prox-regular setting, for both the Lipschitz and the bounded retraction case. The latter case can be explored adding a perturbation and adapting the geodesic accordingly, while the former case can be recreated in the case where the set moves in a Lipschitzian way, but with respect to the excess instead of the Hausdorff distance. This way, the existence of solutions for the sweeping process driven by the geodesics in Chapter 3 could be proven without the need of the catching-up algorithm.

Last, but not least, the Mirror Sweeping Process was born from the need of bridging the worlds of differential inclusions and online optimization, and therefore solving real-life applications of online optimization via the Mirror Sweeping Process is something we are one step closer to, and can be exploited in different scenarios, as proven throughout this thesis.

Bibliography

[1] V. Acary, O. Bonnefon, and B. Brogliato. *Nonsmooth modeling and simulation for switched circuits*, volume 69 of *Lecture Notes in Electrical Engineering*. Springer, Dordrecht, 2011.

[2] 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.

[3] H.-H. Bauschke and P. L. 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.

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

[5] J. M. Borwein and A. S. Lewis. *Convex analysis and nonlinear optimization*, volume 3 of *CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC*. Springer, New York, second edition, 2006. Theory and examples.

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

[7] M. Bounkhel. *Regularity concepts in nonsmooth analysis*, volume 59 of *Springer Optimization and Its Applications*. Springer, New York, 2012. Theory and applications.

[8] L.M. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Computational Mathematics and Mathematical Physics*, 7(3):200–217, 1967.

[9] B. Brogliato. *Nonsmooth mechanics*. Communications and Control Engineering Series. Springer, [Cham], third edition, 2016. Models, dynamics and control.

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

[11] S. Bubeck, M.B. Cohen, Y. T. Lee, J. R. Lee, and A. Madry. 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.

[12] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski. *Nonsmooth analysis and control theory*, volume 178 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

[13] 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.

[14] N. Dinculeanu. *Vector Measures*. Hochschulbücher für Mathematik. Pergamon Press, 1966.

[15] I. Ekeland and R. Témam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999. Translated from the French.

[16] A. Gutiérrez. A Numerical Algorithm for Mirror Sweeping Processes. Master's thesis, Universidad de Chile, Chile, 2023.

[17] 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.

[18] 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.

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

[20] 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.

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

[22] B. Maury and J. Venel. Un modèle de mouvements de foule. In *Paris-Sud Working Group on Modelling and Scientific Computing 2006–2007*, volume 18 of *ESAIM Proc.*, pages 143–152. EDP Sci., Les Ulis, 2007.

[23] 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.

[24] J. J. Moreau. Rafle par un convexe variable. II. In *Travaux du Séminaire d'Analyse Convexe, Vol. II*, Secrétariat des Mathématiques, Publication, No. 122, pages Exp. No. 3, 36. Univ. Sci. Tech. Languedoc, Montpellier, 1972.

[25] 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.

[26] 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.

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

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

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

[30] 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.

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

[32] V. Recupero. *BV* solutions of rate independent variational inequalities. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 10(2):269–315, 2011.

[33] V. Recupero. A continuity method for sweeping processes. *J. Differential Equations*, 251(8):2125–2142, 2011.

[34] V. Recupero. *BV* continuous sweeping processes. *J. Differential Equations*, 259(8):4253–4272, 2015.

[35] V. Recupero. Sweeping processes and rate independence. *J. Convex Anal.*, 23(3):921–946, 2016.

[36] V. Recupero. Convex valued geodesics and applications to sweeping processes with bounded retraction. *J. Convex Anal.*, 27(2):537–558, 2020.

[37] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series. Princeton University Press, Princeton, N. J., 1970.

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

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

[40] E. Vilches. Regularization of perturbed state-dependent sweeping processes with nonregular sets. *J. Nonlinear Convex Anal.*, 19(4):633–651, 2018.