INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2022IPPAE016

Thèse de doctorat

ENSTA

IP PARIS

Viscosity theory of first order Hamilton Jacobi equations in some metric spaces

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'Ecole Nationale Supérieure de Techniques Avancées

École doctorale n°574 Mathematiques Hadamard (EDMH)
Spécialité de doctorat: mathématiques appliquées

Thèse présentée et soutenue à Palaiseau, France, le 15/12/2022, par

Othmane Jerhaoui

Composition du Jury :

Hasnaa Zidani Professeure, INSA Rouen Normandie, France.	Directrice de thèse
Ludovic Rifford Professeur des universités, Université Côte d'Azur, France.	Rapporteur
Fabio Camilli Professeur, Sapienza, Università di Roma, Italie.	Rapporteur
Stéphane Gaubert Directeur de recherche, INRIA Saclay, France.	Examinateur
Emmanuel Chasseigne Maître de conférences, Université de Tours, France.	Examinateur
Marc Quincampoix Professeur des universités, Université de Bretagne Occidentale, France.	Examinateur
Yves Achdou Professeur des universités, Université Paris-Diderot, France.	Président du jury
Frédéric Jean Professeur, ENSTA Paris, France.	Invité

Acknowledgments

> "90 % of success is just showing up!"
>
> Joseph James Rogan, The experience.

This is for me the hardest part I had to write in this manuscript...

The present manuscript is the culmination of four years of hard work and collaborative effort. It is then extremely important for me to thank the people who contributed to this work, both through the many exchanges we had and through the emotional support I received from them.

First, my deepest gratitude goes to my advisor Professor Hasnaa Zidani. She is in every way imaginable my mentor and my "mathematical godmother". I remember distinctively our first encounter, exactly four years ago, during a course she was giving about the Hamilton-Jacobi approach in optimal control problems. I was instantly hooked on the concept and wanted to learn more about it. I asked Hasnaa for an internship and a potential Ph.D. opportunity, and she gave me four topics to choose from! I am thankful for her valuable insight and the numerous (and sometimes heated) discussions regarding the content and the presentation of this manuscript and the papers submitted. I will forever be grateful for her hospitality, her generosity, her patience with me, her kindhearted nature and for believing in me in every step of the way.

I would also like to thank my co-director, Professor Frédéric Jean, a.k.a the director of the lab! My first encounter with Frédéric was during a course he was giving on geometric control. He promptly accepted to join us in this adventure during the second year of my Ph.D. I am thankful for his extremely relevant advices regarding the presentation of this manuscript. It is truly a pleasure and an honor to work with him.

I would like to thank Professor Fabio Camilli and Professor Ludovic Rifford, the two reviewers of this dissertation, who have been so kind to spend time to read and evaluate my works and to give me some valuable comments and encouragements to continue this work in the future. I am also thankful to Professors Stéphane Gaubert, Yves Achdou, Marc Quincampoix and Emmanuel Chasseigne who accepted to be in my thesis committee.

I am immensely grateful to Professors Paolo Vannucci and Laurent Dumas for endorsing my application to transition from a master of mechanics to a master of mathematics and for obtaining the Sophie Germain scholarship during my last master year. A special thanks goes to Professor Olivier Bokanowski, for helping me during the beginning of this adventure, at the internship level.

I would like to thank all the members of the UMA lab at ENSTA Paris, where I spent my internship and my Ph.D. thesis. Specifically, I would like to thank Corinne Chen for her help regarding all the paperwork and the organization of the Ph.D. student seminar. I would also like to thank all the researchers and the administrators of the lab for their hospitality and the numerous meals shared together.

I would like to thank all the members of the LMI lab at INSA Rouen Normandie for their hospitality during my numerous visits to their facilities. In particular, I would like to thank Brigitte Diara for facilitating much of the administrative procedures to me.

I would like to thank Professor Cristopher Hermosilla and all the members of the department of mathematics at University Técnica Federico Santa Marìa in Valparaìso, Chile for their warm welcoming and hospitality during my short stay there. I hope that more collaborations with our Chilean partners will come to fruition in the near future.

A special thanks must go to all my friends and partners in crime in the labs of Rouen and Palaiseau for the great atmosphere. A special tribute goes to Mahran for his precious help regarding LaTeX, Jean-François for his kind soul, Damien for all the interesting theological discussions, the newcomer Thibaut for his sense of humor, the Chilean squad Natalia, Cristian, Piero and Fernanda, Quentin and Alice for co-organising the Ph.D. student seminar with me, Louise for the heart-to-heart conversations we had and the multiple rides she kindly offered me, and finally Averil, the newcomer in Rouen, for her corrections regarding the last chapter of this manuscript.

Last but not least, I would like to thank my family for their constant support. I would like to thank my parents for their encouragements, my brother Ziko, for his unique artistic spirit, my auntie Khalti Mina for providing me with a safe refuge whenever I felt lost, my auntie Aamti Karima and her beautiful children for the sincere love and support and my dear friend Nhi for her constant help during these four incredible years. Finally, I would like to thank my muse, Kaoutar, for putting up with me during this adventure and hopefully for others to come.

Synthèse (en français)

Le sujet principal de cette thèse est l'étude des équations de Hamilton Jacobi posées sur certains espaces métriques. De plus, le Hamiltonien de ces équations pourrait présenter certaines discontinuités bien structurées.

La première partie de cette thèse est consacrée à l'étude d'une équation de Hamilton Jacobi Bellman discontinue, définie sur une stratification de $\mathbb{R}^N$. Cette dernière est le résultat d'une union d'une collection finie de sous-variétés lisses et disjointes de $\mathbb{R}^N$, que l'on nomme *les sous-domaines*. Sur chaque sous-domaine, un Hamiltonien continu y est défini. Cependant, le Hamiltonien global sur $\mathbb{R}^N$ présente des discontinuités lorsque l'on passe d'un sous-domaine à l'autre. On donne une interprétation commande optimale de ce problème et on utilise les techniques de l'analyse non lisse pour montrer que la fonction valeur est l'unique solution de viscosité de l'équation de Hamilton Jacobi Bellman définie dans ce chapitre. L'unicité de la solution est garantie par un principe de comparaison fort, valable pour toute sur-solution semicontinue inférieurement et toute sous-solution semicontinue supérieurement. En ce qui concerne l'éxistence de la solution, on utilise le principe de la programmation dynamique vérifiée par la fonction valeur pour montrer que cette dernière est une solution de viscosité du problème considéré. De plus, on prouve quelques résultats de stabilité en présence de perturbations sur le Hamiltonien discontinu. Finalement, en vertu du principe de comparaison, on montre un résultat de convergence général pour les schémas numériques monotones qui approchent ce problème.

La deuxième partie de cette thèse est consacrée au dévelopement d'une nouvelle notion de viscosité pour les équations de Hamilton Jacobi du premier ordre définies sur les espaces CAT(0) propres. Un espace métrique est dit CAT(0), s'il est un espace géodésique et si ses triangles géodésiques sont plus "minces" que les triangles du plan Euclidien. Les espaces CAT(0) peuvent être considérés comme une généralisation des espaces de Hilbert ou les variétés de Hadamad. Des exemples types des espaces CAT(0) sont les espaces de Hilbert, les arbres métriques et les networks obtenus en collant un nombre fini de demi-espaces selon leur frontière commune. On exploite la strucutre de ces espaces pour étudier les equations de Hamilton Jacobi du premier ordre stationnaires et dépendantes du temps. En particulier, le but du chapitre est de retrouver les principaux résultats de la théorie de la viscosité : *le principe de comparaison et la méthode de Perron*. On définit la notion de viscosité en utilisant des fonctions test qui sont Lipschitz et qui peuvent être représentées comme une différence de deux fonctions semiconvexes. On montre que cette notion de viscosité coïncide avec la notion classique dévelopée sur $\mathbb{R}^N$ en étudiant quelques exemples d'équations classiques. De surcroît, on prouve l'existence et l'unicité de la solution de certaines équations du type Eikonal posées sur des networks qui peuvent résulter du collage de demi-espaces ayant différentes dimensions de Hausdorff.

La troisième partie de la thèse se focalise sur l'étude d'un problème de commande optimale de Mayer sur l'espace des mesures Boréliennes de probabilité sur une variété compacte M. L'étude de ce problème est motivé par certaines situations où un planificateur central d'un système contrôlé n'a qu'une information imparfaite sur l'état initial du système considéré. Le manque d'information est spécifique dans ce problème. Il est décrit par une mesure de probabilité Borélienne selon laquelle l'état initial est distribué. On définit la notion de viscosité sur cet espaces de la même manière que dans la deuxième partie de la thèse en considérant des fonctions test qui sont Lipschitz et qui peuvent être représentées par une différence de deux fonctions semiconvexes. Avec ce choix de fonctions test, on étend la notion de viscosité aux équations de Hamilton Jacobi Bellman définies sur l'espace de Wasserstein et on établit que la fonction valeur associée au problème de commande optimale et l'unique solution de viscosité sur l'espace de Wasserstein sur M.

Contents

Chapter 1

Introduction

This thesis is concerned with giving a new insight in the study of first order Hamilton Jacobi equations in certain classes of metric spaces, potentially in the presence of structured discontinuities on the Hamiltonian.

The study of nonlinear partial differential equations has led to many new innovative approaches, offering an interesting insight on how to solve them. An important class of nonlinear equations is the class of first order Hamilton Jacobi equations. Hamilton Jacobi equations were extensively studied in the literature when posed in the Euclidean space. In full generality, the equation has the following form

$$H(x, v(x), D_x v) = 0, \quad x \in \mathcal{O}, \tag{1.1}$$

where $\mathcal{O}$ is a subdomain of $\mathbb{R}^N$, $H : \mathcal{O} \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ is called the *Hamiltonian* and $v : \mathcal{O} \to \mathbb{R}$ is a continuously differentiable function. In general, equation (1.1) fails to have smooth solutions on a given domain. The most obvious way to circumvent this problem is by relaxing the continuous differentiability requirement. Therefore, we introduce an appropriate notion of generalized solutions, meaning solutions that verify equation (1.1) in a certain weak sense. The present manuscript is concerned with a contribution to the important notion of *viscosity solutions*. This notion was introduced in the late 1970's and the 1980's by Crandall and Lions in the papers [1, 2]. Under mild assumptions on the Hamiltonian, the definition of viscosity solutions is the following.

Definition 1.1.

- We say that an upper semicontinuous function $v : \mathcal{O} \to \mathbb{R}$ is a *viscosity subsolution* to equation (1.1) at a point $x \in \mathcal{O}$ if for any continuously differentiable function $\phi : \mathcal{O} \to \mathbb{R}$ such that $v - \phi$ attains a local maximum at x, we have

$$H(x, v(x), D_x \phi) \leq 0. \tag{1.2}$$

- We say that a lower semicontinuous function $v : \mathcal{O} \to \mathbb{R}$ is a *viscosity supersolution* to equation (1.1) at a point $x \in \mathcal{O}$ if for any continuously differentiable function $\phi : \mathcal{O} \to \mathbb{R}$ such that $v - \phi$ attains a local minimum at x, we have

$$H(x, v(x), D_x\phi) \geq 0. \tag{1.3}$$

- A continuous function $v : \mathcal{O} \to \mathbb{R}$ is said to be a *viscosity solution* to equation (1.1) if it is both a viscosity subsolution and a viscosity supersolution at every point of $\mathcal{O}$.

In the theory of linear partial differential equations, we move the derivative to the test functions using integration by part. In the present setting the derivatives are moved using the maximum principle. The viscosity notion has been extensively studied and refined in the literature by many authors. We would like to refer to Crandall and Lions [1, 2], Crandall, Ishii and Lions [3], Barles [4], Fleming and Soner [5] and Bardi and Capuzzo-Dolcetta [6] among various manuscripts on this topic.

The theory of viscosity was developed initially to find continuous solutions of Hamilton Jacobi equations of the form (1.1). However, it was extended to cover discontinuous frameworks by Ishii in [7] by replacing the Hamiltonian H in inequality (1.2) by its lower semicontinuous envelope and by replacing the Hamiltonian H in inequality (1.3) by its upper semicontinuous envelope. A different notion of discontinuous viscosity solution was introduced by Barron and Jensen [8], known as the *bilateral viscosity solution*.

Another point of view for defining the notion of viscosity solution is given by *nonsmooth analysis* theory. Nonsmooth analysis studies relaxed notions of differentiation when the classical notion of continuous differentiability might not be well defined due the lack of regularity of the functions considered. Many notions of generalized derivatives are developed (Dini, proximal...) and result in various concepts of set-valued operators that allow to give a definition of weak solutions, equivalent to Definition 1.1, of first order Hamilton Jacobi equations of the form (1.1). We refer to the authors Clarke [9, 10], Aubin and Frankowska [11], Bardi and Capuzzo-Dolcetta [6] and Vinter [12] for more details on nonsmooth analysis and the treatment of Hamilton Jacobi equations through these tools. As far as the discontinuous framework is concerned, Frankowska used nonsmooth analysis techniques in [13] to show that the bilateral viscosity solution is intrinsically related to some geometric properties of the discontinuous solution.

Nonsmooth analysis techniques provide a relevant framework to study a special class of Hamilton Jacobi equations coming from optimal control theory. The latter has its roots in the calculus of variations, first started in the 17th century. A standard optimal control problem refers to the problem of finding a control time function that minimizes a certain performance criterion, under the constraint of a parametrized differential equation called *the controlled system*. The mathematical formulation of an optimal control problem is the following. We consider the following controlled system on $\mathbb{R}^N$:

$$\begin{cases} \dot{y}(s) = f(y(s), u(s)), \ a.e. \ \ s \in [t, T], \\ y(t) = x, \end{cases} \tag{1.4}$$

where $T > 0$ is the *final time*, $f : \mathbb{R}^N \times U \to \mathbb{R}^N$ is *the dynamics* or *velocity* of the system, $x \in \mathbb{R}^N$ is the *initial position or initial state* and $t \in [0, T]$ is the *initial time*. The set U is the set of admissible control values which is assumed to be a compact subset of some metric space. *The control function* $u : [t, T] \to U$ is a Borel measurable function. Under mild assumptions on the dynamics, for any measurable control function $u(.)$, the controlled system (1.4) admits a unique absolutely continuous solution. Another point of view which is useful when nonsmooth analysis techniques are used considers equation (1.4) as a *differential inclusion*:

$$\begin{cases} \dot{y}(s) \in F(y(s)), \ a.e. \ \ s \in [t, T], \\ y(t) = x. \end{cases}$$

The set-valued map $F : \mathbb{R}^N \rightsquigarrow \mathbb{R}^N$ is still called *the set of velocities* or *dynamics* of the system. The relation between f and F can be easily deduced in the following way:

$$F(.) := \{f(., u) \ : \ u \in U\}.$$

To emphasize the dependency on the initial data of trajectories solution to equation (1.4), we sometimes denote them by $y_{(t,x)}(.)$. We consider the following optimal control problem with the final cost $\ell : \mathbb{R}^N \to \mathbb{R}$:

$$\inf \{\ell(y_{(t,x)}(T)) \ : \ y_{(t,x)}(.) \text{ is a solution of equation (1.4)}\}. \tag{1.5}$$

In the above optimal control problem, known as the *Mayer problem*, the performence criterion to optimize depends only on the final cost $\ell(.)$. It represents the price to pay to arrive to the final state $y_{(t,x)}(T)$. We will focus solely on this formulation in parts of this manuscript.

Among many approaches to study the above optimal control problem, we will focus on ones involving Hamilton Jacobi equations, which are based on the use of the *dynamic programming principle*, formulated first by Bellman in the 1950's. The dynamic programming principle considers the mapping that associates the initial data (t, x) to the optimal value of the optimization problem (1.5). We denote this mapping by $\vartheta(t, x)$ and we call it the *value function*. More precisely,

$$\vartheta(t, x) := \inf \{\ell(y_{(t,x)}(T)) \ : \ y_{(t,x)}(.) \text{ is a solution of equation (1.4)}\}.$$

The value function satisfies the following functional equation. for any $h \in [t, T - t]$ we have

$$\vartheta(t, x) = \inf \{\vartheta(t + h, y_{(t,x)}(t + h)) \ : \ y_{(t,x)}(.) \text{ is a solution of equation (1.4)}\},$$

which is precisely the dynamic programming principle. It allows to break the optimization problem (1.5) into sub-problems that can be solved in a recursive way. From the dynamic programming principle, if the value function $(t,x) \mapsto \vartheta(t,x)$ is continuously differentiable, then it is the solution of the following *Hamilton Jacobi Bellman* equation

$$-\partial_t v + H(x, D_x v) = 0, \quad (t,x) \in \mathcal{O} = (0,T) \times \mathbb{R}^N, \tag{1.6}$$

satisfying the boundary condition $v(T,x) = \ell(x)$, where the function H is the *Bellman Hamiltonian* defined by

$$H(x,p) = \sup_{u \in U} \{-\langle p, f(x,u) \rangle\}.$$

In general ϑ is not continuously differentiable. However, viscosity theory (or tools of nonsmooth analysis) asserts that if the Bellman Hamiltonian is continuous, then the value function is the unique solution to equation (1.6) in the sense of Definition 1.1 when it is merely continuous.

To solve optimal control problems via the Hamilton Jacobi approach, the regularity of the value function and the Hamiltonian play a very important role. In particular, they both need to be continuous. If the dynamics f is Lipschitz continuous with respect to the state variable x and continuous with respect to the control variable u, then the Bellman Hamiltonian in equation (1.6) is Lipschitz continuous. If furthermore, the final cost ℓ is Lipschitz continuous, then the value function is continuous. When the value function is not continuous, one can use the notion of bilateral viscosity solution to deal with this case. However, when the Hamiltonian H is not continuous, either due to discontinuities in the dynamics f in the case of a Bellman Hamiltonian, or more generally when the Hamiltonian in equation (1.1) is not continuous, then the problem of finding a notion of weak solution that guarantees existence and uniqueness of a solution of equations (1.1) or (1.6) is much more complicated.

There has been a growing interest in studying Hamilton Jacobi equations with discontinuous Hamiltonians both from the mathematical point of view or the potential real-world applications. In the last decades, particular Hamilton Jacobi Bellman equations have been considered with *structured discontinuities* on the Hamiltonian. Bressan and Hong [14] introduced a class of control problems known as stratified domain control problems where the state space $\mathbb{R}^N$ is decomposed into a finite collection of submanifolds, each of which has its own Lipschitz continuous dynamics and its associated continuous Bellman Hamiltonian. However, while on each submanifold the Hamiltonian is continuous, the global Hamiltonian defined on the whole space presents discontinuities once one goes from one submanifold to another. This problem was investigated by several authors in the literature. See for example Barles, Briani and Chasseigne in [15, 16], Barles and Chasseigne in [17], Rao and Zidani in [18] and Rao, Siconolfi and Zidani in [19]. Despite the fact that the discontinuities of the Hamiltonian are structured, the characterization of the value function of the optimal control problem as the unique viscosity solution to an associated Hamilton Jacobi Bellman equation of the form (1.6) presents several difficulties.

Another interesting case of discontinuous Hamilton Jacobi equations comes from modelling problems related to traffic flow. In this problem, the state space $\mathcal{O}$ is a network consisting of a finite collection of half-spaces, of the same dimension, glued along their common boundary. The simplest example of this setting is a one dimensional network resulting from a finite collection of half-lines glued along their origin point called the *junction*. On each half-line a continuous Hamiltonian is considered, which contains the information related to the flow of the traffic. However, the Hamiltonian on the whole network is discontinuous at the junction point.

Two main strategies exist in the literature to define an appropriate notion of viscosity that encompasses the singular nature of this problem. The first one consists in exploiting the piecewise differential structure of the network to define a viscosity notion by considering test functions that are continuously differentiable on each half-line. See Schieborn in [20], Camilli, Marchi and Schieborn in [21], Imbert and Monneau [22, 23], Imbert, Monneau and Zidani [24], Achdou, Camilli, Cutrì and Tchou [25], Achdou and Tchou [26] and Lions and Souganidis [27] for a detailed discussion on this line of thought. The second approach takes a considerable conceptual jump by considering the network as a metric space and aims at developing a theory of viscosity solutions in a more general class of metric spaces, that includes the case of the network.

The theory of viscosity was developed in more general spaces shortly after its first appearance in the Euclidean space. It was first extended to Banach spaces by Crandall and Lions in the series of papers [28, 29, 30, 31]. As far as metric spaces are concerned, the papers by Giga, Hamamuki and Nakayasu [32], Gangbo and Święch [33] and Ambrosio and Feng [34] treat special cases of Hamilton Jacobi equations defined in general metric spaces. The question of defining and treating more general Hamilton Jacobi equations in a general metric space remains widely open in the current state of the literature.

Another particular class of metric spaces that gained a substantial amount of attention in the last few years is the space of probability measures over a base space, typically over the Euclidean space or a Riemannian manifold. Hamilton Jacobi equations in the space of probability measures arise in the modelling of systems consisting of a large number of interacting agents or particles in motion, considered to be indistinguishable from one another. If the total number of particles stays constant at all times, then a convenient way to model this problem is by considering the number of particules to be infinitely large and by looking at the system as one normalized density evolving through time. The space of probability measures in this setting is often endowed with the *Wasserstein distance* coming from optimal transport theory [35]. The various viscosity notions developed in [32, 33, 34] for Hamilton Jacobi equations defined in a general metric space do not cover a wide range of Hamilton Jacobi equations defined in the space of probability measures. This led to a further investigation to define a suitable notion of viscosity solution in this space.

When the space of probability measures is equipped with the Wasserstein distance, then, roughly speaking, it has a structure resembling that of an infinite dimensional Riemannian manifold. This fact has led to the development of various notions of viscosity relying on both nonsmooth analysis tools and viscosity techniques. See Marigonda and Quincampoix [36], Marigonda and Cardaliaguet [37], Jimenez, Marigonda and Quincampoix [38], Cardaliaguet, Delarue, Lasry and Lions in [39] and Gangbo and Tudorascu [40]. Yet, viscosity theory in the space of probability measures is still a very active area of research. Many challenging questions regarding well posedness of Hamilton Jacobi equations defined in this space are still open.

In this thesis, we give a new insight on these challenging problems by extending the notion of viscosity solutions to more general classes of metric spaces with Hamiltonians that can potentially present some structured discontinuities. The thesis is organized into three independent chapters. In Chapter 2, we study well-posedness of Hamilton Jacobi equations coming from stratified domain optimal control problems. We define a new notion of viscosity that encodes the discontinuous nature of the Hamiltonian and we prove that the value function is the unique viscosity solution of the problem using mainly tools of nonsmooth analysis. In Chapter 3, we develop a new viscosity notion for first order Hamilton Jacobi equations defined in a class of metric spaces called spaces of curvature not greater than 0 in the sense of Alexandrov. This class of metric spaces includes Euclidean spaces and networks that can be resulting from gluing half-spaces of different Hausdorff dimension. Chapter 4 aims at introducing a new notion of viscosity in the space of Borel probability measures over a compact Riemannian manifold by studying a Hamilton Jacobi Bellman equation associated to an optimal control problem posed in this space. In the remainder of this introduction, we give a summary of the main contributions presented in the chapters of this thesis.

Chapter 2 : A general comparison principle for Hamilton Jacobi Bellman equations in stratified domains

In this chapter, we are concerned with a Hamilton Jacobi Bellman equation coming from a stratified optimal control problem in $\mathbb{R}^N$, introduced by Bressan and Hong in [14]. We consider n open sets $\mathcal{M}_i$ of $\mathbb{R}^N$ such that

$$\mathbb{R}^N = \cup_{i=1}^n \overline{\mathcal{M}}_i, \quad \mathcal{M}_i \cap \mathcal{M}_j = \emptyset, \text{ for } i \neq j.$$

We suppose that the interface $\Lambda := \mathbb{R}^N \setminus \cup_{i=1}^n \mathcal{M}_i$ separating the open sets $\mathcal{M}_i$ is in the form of a finite and disjoint union of lower dimensional embedded submanifolds.

On each $\mathcal{M}_i$, we define a continuous Bellman Hamilltonian H_i and we consider the following Hamilton Jacobi equation

$$\begin{cases} -\partial_t v(t,x) + H_i(x, \partial_x v(t,x)) = 0, & \text{for } (t,x) \in (0,T) \times \mathcal{M}_i, \\ v(T,x) = \psi(x), \end{cases} \tag{1.7}$$

where $\psi : \mathbb{R}^N \to \mathbb{R}$ is the final cost supposed to be Lipschitz and bounded. Note that we choose to adopt the notation $\partial_x v$ instead of $D_x v$ in this chapter since we are going to use mainly nonsmooth analysis techniques. The Hamiltonians H_i are of the form

$$H_i(x,p) = \sup_{\nu \in F_i(x)} \{-\langle p, \nu \rangle\},$$

where $F_i(.)$ is a set-valued map representing the set of velocities on $\mathcal{M}_i$. Due to the singular nature of our problem, one cannot expect to find a continuous Hamiltonian H defined in all $\mathbb{R}^N$, such that its restriction to each $\mathcal{M}_i$ coincides with H_i. The goal of this chapter is to define the Hamiltonian H on the interface Λ in order to guarantee well-posedness of the Hamilton Jacobi equation (1.7) in a suitable viscosity sense that takes into account the discontinuous nature of H.

When the Hamiltonian H is continuous, nonsmooth analysis tools provide a geometric interpretation of the definition of viscosity supersolutions and subsolutions, known as *weak and strong invariance properties*. It states that a function v is a viscosity supersolution if and only only if there exists at least one trajectory, solution of the controlled system associated to H, starting from a point of the *epigraph* of v that stays confined in the epigraph. In this case, we say that the epigraph of v is *weakly invariant*. Similarly, a function v is a viscosity subsolution if and only if every trajectory that starts from a point in the *hypograph* of v stays confined in the hypograph. We say that the hypograph of v is *strongly invariant*. The invariance principles are powerful nonsmooth analysis techniques that allow to deduce a key result in viscosity theory called *the comparison principle*, which states that any upper semicontinuous subsolution must lie below any lower semicontinuous supersolution if their boundary conditions do. Comparison results guarantee uniqueness of the viscosity solution if it exists.

For the stratified case, the Hamiltonian is necessarily discontinuous and the picture is less clear. Barnard and Wolenski tried to fill the gap in [41] by studying weak and strong invariance properties in the stratified setting. However, their statement regarding the strong invariance property needed further investigation. This is due to the fact that their choice regarding the sub/super- differentials, or equivalently the choice of the test functions, in their definition of viscosity solution, did not take into account the singular nature of the problem (see Chapter 2 for a detailed discussion on this fact). In this chapter, we will prove the invariance principles in the present stratified setting.

Defining a good notion of viscosity when the Hamiltonian H is discontinuous is a difficult task. As mentioned above, one of the key results that the new notion of viscosity needs to provide is a comparison principle between any upper semicontinuous subsolution and any lower semicontinuous supersolution. A comparison principle was obtained by Camilli and Siconolfi in [42] in the case where the Hamiltonian H is measurable with respect to the state variable, under a restrictive assumption called the "transversality condition" on the Hamiltonian. The latter implies that the behavior of the Hamiltonian on the interface could be ignored. Barles and Chasseigne considered a more general setting than the one considered in the present chapter [17]. The authors used the Ishii's extension on the interface and provided a very thorough analysis on the notion of viscosity solution of a stratified system. Furthermore, different conditions under which the comparison principle is satisfied have been investigated in their monograph [43]. Rao and Zidani [18] studied the same stratified domain optimal control problem as in the present setting. They used nonsmooth analysis techniques to prove a comparison principle that holds for any lower semicontinuous supersolution and any upper somicontinuous subsolution that is *Lipschitz continuous* on the interface. Furthermore, in order to define the Hamiltonian H on the interface, they gave an optimal control interpretation of the above Hamilton Jacobi equation (1.7) by considering the notion of *essential dynamics* taken from [41]. The essential dynamics represent the velocities that are actually taken by the trajectories of the stratified controlled system. However, the lack of an appropriate strong invariance result was the limiting factor to prove a comparison principle that holds for any lower semicontinuous supersolution and any upper semicontinuous subsolution in the paper by Rao and Zidani [18].

Concerning numerical schemes approximating this problem, to the best of our knowledge, there are no results of convergence of numerical schemes in the present setting. The only known results of convergence of finite differences numerical schemes are proved in the setting of a one dimensional network due to Guérand and Koumaiha in [44], Carlini, Festa and Forcadel [45] and Morfe in [46].

The novelties of this chapter are the following. We will first define a Hamiltonian on the interface Λ, denoted by H_Λ, and consider the following Hamilton Jacobi Bellman equation:

$$\begin{cases} -\partial_t u(t,x) + H_i(x, \partial_x u(t,x)) = 0 \quad \text{for } (t,x) \in (0,T) \times \mathcal{M}_i, \\ -\partial_t u(t,x) + H_\Lambda(x, \partial_x u(t,x)) = 0 \quad \text{for } (t,x) \in (0,T) \times \Lambda, \\ u(T,x) = \psi(x) \quad \text{for } x \in \mathbb{R}^N. \end{cases}$$

The Hamiltonian H_Λ defined on the interface Λ has the form of a maximum of a finite number of lower semicontinuous *essential Hamiltonians* obtained along the same lines of [41]. Then, we will define a new notion of viscosity that encompasses the singular nature of this problem. This new definition of viscosity will allow us to prove a strong comparison principle, valid for any lower semicontinuous supersolution and any upper semicontinuous subsolutions. The proof of the latter is based on nonsmooth analysis techniques. In particular, we introduce an optimal control problem whose controlled system is defined via the essential dynamics and whose value function will be the solution to the above Hamilton Jacobi Bellman equation. Furthermore, we will prove the weak and the strong invariance principles in this stratified setting. We stress on the fact that the invariance principles are novelties of this chapter.

From the strong comparison principle, we will prove some stability results for the above Hamilton Jacobi Bellman equation in the presence of perturbations on the Hamiltonians H_i and H_Λ. Finally, we will extend the classical result due to Barles and Souganidis [47] regarding convergence of monotone numerical schemes to the stratified setting.

We would like to point out that the interface condition considered in this work is different from the multiple interface conditions presented in the book by Barles and Chasseigne [43]. Moreover, we use mainly nonsmooth analysis techniques whereas in Barles and Chasseigne's book, the focus is more on viscosity techniques. We will discuss a comparison between the results of this chapter and some of the settings considered in [43].

Chapter 3 : Viscosity solutions of Hamilton Jacobi equations in proper CAT(0) spaces

In this chapter, we combine techniques coming from viscosity theory with techniques related to metric geometry to develop a first order viscosity theory in proper geodesic metric spaces of curvature not greater than 0 in the sense of Alexandrov.

Metric geometry, at its core, is a branch of mathematics that aims to study geometric notions such as length, angles and curvature using purely metric distances. The fundamental object of study in metric geometry is the concept of geodesic spaces. A geodesic space is a metric space with the property that distances can be understood as lengths of paths between the points of the space. An important class of geodesic spaces of particular importance in metric geometry are spaces with curvature either bounded from above or bounded from below, studied mainly by the Russian school founded by Alexandrov in the last century. In this chapter, we are interested in geodesic spaces with curvature not greater than 0. These metric spaces can be regarded as a generalization of Riemannian manifolds of nonpositive sectional curvature. Their study started with the work of Hadamard and Cartan on hyperbolic spaces in the beginning of the last century. The notion of geodesic spaces of curvature not greater than 0 can be traced back to the work of Alexandrov published in Russian in 1951. He later summarized his ideas in the paper [48]. Actually, Alexandrov gave a meanning of what it means for a geodesic space to have curvature bounded from above by any real number $\kappa \in \mathbb{R}$. Alexandrov's work was popularized by the work of Gromov in the last decades. Notably, in his lectures at Collège de

France in 1981 [49], Gromov explained the main properties of the global geometry of manifolds of nonpositive sectional curvature based on Alexandrov's definition of curvature not greater than 0, which Gromov named the CAT(0) inequality. Note that the initials "C", "A" and "T" stand for Cartan, Alexandrov and Toponogov, each of whom contributed to the understanding of curvature via inequalities involving distances. Examples of CAT(0) spaces include Hilbert spaces, simply connected Riemannian manifolds with nonpositive sectional curvature and networks. For more on the topic of CAT(0) spaces, we refer to the books by Bridson and Haefliger [50], Alexander, Kapovitch and Petrunin [51], D. Burago, Y. Burago and Ivanov [52] and Bacák [53].

CAT(0) spaces enjoy a rigid structure that makes possible a first order calculus on them. Indeed, even though CAT(0) spaces are not manifolds in general, they resemble manifolds of nonpositive sectional curvature. For example, the CAT(0) inequality allows one to define the notion of *tangent cone* at every point, which is the metric analogue of the tangent space in differential geometry or the Bouligand tangent cone in convex analysis. Furthermore, the CAT(0) inequality implies that the distance function is *geodesically convex* (or simply convex in this manuscript) and the squared distance function is *semiconvex*. Real valued semiconvex functions and semiconcave functions, or more generally functions that could be represented as a difference of two semiconvex functions, called *DC functions*, exist in abundance in CAT(0) spaces. If furthermore they are Lipschitz, then they admit directional derivatives along geodesics at every point. These facts are exploited to define a notion of *differential* of a DC function.

DC functions were used in an unpublished paper by Perelman [54] to show the existence of a DC atlas in a dense set (with respect to the Hausdorff measure) of a finite dimensional locally compact geodesic space of curvature bounded from below in the sense of Alexandrov. Furthermore, he proved that DC functions admit a second order expansion at almost every point in this space. Recently, Perelman's results were extended by Ambrosio and Bertrand in [55]. DC functions are also used in the theory of gradient flows in geodesic spaces with one curvature bound in the sense of Alexandrov. See the manuscripts by Petrunin [56], Lytchak [57], Ohta [58], Ohta and Pàlfia [59] Ambrosio, Gigli and Savaré [60] and Mayer [61] for more details on the subject.

Recall that viscosity theory was first extended to Banach spaces by Crandall and Lions in a series of papers [28, 29, 30, 31] in the 1980s. The extension to more general classes of metric spaces is motivated by numerous applications involving traffic management issues, data transmission, geometric optics, wave front propagation, etc. Giga, Hamamuki and Nakayasu [32] treat the case of Hamilton Jacobi equations of Eikonal type defined in a general metric space. Gangbo and Święch [33] and Ambrosio and Feng [34] treat a class of Hamilton Jacobi equations on complete geodesic metric spaces where the Hamiltonian depends on the derivative of the unknown function only through its local Lipschitz constant, also known as the local slope. The generalization of the viscosity notion to abstract metric spaces for a more general class of Hamilton Jacobi equations is far from being straightforward. This is due to the lack of structure in these spaces. In particular, the notions of directions of curves, differential of functions and scalar product are not well defined.

Several contributions have restricted their analysis to special cases of metric spaces that have additional structural properties. For instance, Schieborn in [20] and Camilli, Marchi and Schieborn in [21] studied the Eikonal equation in a special case of metric spaces called ramified spaces. The latter can be visualized as a locally finite collection of manifolds (branches) of the same dimension, glued together along parts of their boundary (the junction). The simplest example of such setting is a one dimensional network, obtained by gluing a finite collection of half lines along their origin. The approach used in their setting was to exploit the differential structure that each branch enjoys, to define the Hamiltonian and the notion of viscosity. This approach was considered by Imbert and Monneau [22, 23], Imbert, Monneau and Zidani [24], Achdou and Tchou [26] and Lions and Souganidis [27] to study more general Hamilton Jacobi equations on networks. On each branch, they considered a Hamiltonian that is supposed to be continuous, yet the Hamiltonian on the whole network is discontinuous at the junction. They proved well-posedness of the problem by taking test functions that are continuously differentiable on each branch. Although this approach allows to treat more general Hamilton Jacobi equations, it relies heavily on the piecewise differential structure of the underlying space, which makes it inconvenient for other classes of metric spaces.

In this chapter, we aim at developing a novel notion of viscosity solution of first order Hamilton Jacobi equations defined in CAT(0) spaces. To do so, we are going to use the metric structure of CAT(0) spaces to define the Hamiltonian and we are going to use Lipschitz and DC functions to define the viscosity notion. The notion of viscosity we propose in this chapter in CAT(0) spaces offers many advantages with respect to the two main lanes that exist in the literature. Indeed, compared to [33, 34, 32], we can treat a more general class of Hamilton Jacobi equations. Furthermore, compared to [20, 21, 22, 23, 24, 27] we can treat Hamilton Jacobi equations on structures more complex than networks. Let us summarize the results of this chapter.

Let (X, d) be a proper CAT(0) space, i.e. a CAT(0) space whose closed bounded sets are compact. We propose to study the following boundary value problem

$$\begin{cases} H(v(x), x, D_x v) = 0, \quad x \in \Omega, \\ v(x) = \ell(x), \quad x \in \partial\Omega, \end{cases} \tag{1.8}$$

and its time dependent variant

$$\begin{cases} \partial_t v + H(x, D_x v) = 0, \quad \forall\, (t, x) \in (0, +\infty) \times X, \\ v(0, x) = \ell(x), \quad x \in X, \end{cases} \tag{1.9}$$

where Ω is an open subset of X and $\ell : X \to \mathbb{R}$ is a continuous and bounded function. The expression $D_x v$ is the differential of v at x, well defined if v is a Lipschitz and DC function. The function H is the Hamiltonian depending on $v(x) \in \mathbb{R}$, $x \in X$ and the differential function of v at x.

The main novelties of this chapter are the following. First, we give a precise definition of the Hamilton Jacobi equations (1.8) and (1.9) and the exact hypotheses required for the Hamiltonians in this setting. Moreover, we define the notion of viscosity adopted in this chapter using Lipschitz and DC functions. Then we prove the comparison principle for the stationary and the time dependent cases. The proof of the comparison principles is done in the exact same way as in the classical theory of viscosity. It relies essentially on the *variable doubling* technique using the squared distance function. Moreover, we prove existence of the solution by virtue of *Perron's method* in a similar manner as in the classical theory of viscosity. Finally, we give several examples covered by this setting which shows its degree of generality. In particular, we show that this setting coincides with the classical setting in $\mathbb{R}^N$ by treating several examples of Hamilton Jacobi equations defined in $X = \mathbb{R}^N$. The major difference between the current setting and the classical setting in $\mathbb{R}^N$ is that we use different sets of test functions. Moreover, we give several examples of Eikonal type equations defined on proper CAT(0) spaces of the form:

- the proper CAT(0) space obtained by gluing together three half-lines of $\mathbb{R}^2$

$$\begin{cases} X_1 := [0, +\infty)e_1, \\ X_2 := [0, +\infty)e_2, \\ X_3 := [0, +\infty)e_3, \end{cases}$$

  along the origin point $A = \{0\}$;

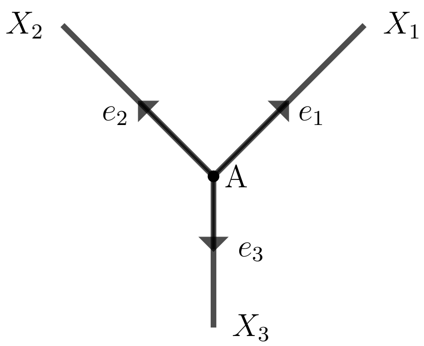


Figure 1.1: The space obtained by gluing X_1, X_2 and X_3 along A.

- the proper CAT(0) space obtained by gluing together the sets

$$\begin{cases} X_1 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_3 = 0\}, \\ X_2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_1 = x_2 = 0\}, \end{cases}$$

along the origin point $A = \{0\}$;

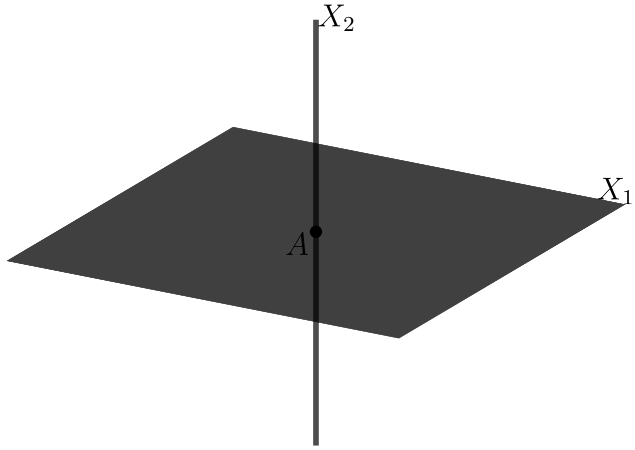


Figure 1.2: The space obtained by gluing X_1 and X_2 along A.

- the CAT(0) space obtained by gluing together the sets

$$\begin{cases} X_1 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_3 = 0\}, \\ X_2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_1 = x_2 = 0, \ x_3 \geq 0\}, \end{cases}$$

along the origin point $A = \{0\}$;

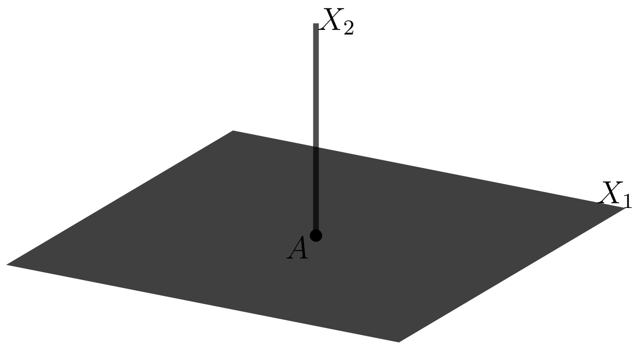


Figure 1.3: The space obtained by gluing X_1 and X_2 along A.

Finally, we would like to mention that in this chapter, we chose clarity over generality. Indeed, all the proofs given in this chapter are local in nature, hence they can be extended to any proper CAT(κ) space for any $\kappa \in \mathbb{R}$. Furthermore, in the next chapter we will explore the possibility of transposing the notion of viscosity defined in this chapter to the specific example of a Hamilton Jacobi Bellman equation defined on Wasserstein spaces over compact Riemannian manifolds, which could be regarded as a geodesic space with curvature bounded from *below* in the sense of Alexandrov.

Chapter 4 : Deterministic optimal control problem in Riemannian manifolds under probability knowledge of the initial condition

In this chapter, we introduce a novel notion of viscosity solution to study well-posedness of a Hamilton Jacobi Bellman equation associated to an optimal control problem defined on the Wasserstein space over a compact Riemannian manifold equipped with the *Wasserstein distance of order* 2.

For a compact Riemannian manifold (M, d) equipped with its Riemannian distance, the Wasserstein space over it is the space of distributions of mass over M, whose total mass is constant for all distributions. By convention, we normalize with the total mass and the Wasserstein space can be identified with the space of probability measures over M and denoted by $\mathcal{P}(M)$. It is endowed with the Wasserstein distance of order 2. The Wasserstein distance $d_W(\mu_1, \mu_2)$ between two measures μ_1 and μ_2 represents the minimum cost to move the total mass from one configuration represented by μ_1 to the configuration represented by μ_2, where the unit cost to move a unit mass from a point $x_1 \in M$ to the point $x_2 \in M$ is $d^2(x_1, x_2)$.

There has been an increasing interest in developing an optimal control and viscosity theories in Wasserstein spaces in the last decade. It stems from the numerous potential applications involving modelling certain specific uncertainties of an otherwise deterministic system or the modelling the collective behaviour of a large number of interacting agents supposed to be indistinguishable from one another. The potential real-world applications include modelling pedestrian motion, traffic flow, autonomous vehicle motion, aggregation phenomena in biology and problems related to fluid mechanics.

At the individual level, the behavior of each agent is dictated not only by local interactions between the agents but also by the *non local* interactions that depend on the distribution of all agents. When the number of agents is assumed to be very large, the complexity of the system grows extremely fast. To model a multi-agent system on M, a macroscopic point of view is considered. It consists in taking an infinite dimensional approximation of the problem and in considering the collection of the agents as a density that evolves in time. Furthermore, if the number of agents is constant at all times, then we can normalize the density and assume that the density of the system has a mass equal to 1 at all times. Hence, the evolution of the multi-agent system, seen as normalized spatial density in M is described by a curve $t \mapsto \mu_t \in \mathcal{P}(M)$ where μ_t represents the spatial density of the multi-agent system at a given time $t \geq 0$. The conservation of the mass along the trajectory $t \mapsto \mu_t$ is described by the continuity equation

$$\partial_t \mu_t + \operatorname{div}(w_t \mu_t) = 0,$$

where $w_t(.)$ is a time-dependent Borel vector field, and the equation is understood in the sense of distributions.

In this chapter, we propose to study a simple model of multi-agent systems, where the non local interactions between the agents are not considered. This problem can be interpreted as a deterministic controlled system with imperfect information on the initial condition, i.e. the initial condition is not known precisely by the controller, but they only know that the initial condition follows a probability distribution $\mu_0 \in \mathcal{P}(M)$. More precisely, consider the following controlled equation:

$$\begin{cases} \dot{y}(t) = f(y(t), u(t)), \quad t \in [t_0, T], \\ y(t_0) = x_0, \quad u(t) \in U, \end{cases} \tag{1.10}$$

where the set U is the set of admissible control values which is assumed to be a compact subset of some metric space. The dynamics $f : M \times U \to TM$, is assumed to be Lipschitz with respect to the first variable, continuous with respect to the second variable and the following set of functions

$$\{x \mapsto f(x,u) \ : \ u \in U\}$$

is convex. The initial position is $x_0 \in M$ and the initial time is $t_0 \in [0,T]$. The control function $u(.) \subset U$ is a Borel measurable function $u : [t_0,T] \to U$. The main feature of this problem is that the initial position x_0 is not perfectly known, but rather distributed along the probability measure μ_0. Notice that since $f(., u(t))$ is Lipschitz continuous and bounded, the evolution curve of the uncertainty, $t \mapsto \mu_t$ starting from μ_0, is the unique solution to the equation

$$\begin{cases} \partial_t \mu_t + div(f(., u(t))\mu_t) = 0, \ t \in (t_0, T), \\ \mu_{t_0} = \mu_0, \end{cases}$$

in the distributional sense. The measures μ_t are obtained by the pushforward of μ_0 by the flow at time t of the controlled equation (1.10). The controller aims at minimizing the following final cost:

$$L(\mu) = \int \ell(y) d\mu(y),$$

where $\ell : M \to \mathbb{R}$ is a Lipschitz function. The quantity $L(\mu_T)$ represents the expectation of the deterministic final cost with respect to the measure μ_T. To this optimal control problem, we associate the following value function:

$$\vartheta(t_0, \mu_0) = \inf_{u(.) \in U} L(\mu_T).$$

The first main goal of this chapter is to study the properties and the regularity of the value function. In particular, we will show that the value function is Lipschitz continuous with respect to both variables and that it verifies the dynamic programming principle. The second goal of the chapter is to prove that the value function can be characterized as the unique viscosity solution of a suitable Hamilton Jacobi Bellman equation of the form

$$\begin{cases} \partial_t v + H(\mu, D_\mu v) = 0, & (t,\mu) \in [0,T) \times \mathcal{P}_2(M), \\ v(T,\mu) = L(\mu), \end{cases} \tag{1.11}$$

in the Wasserstein space $\mathcal{P}(M)$.

A similar problem was studied by Cardaliaguet and Quincampoix in [37] in the context of differential games and by Marigonda and Quincampoix in [36] for Mayer optimal control problems. A more general optimal control problem of multi-agent systems, where the non local interactions are taken into account by making the dynamics f depend also on the measure variable was considered by Jimenez, Marigonda and Quincampoix [38]. In [37, 38, 36] the authors proved that value function is Lipschitz continuous and it is a viscosity solution to a Hamilton Jacobi equation in a certain weak sense using techniques coming from nonsmooth analysis applied to the Wasserstein space over the Euclidean space.

Regarding the viscosity notion, the study of Hamilton Jacobi equations in Wasserstein spaces over the Euclidean space is based on two different strategies. The first strategy consists in considering a suitable notion of sub/super- differentials to define the notion of viscosity, see for example Cardaliaguet and Quincampoix, [37] and Jimenez, Marigonda and Quincampoix [62]. The second strategy is based on the so-called Lions calculus introduced by Lions in 2006 at Collège de France [63]. The idea of this line of thought is to “lift” the Hamilton Jacobi equation defined in the Wasserstein space to a Hamilton Jacobi equation defined in a Hilbert space. One then uses the viscosity theory techniques developed in Hilbert spaces to define a suitable notion of viscosity in the Wasserstein space, in an extrinsic way, through this lift. For more details we refer to the paper by Gangbo and Tudorascu [40]. Both of these notions were proven to be equivalent in the recent preprint by Jimenez, Marigonda and Quincampoix [62] in the case of certain Hamilton Jacobi equations coming from modelling multi-agent systems. In this chapter, our approach is different. We propose a different notion of viscosity solution, based on exploiting tools of metric geometry applied to the Wasserstein space $\mathcal{P}(M)$. In particular, We define the viscosity notion by considering test functions that are Lipschitz and DC in the same manner as in the previous chapter.

The Wasserstein space $(\mathcal{P}(M), d_W)$ is a geodesic space. This was first noticed by McCann in his PhD thesis [64] in the case of the Wasserstein space over the Euclidean space. He used geodesics of the Wasserstein space to prove uniqueness of minimizers of certain functions that are geodesically convex (or simply convex in this manuscript). This property is also known in the literature by the name of *displacement convexity*. Later, Otto introduced a formal Riemannian structure on the Wasserstein space over the Euclidean space on a purely heuristic level in order to prove that the heat equation in the Euclidean space could be interpreted as a gradient flow in the Wasserstein space over it [65]. Furthermore, he obtained some formal computations that indicated that the Wasserstein space possesses a nonnegative curvature. This was made precise by several authors.

If M has nonnegative sectional curvature, Lott and Villani [66] and Sturm [67] showed that the Wasserstein space over it $\mathcal{P}(M)$ has nonnegative curvature in the sense of Alexandrov. More generally, if M is any compact Riemannian manifold, Ohta showed in [58] that $\mathcal{P}(M)$ has a "2-uniform concavity" sturucture, which can be regarded as a generalization of curvature bounded from below in the sense of Alexandrov. Moreover, in [68] Gigli showed that the notion of *tangent cone* in the sense of metric geometry is well defined at every point of $\mathcal{P}(M)$ and gave an explicit isometric representative of it. Furthemore, he showed that the squared Wasserstein distance is Lipschitz and semiconcave and gave an explicit expression of its directional derivatives at every point.

We will exploit all these results concerning the geometry of $\mathcal{P}(M)$ in this chapter. In particular, we use the tangent cone to give a precise definition of the Hamiltonian we are going to work with. We define a notion of a *differential* for a Lipschitz and DC function in the same manner as in the previous chapter so that the notation $D_\mu v$ in equation (1.11) will become precise. Moreover, we use the explicit expression of the differential of the squared Wasserstein distance to prove a comparison principle by means of the variable doubling technique. Finally, we prove that the value function is the unique solution of a Hamilton Jacobi Bellman equation of the form (1.11).

The results of this chapter are summarized as following. First, we show that the value function of the optimal control problem defined above is Lipschitz continuous with respect to both variables. Then, we show that it verifies the dynamic programming principle in same manner as in the classical case. Furthermore, we prove that the value function is the unique viscosity solution of a suitable Hamilton Jacobi Bellman equation defined in $\mathcal{P}(M)$. The viscosity notion is defined by means of test functions that are Lipschitz and DC in $\mathcal{P}(M)$. The uniqueness of the solution is established by proving a comparison principle valid for any bounded upper semicontinuous subsolution and any bounded lower semicontinuous supersolution. Existence of the solution is proved by means of the dynamic programming principle verified by the value function.

The results of this work will be the subject of the following publications:

1. F. Jean, O. Jerhaoui and H. Zidani. A Mayer optimal control problem on Wasserstein spaces over Riemannian manifolds. Proceedings of the 18th IFAC Workshop on Control Applications of Optimization. `Accepted`.

2. O. Jerhaoui and H. Zidani. A general comparison principle for Hamilton Jacobi Bellman equations on stratified domains. `Submitted`.
3. F. Jean, O. Jerhaoui and H. Zidani. Deterministic optimal control problem in Riemannian manifolds under probability knowledge of the initial condition. `Submitted`.
4. O. Jerhaoui and H. Zidani. Viscosity solutions of Hamilton Jacobi equations in proper CAT(0) spaces, `In preparation`.

Chapter 2

A general comparison principle for Hamilton Jacobi Bellman equations in stratified domains

Contents

2.1 Introduction

In this chapter, we study well-posedness of a system of Hamilton Jacobi Bellman equations (HJB in short) defined on a stratification of $\mathbb{R}^N$. This problem was first indroduced in [14] and [41]. A stratification of $\mathbb{R}^N$ is a finite collection of disjoint open sets of $\mathbb{R}^N$ denoted $(\mathcal{M}_i)_{i=1,\dots,n}$ such that

$$\mathbb{R}^N = \bigcup_{i=1}^{n} \overline{\mathcal{M}}_i, \quad \text{and} \quad \mathcal{M}_i \cap \mathcal{M}_j = \emptyset, \text{ whenever } i \neq j.$$

The union of the open sets $\cup_{i=1}^{n}\mathcal{M}_i$ is called the regular part of the stratification. The singular part of the stratification is the union of all the interfaces between the open sets $(\mathcal{M})_{i=1,\dots,n}$. It is the set

$$\Lambda := \mathbb{R}^N \setminus \bigcup_{i=1}^{n} \mathcal{M}_i.$$

We consider the following HJB system

$$\begin{cases} -\partial_t u(t,x) + H_{F_i}(x, \partial_x u(t,x)) = 0 \quad \text{for } (t,x) \in (0,T) \times \mathcal{M}_i \\ u(T,x) = \psi(x), \quad x \in \mathcal{M}_i \end{cases} \tag{2.1}$$

where $T > 0$ is the final time and $\psi : \mathbb{R}^N \to \mathbb{R}$ is the final cost, assumed to be Lipschitz continuous and bounded. $H_{F_i} : \mathcal{M}_i \times \mathbb{R}^N \to \mathbb{R}$ are Bellman Hamiltonians defined the following way

$$H_{F_i}(x,p) = \sup_{q \in F_i(x)} \{-\langle p, q \rangle\}.$$

$F_i : \mathcal{M}_i \rightsquigarrow \mathbb{R}^N$ are set-valued maps, called the dynamics, that satisfy standard hypotheses. The HJB system (2.1) is not defined on the singular set Λ. We propose to study, under which condition to add at the singular set Λ, that will guarantee the well-posedness of the HJB system (2.1).

The most natural way to add a condition to the HJB system (2.1) on the singular set is to define a Hamiltonian $H : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ on all $\mathbb{R}^N$ such that

$$\forall p \in \mathbb{R}^N, \quad H(x,p) = H_{F_i}(x,p), \quad \text{whenever } x \in \mathcal{M}_i,$$

and the HJB system becomes

$$\begin{cases} -\partial_t u(t,x) + H(x, \partial_x u(t,x)) = 0 \quad \text{for } (t,x) \in (0,T) \times \mathbb{R}^N \\ u(T,x) = \psi(x), \quad x \in \mathbb{R}^N. \end{cases} \tag{2.2}$$

Obviously, H cannot be assumed to be continuous on $\mathbb{R}^N$ since the Hamiltonians H_{F_i}, $i = 1, \dots, n$ are different and, a priori, don't have any connection with one another. Therefore, the expression of H on the singular set Γ is not obvious at first sight.

In the case when H is continuous, the well-posedness of equation (2.2) was first studied in [1]. In particular, a notion of generalized solutions, called the *viscosity* notion, was introduced to guarantee existence, uniqueness and stability of the solution. This notion was extended to the discontinuous case by Ishii in [69] by replacing the Hamilton Jacobi equation (2.2) with two Hamilton Jacobi inequalities involving the lower semicontinuous envelope of H, denoted H_*, and the upper semicontinuous envelope of H, denoted H^*, the following way

$$\begin{cases} -\partial_t u(t,x) + H^*(x, \partial_x u(t,x)) \geq 0 & \text{for } (t,x) \in (0,T) \times \mathbb{R}^N \quad (2.3a) \\ -\partial_t u(t,x) + H_*(x, \partial_x u(t,x)) \leq 0 & \text{for } (t,x) \in (0,T) \times \mathbb{R}^N. \quad (2.3b) \end{cases}$$

When $u : (0,T] \times \mathbb{R}^N \to \mathbb{R}$ verifies inequality (2.3a) in the viscosity sense, then it is called a *supersolution* and when it verifies inequality (2.3b) in the viscosity sense, then it is called a *subsolution*. In our particular case of HJB equation (2.1), Ishii's extention to the singular set Λ has the following form

$$\begin{cases} -\partial_t u(t,x) + \max\limits_{i=1,\ldots n} \{H_{F_i}(x, \partial_x u(t,x))\} \geq 0, & (t,x) \in (0,T) \times \Lambda \\ -\partial_t u(t,x) + \min\limits_{i=1,\ldots n} \{H_{F_i}(x, \partial_x u(t,x))\} \leq 0, & (t,x) \in (0,T) \times \Lambda. \end{cases}$$

However, using Ishii's extension to the singular set does not guarantee uniqueness of the viscosity solution in general. In viscosity theory, the uniqueness of the solution comes from the so-called *comparison principle*. It asserts that if an upper semicontinuous subsolution u is inferior to a lower semicontinuous supersolution v on $\{T\} \times \mathbb{R}^N$, then u is inferior to v on $(0,T] \times \mathbb{R}^N$. Such a comparison result can no longer be obtained using Ishii's extension of the Hamiltonian to the singular set. Therefore, in this chapter, we look to impose a stronger condition on Λ that will allow us to obtain a comparison result that holds for any upper semicontinuous subsolution and any lower semicontinuous super solution.

In the literature, a comparison principle was obtained in [42] for equation (2.2) under the assumption that the Hamiltonian H is measurable with respect to the space variable. Ishii's viscosity notion was not used. Instead, it was assumed that the Hamiltonian H satisfies a "transversality condition" that essentially boils down to the fact that the behavior of the Hamiltonian at the singular set can be ignored. The authors in [14] were the first to study discontinuous HJB equations in a similar layout than ours. They defined suitable Hamilton Jacobi inequalities at the singular set Λ and showed that a comparison result holds under the assumption that the subsolution is continuous. In [70], the authors extended the results of [14] and provided a very thorough analysis on the notion of viscosity solution of a stratified system. Furthermore, different conditions under which the comparison principle is satisfied have been investigated in their monograph [43]. In [18, 71], the authors have a similar layout to ours. They defined the Hamiltonian H on the singular part by using the notion of *essential Hamiltonian* first introduced in [41]. The essential Hamiltonian is defined from an optimal control interpretation of the HJB system (2.1). It comes from the set-valued map that represents the "essential velocities" of the system, meaning the velocities that are taken by the trajectories of the optimal control problem associated to the HJB system (2.1). This approach is the one we are going to follow in this chapter as well. We also mention similar publications of Hamilton Jacobi equations on one dimensional networks that share the same kind of difficulties as our layout [22, 24, 25, 72, 27].

As for numerical schemes approximating this problem, the only known results of convergence of finite differences numerical schemes are in the setting of a one dimensional network [44, 46, 45]. However, to the best of our knowledge, there aren't any known convergence results in our setting.

HJB equations are also related to a geometric notion known as flow invariance in the theory of differential inclusions [9, chapter 12]. There are two types of invariances, weak and strong invariance. Weak/strong invariance are geometric properties that link the controlled system of the optimal control problem associated to the HJB system, with the epigraph/hypograph of the supersolution/subsolution respectively. The classical case, meaning the absence of a stratification, has been treated thoroughly in the literature [9, 6]. In the case of stratified domains, the authors in [41] analyzed the characterization of weak and strong invariance principles using *the essential Hamiltonian*. However their statement of the strong invariance property was inaccurate. Despite their correct intuition regarding the choice of the Hamiltonian, their choice of the "test functions" (in analogy with the viscosity theory) did not take into account the singular geometry of the problem which turns out to be crucial for proving this property.

In this present work, we aim to prove the well-posedness of the HJB equation on stratified domains. We will first define a Hamiltonian H_Λ on the singular set Λ and consider the HJB equation:

$$\begin{cases} -\partial_t u(t,x) + H_{F_i}(x, \partial_x u(t,x)) = 0 \quad \text{for } (t,x) \in (0,T) \times \mathcal{M}_i, \\ -\partial_t u(t,x) + H_\Lambda(x, \partial_x u(t,x)) = 0 \quad \text{for } (t,x) \in (0,T) \times \Lambda, \\ u(T,x) = \psi(x) \quad \text{for } x \in \mathbb{R}^N. \end{cases} \tag{2.4}$$

The Hamiltonian H_Λ is defined on the interface Λ in the form of a maximum of lower semicontinuous Hamiltonians obtained along the same lines of [41]. Then, we will revisit the definition of viscosity solution and give a new one that encodes the nature of the singular geometry of the problem. This new definition of viscosity will allow us to extend the strong comparison type results, known when the Hamiltonian is Lipschitz continuous, to the present setting. More precisely, we prove the following result:

Let u and v be respectively upper semi-continuous and lower semi-continuous functions on $(0,T]\times\mathbb{R}^N$. If u is a subsolution of (2.4), and if v is a supersolution of (2.4), then $u\le v$ on $(0,T]\times\mathbb{R}^N$.
The proof of this result relies on nonsmooth analysis techniques. In particular, we introduce an optimal control problem whose value function will be the solution of the HJB equation. In the classical case, the nonsmooth analysis approach consists in interpreting the subsolution property of the value function as the strong invariance of the hypograph of the value function and the supersolution property as the weak invariance of its epigraph. We will extend the weak and strong invariance results to the stratified setting. We would like to emphasize that the extension of invariance principles is also a contribution of this chapter.

The strong comparison principle will have two major consequences. First, it will allow to obtain some stability results in this setting in the presence of perturbations on the dynamics. We prove that if there exist sequences $\left(F_i^j\right)_j$ of set-valued maps such that $F_i^j \longrightarrow F_i$ with respect to the Hausdorff distance, and a sequence $(v^j:\mathbb{R}^N\to\mathbb{R})_j$ of lower semicontinuous (respectively upper semicontinuous) functions such that $v^j\to v$ locally uniformly in $\mathbb{R}^N$ and suppose for all j, v^j is a supersolution (respectively subsolution) of

$$\begin{cases} -\partial_t u(t,x)+H_{F_i^j}(x,\partial_x u(t,x))=0 \quad \text{for } (t,x)\in(0,T)\times\mathcal{M}_i,\\ -\partial_t u(t,x)+H_\Lambda^j(x,\partial_x u(t,x))=0 \quad \text{for } (t,x)\in(0,T)\times\Lambda,\\ u(T,x)=\psi(x) \quad \text{for } x\in\mathbb{R}^N. \end{cases}$$

then v is a supersolution (respectively subsolution) of (2.4).

Finally, we will extend the known result due to Barles and Souganidis [47] for the convergence of monotone numerical schemes to the stratified setting. The numerical scheme has the following form in each demain:

$$\begin{cases} S_i^h(t_h,x_h,u^h(t_h,x_h),[u^h]_{(t_h,x_h)})=0 \quad \text{for } (t_h,x_h)\in(\Pi^{\Delta t}\times\mathcal{G}^{\Delta x})\bigcap((0,T)\times\mathcal{M}_i),\\ S_\Lambda^h(t_h,x_h,u^h(t_h,x_h),[u^h]_{(t_h,x_h)})=0 \quad \text{for } (t_h,x_h)\in(\Pi^{\Delta t}\times\mathcal{G}^{\Delta x})\bigcap((0,T)\times\Lambda),\\ u^h(T,x_h)=\psi(x_h), \quad \text{for } (t_h,x_h)\in(\Pi^{\Delta t}\times\mathcal{G}^{\Delta x})\cap\{t_h=T\}, \end{cases}$$

where $\Pi^{\Delta t}$ is a time grid, $\mathcal{G}^{\Delta x}$ is a spatial grid, $h=(\Delta t,\Delta x)$ is the step of the grid and $[u^h]_{(t_h,x_h)}$ are all the values of of u^h on $\mathcal{G}^{\Delta x}$ at other points than (t_h,x_h) on the grid. We show that under the usual hypotheses of monotonicity, stability and consistency, the numerical scheme converges locally uniformly to the viscosity solution of (2.4).

The chapter is organized as follows: in Section 2.2, we define the notations and conventions used throughout the chapter. We also define the geometry of the problem, the dynamics of the HJB equation and we state the main results. Section 2.3

is devoted to the invariance principles, a nonsmooth analysis point of view of the HJB equation. In Section 2.4, we first define the optimal control problem associated to the HJB equation, we introduce the value function and we prove that the super-optimality and sub-optimality properties of the value function are equivalent to it being a viscosity supersolution and subsolution respectively. Then we prove the strong comparison result. Section 2.5 is devoted to the proofs of the stability results. Finally, we prove in section 2.6 a general convergence result for monotone numerical schemes.

2.2 Main results

2.2.1 Notations and conventions

Throughout the chapter, we denote by $\mathbb{R}^N$ the Euclidean space where the stratification is defined, $\mathbb{B}$ the unit ball of center 0 of $\mathbb{R}^N$ and $\mathbb{B}(x,r) = x + r\mathbb{B}$.

For any set $S \subset \mathbb{R}^N$, we denote by $\overline{S}$ and ∂S its closure and topological boundary. We denote by $\mathrm{co}(S)$ the convex hull of S and by $\mathscr{L}$, the Lebesgue measure on $\mathbb{R}$.

The distance function associated to S is denoted by $d_S(x) = \inf\{|x-y| \;:\; y \in S\}$ and the set of solutions where the infimum is attained is called the projection of x on S and denoted by $proj_S(x)$ (note that it might be empty).

The Bouligand tangent cone of S at x, denoted $\mathcal{T}_S(x)$ is defined the following way:

$$\mathcal{T}_S(x) = \left\{ v \in \mathbb{R}^N \;:\; \liminf_{t \to 0^+} \frac{d_S(x+tv)}{t} = 0 \right\}.$$

If A and B are two sets of $\mathbb{R}^N$, we define a distance between them by

$$d(A,B) = \inf \{ |a-b| : \; (a,b) \in A \times B \},$$

with the convention $d(\emptyset,\emptyset) = 0$ and $d(\emptyset, B) = +\infty$ if $B \neq \emptyset$.

For K_1 and K_2 two compact sets of $\mathbb{R}^N$, the Hausdorff distance between them is given by

$$d_{\mathcal{H}}(K_1,K_2) = max \left\{ \sup_{x \in K_2} d_{K_1}(x) \;,\; \sup_{x \in K_1} d_{K_2}(x) \right\},$$

with the convention $d_{\mathcal{H}}(\emptyset,\emptyset) = 0$ and $d_{\mathcal{H}}(\emptyset, S) = +\infty$ if $S \neq \emptyset$.

For a given function $f : \mathbb{R}^N \to \mathbb{R}$, we denote by $epi(f)$ and $hyp(f)$ respectively its epigraph and hypograph, and defined the following way:

$$epi(f) = \Big\{ (x,r) \in \mathbb{R}^N \times \mathbb{R} \;:\; f(x) \leq r \Big\}, \quad hyp(f) = \Big\{ (x,r) \in \mathbb{R}^N \times \mathbb{R} \;:\; f(x) \geq r \Big\}.$$

If $\Gamma : \mathbb{R}^N \rightsquigarrow \mathbb{R}^N$ is a set-valued map, then we denote by $dom(\Gamma)$ the set of points $x \in \mathbb{R}^N$ such that $\Gamma(x) \neq \emptyset$.

For $T > 0$, we define the differential inclusion associated to Γ, with the initial condition $(t, x) \in (0, T) \times \mathbb{R}^N$, by

$$(DI)_\Gamma(t,x) : \begin{cases} \dot{y}(s) \in \Gamma(y(s)) & a.e.\ s \in [t,T] \\ y(t) = x. \end{cases}$$

Finally, the abreviations 'u.s.c.', 'l.s.c', 'HJB' and 'w.r.t' respectively stand for: 'upper semicontinuous', 'lower semicontinuous', 'Hamilton Jacobi Bellman' and 'with respect to'.

2.2.2 Stratification

Let $N, n \geq 1$ be two strictly positive integers. Let $\mathcal{M}_i$, $i = 1, ..., n$ be pairwise disjoint, connected open sets of $\mathbb{R}^N$. We suppose that $\mathbb{R}^N = \cup_{i=1}^n \overline{\mathcal{M}}_i$ and we denote by $\Lambda := \mathbb{R}^N \setminus \cup_{i=1}^n \mathcal{M}_i$ the singular set. Furthermore, we suppose that Λ is equal to a union of l, pairwise disjoint, C^2 embedded sumbanifolds $\mathcal{M}_{n+1}, \ldots, \mathcal{M}_{n+l}$ of lower dimension than N and with empty boundary, so that we have

$$\mathbb{R}^N = \bigcup_{i=1}^{n} \overline{\mathcal{M}}_i = \left(\bigcup_{i=1}^{n} \mathcal{M}_i \right) \bigcup \Lambda = \bigcup_{i=1}^{n+l} \mathcal{M}_i.$$

Finally, we suppose that each $\overline{\mathcal{M}}_i$, $i = 1, \ldots, n + l$, is proximally smooth and relatively wedged. All these assumptions on the stratification are summarized as following:

$$(H_1) \begin{cases} (i) & \text{Each } \mathcal{M}_i \text{ is a } C^2 \text{ embedded submanifold, with empty boundary,} \\ (ii) & dim(\mathcal{M}_1) = ... = dim(\mathcal{M}_n) = N \text{ and } dim(\mathcal{M}_{n+1}), ..., dim(\mathcal{M}_{n+l}) < N, \\ (iii) & \mathbb{R}^N = \displaystyle\bigcup_{i=1}^{n} \overline{\mathcal{M}}_i = \bigcup_{i=1}^{n+l} \mathcal{M}_i, \\ (iv) & \forall i, j = 1, ..., n+l, \quad \mathcal{M}_i \cap \mathcal{M}_j = \emptyset, \text{ if } i \neq j, \\ (v) & \text{if } \mathcal{M}_i \cap \overline{\mathcal{M}}_j \neq \emptyset, \text{ then } \mathcal{M}_i \subset \overline{\mathcal{M}}_j \text{ or } \mathcal{M}_j \subset \overline{\mathcal{M}}_i, \\ (vi) & \text{each } \overline{\mathcal{M}}_i \text{ is proximally smooth and relatively wedged.} \end{cases}$$

We call $\bigcup_{i=1}^{n} \mathcal{M}_i$ the regular part of the stratification and $\Lambda := \bigcup_{i=1}^{l} \mathcal{M}_{n+i}$ the singular part or the interfaces.

Comments on the Hypothesis (H_1)

Hypotheses $(H_1)(i)$ to $(H_1)(v)$ are standard for a stratification of $\mathbb{R}^N$. As for $(H_1)(vi)$, a closed set $X \subset \mathbb{R}^N$ is said to be proximally smooth if there exists $r > 0$ such that the projection map $proj_X(.)$ is a singleton on the tube $\{x \in X, d_X(x) < r\}$ [73]. The class of proximally smooth sets includes convex subsets of $\mathbb{R}^N$ and C^2 compact submanifolds of $\mathbb{R}^N$.

Relative wedgeness hypothesis was introduced in [41] for C^2 submanifolds of $\mathbb{R}^N$ such that their closure is proximally smooth. Roughly speaking, relative wedgeness of $\overline{\mathcal{M}}_i$, with $i \in \{1, ..., n+l\}$, means that the dimension of the Bouligand tangent cone at every point of $\overline{\mathcal{M}}_i$ is equal to the dimension of the manifold $\mathcal{M}_i$ [41]. The precise definition of this property is presented in Section 2.7.1.

Example 2.1. Figure 2.1 shows an example of the stratified setting, where $N = 1$, $n = 2$, $l = 1$.

$$\mathcal{M}_1 = (0, +\infty)e_1, \quad \mathcal{M}_2 = (0, +\infty)e_2, \quad \mathcal{M}_3 = \{0\}.$$

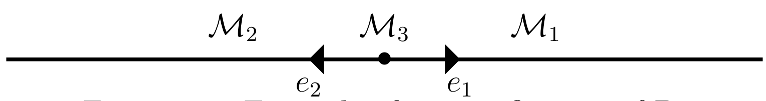


Figure 2.1: Example of a stratification of $\mathbb{R}$.

Example 2.2. This example shows a stratification of $\mathbb{R}^2$, where $N = 2$, $n = 2$, $l = 3$.

$$\mathcal{M}_1 = \{(x_1, x_2) \in \mathbb{R}^2 \ : \ x_2 > 0\}, \quad \mathcal{M}_2 = \{(x_1, x_2) \in \mathbb{R}^2 \ : \ x_2 < 0\}, \quad \mathcal{M}_5 = \{(0, 0)\}$$

$$\mathcal{M}_3 = \{(x_1, x_2) \in \mathbb{R}^2 \ : \ x_1 < 0, \ x_2 = 0\}, \quad \mathcal{M}_4 = \{(x_1, x_2) \in \mathbb{R}^2 \ : \ x_1 > 0, \ x_2 = 0\}$$

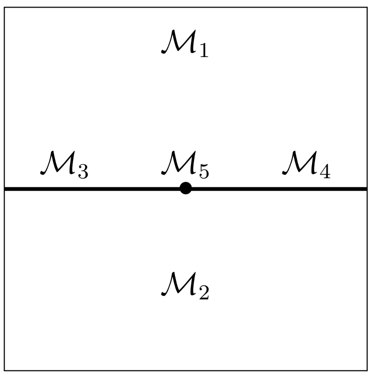


Figure 2.2: Example of a stratification of $\mathbb{R}^2$.

Example 2.3. Figure 2.3 shows an example of the stratified setting, where $N = 2$, $n = 4$, $l = 5$.

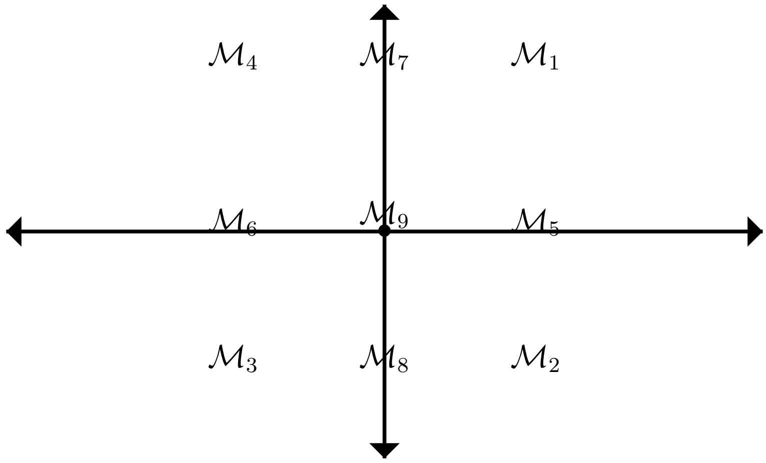
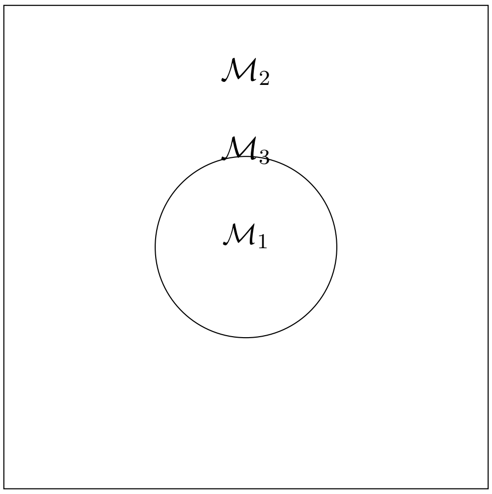


Figure 2.3: Example of a stratification of $\mathbb{R}^2$.

Example 2.4. The following example shows a stratification of $\mathbb{R}^2$, with $n = 2$ and $l = 1$. $\mathcal{M}_1$ is the unit open disc, $\mathcal{M}_2$ is the complement of the unit closed disc and $\mathcal{M}_3$ is the unit circle.



For any $x \in \mathbb{R}^N$ we define the index set-valued map

$$I(x) := \{\, i \in \{1, ..., n + l\} \;:\; x \in \overline{\mathcal{M}_i} \,\}.$$

Remark 2.2.1. It is clear from the definition of the stratification that for $x \in \mathbb{R}^N$ fixed, and $y \in \mathbb{R}^N$ close enough to x, we have $I(y) \subseteq I(x)$.

2.2.3 Setting of the problem

We begin by defining the dynamics for the Hamiltonians presented in the introduction. On each $\mathcal{M}_i$ with $i = 1, ..., n$, we are given a set-valued map $F_i : \mathcal{M}_i \rightsquigarrow \mathbb{R}^N$

that satisfies the standard hypotheses

$$(SH)\begin{cases} (i) & x \rightsquigarrow F_i(x) \text{ has non empty convex and compact images,}\\ (ii) & \exists\,\lambda>0 \text{ such that } \max\{\,|p|,\ p\in F_i(x)\}\le \lambda(1+|x|),\\ (iii) & F_i \text{ is Lipschitz continuous on bounded sets of } \mathcal{M}_i \text{ w.r.t the Hausdorff}\\ & \text{metric, i.e. for each } R>0, \text{ there are constants } K_{1,R},\dots,K_{n,R}>0:\\ & d_{\mathcal{H}}(F_i(x),F_i(y))\le K_{i,R}|x-y| \quad \text{if } x,y\in\mathbb{B}(0,R)\cap\mathcal{M}_i,\ i\in\{1,\dots,n\}. \end{cases}$$

We are interested in studying the well-posedness of the following HJB equation.

$$\begin{cases} -\partial_t u(t,x)+\sup\limits_{\nu\in F_i(x)}\{\,-\langle \nu,\partial_x u(t,x)\rangle\,\}=0 \quad \text{for } (t,x)\in(0,T)\times\mathcal{M}_i,\ i=1,\dots,n,\\ u(T,x)=\psi(x), \end{cases} \tag{2.5}$$

where $T>0$ is the final time and $\psi:\mathbb{R}^N\to\mathbb{R}$ is the final cost. We assume the following hypothesis on the final cost:

$$(H\psi)\ :\qquad \psi \text{ is Lipschitz continuous and bounded.}$$

The study of HJB equations is done using a weak notion of solutions, called viscosity solutions. This setting requires the HJB equation to be defined at every point. Hence, we need to find suitable interfaces conditions in order to guarantee the well-posedness of the system. To do so, we aim at defining some appropriate dynamics to consider at the interfaces.

Notice first that since the dynamics F_i, $i=1,...,n$ verify hypothesis (SH), then they can be extended to $\overline{\mathcal{M}}_i$ while verifying the same hypothesis (SH). *We denote this extension by F_i as well.* In order to define the dynamics on the whole space, a classical idea is to consider the Filippov regularization of $(F_i)_{i=1,\dots,n}$, denoted $F:\mathbb{R}^N\rightsquigarrow\mathbb{R}^N$ and defined by [74]

$$F(x)=\bigcap_{\varepsilon>0}\overline{co}\bigcup_{y}\{\cup_{i\in\{1\dots n\}}F_i(y)\ :|x-y|\le\varepsilon\}.$$

It is straightforward to check that F has a linear growth. However, it might not be Lipschitz in general. By the nature of our problem, the Filippov regularization is equal to

$$F(x)=co\,\{F_i(x)\ :\ i\in\{1\dots n\}\,\}.$$

For $(x,p)\in\mathbb{R}^N\times\mathbb{R}^N$, we define the Hamiltonian associated to F by

$$H_F(x,p)=\sup_{q\in F(x)}\{-\langle p,q\rangle\}\,.$$

Since F is only upper semicontinuous, the Hamiltonian $H_F(.,p)$ is also only upper semicontinuous. If $H_F(.,p)$ were to be Lipschitz continuous, we would have defined our HJB equation using the Hamiltonian associated to F and the well-posedness of the HJB system would follow from the classical theory, see [9, 75]. This is generally not the case in a stratified domain.

Nevertheless, the next step is to use F to define the dynamics at the interfaces. We define the dynamics $F_{n+i} : \mathcal{M}_{n+i} \rightsquigarrow \mathbb{R}^N$, for $i = 1, ..., l$ on each interface $\mathcal{M}_{n+i}$ by

$$F_{n+i}(x) = F(x) \cap \mathcal{T}_{\mathcal{M}_{n+i}}(x),$$

where $\mathcal{T}_{\mathcal{M}_{n+i}}(x)$ is the Bouligand tangent cone which coincides with the classical tangent space of $\mathcal{M}_{n+i}$ at x since it is a C^2 manifold. Furthermore, we suppose that all the interface dynamics are Lipschitz continuous on bounded sets as well:

$$(H_D) \quad \text{for } i = 1, ..., l, \; F_{n+i}(.) \text{ is Lipschitz continuous on bounded sets of } \mathcal{M}_i.$$

We point out that since we have the conventions $d(\emptyset, S) = +\infty$ if $S \neq \emptyset$ and $d(\emptyset, \emptyset) = 0$, it follows that (H_D) implies that F_{n+i} is either identically the empty set or nonempty on the whole domain $\mathcal{M}_{n+i}$. Under assumption (H_D), each $F_i : \mathcal{M}_i \rightsquigarrow \mathbb{R}^N$, $(i = n+1, ..., n+l)$ satisfies (SH). Thus each F_i can be extended to $\overline{\mathcal{M}}_i$ while verifying the same hypothesis (SH). *We denote this extension by F_i as well.*

A sufficient condition for (H_D) to be satisfied is full controllability near Λ. We mean by full controllability the following assumption:

$$(CH) \;\; \exists r > 0 \;:\quad \text{for all } \; i = 1, ..., n, \text{ and } x \in \Lambda \cap \overline{\mathcal{M}}_i \;:\; \mathbb{B}(0, r) \subseteq F_i(x).$$

Proposition 2.4.1. *[18, Lemma 2.2.] Assume (H_1) and (CH). Then, (H_D) holds.*

For $x \in \overline{\mathcal{M}}_i$, $i = 1, ..., n+l$, and $p \in \mathbb{R}^N$, we define the Hamiltonian

$$H_{F_i}(x, p) := \sup_{q \in F_i(x)} \{-\langle p, q\rangle\}.$$

At this point, we are tempted to define the HJB equation on the singular set using the dynamics $F_i(.)$ defined above, in the following way:

$$\begin{cases} -\partial_t u(t,x) + H_{F_i}(x, \partial_x u(t,x)) = 0 \quad \text{for } (t,x) \in (0,T) \times \mathcal{M}_i, \; i = 1, \ldots, n, \\ -\partial_t u(t,x) + \max_{i \in I(x)} \{H_{F_i}(x, \partial_x u(t,x))\} = 0 \quad \text{for } (t,x) \in (0,T) \times \Lambda, \\ u(T,x) = \psi(x). \end{cases} \tag{2.6}$$

However, it turns out that the set of dynamics in equation (2.6) is too large. These dynamics may contain velocities that are not *useful* for the evolution of the solution at the interface. This claim is analyzed in the next subsection.

2.2.4 The essential dynamics

We define the notion of *essential dynamics* for each domain, introduced in [41] for stratified Euclidean spaces. For $i = 1, ...n + l$, we define the essential dynamics on each $\overline{\mathcal{M}}_i$ by

$$F_i^\sharp(x) = F_i(x) \cap \mathcal{T}_{\overline{\mathcal{M}}_i}(x), \quad \text{for all } x \in \overline{\mathcal{M}}_i,$$

where $\mathcal{T}_{\overline{\mathcal{M}}_i}(x)$ is the Bouligand tangent cone of $\overline{\mathcal{M}}_i$ at x. Notice that if $x \in \mathcal{M}_i$, we have $F_i^\sharp(x) = F_i(x)$. The associated Hamiltonian is defined as

$$H_{F_i^\sharp}(x,p) = \sup_{q \in F_i^\sharp(x)} \{-\langle p, q\rangle\}.$$

The essential dynamics $F_i^\sharp$ on each domain represent the *inward pointing* velocities of F_i on $\overline{\mathcal{M}}_i$. We suppose that each $F_i^\sharp$ is l.s.c.

$$(H_{ESS}) \quad \text{for all } i = 1, \dots, n+l, \quad F_i^\sharp \text{ is } l.s.c.$$

Hypothesis (H_{ESS}) holds for many number of cases. In particular, if we assume the controllability assumption (CH) to hold for the dynamics, then (H_{ESS}) holds for all stratifications presented in Examples 2.1, 2.2 and 2.3. A discussion about sufficient conditions to ensure (H_{ESS}) is given in Section 2.7.2.

The essential dynamics on $\mathbb{R}^N$ is defined as the union of the essential dynamics on each domain.

$$\forall x \in \mathbb{R}^N, \quad F^\sharp(x) = \bigcup_{i=1}^{n+l} \{\, F_i(x) \cap \mathcal{T}_{\overline{\mathcal{M}}_i}(x) \ : \ x \in \overline{\mathcal{M}}_i \,\}.$$

Its associated Hamiltonian is also defined as usual. For $(x,p) \in \mathbb{R}^N \times \mathbb{R}^N$, we have

$$H_{F^\sharp}(x,p) = \sup_{q \in F^\sharp(x)} \{-\langle p, q\rangle\}.$$

Example 2.5. We consider the stratification of $\mathbb{R}$ defined in Example 2.1. Let $c_i \geq 0$ with $i = 1, 2$ be real positive constants. We define the following dynamics on each branch

$$F_i(x) = [-c_i, c_i] \ \ , i = 1, 2.$$

The resulting HJB system is the *Eikonal equation* on the stratification 2.1. The dynamics at the interface $\mathcal{M}_3$ and the essential dynamics are respectively equal to

$$F_3(.) \equiv \{0\} \quad , \quad F^\sharp(x) = \begin{cases} [-c_i, c_i] & x \in \mathcal{M}_i \ \ i = 1, 2, \\ [-c_2, c_1] & x = 0. \end{cases}$$

Let $T > 0$ be a given time horizon. We consider the following HJB system associated to the dynamics $F_i^\sharp$

$$\begin{cases} -\partial_t u(t,x) + \max\limits_{i \in I(x)} \{ H_{F_i^\sharp}(x, \partial_x u(t,x)) \} = 0 & \text{for } (t,x) \in (0,T) \times \mathbb{R}^N, \\ u(T,x) = \psi(x), \end{cases} \tag{2.7}$$

where $\psi : \mathbb{R}^N \to \mathbb{R}$ is the final cost that satisfies $(H\psi)$.

Notice that in the HJB equation (2.7), if x belongs to the regular part of the stratification (i.e. $x \in \bigcup_{i=1}^{n} \mathcal{M}_i$), then it is the same as the HJB equation (2.5). So equation (2.7) has the following form

$$\begin{cases} -\partial_t u(t,x) + H_{F_i}(x, \partial_x u(t,x)) = 0 \quad \text{for } (t,x) \in (0,T) \times \mathcal{M}_i, \ i = 1, \cdots, n, \\ -\partial_t u(t,x) + \max_{i \in I(x)} \{ H_{F_i^\sharp}(x, \partial_x u(t,x)) \} = 0 \quad \text{for } (t,x) \in (0,T) \times \Lambda, \ i = n+1, \cdots, n+l, \\ u(T,x) = \psi(x). \end{cases} \tag{2.8}$$

Given that we consider a stratified setting, one cannot use the classical notion of viscosity solution. We are going to define a new one that encapsulates the singular nature of the problem.

Definition 2.1. (Viscosity supersolution). Let $u : (0,T] \times \mathbb{R}^N \to \mathbb{R}$ be a l.s.c function. We say that u is a **supersolution** of (2.7) at $(t,x) \in (0,T) \times \mathbb{R}^N$ if and only if there exists $i \in I(x)$ such that for all $\phi \in C^1((0,T) \times \mathbb{R}^N)$, $u - \phi$ attains a local minimum in $(0,T) \times \overline{\mathcal{M}}_i$ at (t,x), we have

$$-\partial_t \phi + H_{F_i^\sharp}(x, \partial_x \phi) \geq 0.$$

Definition 2.2. (Viscosity subsolution). Let $u : (0,T] \times \mathbb{R}^N \to \mathbb{R}$ be a u.s.c function. We say that u is a **subsolution** of (2.7) at $(t,x) \in (0,T) \times \mathbb{R}^N$ if and only if for all $i \in I(x)$, for all $\phi \in C^1((0,T) \times \mathbb{R}^N)$, $u - \phi$ attains a local maximum in $(0,T) \times \overline{\mathcal{M}}_i$ at (t,x), we have

$$-\partial_t \phi + H_{F_i^\sharp}(x, \partial_x \phi) \leq 0.$$

Definition 2.3. (Viscosity solution). u is a viscosity solution of (2.7) if and only if it is both a supersolution and a subsolution and satisfies the final condition

$$u(T,x) = \psi(x), \quad \forall x \in \mathbb{R}^N.$$

The above definitions of viscosity super- and subsolutions can be rewritten using the viscosity sub-gradient and super-gradient (also known as the semijets [3, Page 10] or Dini sub/super gradient [9, Definition 11.18]).

Definition 2.4. (Viscosity sub/super-gradient). We define the viscosity sub- and super-gradient the following way.

- Let $u : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ be a l.s.c function. The viscosity sub-gradient (or subjet) at a point $x \in dom(u)$ is the set,

$$D^- u(x) := \left\{ p \in \mathbb{R}^N \ : \ \liminf_{y \to x} \frac{u(y) - u(x) - \langle p, y - x \rangle}{|y - x|} \geq 0 \right\}.$$

- Similarly, for an u.s.c function $u : \mathbb{R}^N \to \mathbb{R} \cup \{-\infty\}$, the viscosity super-gradient (or superjet) at a point $x \in dom(u)$ is the set,

$$D^+ u(x) := -D^-(-u)(x) = \left\{ p \in \mathbb{R}^N \ : \ \limsup_{y \to x} \frac{u(y) - u(x) - \langle p, y - x \rangle}{|y - x|} \leq 0 \right\}.$$

Remark 2.2.2.

- Let $u : (0, T] \times \mathbb{R}^N \to \mathbb{R}$ be a l.s.c function. We say that u is a **supersolution** of (2.7) at $(t, x) \in (0, T) \times \mathbb{R}^N$ if and only if there exists $i \in I(x)$, such that

$$-\theta + H_{F_i^\sharp}(x, \xi_i) \geq 0 \quad \forall (\theta, \xi_i) \in D^- u_i(t, x),$$

  with $u_i \equiv u$ on $\overline{\mathcal{M}}_i$ and $u_i \equiv +\infty$ elsewhere.

- Let $u : (0, T] \times \mathbb{R}^N \to \mathbb{R}$ be an u.s.c function. u is a **subsolution** of (2.7) at $(t, x) \in (0, T) \times \mathbb{R}^N$ if and only if for all $i \in I(x)$, we have

$$-\theta + H_{F_i^\sharp}(x, \xi) \leq 0 \qquad \forall (\theta, \xi) \in D^+ u_i(t, x),$$

  with $u_i \equiv u$ on $\overline{\mathcal{M}}_i$ and $u_i \equiv -\infty$ elsewhere.

Indeed, if for example $u : (0, T] \times \mathbb{R}^N \to \mathbb{R}$ is a l.s.c function, we set: $u_i \equiv u$ on $\overline{\mathcal{M}}_i$ and $u_i \equiv +\infty$ elsewhere, for any $i = 1, ..., n + l$. Then, for $(t, x) \in (0, T) \times \overline{\mathcal{M}}_i$, we have

$$(\theta, \xi_i) \in D^- u_i(t, x) \iff \exists\, \phi^{(i)} \in C^1((0, T) \times \mathbb{R}^N), \text{ such that } u_i - \phi^{(i)} \text{ attains a local minimum at } (t, x).$$

Since $u_i - \phi^{(i)} \equiv +\infty$ whenever $x \notin \overline{\mathcal{M}}_i$, we get that $\phi^{(i)}$ satisfies the requirements of Definition 2.1. Conversely, if there exists such function ϕ as in Definition 2.1, then $u_i - \phi$ attains a local maximum in $\mathbb{R}^N$ at (t, x). The exact same reasoning holds for subsolutions.

Next we state the main results. In particular, we will show that equation (2.7) has a unique viscosity solution (following Definition 2.3).

2.2.5 Statement of the main results

Theorem 2.6. *Assume* (H_1)*,* (SH)*,* $(H\psi)$*,* (H_{ESS}) *and* (CH)*. Then the HJB equation* (2.7) *has a unique continous solution in the sense of Definition 2.3.*

Theorem 2.7. *(Strong comparison principle). Assume* (H_1)*,* (SH)*,* (H_D) *and* (H_{ESS})*. Let* $u_1, u_2 : (0,T] \times \mathbb{R}^N \to \mathbb{R}$ *be respectively a l.s.c supersolution and an u.s.c subsolution in the sense of Definition 2.3 with* $u_2(T,.) \leq u_1(T,.)$*. Then*

$$u_2(t,x) \leq u_1(t,x) \quad \forall (t,x) \in (0,T] \times \mathbb{R}^N.$$

It is worth-noticing that, unlike the previous literature on the subject [19, 18, 71] or [70, 16, 15], the strong comparison principle stated in the above theorem does not require the subsolution to be continuous nor to have any particular behavior on the interface. The proof of this result will clearly show the importance of the use of essential dynamics with the notion of viscosity as it is defined in Definitions 2.1-2.2 (and more precisely the choice of the test functions in those definitions). Furthermore, the unique viscosity solution in Theorem 2.6 is the *value function* associated to the Mayer optimal control problem with the dynamics $F(.)$. A study of the value function and the associated optimal control problem is presented in Section 2.4. Hypothesis (CH) in Theorem 2.6 is only used to give sufficient conditions for the value function to be continuous. Theorefore, Theorem 2.6 holds if one assumes that the value function is continuous instead of assuming (CH).
The proofs of Theorems 2.6, 2.7 are given in Section 2.4. The proofs will rely on invariance theorems, particularly the strong invariance property. Both invariance theorems are stated in Section 2.3 and proven in Section 2.7.3. Furthermore, we will establish stability results for supersolutions and subsolutions in presence of perturbations of the Hamiltonian in Section 2.5. Section 2.6 is devoted to stating and proving a general convergence result of monotone numerical schemes. The numerical scheme has the following form in each $\overline{\mathcal{M}}_i$, $i = 1, \ldots, n+l$,

$$S_i^h(t_h, x_h, u^h(t_h, x_h), [u^h]_{(t_h,x_h)}\,\} = 0 \quad \text{for } (t_h, x_h) \in (\Pi^{\Delta t} \times \mathcal{G}_i^{\Delta x}),$$

where $\Pi^{\Delta t}$ is time grid, $\mathcal{G}_i^{\Delta x}$ is a spatial grid of $\overline{\mathcal{M}}_i$ and $h = (\Delta t, \Delta x)$ is the step of the grid. We show that under the usual hypotheses of monotonicity, stability and consistency, the numerical scheme converges. This result generalizes the famous convergence theorem of monotone numerical schemes in the classical case, du to Barles and Souganidis [47].

2.2.6 Comparison with existing literature

Recently, control problems and Hamilton Jacobi equations on stratified structures have been investigated in several works. A similar setting to the one considered in this chapter can be found in [16, 15, 18, 19, 71]. The techniques used in [18, 19, 71] are also all based on invariance principles and on the use of the essential

Hamiltonian to describe the behavior of the value function. Here, we investigate further the essential dynamics and its corresponding Hamiltonian. In particular, we show that the principles of invariances (weak and strong) can be fully characterized by using the essential Hamiltonian (for both principles). This result is new, it generalizes to the stratified case the invariance principles known in the literature for a Lipschitz dynamics. As consequence of the invariance principles, particularly the strong invariance principle, we obtain a strong comparison principle for equation (2.7) by assuming further that the essential dynamics $F_i^\sharp(.)$ are l.s.c in their domains. The comparison principle states that for any u_1 u.s.c subsolution of (2.7) and for any u_2 l.s.c supersolution of (2.7), we have $u_1 \leq u_2$ in $\mathbb{R}^N \times (0, T]$. Unlike the results established in [16, 15, 18, 19, 71], the comparison principle does not require any additional controllability assumption nor the continuity of the subsolution around the interface.

The setting of control problems considered in [14, 70] is very close to ours. However, in those papers, the HJB equation considered on the singular set Λ is different from the one we use in (2.7). Indeed, in [14, 70], the Hamiltonian in each stratum is built by using only *local* information with the dynamics defined in the stratum without taking into account the behaviour of the dynamics at the boundary of the stratum. Therefore, the Hamiltonian at the interface does not take into account the information coming from neighboring strata. As consequence, the comparison principle in [14] requires an additional controllability condition and the continuity of subsolutions. The work of [70] is more general, it requires a weaker controllability assumption and gives a comparison between u.s.c subsolutions and l.s.c supersolutions if the subsolution satisfies Ishii's condition or a weak continuity requirement.

To compare our results with those presented in [70], we will analyze different examples. In the first case, we start from a control problem as considered in [70], then we compare on this problem our assumptions and results to those in [70]. Then, we consider two specific cases.

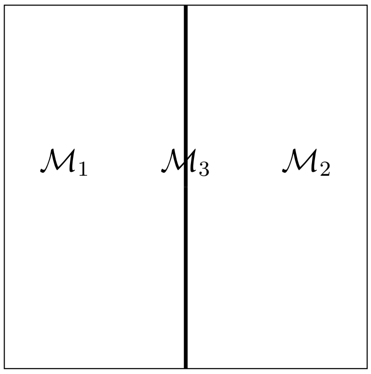
Example 2.8. Notice that the setting in [70] concerns optimal control problems with final and running costs, on a general stratifications of $\mathbb{R}^N$. For a simple comparison, we will restrict ourselves to a framework without a distributed cost. We also suppose that $N = 2$ and the stratification is composed simply of two half-spaces of $\mathbb{R}^2$ separated by a line ($n = 2$ and $l = 1$)

$$\mathbb{R}^N = \mathcal{M}_1 \cup \mathcal{M}_2 \cup \mathcal{M}_3$$

with

$$\mathcal{M}_1 := \{(x_1, x_2) \in \mathbb{R}^2, x_1 < 0\}, \quad \mathcal{M}_2 := \{(x_1, x_2) \in \mathbb{R}^2,\ x_1 > 0\},$$

$$\mathcal{M}_3 := \{(x_1, x_2) \in \mathbb{R}^2,\ x_1 = 0\}.$$



This stratification satisfies (H_1) and it is a *regular* stratification (in the sense of [70]) as well. Let $F(.)$ be an u.s.c dynamics defined on $\mathbb{R}^N$, with

$$F_i(x) = F(x) \cap \mathcal{T}_{\mathcal{M}_i}(x) \quad \text{for } x \in \mathcal{M}_i, \text{ and for } i = 1, 2, 3. \tag{2.9}$$

For $t \in [0, T]$ and $x \in \mathbb{R}^N$, consider the set of trajectories $S_{(t,T)}(x)$ defined as

$$S_{(t,T)}(x) := \{y : [t, T] \to \mathbb{R}^N \ : \ \dot{y}(s) \in F(y(t)) \text{ for a.e. } s \in [t, T], \ y(t) = x\}.$$

We define the *value function* the following way

$$\vartheta(x, t) := \inf\{\psi(y(T)), \ y(.) \in S_{(t,T)}(x)\}.$$

We will discuss the properties of the value function in Section 2.4. In [70], the cost function ψ, the dynamics $F(.)$ and the dynamics $F_i(.)$ satisfy

$$(H_{D,[5]}) \begin{cases} (i) \text{ The dynamics } F \text{ is uniformly bounded on } \mathbb{R}^N. \\ (ii) \text{ For } i = 1, 2, 3, \quad F_i \text{ is uniformly continuous on } \mathcal{M}_i. \\ (iii) \text{ The cost function } \psi \text{ is bounded and uniformly continuous on } \mathbb{R}^N. \end{cases}$$

Besides, a normal controllability assumption is introduced in [70]. In the simple setting of this example, this normal controllability is the following

$(H_{N,[5]})$ For every $x \in \mathcal{M}_3$, and $r > 0$, there exist $C > 0$ and $\delta > 0$ such that

$$H_{F_3}(y, p) \geq \delta|p_2| - C(1 + |p_1|), \ \ \forall y \in B(x, r) \cap \mathcal{M}_3, \ p = (p_1, p_2) \in \mathcal{T}_{\mathcal{M}_3}(x) \times \mathcal{T}^{\perp}_{\mathcal{M}_3}(x).$$

The set $\mathcal{T}^{\perp}_{\mathcal{M}_3}(x)$ is the orthogonal space to the tangent space $\mathcal{T}_{\mathcal{M}_3}(x)$ in $\mathbb{R}^N$. So $\mathbb{R}^N$ is the direct sum of the tangent space and the normal space

$$\mathbb{R}^N = \mathcal{T}_{\mathcal{M}_3}(x) \oplus \mathcal{T}^{\perp}_{\mathcal{M}_3}(x).$$

According to [70], under assumptions $(H_{D,[5]})$ and $(H_{N,[5]})$, the value function ϑ is bounded and continuous. Moreover, ϑ is a supersolution of the equation

$$-\partial_t v(x, t) + H_F(x, \partial_x v(x, t)) \geq 0, \tag{2.10a}$$

and ϑ is a subsolution to the system of equations

$$-\partial_t v(x,t) + H_{F_i}(x, \partial_x v(x,t)) \le 0 \quad \forall x \in \mathcal{M}_i, \text{ and for } i = 1,2,3. \tag{2.10b}$$

Furthermore, let v_2 be a l.s.c supersolution of (2.10a), and let v_1 be a u.s.c subsolution of (2.10b) satisfying one of the following conditions:

- (i) v_1 is continuous on $\mathcal{M}_3$,
- (ii) v_1 satisfies Ishii's condition, i.e, u is subsolution to

  $$-\partial_t v(x,t) + H_*(x, \partial_x v(x,t)) \le 0, \quad \text{for } x \in \mathcal{M}_3,$$

  where H_* is the l.s.c envelope of H:

  $$H_*(x,p) := \liminf_{(y,q)\to(x,p)} H_F(y,q).$$

Then, under assumptions $(H_{D,[5]})$ and $(H_{N,[5]})$, and by [70, Theorem 4.1] we have $v_1 \le v_2$ on $\mathbb{R}^N \times [0,T]$. The result is even more precise and provides a *local* strong comparison result.

Now, let us see how our work differs from [70]. Our assumptions (SH), (H_D) and (H_{ESS}) require the dynamics $F_i(.)$ to be Lipschitz continuous on bounded sets of $\mathcal{M}_i$ with a linear growth. No boundedness is required. Furthermore, the essential dynamics $F_i^\sharp(.)$ are l.s.c. on $\overline{\mathcal{M}}_i$. Under assumption (H_{ESS}), the result of Theorem 2.7 provides a comparison between subsolutions and supersolutions of the following HJB system

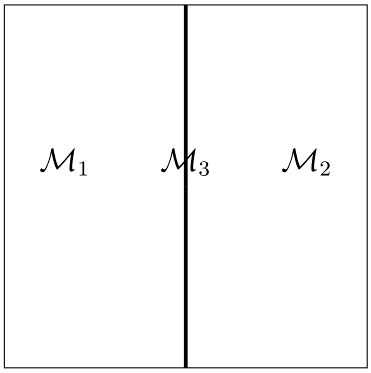
$$\begin{cases} -\partial_t u(t,x) + H_{F_i}(x, \partial_x u(t,x)) = 0 \quad \text{for } (t,x) \in (0,T) \times \mathcal{M}_i, \ i = 1,2, \\ -\partial_t u(t,x) + \max\limits_{i=1,2,3} \{ H_{F_i^\sharp}(x, \partial_x u(t,x)) \} = 0 \quad \text{for } (t,x) \in (0,T) \times \mathcal{M}_3, \\ u(T,x) = \psi(x). \end{cases} \tag{2.11}$$

This result does not require a controllability assumption and it states that for every u.s.c subsolution u_1 to (2.11) and for every l.s.c supersolution of (2.11), we have $u_1 \le u_2$ on $(0,T] \times \mathbb{R}^N$.

Note that in the system (2.11), we use the same Hamiltonian $\max\limits_i H_{F_i^\sharp}$ on $\mathcal{M}_3$ to describe the sub and super optimality. This Hamiltonian includes all the dynamics of trajectories starting from a position in $\mathcal{M}_3$. Therefore, the Hamiltonian $\max\limits_i H_{F_i^\sharp}$ *contains more information* than Hamiltonian H_{F_3}. Theorem 2.7 gives a comparison principle for (2.11) without requiring additional information on subsolutions.

The controllability assumption (CH), which is stronger than $(H_{N,[5]})$ is used only to prove that the value function is continuous (see Section 2.4). If it is continuous then we prove that it is the viscosity solution of (2.11). The continuity of the value function can be obtained in some cases without assumption $(H_{N,[5]})$ or (CH).

Example 2.9. Consider the same stratification as in Example 2.8.



The dynamics $F_1(.)$ and $F_2(.)$ defined in $\mathcal{M}_1$ and $\mathcal{M}_2$ respectively, with

$$F_1(x) := \mathbb{B}(0,1) \text{ in } \mathcal{M}_1, \quad F_2(x) := \mathbb{B}(0,2) \;\text{ in } \mathcal{M}_2.$$

We define $F(.)$ as the Filippov regularization of $F_1(.)$ and $F_2(.)$. It is given by

$$F(x) = \begin{cases} F_1(x) = \mathbb{B}(0,1) & \text{if } x \in \mathcal{M}_1, \\ F_2(x) = \mathbb{B}(0,2) & \text{if } x \in \mathcal{M}_2, \\ \mathbb{B}(0,2) & \text{if } x \in \mathcal{M}_3. \end{cases}$$

The essential dynamics are given by

$$F_1^\sharp(.) = F_1(.) \;\text{ in } \mathcal{M}_1 \quad \text{and} \quad F_1^\sharp(.) = [-1,0]\times[-1,1] \;\text{ in } \mathcal{M}_3 = \overline{\mathcal{M}}_1 \setminus \mathcal{M}_1,$$

$$F_2^\sharp(.) = F_2(.) \;\text{ in } \mathcal{M}_2 \quad \text{and} \quad F_2^\sharp = [0,2]\times[-2,2] \;\text{ in } \mathcal{M}_3 = \overline{\mathcal{M}}_2 \setminus \mathcal{M}_2,$$

$$F_3^\sharp(.) = \{0\}\times[-2,2] \;\text{ in } \overline{\mathcal{M}}_3 = \mathcal{M}_3.$$

The dynamics of this example satisfy assumptions $(SH), (H_D)$, (H_{ESS}) and (CH). Furthermore, it satisfies hypotheses $(H_{D,[5]})$ and $(H_{N,[5]})$ from [70]. Hence, our results give a comparison principle for any u.s.c subsolution u_1 and any l.s.c.supersolution u_2, in the sense of Theorem 2.7, of the following equation

$$\begin{cases} -\partial_t v(t,x) + |\partial_x v(t,x)| = 0 \quad \text{in } (0,T)\times\mathcal{M}_1, \\ -\partial_t v(t,x) + 2|\partial_x v(t,x)| = 0 \quad \text{in } (0,T)\times\mathcal{M}_2, \\ -\partial_t v(t,x) + \max\Big(\max\limits_{\theta\in[-\frac{\pi}{2},\frac{\pi}{2}]} \langle \partial_x v(t,x), \begin{pmatrix}\cos(\theta)\\ \sin(\theta)\end{pmatrix}\rangle, \quad 2 \max\limits_{\theta\in[\frac{\pi}{2},\frac{3\pi}{2}]} \langle \partial_x v(t,x), \begin{pmatrix}\cos(\theta)\\ \sin(\theta)\end{pmatrix}\rangle \Big) = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0,T)\times\mathcal{M}_3, \\ v(T,x) = \psi(x). \end{cases}$$

Moreover, by Theorem 2.6, the above equation admits a unique continuous viscosity solution. The viscosity solution is the value function ϑ associated to the optimal control problem defined using the dynamics $F(.)$ and the final cost ψ (see Section 2.4). By [70, Theorem 4.1], ϑ is also the unique continuous function that satisfies $v(T,x) = \psi(x)$ and is both a viscosity supersolution of

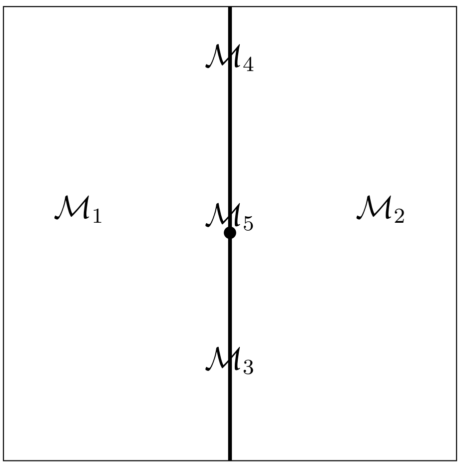
$$\begin{cases} -\partial_t v(t,x) + |\partial_x v(t,x)| \geq 0 & \text{in } \mathcal{M}_1, \\ -\partial_t v(t,x) + 2|\partial_x v(t,x)| \geq 0 & \text{in } \mathcal{M}_2 \cup \mathcal{M}_3, \end{cases} \tag{2.12}$$

and a viscosity subsolution of

$$\begin{cases} -\partial_t v(t,x) + |\partial_x v(t,x)| \leq 0 & \text{in } \mathcal{M}_1, \\ -\partial_t v(t,x) + 2|\partial_x v(t,x)| \leq 0 & \text{in } \mathcal{M}_2, \\ -\partial_t v(t,x) + \max\Big(\langle \partial_x v(t,x), \begin{pmatrix} 0 \\ 2 \end{pmatrix} \rangle, \langle \partial_x v(t,x), \begin{pmatrix} 0 \\ -2 \end{pmatrix} \rangle\Big) \leq 0 & \text{in } \mathcal{M}_3. \end{cases} \tag{2.13}$$

The comparison theorem in [70] allows also to compare any l.s.c supersolution of (2.12) and any u.s.c subsolution of (2.13).

Example 2.10. In this example, we consider a different stratification of $\mathbb{R}^2$ in the following way



$$\mathcal{M}_1 := \{x = (x_1, x_2) \in \mathbb{R}^2, x_1 < 0\}, \quad \mathcal{M}_2 := \{x = (x_1, x_2) \in \mathbb{R}^2, \ x_1 > 0\},$$

$$\mathcal{M}_3 := \{0\} \times]-\infty, 0[, \quad \mathcal{M}_4 := \{0\} \times]0, +\infty[, \ \text{and } \mathcal{M}_5 := \{0\}.$$

The stratification satisfies (H_1). Consider the dynamics $F_1(.)$ and $F_2(.)$ on $\mathcal{M}_1$ and $\mathcal{M}_2$ respectively defined as follows:

$$F_1(x) := c_1 \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix} \text{ on } \mathcal{M}_1, \quad F_2(x) := c_2 \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix} \text{ on } \mathcal{M}_2.$$

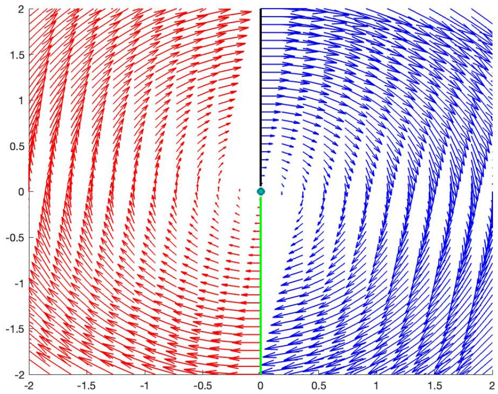


Figure 2.4: In red the dynamics $F_1(.)$ and in blue the dynamics $F_2(.)$.

The convexified dynamics $F(.)$ defined on $\mathbb{R}^2$ that coincide with $F_1(.)$ on $\mathcal{M}_1$ and with $F_2(.)$ on $\mathcal{M}_2$ is equal to

$$F(x) = \begin{cases} F_1(x), & x \in \mathcal{M}_1, \\ F_2(x), & x \in \mathcal{M}_2, \\ [\min(c_1, c_2), \max(c_1, c_2)]e_{x_1} & \text{elsewhere.} \end{cases}$$

Assume that $c_1, c_2 > 0$. Then we have

$$F_3(.) = \emptyset \text{ in } \overline{\mathcal{M}_3}, \quad F_4(.) = \emptyset \text{ in } \overline{\mathcal{M}_4}, \quad F_5(.) = \{0\} \text{ in } \overline{\mathcal{M}_5}.$$

The essential dynamics are defined by

$$F_1^\sharp(x) = \begin{cases} F_1(x), & \text{in } \mathcal{M}_1, \\ F_1(x), & \text{in } \mathcal{M}_3, \\ \emptyset, & \text{in } \mathcal{M}_4, \\ \{0\}, & \text{in } \mathcal{M}_5. \end{cases} \quad F_2^\sharp(x) = \begin{cases} F_2(x) & \text{in } \mathcal{M}_2, \\ \emptyset, & \text{in } \mathcal{M}_3, \\ F_2(x), & \text{in } \mathcal{M}_4, \\ \{0\}, & \text{in } \mathcal{M}_5. \end{cases}$$

$$F_3^\sharp(x) = \begin{cases} \emptyset, & \text{in } \mathcal{M}_3, \\ \{0\} & \text{in } \mathcal{M}_5. \end{cases} \quad F_4^\sharp(x) = \begin{cases} \emptyset, & \text{in } \mathcal{M}_4, \\ \{0\} & \text{in } \mathcal{M}_5. \end{cases} \quad F_5^\sharp(0) = \{0\}.$$

We consider the following HJB system

$$\begin{cases} -\partial_t u(t,x) + \sup\limits_{\nu \in F_i(x)} \{-\langle \partial_x u(t,x), \nu \rangle\} = 0 & \text{for } (t,x) \in (0,T) \times \mathcal{M}_i,\ i = 1,2, \\ -\partial_t u(t,x) + \sup\limits_{\nu \in F_1(x)} \{-\langle \partial_x u(t,x), \nu \rangle\} = 0 & \text{for } (t,x) \in (0,T) \times \mathcal{M}_3, \\ -\partial_t u(t,x) + \sup\limits_{\nu \in F_2(x)} \{-\langle \partial_x u(t,x), \nu \rangle\} = 0 & \text{for } (t,x) \in (0,T) \times \mathcal{M}_4, \\ -\partial_t u(t,0) = 0 \quad \text{for } t \in (0,T), \\ u(T,x) = \psi(x) \quad \text{for } x \in \mathbb{R}^N. \end{cases}$$

The dynamics of this example satisfy the assumptions (SH), (H_D) and (H_{ESS}) but the controllability assumptions are not satisfied. However, whenever we choose ψ Lipschitz continuous and bounded, we can prove there exists a unique continuous solution to the above HJB system and it is the value function associated to the optimal control problem defined with $F(.)$ (see Section 2.4). Moreover, Theorem 2.7 provides a strong comparison principle for the above HJB system.

2.3 Invariance Principles

The following results are known as weak and strong invariance properties. They are known in the classical case, when the dynamics $F(.)$ is Lipschitz continuous, see [9, Chapter 11]. For the stratified case, the first attempt to prove these results was in [41], using the essential Hamiltonian strategy. Although their intuition was correct, the proximal normal cone used is the same as the one in the classical case (see [9, Chapter 11]), which does not take into account the geometry of the problem. We start by recalling the definitions of some nonsmooth analysis tools and the definitions of weak and strong invariance.

Definition 2.5. (Proximal sub-gradient and super-gradient).

- Let $u : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ be a l.s.c function. We say that ζ is proximal sub-gradient at a point $x \in dom(u)$ for some $\sigma = \sigma(x, \zeta)$ and some neighborhood $V = V(x, \zeta)$ of x if we have

$$u(y) - u(x) + \sigma|y - x|^2 \geq \langle \zeta, y - x \rangle, \ \ \forall\, y \in V.$$

  The collection of such ζ form the proximal sub-gradient. It is denoted $\partial_p u(x)$.

- Similarly, Let $u : \mathbb{R}^N \to \mathbb{R} \cup \{-\infty\}$ be an u.s.c function. We say that ζ is proximal super-gradient at a point $x \in dom(u)$ for some $\sigma = \sigma(x, \zeta)$ and some neighborhood $V = V(x, \zeta)$ of x if we have

$$u(y) - u(x) + \sigma|y - x|^2 \leq \langle \zeta, y - x \rangle, \ \ \forall\, y \in V.$$

  The collection of such ζ forms the proximal super-gradient. It is denoted $\partial^p u(x)$. We also have the property $\partial^p u(x) = -\partial_p(-u)(x)$.

Definition 2.6. (Proximal normal cone).
Let $S \subseteq \mathbb{R}^N$ be a closed set and $x \in S$. A vector ζ is a proximal normal to the closed set S at the point x if there exists $\sigma > 0$ such that

$$\langle \zeta, y - x \rangle \leq \frac{|\zeta|}{2\sigma}|y - x|^2 \ \ \forall y \in S.$$

The set of all proximal normal vectors at x is denoted by $N_S^p(x)$.

Definition 2.7. (Weak invariance).
Let $\Gamma : \mathbb{R}^N \rightsquigarrow \mathbb{R}^N$ be a set-valued map. Let $S \subseteq \mathbb{R}^N$ be a closed set. We say that (S, Γ) is weakly invariant provided that for $x \in S$ and $t \in [0, T]$, there exists $y(.)$ a solution of $(DI)_\Gamma(t, x)$ such that $y(\tau) \in S$ for all τ in $[t, T]$.

Definition 2.8. (Strong invariance).
Let $\Gamma : \mathbb{R}^N \rightsquigarrow \mathbb{R}^N$ be a set-valued map. Let $S \subseteq \mathbb{R}^N$ be a closed set. We say that (S, Γ) is strongly invariant provided that for $x \in S$, $t \in [0, T]$ and every $y(.)$ a solution of $(DI)_\Gamma(t, x)$, we have $y(\tau) \in S$ for all τ in $[t, T]$.

Theorem 2.11. *Assume (H_1), (SH) and (H_D). Let S be a closed set of $\mathbb{R}^N$. We denote by $S_i := \overline{\mathcal{M}}_i \cap S$. The following assertions are equivalent:*

- *(i) (S, F) is weakly invariant;*
- *(ii) $\forall x \in S,\ \exists\, i \in I(x)\ :\ \forall \eta_i \in N^p_{S_i}(x), \quad H_{F_i}(x, \eta_i) \geq 0$;*
- *(iii) $\forall x \in S,\ \exists\, i \in I(x)\ :\ \forall \eta_i \in N^p_{S_i}(x), \quad H_{F_i^\sharp}(x, \eta_i) \geq 0$.*

Theorem 2.12. *Assume (H_1), (SH) and (H_D). Let S be a closed set of $\mathbb{R}^N$. We denote by $S_i := \overline{\mathcal{M}}_i \cap S$. The following assertions are equivalent:*

- *(i) (S, F) is strongly invariant;*
- *(ii) $\forall x \in S, \quad \forall i \in I(x), \quad \forall \eta_i \in N^p_{S_i}(x), \quad H_{F_i^\sharp}(x, -\eta_i) \leq 0$.*

The complete proofs of Theorems 2.11 and 2.12 are given in Section 2.7.3.

In [41], the authors tried to establish similar invariance principles in the stratified case using the essential Hamiltonian. In particular, for the strong invariance principle, if we assume that (H_1), (SH) and (H_D) hold, [41, Theorem 5.1] states that (S, F) is strongly invariant for some closed set $S \subseteq \mathbb{R}^N$ if and only if

$$\forall x \in S,\ \forall \xi \in N^p_S, \quad H_{F^\sharp}(x, -\xi) \leq 0. \tag{2.14}$$

However, the sufficient implication fails to be true in general. Here is a counterexample.

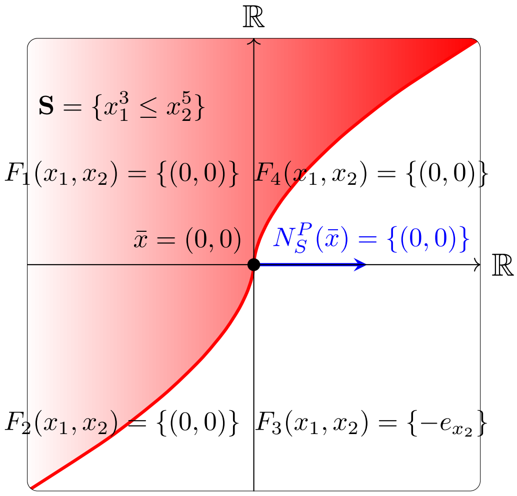


Figure 2.5: Counterexample with a stratification in $\mathbb{R}^N$, with $N = 2$, $n = 4$ and $l = 5$.

We are given a stratification as follows

$$\mathcal{M}_1 = \{(x_1,x_2) \in \mathbb{R}^2 \ : \ x_1 < 0 \ \& \ x_2 > 0\} \quad \mathcal{M}_2 = \{(x_1,x_2) \in \mathbb{R}^2 \ : \ x_1 < 0 \ \& \ x_2 < 0\},$$

$$\mathcal{M}_3 = \{(x_1,x_2) \in \mathbb{R}^2 \ : \ x_1 > 0 \ \& \ x_2 < 0\} \quad \mathcal{M}_4 = \{(x_1,x_2) \in \mathbb{R}^2 \ : \ x_1 > 0 \ \& \ x_2 > 0\},$$

$$\mathcal{M}_5 = (0,+\infty)e_{x_1} \quad \mathcal{M}_6 = (-\infty,0)e_{x_1} \quad \mathcal{M}_7 = (0,+\infty)e_{x_2} \quad \mathcal{M}_8 = (-\infty,0)e_{x_2} \quad \mathcal{M}_9 = \{0\}.$$

Take S to be the closed set

$$S = \{(x_1,x_2) \in \mathbb{R}^2 \ : x_1^3 \leq x_2^5\},$$

represented in red in Figure 2.5 and consider the following dynamics

$$F_1(x_1,x_2) = F_2(x_1,x_2) = F_4(x_1,x_2) = F_5(x_1,x_2) = \{(0,0)\},$$

$$F_6(x_1,x_2) = F_7(x_1,x_2) = F_9(x_1,x_2) = \{(0,0)\}, \quad F_3(x_1,x_2) = F_8(x_1,x_2) = \{-e_{x_2}\}.$$

Clearly, (S,F) is not strongly invariant since $F(0,x_2) = co\{0,-e_{x_2}\}$ if $y \leq 0$ and the trajectory

$$\tilde{Z}(s) = (0, t-s) \in S_{(t,T)}(0,0)$$

is a trajectory of F that starts at $\bar{x} = (0,0) \in S$, but $\tilde{Z}(.) \not\subset S$. On the other hand, since the proximal normal cone to S at $\bar{x} = (0,0)$ is equal to $N_S^p(\bar{x}) = \{(0,0)\}$, the Hamiltonian inequality (2.14) is therefore verified at $\bar{x} = (0,0)$. For the remaining points of S, the Hamiltonian inequality (2.14) is trivially verified since the dynamics F is reduced to $\{(0,0)\}$. In conclusion, we have shown in this example that the Hamiltonian inequality

$$\forall x \in S, \ \forall \xi \in N_S^p, \quad H_{F^\sharp}(x,-\xi) \leq 0,$$

is verified for all the points in S. However, (S, F) is not strongly invariant.

We can also verify that (S, F) is not strongly invariant using the new characterization of strong invariance given in Theorem 2.12. Indeed we have

$$S_8 := S \cap \overline{\mathcal{M}}_8 = S \cap (-\infty, 0]e_{x_2} = \{(0,0)\}.$$

Therefore, we have that the proximal normal to S_i at $\bar{x} := (0,0)$ is equal to $\mathbb{R}^2$. Moreover we have $F_8^\sharp(\bar{x}) = co\{0, -e_{x_2}\} = [0, -e_{x_2}]$. Thus

$$H_{F_8^\sharp}(\bar{x}, -N_{S_8}^p(\bar{x})) \geq \langle -e_{x_2}, -e_{x_2} \rangle = 1 > 0.$$

This shows that the sufficient implication in Theorem 2.12 is not satisfied in this example.

2.4 Proof of Theorems 2.6 and 2.7.

2.4.1 Optimal control problem and the value function

In this section, we consider the differential inclusion assotiated with the set-valued map $F(.)$

$$(DI)_F(t,x) \ : \left\{ \begin{array}{l} \dot{y}(s) \in F(y(s)), \ s \in [t,T] \text{ a.e.,} \\ y(t) = x. \end{array} \right.$$

Since F is u.s.c with nonempty, convex and compact images, then the above differential inclusion admits Lipschitz solutions for any $(t,x) \in [0,T] \times \mathbb{R}^N$. Furthermore, the set of solutions is compact in the topology of uniform convergence [76, Theorem 1, pp 60]. We denote by $S_{(t,T)}(x)$ the set of solutions of the differential inclusion associated to $F(.)$:

$$S_{(t,T)}(x) := \left\{ y_{(t,x)}(.) \in W^{1,1}([t,T];\mathbb{R}^N) \ : \left\{ \begin{array}{l} \dot{y}_{(t,x)}(s) \in F(y_{(t,x)}(s)), \ s \in [t,T], \text{ a.e.,} \\ y(t) = x. \end{array} \right. \right\}.$$

We consider the following Mayer optimal control problem defined for $(t,x) \in [0,T] \times \mathbb{R}^N$ by

$$\left\{ \begin{array}{ll} \inf & \psi(y_{(t,x)}(T)) \\ \text{such that} & \dot{y}_{(t,x)}(s) \in F(y_{(t,x)}(s)), \ s \in [t,T], \text{ a.e.,} \\ & y_{(t,x)}(t) = x, \end{array} \right. \tag{2.15}$$

where the infimum is taken over all trajectories $y_{(t,x)}(.) \in S_{(t,T)}(x)$ and it is reached. Next, we consider the value function associated to the optimal control problem defined on $(t,x) \in [0,T] \times \mathbb{R}^N$ by

$$\vartheta(t,x) = \inf\{ \ \psi(y_{(t,x)}(T)) \ , \ \ y_{(t,x)}(.) \in S_{(t,T)}(x) \ \}.$$

We now proceed to define some properties of the value function.

Definition 2.9. Let $u : (0,T] \times \mathbb{R}^N \to \mathbb{R}$ be a function. u is said to enjoy

- **the super-optimality property** if for all $(t,x) \in (0,T] \in \mathbb{R}^N$, there exists $y_{(t,x)}(.) \in S_{(t,T)}(x)$ such that

$$u(t, y_{(t,x)}(t)) \geq u(s, y_{(t,x)}(s)),\ \forall s \in [t,T];$$

- **the sub-optimality property** if for all $(t,x) \in (0,T] \in \mathbb{R}^N$, for all $y_{(t,x)}(.) \in S_{(t,T)}(x)$ we have

$$u(t, y_{(t,x)}(t)) \leq u(s, y_{(t,x)}(s)),\ \forall s \in [t,T].$$

As in the classical case, the value function ϑ satisfies the Dynamic programming principle, which corresponds to the super-optimality and the sub-optimality properties.

Proposition 2.12.1. *([18, Proposition 3.1]). Assume (H_1) and $(H\psi)$. Then the value function satisfies the super-optimality and the sub-optimality properties.*

The next proposition states that the controllability hypothesis (CH) is a sufficient condition to ensure that the value function is locally Lipschitz continuous.

Proposition 2.12.2. *Suppose (H_1), (CH) and $(H\psi)$ hold. Then, $\vartheta : [0,T] \times \mathbb{R}^N \longrightarrow \mathbb{R}$ is locally Lipschitz continuous.*

Proof. From the controllability assumption (CH), there exists a neighborhood of Λ (the interfaces), denoted by $V := \Lambda + \varepsilon\mathbb{B}$, and there exists $r > 0$, such that for all $x \in V$, we have $r\mathbb{B} \subset F(x)$.
First, we prove that $\vartheta(t,.)$ is locally Lipschitz. Let $x, z \in \mathbb{R}^N$. Suppose that $x, z \in V$. Let M be the local supremum bound of F and L_ψ the Lipschitz constant of the final cost ψ. Without loss of generality, we suppose $\vartheta(t,x) \geq \vartheta(t,z)$.
Let $y_{t,z}(.) \in S_{(t,T)}(z)$ such that $\vartheta(t,z) = \psi(y_{t,z}(T))$. Set

$$h = \frac{\mid x - z \mid}{r} \quad \text{and} \quad \xi(s) = x + r\frac{z - x}{\mid x - z \mid}(s - t) \quad \text{for } s \in [t, t+h].$$

We define :

$$\tilde{y}(s) = \begin{cases} \xi(s) & \text{for } s \in [t, t+h] \\ y_{t,z}(s-h) & \text{for } s \in [t+h, T] \end{cases}$$

It is easy to see that $\tilde{y}(.)$ is an F-trajectory. So we get

$$\begin{aligned} \vartheta(t,x) - \vartheta(t,z) &\leq \psi(\tilde{y}(T)) - \psi(y_{t,z}(T)) \\ &\leq L_\psi |\tilde{y}(T) - y_{t,z}(T)| \\ &= L_\psi |y_{t,z}(T-h) - y_{t,z}(T)| \\ &\leq L_\psi M h \;=\; L_\psi \frac{M}{r}|x - z|. \end{aligned}$$

Suppose now for example $z \notin V$ (we can do the same reasoning on x instead). Then by taking x and z close enough to each other, there exists $i \in \{1, ..., n\}$ such that $x, z \in \mathcal{M}_i$. Let $y_{t,z}(.) \in S_{(t,T)}(z)$ such that $\vartheta(t, z) = \psi(y_{t,z}(T))$.

Suppose $y_{t,z}(.)$ crosses the boundary of $\mathcal{M}_i$. Let $t_0 \in [t, T]$ be such that

$$y_{t,z}([t, t_0]) \subset \mathcal{M}_i, \text{ and } y_{t,z}(t_0) \in V.$$

Let $y_{t,x}(.) \in S_{(t,T)}(x)$ such that $y_{t,x}([t, t_0]) \subset \mathcal{M}_i$. We have

$$|y_{t,x}(t_0) - y_{t,z}(t_0)| \leq e^{Mt_0}|x - z| \leq e^{MT}|x - z|$$

Since they are both F_i-trajectories, see [77, Theorem 4.3.11]. Furthermore, we can also suppose that

$$y_{t,x}(t_0) \in V \cap \mathcal{M}_i.$$

We can always find such a trajectory when x and z are close enough and $y_{t,z}(t_0) \in V$. Set $h = \dfrac{| y_{t,z}(t_0) - y_{t,x}(t_0) |}{r}$ and $\xi(s) = y_{t,x}(t_0) + r\dfrac{y_{t,z}(t_0) - y_{t,x}(t_0)}{| y_{t,z}(t_0) - y_{t,x}(t_0) |}(s - t_0)$ for $s \in [t_0, t_0 + h]$. We define

$$\tilde{y}(s) = \begin{cases} y_{t,x}(s) & \text{for } s \in [t, t_0] \\ \xi(s) & \text{for } s \in [t_0, t_0 + h] \\ y_{t,z}(s - h) & \text{for } s \in [t_0 + h, T] \end{cases}$$

It is easy to see that $\tilde{y}(.)$ is an F-trajectory. So we get

$$\begin{aligned}\vartheta(t, x) - \vartheta(t, z) \leq \psi(\tilde{y}(T)) - \psi(y_{t,z}(T)) &\leq L_\psi|\tilde{y}(T) - y_{t,z}(T)| = L_\psi|y_{t,z}(T - h) - y_{t,z}(T)| \\ &\leq L_\psi M h \\ &= L_\psi \frac{M}{r}|y_{t,z}(t_0) - y_{t,x}(t_0)|, \\ &\leq L_\psi e^{MT}\frac{M}{r}|x - z|.\end{aligned}$$

If the trajectory $y_{t,z}(.)$ does not cross the boundary of $\mathcal{M}_i$, then, it is an F_i-trajectory. Furthermore, we can always find a F-trajectory $y_{t,x}(.)$ that stays in $\mathcal{M}_i$ by the controllability assumption. Indeed one can always choose a trajectory with zero velocity once it reaches the neighborhood V. So $y_{t,x}(.)$ is also an F_i-trajectory. Hence the result follows again from the classical case, see [77, theorem 4.3.11].

This finishes the proof of $\vartheta(t, .)$ is locally Lipschitz.

Now we prove that ϑ is locally Lipschitz w.r.t the time variable. Let $x \in \mathbb{R}^N$. Let $t, s \in [0, T]$ such that $t < s$. By the super-optimality property, there exists $y(.) \in S_{(t,T)}(x)$ such that $\vartheta(t, x) = \vartheta(s, y(s))$. Then

$$|\vartheta(t, x) - \vartheta(s, x)| = |\vartheta(s, y(s)) - \vartheta(s, x)| \leq |\vartheta(s, y(s)) - \vartheta(s, y(t))|.$$

Since both $\vartheta(s, .)$ and $y(.)$ are locally Lipschitz, then from the expression above, $\vartheta(., x)$ is locally Lipschitz. This ends the proof. ■

The next proposition shows that F-trajectories are the same as $F^\sharp$-trajectories. This implies that the essential dynamics completely characterize the optimal control problem 2.15.

Proposition 2.12.3. *([41, Proposition 2.1]).*
Let $(t,x) \in [0,T] \times \mathbb{R}^N$ and let $y : [t,T] \to \mathbb{R}^N$ be an absolutely continuous arc. Then the following statements are equivalent.

- *(i) $y(.)$ verifies*
$$\begin{cases} \dot{y}(s) \in F(y(s)), \ s \in [t,T] \ a.e., \\ y(t) = x. \end{cases}$$
- *(ii) $y(.)$ verifies*
$$\begin{cases} \dot{y}(s) \in F^\sharp(y(s)), \ s \in [t,T] \ a.e., \\ y(t) = x. \end{cases}$$
- *(iii) For each $k \in \{1, ..., n+l\}$, $y(.)$ satisfies $y(t) = x$ and*
$$\dot{y}(s) \in F_k^\sharp(y(s)), \ s \in [t,T] \ a.e., \quad \text{whenever } y(s) \in \mathcal{M}_k.$$

Proof. It is clear that the implications $(iii) \implies (ii) \implies (i)$ are immediately verified since $F_k(.) \subseteq F^\sharp(.) \subseteq F(.)$ on each $\mathcal{M}_k$.

Now, suppose (i) is verified and let us show that it implies (iii). Let $y(.) \in S_{(t,T)}(x)$. For $k \in \{1, ..., n+l\}$, let

$$J_k := \{s \in [t,T] \ : \ y(s) \in \mathcal{M}_k\}.$$

Suppose that $\mathscr{L}(J_k) > 0$, with $\mathscr{L}$ being the Lebesgue measure on $\mathbb{R}$. We set

$$\tilde{J}_k := \{s \in J_k \ : \ \dot{y}(s) \text{ exists in } F(y(s)) \text{ and } s \text{ is a Lebesgue point of } J_k\}.$$

Clearly $\mathscr{L}(J_k) = \mathscr{L}(\tilde{J}_k)$. Let $s \in \tilde{J}_k$. Since s is a Lebesgue point, then there exists a sequence $(s_n)_n \subset J_k$ such that $s_n \to s$ and $s_n \neq s$ for all n. Since $y(s_n) \in \mathcal{M}_k$, we have
$$\dot{y}(s) = \lim_{s_n \to s} \frac{y(s_n) - y(s)}{s_n - s} \in \mathcal{T}_{\mathcal{M}_k}(y(s)) = F_k^\sharp(y(s)).$$
Therefore, we get
$$\dot{y}(s) \in F(y(s)) \cap \mathcal{T}_{\mathcal{M}_k}(y(s)),$$
which is the required result. ■

The above proposition shows in particular that the optimal control problem could be defined using F or $F^\sharp$ or $F_i^\sharp$, $i = 1, \dots, n+l$. The latter dynamics are more suitable to our setting since it gives us the link with the HJB equation (2.7) we want to solve.

2.4.2 The super-optimality and supersolution property

In this section, we characterize functions that are supersolutions of equation (2.7) with the super-optimality property (Definition 2.9). The characterization using the Hamiltonian H_F is standard in the literature since the set-valued map F satisfies the usual hypotheses (upper semi-continuity with nonempty, convex and compact images). In here, we will prove a more general result. We show that supersolutions are characterized using the Hamiltonians $H_{F_i^\sharp}$, $i = 1, \dots, n+l$, in the viscosity sense given in Definition 2.1. We recall that for a l.s.c function $u : (0,T] \times \mathbb{R}^N \to \mathbb{R}$ and $i = 1, \dots, n+l$, we define the function $u_i : (0,T] \times \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ by

$$u_i \equiv u \text{ on } \overline{\mathcal{M}}_i \text{ and } u_i \equiv +\infty \text{ elsewhere.}$$

Theorem 2.13. *Suppose (H_1), (SH) and (H_D) hold. Let $u : (0,T] \times \mathbb{R}^N \to \mathbb{R}$ be l.s.c. The following assertions are equivalent:*

- *(i) u is a supersolution of (2.7) in the sense of Definition 2.1,*
- *(ii) u satisfies the super-optimality property.*

Proof. The fact that (i) is equivalent to

$$-\theta + H_F(x,\xi) \geq 0 \qquad \forall (\theta,\xi) \in D^- u(t,x), \tag{2.16}$$

is well known since $F(.)$ is u.s.c with nonempty, convex and compact images. For more on this, see [18, Proposition 3.5] or [9, Chapter 19].

Moreover, it is obvious from this that $(i) \Longrightarrow (ii)$ since $H_{F_i^\sharp}(.,.) \leq H_{F_i}(.,.) \leq H_F(.,.)$ and $D^- u(.,.) \subset D^- u_i(.,.)$, for all $i \in [1, n+l]$.

It remains to prove $(ii) \Longrightarrow (i)$. Let $y(.) : [t,T] \to \mathbb{R}^N$ be a trajectory solution of $(DI)_F(t,x)$ such that the super-optimality property holds in $y(.)$. We claim the following.

Claim: there exists $j \in I(x)$ such that there exists a sequence $(t_n)_n$, $t_n \downarrow t$ and $x_n := y(t_n) \in \mathcal{M}_j$, so that $\dfrac{x_n - x}{t_n - t} \to \nu$ and $\nu \in F_j^\sharp(x)$.

Deferring the proof of the claim, let $\phi \in C^1((0,T) \times \mathbb{R}^N)$ such that $u_j - \phi$ attains a local minimum at (t,x) in $(0,T) \times \overline{\mathcal{M}}_j$. For n big enough the super-optimality property gives

$$u_j(t,x) - u_j(t_n,x_n) \geq 0.$$

This inequality combined with the fact that $u_j(t_n,x_n) - \phi(t_n,x_n) \geq u_j(t,x) - \phi(t,x)$ lead to

$$\frac{1}{t_n - t}(\phi(t,x) - \phi(t_n,x_n)) \geq 0.$$

By letting n tend to $+\infty$, we obtain

$$-\partial_t\phi(t,x) - \langle \nu, \partial_x\phi(t,x)\rangle \geq 0$$

and then

$$-\partial_t\phi(t,x) + H_{F_j^\sharp}(x, \partial_x\phi(t,x)) \geq -\partial_t\phi(t,x) - \langle \nu, \partial_x\phi(t,x)\rangle \geq 0.$$

This concludes the proof.

Now we turn our attention to the proof of the claim. We distinguish two cases: either there exists $r > 0$ such that $y([t, t+r])$ stays in one domain $\mathcal{M}_j$ for some $j \in \{1, \ldots, n\} \cap I(x)$, almost everywhere, or it touches or crosses the singular set Λ infinitely many times no matter how we are close to x.

We begin with the first case. Suppose there exists $r > 0$ such that $y([t, t+r]) \subset \mathcal{M}_j$ for some $j \in \{1, \ldots, n\} \cap I(x)$ almost everywhere. So, there exists a sequence $(t_n)_n$, $t_n \downarrow t$ and $x_n := y(t_n) \in \mathcal{M}_j$, such that $\dfrac{x_n - x}{t_n - t} \to \nu$.

Notice that $\nu = \lim\limits_{n\to+\infty} \dfrac{x_n - x}{t_n - t} \in \mathcal{T}_{\overline{\mathcal{M}}_j}(x)$, since $x_n \in \mathcal{M}_j$. It remains to prove that ν belongs to $F_j(x)$. Denote by κ and M respectively the Lipschitz constant of $F_j(.)$ and the Lipschitz constant of $y(.)$. We have

$$\begin{aligned}
\nu &= \lim_{n\to+\infty} \frac{1}{t_n - t}\int_{[t,t_n]} \dot{y}(s)\, ds \\
&\in \lim_{n\to+\infty} \Big(\frac{1}{t_n - t}\int_{[t,t_n]} proj_{F_j(x)}(\dot{y}(s))\, ds + \frac{\kappa}{t_n - t}\int_{[t,t_n]} |y(s) - x|\mathbb{B}\, ds\Big) \\
&\subseteq \lim_{n\to+\infty} \Big(F_j(x) + \frac{\kappa M}{t_n - t}\Big[\int_{[t,t_n]} (s-t)\, ds\Big]\mathbb{B}\Big) \\
&\subseteq \lim_{n\to+\infty} \Big(F_j(x) + \kappa M\frac{|t_n - t|}{2}\mathbb{B}\Big) = F_j(x).
\end{aligned}$$

In conclusion, we get $\nu \in F_j(x) \cap \mathcal{T}_{\overline{\mathcal{M}}_j}(x) = F_j^\sharp(x)$.
Now we get to the second case. Since $y(.)$ touches or crosses Λ infinitely many times no matter how we are close to x, then there exists $j \in \{n+1, \ldots, n+l\} \cap I(x)$, a sequence $(t_n)_n$, $t_n \downarrow t$ and $x_n := y(t_n) \in \mathcal{M}_j$, so that $\dfrac{x_n - x}{t_n - t} \to \nu$.

Notice that $\nu = \lim\limits_{n\to+\infty} \dfrac{x_n - x}{t_n - t} \in \mathcal{T}_{\overline{\mathcal{M}}_j}(x)$, since $x_n \in \mathcal{M}_j$. It remains to prove that ν belongs to $F_j(x)$. For $k = 1, \ldots, n+l$, we set

$$J_n^k := \{\, s \in [t, t+t_n] \;:\; y(s) \in \mathcal{M}_k \,\},\ \mu_n^k := \mathscr{L}(J_n^k),\ \mathbb{K}(x) := \{k \;:\; \mu_n^k > 0,\ \forall n \in \mathbb{N}\},$$

where we recall that $\mathscr{L}$ is the Lebesgue measure on $\mathbb{R}$. We obviously have $\mathbb{K}(x) \subset I(x)$. Furthermore, up to a subsequence, there exist $0 \leq \lambda_k \leq 1$ and $p_k \in \mathbb{R}^N$ such that

$$\frac{\mu_n^k}{t_n - t} \to \lambda_k, \sum_{k \in \mathbb{K}(x)} \lambda_k = 1, \ \frac{1}{\mu_n^k} \int_{J_n^k} \dot{y}(s) ds \to p_k, \quad \text{as } n \to \infty.$$

Denote by κ and M respectively the Lipschitz constant of $F_k(.)$ and the Lipschitz constant of $y(.)$, we get

$$\begin{aligned}
p_k &= \lim_{n \to +\infty} \frac{1}{\mu_n^k} \int_{J_n^k} \dot{y}(s)\, ds \\
&\in \lim_{n \to +\infty} \Big(\frac{1}{\mu_n^k} \int_{J_n^k} proj_{F_k(x)}(\dot{y}(s))\, ds + \frac{\kappa}{\mu_n^k} \int_{J_n^k} |y(s) - x| \mathbb{B}\, ds \Big) \\
&\subseteq \lim_{n \to +\infty} \Big(F_k(x) + \frac{\kappa M}{\mu_n^k} \Big[\int_{J_n^k} (s-t)\, ds \Big] \mathbb{B} \Big) \\
&\subseteq \lim_{n \to +\infty} \Big(F_k(x) + \kappa M |t_n - t| \mathbb{B} \Big) = F_k(x).
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\nu = \lim_{n \to +\infty} \frac{1}{t_n - t} \int_{[t, t_n]} \dot{y}(s)\, ds = \sum_{k \in \mathbb{K}(x)} \lim_{n \to +\infty} \frac{\mu_n^k}{t_n - t} \left[\frac{1}{\mu_n^k} \int_{J_n^k} \dot{y}(s)\, ds \right] \\
\subset co \{ F_k(x) \ : \ k \in \mathbb{K}(x) \}.
\end{aligned}$$

Finally, we get

$$\nu \in co \Big\{ F_k(x) \ : \ k \in \mathbb{K}(x) \Big\} \cap \mathcal{T}_{\overline{\mathcal{M}}_j}(x) \subset co \Big\{ F_k(x) \ : \ k \in I(x) \Big\} \cap \mathcal{T}_{\overline{\mathcal{M}}_j}(x) = F_j^\sharp(x).$$

This ends the proof of the claim. ■

Remark 2.4.1. By the arguments presented at the beginning of the above proof, it is easy to see that under the same assumptions of Theorem 2.13, the following statements are equivalent:

- u satisfies the super-optimality property,
- u is a supersolution of (2.7) in the sense of Definition 2.1,
- u is a supersolution of (2.6) in the sense of Definition 2.1,
- u verifies inequality (2.16).

Remark 2.4.2. It was already known from [18], that inequality (2.16) is equivalent to the super-optimality property if we only use the classical definition of viscosity. The importance of this result lies in the fact that the equivalence is valid even if we take the notion of viscosity stated in Definition 2.1.

2.4.3 The sub-optimality and subsolution property

This section aims at establishing the link between the sub-optimality property and the the subsolution property of the HJB equation (2.7). The subsolution property is characterized by the Hamiltonians associated to the dynamics $F_i^\sharp$, which represent the velocities that are taken by the trajectories of the controlled system $(DI)_F(t,x)$. The importance of the essential dynamics are displayed in the following result, where its proof can be found in [18, Lemma 3.9].

Lemma 2.14. *Let $i \in \{1,\dots,n+l\}$ and let $x \in \overline{\mathcal{M}}_i$. Then for any $p \in F_i^\sharp(x)$, there exists $t,\tau \in \mathbb{R}$ with $t<\tau<T$ and a trajectory $y(.) \in S_{(t,T)}(x)$ such that it is C^1 on $[t,\tau]$, $\dot{y}(t)=p$ and $y([t,\tau]) \subset \overline{\mathcal{M}}_i$.*

Theorem 2.15. *Suppose that (H_1), (SH), (H_D) and (H_{ESS}) are verified. Let $u : [0,T]\times\mathbb{R}^N \to \mathbb{R}$ be an u.s.c function. The following assertions are equivalent:*

- *(i) u satisfies the sub-optimality principle,*
- *(ii) u is a subsolution of the HJB equation (2.7).*

Proof. We prove $(i) \Longrightarrow (ii)$ first. Let $i \in \{1,...,n+l\}$ and $(t,x) \in [0,T]\times\overline{\mathcal{M}}_i$. By Lemma 2.14, for every $p \in F_i^\sharp(x)$, there exists a C^1 trajectory $y(.)$ defined on some interval $[t,t+\varepsilon]$, with $\varepsilon>0$, such that $y(t)=x$, $\dot{y}(t)=p$ and $y(.) \subseteq \overline{\mathcal{M}}_i$.

Let $u_i \equiv u$ on $\overline{\mathcal{M}}_i$ and $u_i \equiv -\infty$ otherwise. Let (θ,ξ) be in $D^+u_i(t,x)$. For any sequence $((t_n,x_n))_n$ such that $(t_n,x_n) \in dom(u_i)$ and $(t_n,x_n) \to (t,x)$, we have

$$\limsup_{n\to\infty} \frac{u(t_n,x_n)-u(t,x)-\theta(t_n-t)-\langle \xi, x_n-x\rangle}{|x_n-x|+|t_n-t|} \le 0.$$

Setting $x_n = y(t+\dfrac{\varepsilon}{n})$ and $t_n := t+\dfrac{\varepsilon}{n}$, we get by sub-optimality of u

$$\frac{-\theta(t_n-t)-\langle \xi, x_n-x\rangle}{|x_n-x|+|t_n-t|} \le \frac{u(t_n,x_n)-u(t,x)-\theta(t_n-t)-\langle \xi, x_n-x\rangle}{|x_n-x|+|t_n-t|},$$

By letting $n\to\infty$ we get

$$\frac{-\theta-\langle \xi,p\rangle}{|p|+1} \le 0 \implies -\theta-\langle \xi,p\rangle \le 0.$$

Since p is arbitrary, we get the result by taking the supremum over $F_i^\sharp(x)$.

It remains to prove $(ii) \Longrightarrow (i)$. We define the augmented stratification by

$$M_i := \mathbb{R}\times\mathcal{M}_i\times\mathbb{R}.$$

Furthermore, for all $i=1,...,n+l$, we define $v:=-u$ (so v is l.s.c) and we denote by $v_i \equiv v$ on $\overline{\mathcal{M}}_i$ and $v_i \equiv +\infty$ otherwise. Next, we divide the proof into 2 steps.

Step 1.. We show that

$$\forall i \in \{1, ..., n+l\},\ epi(v_i) = epi(v) \cap \overline{M}_i.$$

Let $(t, x, r) \in epi(v_i)$. So $v_i(t, x) \leq r$. Hence $x \in \overline{\mathcal{M}}_i$ and $v(t, x) \leq r$. Thus we get $(t, x, r) \in epi(v) \cap \overline{M}_i$. Conversely, if $(t, x, r) \in epi(v) \cap \overline{M}_i$, then $v(t, x) \leq r$ and $x \in \overline{\mathcal{M}}_i$. So $v_i(t, x) = v(t, x)$, whence $v_i(t, x) \leq r$, which finishes the proof of step 1.

Step 2. (Augmented dynamics). Let us first point out the fact that assertion (ii) is equivalent to

$$-\theta + H_{F_i^\sharp}(x, \nu) \leq 0 \quad \text{for all } (t, x) \in (0, T) \times \overline{\mathcal{M}}_i,\ (\theta, \nu) \in -D^- v_i(t, x),$$

since $D^+ u_i(t, x) = -D^-(-u_i)(t, x) = -D^- v_i(t, x)$.

We establish the following claim.

Claim. Let $G_i^\sharp$ be the augmented dynamics defined by

$$G_i^\sharp(t, x, z) := \{1\} \times F_i^\sharp(x) \times \{0\}, \quad \text{for any } (t, x, z) \in \overline{M}_i.$$

If we have

$$-\theta + H_{F_i^\sharp}(x, \nu) \leq 0 \quad \text{for all } (t, x) \in (0, T) \times \overline{\mathcal{M}}_i,\ (\theta, \nu) \in -D^- v_i(t, x),$$

Then it holds

$$\sup_{\nu \in G_i^\sharp(t,x,z)} \{ \langle \eta, \nu \rangle \} \leq 0 \qquad \forall (t, x, z) \in epi(v_i),\ \eta \in N^p_{epi(v_i)}(t, x, z). \tag{2.17}$$

Let $(t, x, z) \in epi(v_i)$. If $F_i^\sharp(x) = \emptyset$ then the result holds by vacuity. Otherwise, let $(\xi, -\lambda) \in N^p_{epi(v_i)}(t, x, z)$. So we have $\lambda \geq 0$ because $(\xi, -\lambda)$ belongs to the proximal normal cone of the epigraph of v_i. If $\lambda > 0$. Then we have $z = v_i(t, x)$ and there exists $(\theta, \zeta) \in -\partial_p v_i(t, x) \subset -D^- v_i(t, x)$ such that $\xi = (-\lambda\theta, -\lambda\zeta)$. Hence, by [9, Theorem 11.32], for any $\nu \in G_i^\sharp(t, x, z)$ we have, for some $p \in F_i^\sharp(x)$:

$$\langle (\xi, -\lambda), \nu \rangle = -\lambda(\theta + \langle \zeta, p \rangle) \leq \lambda(-\theta + H_{F_i^\sharp}(x, \zeta)) \leq 0$$

We take the supremum over ν and we get the result. Now, if $\lambda = 0$, then by [9, Theorem 11.31], there exist sequences $((t_n, x_n))_n \subseteq [0, T] \times \overline{\mathcal{M}}_i$, $(\xi_n)_n \subseteq \mathbb{R}^{N+1}$ and $(\lambda_n)_n \subseteq (0, \infty)$ such that

$$(t_n, x_n, \lambda_n) \to (t, x, 0), \qquad v(t_n, x_n) \to z \qquad \xi_n \to \xi, \qquad \frac{1}{\lambda_n}\xi_n \in -\partial_p v_i(t_n, x_n).$$

Thus the argument above shows that

$$\langle (\xi_n, -\lambda_n), \nu_n \rangle \leq 0 \qquad \forall \nu_n \in G_i^\sharp(t_n, x_n, u_i(t_n, x_n)),\ \forall n \in \mathbb{N}.$$

Furthermore, by Hypothesis (H_{ESS}) we have that $G_i^\sharp(.)$ is lower semicontinuous. So for any $\nu \in G_i^\sharp(t,x,z)$, there exists a sequence $(\nu_n)_n \to \nu$ such that $\nu_n \in G_i^\sharp(t_n, x_n, v_i(t_n, x_n))$. By evaluating the last inequality at this sequence and letting $n \to +\infty$, then taking the supremum over ν, we get the result.

Consequently, since equation (2.17) holds from (ii), then We can apply Theorem 2.12 to the augmented dynamics $G(.) := \{1\} \times F(.) \times \{0\}$, the stratification $\mathbb{R}^{N+2} = \cup_{i=1}^{n+l} M_i$ and the set $S = epi(v)$. Thus we get that $(epi(v), G)$ is strongly invariant. Let $(t,x) \in (0,T) \times \mathbb{R}^N$ and $y(.)$ be a solution of $(DI)_F(t,x)$. So we have that

$$Y(s) = (s, y(s), v(t, y(t)) \qquad s \in [t,T],$$

is a solution of the differential inclusion with the augmented dynamics $G(.)$ and initial condition $(t,x,v(t,x)) = (t, y(t), v(t, y(t)))$
we have $(t, y(t), v(t, y(t)) \in epi(v)$. Thus, by Theorem 2.12, we get

$$(t+h, y(t+h), v(t, y(t)) \in epi(v)$$

for all $h \in [0, T-t]$. Hence

$$v(t+h, y(t+h)) \le v(t, y(t)) \Longleftrightarrow u(t, y(t)) \le u(t+h, y(t+h)) \quad (\text{since } u = -v).$$

This ends the proof of $(ii) \implies (i)$ and Theorem 2.15. ■

2.4.4 Proof of Theorems 2.6 and 2.7.

Proof. (Theorem 2.6). Proposition 2.12.2 shows that the value function ϑ is locally Lipschitz continuous. Furthermore, we have $\vartheta(T,x) = \psi(x)$. Theorem 2.13 shows that ϑ is a viscosity supersolution since it enjoys the super-optimality property. In addition, Theorem 2.15 shows that ϑ is a viscosity subsolution since it enjoys the sub-optimality property. Finally, uniqueness comes from the comparison result in Theorem 2.7. ■

Remark 2.4.3. Notice that assumption (CH) is only used to ensure that the value function is continuous. If we directly assume that the value function is continuous instead of assuming (CH), then the conclusion of Theorem 2.6 is still valid.

Proof. (Theorem 2.7). Let u_1 and u_2 respectively be a l.s.c supersolution and an u.s.c subsolution of equation (2.7). By Theorem 2.13, we conclude that u_1 satisfies the super-optimality principle which means that for all $(t,x) \in (0,T] \times \mathbb{R}^N$, there exists a trajectory $y(.) \in S_{(t,T)}(x)$ such that

$$u_1(t,x) \ge u_1(T, y(T)).$$

Likewise, by Theorem 2.15, we conclude that u_2 satisfies the sub-optimality principle. Hence, we get that for the same trajectory $y(.)$ we have

$$\forall\, (t,x) \in (0,T] \times \mathbb{R}^N, \quad u_2(t,x) \le u_2(T, y(T)).$$

Therefore, using the fact that $u_2(T,\cdot) \le u_1(T,\cdot)$, we get

$$u_2(t,x) \le u_1(t,x) \text{ for any } (t,x) \in (0,T] \times \mathbb{R}^N,$$

which is the required result. ■

2.5 Stability

Theorem 2.16. *For $i = 1,\dots,n+l$, let $\left(F_i^j : \overline{\mathcal{M}}_i \rightsquigarrow \mathbb{R}^N\right)_j$ be a sequence of set-valued maps satisfying (SH) and such that $F_i^j \longrightarrow F_i$ w.r.t the Hausdorff metric (i.e. uniform convergence). Let $(v^j : \mathbb{R}^N \to \mathbb{R})_j$ be a sequence of l.s.c functions such that $v^j \to v$ locally uniformly in $\mathbb{R}^N$. Suppose in addition that for all j, v^j is a supersolution of*

$$-\partial_t v^j(t,x) + \max_{i\in I(x)} \{H_{F_i^{j\sharp}}(x, \partial_x v^j(t,x))\} = 0 \quad \textit{for all } (t,x) \in (0,T) \times \mathbb{R}^N,$$

in the sense of Definition 2.1. Then v is a supersolution of (2.7).

Proof. Let $(t,x) \in (0,T) \times \mathbb{R}^N$. Using Remark 2.4.1, it suffices to prove that v satisfies the inequality (2.16). Let $\phi \in C^1((0,T) \times \mathbb{R}^N)$ such that $u - \phi$ attains a local minimum at (t,x). Then, there exists $(t^j,x^j) \in (0,T) \times \mathbb{R}^N$ such that $v^j - \phi$ attains local minimum and such that $(t^j,x^j) \to (t,x)$. Since the stratification is finite and v^j is a supersolution of (2.7), then up to a subsequence (not relabelled), there exists $i_0 \in [1, n+l]$ such that for all j, we have

$$-\partial_t \phi(t^j,x^j) + H_{F_{i_0}^{j\sharp}}(x^j, \partial_x \phi(t^j,x^j)) \ge 0.$$

Since $F_{i_0}^{j\sharp}(.) \subseteq F_{i_0}^j(.)$, we get

$$-\partial_t \phi(t^j,x^j) + H_{F_{i_0}^{j}}(x^j, \partial_x \phi(t^j,x^j)) \ge 0.$$

So by letting j tend to infinity, we get

$$-\partial_t \phi(t,x) + H_{F_{i_0}}(x, \partial_x \phi(t,x)) \ge 0.$$

Finally, since $F_{i_0}(.) \subseteq F(.)$, then we get

$$-\partial_t \phi(t,x) + H_F(x, \partial_x \phi(t,x)) \ge 0,$$

which is the required result by Remark 2.4.1. ■

Theorem 2.17. *For $i = 1, ..., n+l$, let $\left(F_i^j : \overline{\mathcal{M}}_i \rightsquigarrow \mathbb{R}^N\right)_j$ be a sequence of set-valued maps satisfying (SH). We denote By*

$$F_i^{j\sharp}(.) = F_i^j(.) \cap \mathcal{T}_{\overline{\mathcal{M}}_i}(.).$$

Suppose that $F_i^{j\sharp} \longrightarrow F_i^{\sharp}$ *w.r.t the Hausdorff metric. Let* $(v^j : \mathbb{R}^N \to \mathbb{R})_j$ *be a sequence of u.s.c functions such that* $v^j \to v$ *locally uniformly in* $\mathbb{R}^N$*. Suppose in addition that for all* $j \in \mathbb{N}$*,* v^j *is a subsolution of*

$$-\partial_t v^j(t,x) + H_{F_i^{j\sharp}}(x, \partial_x v^j(t,x)) = 0 \quad \textit{for all } (t,x) \in (0,T) \times \mathbb{R}^N,\ i \in I(x),$$

in the sense of definition 2.2. Then v *is a subsolution of* (2.7)*.*

Proof. Let $(t,x) \in (0,T) \times \mathbb{R}^N$ and $i \in I(x)$. Let $\phi \in C^1((0,T) \times \mathbb{R}^N)$, such that $u - \phi$ attains a local maximum in $(0,T) \times \overline{\mathcal{M}}_i$, at (t,x). Without loss of generality, we can always suppose that the maximum is strict. Then by [4, Lemma 2.2] there exists $(t^j, x^j) \in (0,T) \times \overline{\mathcal{M}}_i$, such that $v^j - \phi$ attains local maximum in $(0,T) \times \overline{\mathcal{M}}_i$ at (t^j, x^j), and such that $(t^j, x^j) \to (t,x)$. Since v^j is a subsolution, we get

$$-\partial_t \phi(t^j, x^j) + H_{F_i^{j\sharp}}(x^j, \partial_x \phi(t^j, x^j)) \leq 0.$$

Now let $\nu \in F_i^{\sharp}(x)$. Then by the Hausdorff convergence of the sequence $(F_i^{j\sharp})_j$ there exists a sequence $\nu^j \in F_i^{j\sharp}(x^j)$ such that $\nu^j \to \nu$. Finally, we arrive at

$$-\partial_t \phi(t^j, x^j) + \langle -\nu^j, \partial_x \phi(t^j, x^j) \rangle \leq -\partial_t \phi(t^j, x^j) + H_{F_i^{j\sharp}}(x^j, \partial_x \phi(t^j, x^j)) \leq 0.$$

By letting j tend to infinity, we get

$$-\partial_t \phi(t,x) + \langle -\nu, \partial_x \phi(t,x) \rangle \leq 0.$$

Lastly, since ν is arbitray, we take the supremum over ν and we get the required result. ∎

2.6 General convergence result for monotone schemes

In this section, we aim at studying the convergence of monotone numerical schemes approximating the HJB equation (2.7).

Let $\mathcal{G}^{\Delta x} = \bigcup_{i=1}^{n+l} \mathcal{G}_i^{\Delta x}$ be a spatial grid of $\mathbb{R}^N$ of step $\Delta x > 0$, such that each $\mathcal{G}_i^{\Delta x}$ is a subgrid of $\overline{\mathcal{M}}_i$ and $\mathcal{G}^{\Delta x}$ is compatible with the stratification $(\mathcal{M}_i)_{i=1,\dots,n+l}$ in the following sense:

$$\text{(CC):} \begin{cases} (i) \text{ For all } i,j = 1,...,n+l, \text{ such that } \mathcal{M}_j \subset \overline{\mathcal{M}}_i,\ \mathcal{G}_i^{\Delta x} \text{ and } \mathcal{G}_j^{\Delta x} \text{ coincide on } \mathcal{G}_j^{\Delta x}, \\ (ii)\ \forall R > 0,\ \forall i = 1,...,n+l,\ \lim\limits_{\Delta x \to 0} d_{\mathcal{H}}\Big(\overline{\mathcal{M}}_i \cap \overline{\mathbb{B}}(0,R)\ ,\ \mathcal{G}_i^{\Delta x} \cap \overline{\mathbb{B}}(0,R)\Big) = 0. \end{cases}$$

Comments on the hypothesis (CC)

Hypothesis (CC)(i) implies that the grid $\mathcal{G}^{\Delta x}$ is divided into $n+l$ subgrids $(\mathcal{G}_i^{\Delta x})_i$ with a partial order relation that ensures compatibility with the stratification. Hypothesis (CC)(ii) asserts that for each $i = 1, \dots, n+l$, the subgrid $\mathcal{G}_i^{\Delta x}$ approaches $\overline{\mathcal{M}}_i$ in the the sense of Hausdorff convergence. Notice that this implies in particular that the points of a subgrid $\mathcal{G}_i^{\Delta x}$ don't have to belong to $\overline{\mathcal{M}}_i$ meaning that we do not require that

$$\mathcal{G}_i^{\Delta x} \subset \overline{\mathcal{M}}_i.$$

What is important here is that the grid $\mathcal{G}^{\Delta x}$ is divided into $n+l$ subgrids compatible with the stratification, and each subgrid converges in the Hausdorff sense to its corresponding domain.

We define for any $x \in \mathcal{G}^{\Delta x}$, the index set-valued map of the grid

$$I_{\mathcal{G}^{\Delta x}}(x) := \{\, i \in \{1, ..., n+l\} \ : \ x \in \mathcal{G}_i^{\Delta x} \,\}.$$

Let Δt be a constant time step of a regular grid $\Pi^{\Delta t}$ of $[0, T]$. We denote $h = (\Delta t, \Delta x)$. We consider the following numerical scheme:

$$\begin{cases} \max\limits_{i \in I_{\mathcal{G}^{\Delta x}}(x_h)} \{\, S_i^h(t_h, x_h, u^h(t_h, x_h), [u^h]_{(t_h,x_h)}) \,\} = 0 & \text{for } (t_h, x_h) \in (\Pi^{\Delta t} \times \mathcal{G}^{\Delta x}) \setminus \{t_h = T\}, \\ u^h(T, x_h) = \psi(x_h), \quad \text{for } (t_h, x_h) \in (\Pi^{\Delta t} \times \mathcal{G}^{\Delta x}) \cap \{t_h = T\}, & \end{cases}$$

with $u^h : \Pi^{\Delta t} \times \mathcal{G}^{\Delta x} \to \mathbb{R}$ is the approximate solution and $[u^h]_{(t_h,x_h)}$ are all the values of of u^h on $\mathcal{G}^{\Delta x}$ at other points than (t_h, x_h). Each S_i^h, $i = 1, \dots, n+l$, is supposed to verify the following hypotheses:

- **Monotonicity** : $S_i^h(s, z, u, [w_1]^h_{(s,z)}) \leq S_i^h(s, z, u, [w_2]^h_{(s,z)})$, if $w_1 \geq w_2$.
- **Stability** :
  - each u^h is bounded on bounded sets of $\mathbb{R}^N$ independently from h, for h small enough, i.e. for all $\rho > 0$, there exists a $C_\rho > 0$, independent of h, such that

    $$|u^h(t_i^h, x_i^h)| \leq C_\rho \quad \text{if } (t_i^h, x_i^h) \in \Big([0, T] \times \mathbb{B}(0, \rho)\Big) \cap \Big(\Pi^{\Delta t} \times \mathcal{G}^{\Delta x}\Big).$$

  - u^h verifies the following inequality on a neighborhood of the interfaces: there exists $r > 0$ and $C_r > 0$, independent of h, such that

    $$|u^h(t_h, x_h) - u^h(s_h, y_h)| \leq C_r(|t_h - s_h| + |x_h - y_h|),$$
    $$\text{for all } (t_h, x_h), (s_h, y_h) \in \Big([0, T] \times (\Lambda + \mathbb{B}(0, r))\Big) \cap \Big(\Pi^{\Delta t} \times \mathcal{G}^{\Delta x}\Big).$$

- **Consistency** : for all $\phi \in C^1((0,T) \times \mathbb{R}^N)$ and $(t,x) \in (0,T) \times \mathbb{R}^N$, we have

$$\liminf_{\substack{\Pi^{\Delta t} \times \mathcal{G}_i^{\Delta x} \ni (s,z) \to (t,x) \\ h \to 0 \\ \zeta \to 0}} S_i^h(s,z,\phi(s,z)+\zeta,[\phi+\zeta]^h_{(s,z)}) \geq -\partial_t \phi(t,x) + H_{F_i^\sharp}(x,\partial_x \phi(t,x)),$$

$$\limsup_{\substack{\Pi^{\Delta t} \times \mathcal{G}_i^{\Delta x} \ni (s,z) \to (t,x) \\ h \to 0 \\ \zeta \to 0}} S_i^h(s,z,\phi(s,z)+\zeta,[\phi+\zeta]^h_{(s,z)}) \leq -\partial_t \phi(t,x) + H_F(x,\partial_x \phi(t,x)).$$

The following theorem is an extension of the result by Barles and Souganidis [47], to the case of HJB equations defined on stratified domains.

Theorem 2.18. *Suppose that the HJB equation (2.7) admits a continuous viscosity solution u and the comparison principle in Theorem 2.7 holds. Assume that for every $h > 0$ small enough, the numerical scheme admits a solution u^h. Assume further that the spatial grid verifies hypothesis **(CC)** and that each S_i^h verifies the monotonicity, stability and consistency hypotheses.*

then, u^h converges locally uniformly to u.

Proof. First, we begin by defining the following functions

$$\overline{u}(t,x) := \limsup_{\substack{\Pi^{\Delta t} \times \mathcal{G}^{\Delta x} \ni (s,z) \to (t,x) \\ h \to 0}} u^h(s,z), \quad \underline{u}(t,x) := \liminf_{\substack{\Pi^{\Delta t} \times \mathcal{G}^{\Delta x} \ni (s,z) \to (t,x) \\ h \to 0}} u^h(s,z).$$

We aim to prove that

$$\overline{u} = \underline{u} = u.$$

We already have $\underline{u} \leq \overline{u}$. Therefore to prove our result It suffices to prove that $\underline{u}$ is a l.s.c supersolution and $\overline{u}$ is an u.s.c subsolution of the HJB equation (2.7). We first prove that $\overline{u}$ is a subsolution.

Let $(t,x) \in (0,T) \times \mathbb{R}^N$. Without loss of generality, we suppose that $x \in \mathbb{B}(0,\rho)$ for some $\rho > 0$ big enough and we restrict our analysis on this bounded open set. Let $i \in I(x)$ and let $\phi \in C^1((0,T) \times \mathbb{R}^N)$ such that $\overline{u}_i - \phi$ attains its local maximum at $(t,x) \in (0,T) \times \overline{\mathcal{M}}_i$. We recall that $\bar{u}_i = u$ in $(0,T) \times \overline{\mathcal{M}}_i$ and $u_i \equiv -\infty$ otherwise.

Without loss of generality, we can suppose that

$$\overline{u}_i(t,x) = \phi(t,x), \quad \overline{u}_i(s,z) < \phi(s,z) \text{ if } (s,z) \neq (t,x),$$

$$\phi \geq C_\rho + 1 \text{ outside of a neighborhood } \Omega \text{ of } (t,x), \text{ and } \quad \Omega \subsetneq (0,T) \times \mathbb{B}(0,\rho),$$

where C_ρ is defined from the first part of the stability assumption. Furthermore, from the second part of the stability assumption, $\overline{u}$ is Lipschitz continuous in a neighborhood of the interfaces. So we get

$$0 = \overline{u}_i(t,x) - \phi(t,x) = \limsup_{\substack{\Pi^{\Delta t}\times\mathcal{G}^{\Delta x}\ni(s,z)\to(t,x)\\ h\to 0}} u^h(s,z) - \phi(s,z) = \limsup_{\substack{\Pi^{\Delta t}\times\mathcal{G}_i^{\Delta x}\ni(s,z)\to(t,x)\\ h\to 0}} u^h(s,z) - \phi(s,z).$$

The second part of the stability assumption is essential here to get the last equality since the limit sup might not be reached from every subsgrid $\mathcal{G}_i^{\Delta x}$. Moreover, outside of Ω, we have $\overline{u}_i - \phi \leq -1$. So, there exists $r > 0$, such that

$$0 \geq u^h(s,z) - \phi(s,z) \geq -1 \text{ for all } (s,z) \in ([t-r,t+r]\cap\Pi^{\Delta t}) \times (\overline{\mathbb{B}}(0,r)\cap\mathcal{G}_i^{\Delta x}) \subset \Omega.$$

So, the maximum of $u^h - \phi$ is attained in the compact set

$$([t-r,t+r]\cap\Pi^{\Delta t}) \times (\overline{\mathbb{B}}(0,r)\cap\mathcal{G}_i^{\Delta x}) \subset \Omega.$$

Let (t_i^h, x_i^h) be the maximum and let (t_i, x_i) be the limit when $h \to 0$ of a subsequence not relabelled. we have

$$\lim_{h\to 0} u^h(t_i^h, x_i^h) - \phi(t_i^h, x_i^h) \geq \limsup_{\substack{\Pi^{\Delta t}\times\mathcal{G}_i^{\Delta x}\ni(s,z)\to(t,x)\\ h\to 0}} u^h(s,z) - \phi(s,z) = \overline{u}_i(t,x) - \phi(t,x) = 0.$$

On the other hand, since $\overline{u}_i - \phi$ is u.s.c, we get

$$0 \geq \overline{u}_i(t_i,x_i) - \phi(t_i,x_i) \geq \lim_{h\to 0} u^h(t_i^h, x_i^h) - \phi(t_i^h, x_i^h).$$

Thus, we conclude

$$(t_i, x_i) = (t,x), \quad u^h(t_i^h, x_i^h) \to \overline{u}_i(t,x).$$

Let $\zeta_h := u^h(t_i^h, x_i^h) - \phi(t_i^h, x_i^h)$. We get

$$u^h(t_i^h, x_i^h) = \phi(t_i^h, x_i^h) + \zeta_h, \quad u^h(s,z) \leq \phi(s,z) + \zeta_h,\ (t_i^h, x_i^h) \neq (s,z) \in \Pi^{\Delta t} \times \mathcal{G}_i^{\Delta x}.$$

From the monotonicity of the scheme and u^h being a solution, we get

$$S_i^h(t_i^h, x_i^h, u^h(t_i^h, x_i^h), [\phi+\zeta_h]^h_{(t_i^h,x_i^h)}) \leq S_i^h(t_i^h, x_i^h, u^h(t_i^h, x_i^h), [u^h]_{(t_i^h,x_i^h)}) \leq 0,$$

and by the consistency hypothesis, by passing to the infimum limit in the above inequality, we get

$$-\partial_t\phi(t,x) + H_{F_i^\sharp}(x, \partial_x\phi(t,x)) \leq \liminf_{\substack{(t_i^h,x_i^h)\to(t,x)\\ h\to 0\\ \zeta_h\to 0}} S_i^h(t_i^h, x_i^h, \phi(t_i^h, x_i^h) + \zeta_h, [\phi+\zeta_h]^h_{(t_i^h,x_i^h)}) \leq 0.$$

This is true for any $i \in I(x)$, which ends the proof.

Now we prove that $\underline{u}$ is a supersolution. Following Remark 2.4.1, it suffices to prove that $\underline{u}$ is a supersolution of (2.16).

Let $(t,x) \in (0,T) \times \mathbb{R}^N$. Let $\phi \in C^1((0,T) \times \mathbb{R}^N)$ such that $\underline{u} - \phi$ attains its local minimum in $(0,T) \times \mathbb{R}^N$ at (t,x) and $\phi(t,x) = \underline{u}(t,x)$. Using the same arguments as in the first part of the proof, we get a sequence $(h_n)_n$ such that

$$h_n \downarrow 0, \quad u^{h_n} - \phi \text{ attains a local maximum at } (t_n, x_n) \in \Pi^{\Delta t} \times \mathcal{G}^{\Delta x}, \text{ and}$$

$$(t_n, x_n) \to (t,x), \quad u^{h_n}(t_n, x_n) \to \underline{u}(t,x).$$

Furthermore, since the stratification is finite and u^{h_n} is a solution to the numerical scheme, there exists a subsequence (not relabelled) of $(h_n)_n$ such that there exists $i_0 \in I_{\mathcal{G}^{\Delta x}}(x_n)$, for all $n \in \mathbb{N}$ and $(x_n)_n \subseteq \mathcal{G}_{i_0}^{\Delta x}$ and we have

$$\max_{i \in I_{\mathcal{G}^{\Delta x}}(x_n)} S_i^{h_n}(t_n, x_n, u^{h_n}(t_n, x_n), u^{h_n}) = S_{i_0}^{h_n}(t_n, x_n, u^{h_n}(t_n, x_n), u^{h_n}) \geq 0.$$

Let $\zeta_n := u^{h_n}(t_n, x_n) - \phi(t_n, x_n)$. So

$$u^{h_n}(t_n, x_n) = \phi(t_n, x_n) + \zeta_n, \quad u^{h_n}(s,z) \geq \phi(s,z) + \zeta_n, \ (t_i^h, x_i^h) \neq (s,z) \in \Pi^{\Delta t} \times \mathcal{G}^{\Delta x}.$$

From the monotonicity assumption and u^{h_n} being a solution, we get

$$S_{i_0}^{h_n}(t_n, x_n, \phi(t_n, x_n) + \zeta_n, [\phi + \zeta_h]^{h_n}_{(t_n^h, x_n^h)}) \geq S_{i_0}^{h_n}(t_n, x_n, u^{h_n}(t_n, x_n), [u^{h_n}]^{h_n}_{(t_n^h, x_n^h)}) \geq 0.$$

and by the consistency hypothesis and passing to the supremum limit, we get

$$-\partial_t \phi(t,x) + H_F(x, \partial_x \phi(t,x)) \geq \limsup_{\substack{(t_n,x_n)\to(t,x) \\ h_n \to 0 \\ \zeta_n \to 0}} S_{i_0}^{h_n}(t_n, x_n, \phi(t_n, x_n) + \zeta_n, [\phi + \zeta_h]^{h_n}_{(t_n^h, x_n^h)}) \geq 0.$$

By Remark 2.4.1, we get the required result.

Finally, with similar arguments as above, we prove that at time $t = T$, $\overline{u}$ (resp. $\underline{u}$) is subsolution (resp. supersolution) of

$$\min\left(-\partial_t u(t,x) + \max_{i \in I(x)} \{H_{F_i^\sharp}(x, \partial_x u(t,x))\}, u(T,x) - \psi(x)\right) \leq 0,$$

resp.

$$\max\Big(- \partial_t u(t,x) + H_F(x, \partial_x u(t,x)), u(T,x) - \psi(x)\Big) \geq 0,$$

for $(t,x) \in (0,T] \times \mathbb{R}^N$. Finally, by the same reasoning as in [4, Theorem 4.7], we obtain that

$$\overline{u}(T, \cdot) \leq \psi(.) \leq \underline{u}(T, \cdot).$$

In conclusion, by Theorem 2.7, we have $\overline{u} = \underline{u} = u$ and u^h converges locally uniformly to u, which ends the proof. ∎

2.7 Appendices

2.7.1 Relative wedgeness

This section presents the concept of relative wedgeness first introduced in [41]. Let $S \subset \mathbb{R}^N$ be a closed set. S is said to be proximally smooth if there exists $R > 0$ such that the projection map $proj_S(.)$ is a singleton on the set $\{\, x \in \mathbb{R}^N \ : \ d_S(x) < R \,\}$. If S is proximally smooth, then its Clarke tangent cone is equal to its Bouligand tangent cone $\mathcal{T}_S(.)$. It is a known fact that Clarke tangent cone is always closed and convex [9].

Now, let $\mathcal{M}$ be a C^2 embedded manifold in $\mathbb{R}^N$ such that $\overline{\mathcal{M}}$ is proximally smooth and let $d = dim(\mathcal{M})$ be its dimension. Then, for every $x \in \overline{\mathcal{M}}$, the tangent cone $\mathcal{T}_{\overline{\mathcal{M}}}(x)$ is closed and convex, hence it has a relative interior (in the sense of convex analysis), denoted by $r\text{-}int(\mathcal{T}_{\overline{\mathcal{M}}}(x))$.

The set $\overline{\mathcal{M}}$ is said to be relatively wedged if for every $x \in \overline{\mathcal{M}}$, the dimension of $r\text{-}int(\mathcal{T}_{\overline{\mathcal{M}}}(x))$ (in the sense of convex analysis) is equal to the dimension of $\mathcal{M}$:

$$dim(r\text{-}int(\mathcal{T}_{\overline{\mathcal{M}}}(x)) = dim(\mathcal{M}) = d.$$

2.7.2 Lower semicontinuity of the essential dynamics

In this section, we give sufficient conditions for Hypothesis (H_{ESS}) to hold. For $i = 1, \ldots, n+l$, we recall from Section 2.2 that the essential dynamics $F_i^\sharp(.)$ defined on $\overline{\mathcal{M}}_i$ is of the form

$$F_i^\sharp(x) = F_i(x) \cap \mathcal{T}_{\overline{\mathcal{M}}_i}(x), \quad \forall x \in \overline{\mathcal{M}}_i.$$

We suppose that the dynamics $F_i(.)$ verify Hypotheses (SH), (CH) and (H_D). Let $(\mathcal{M}_i)_{i=1,\ldots,n+l}$ be a stratification of $\mathbb{R}^N$ such that any $\overline{\mathcal{M}}_i$ is either vector subspace of $\mathbb{R}^N$ or a half space of $\mathbb{R}^N$. All stratification of $\mathbb{R}^N$ given in Examples 2.1, 2.2 and 2.3 verify this condition. Furthermore, it immediately follows that the stratification verifies (H_1). If $\overline{\mathcal{M}}_i$ is a vector subspace of $\mathbb{R}^N$, then we have $\mathcal{M}_i = \overline{\mathcal{M}}_i$. Consequently we get

$$F_i(.) = F_i^\sharp(.), \quad \forall x \in \overline{\mathcal{M}}_i.$$

Therefore, $F_i^\sharp(.)$ is locally Lipschitz continuous. Hence it is l.s.c.
Suppose now that $\overline{\mathcal{M}}_i$ is a half space of a vector subspace of $\mathbb{R}^N$. For simplicity we denote the vector subspace by $\mathbb{E} \subset \mathbb{R}^N$. Since $\overline{\mathcal{M}}_i$ is a convex subset of $\mathbb{E}$, then by [77, Corollary 3.6.13], the set-valued map

$$\overline{\mathcal{M}}_i \ni x \rightsquigarrow \mathcal{T}_{\overline{\mathcal{M}}_i}(x)$$

is l.s.c as a set-valued map from $\overline{\mathcal{M}}_i$ to $\mathbb{E}$. Furthermore, since $\overline{\mathcal{M}}_i$ is a half space of $\mathbb{E}$, then we have

$$\mathcal{T}_{\overline{\mathcal{M}}_i}(x) = \overline{\mathcal{M}}_i, \quad \forall x \in \overline{\mathcal{M}}_i.$$

Hence, $\mathcal{T}_{\overline{\mathcal{M}}_i}(x)$ is convex with nonempty interior in $\mathbb{E}$ for all $x \in \overline{\mathcal{M}}_i$. On the other hand, by Hypotheses (CH) and (H_D) the set-valued map $x \rightsquigarrow F_i(x)$ is l.s.c as a set-valued map from $\overline{\mathcal{M}}_i$ with images in $\mathbb{E}$ that are convex and have nonempty interior. Therefore, following [78, Theorem B], the set-valued map

$$F_i^\sharp(x) = F_i(x) \cap \mathcal{T}_{\overline{\mathcal{M}}_i}(x)$$

is l.s.c as a set-valued map from $\overline{\mathcal{M}}_i$ to $\mathbb{E}$. Whence, $x \rightsquigarrow F_i^\sharp(x)$ is l.s.c as a set-valued map from $\overline{\mathcal{M}}_i$ to $\mathbb{R}^N$.

2.7.3 Proof of invariance theorems

Proof. (**Theorem 2.11**). Since F is an u.s.c set-valued map with convex, compact and non empty images, and since S is a closed set of $\mathbb{R}^N$, it is a known fact that assertion (i) is equivalent to (see for instance [9, Theorem 12.11])

$$H_F(x,\eta) \geq 0 \quad \forall \eta \in N_S^p(x),\ \forall x \in \mathbb{R}^N.$$

Let $x \in \mathbb{R}^N$. We have $F_i^\sharp(.) \subseteq F_i(.) \subseteq F(.)$. We have

$$H_{F_i}(x,\eta_i) \geq H_{F_i^\sharp}(x,\eta_i), \quad \forall \eta_i \in N_{S_i}^p(x). \tag{2.18}$$

From this inequality, we deduce easily the implication $(iii) \Longrightarrow (ii)$. Moreover, since $N_S^p(x) \subseteq N_{S_i}^p(x)$ for all $i \in I(x)$, then

$$H_F(x,\eta) \geq \max_{i \in I(x)} H_{F_i}(x,\eta) \geq \max_{i \in I(x)} H_{F_i^\sharp}(x,\eta), \quad \forall \eta \in N_S^p(x).$$

Hence, the implication $(ii) \Longrightarrow (i)$ holds. It remains to prove the implication

$$(i) \Longrightarrow (iii).$$

Suppose (S,F) is weakly invariant. Let $x \in \mathbb{R}^N$ and $t \in [0,T]$. So there exists a trajectory $y(.)$ solution of $(DI)_F(t,x)$ such that $y(.) \subset S$. We claim the following:

Claim: $\exists\, j \in I(x)$ such that there exists a sequence $(t_n)_n$, $t_n \downarrow t$ and $x_n := y(t_n) \in \mathcal{M}_j$, so that $\dfrac{x_n - x}{t_n - t} \to \nu$ and $\nu \in F_j^\sharp(x)$.

The proof of the claim is the same as the proof of the same claim in Proposition 2.13. With this claim, we are almost done. Indeed let $\eta_j \in N_{S_j}^p(x)$ be such that the proximal normal inequality in Definition 2.6 is satisfied with $\sigma > 0$. we get

$$\langle \nu, \eta_j \rangle = \lim_{n \to +\infty} \Big\langle \frac{x_n - x}{t_n - t}, \eta_j \Big\rangle \leq \lim_{n \to +\infty} \frac{1}{2\sigma(t_n - t)} |x_n - x|^2 = 0.$$

Thus, we have

$$H_{F_j^\sharp}(x,\eta_j) \geq -\langle \nu, \eta_j \big\rangle \geq 0,$$

Which is the required result. ■

Proof. (**Theorem 2.12**). The implication $(ii) \Longrightarrow (i)$ is proven first. We separate the proof into 3 parts. First we prove the result for every trajectory that lies entirely in one of the domains $\mathcal{M}_i$, $i = 1, ..., n+l$. Then, we prove the result for every trajectory that does not present any chattering phenomenon, also known as Zeno effect, using an induction argument. Finally we prove the result for every trajectory using Filippov's theorem [10, Theorem 3.1.6].

Step 1. (Inspired from [9, Theorem 12.15]). Let $y(.)$ be a trajectory of F such that $y((t,T)) \subset \mathcal{M}_i$ and such that $y(t) = \alpha \in S \cap \overline{\mathcal{M}_i}$. We show that for some $\varepsilon \in (0, T-t)$, we have $y([t, t+\varepsilon]) \subset S$ which is sufficient to conclude.

Let $r > 0$ small enough such that $\mathbb{B}(\alpha, r) \cap \mathcal{M}_i$ is a relative neighborhood in $\mathcal{M}_i$. Let $\kappa > 0$ be the Lipschitz constant of F_i on $\mathbb{B}(\alpha, r)$ and $||F_i|| > 0$ be an upper bound for any velocities that my appear in $\mathbb{B}(\alpha, r)$. So $y(.)$ is Lipschitz continuous in this ball. There exists $\varepsilon \in (0, T-t)$ such that

$$\forall \tau \in [t, t+\varepsilon], \quad s \in proj_{S_i}(y(\tau)) \Longrightarrow y(\tau) \in \mathbb{B}(\alpha, r) \cap \mathcal{M}_i, \quad s \in \mathbb{B}(\alpha, r) \cap \mathcal{M}_i.$$

We define $f(\tau) := d_{S_i}(y(\tau))$. f is Lipschitz continuous on $[t, t+\varepsilon]$. We prove the following Lemma:

Lemma 2.19. $\dfrac{d}{d\tau} f(\tau) = \dot{f}(\tau) \leq \kappa f(\tau)$ *for almost all* $\tau \in (t, t+\varepsilon)$.

Proof. Let $\tau_* \in (t, t+\varepsilon)$ be such that $\dot{f}(\tau_*)$ exists, $\dot{y}(\tau_*)$ exists and $\dot{y}(\tau_*) \in F_i(y(\tau_*))$ (almost all points satisfy those conditions). If $f(\tau_*) = 0$ then f attains a minimum at τ_* and therefore $\dot{f}(\tau_*) = 0$ and the inequality holds. Suppose now $f(\tau_*) > 0$ and let $s \in proj_{S_i}(y(\tau_*))$. Then by [9, Proposition 11.29] we have

$$\eta := \frac{y(\tau_*) - s}{|y(\tau_*) - s|} \in N^p_{S_i}(s).$$

Since $F_i (= F_i^\sharp$ on $\mathcal{M}_i)$ is Lipschitz continuous on $\mathbb{B}(\alpha, r)$ with constant κ, there exists $\nu \in F_i(s)$ such that

$$|\dot{y}(\tau_*) - \nu| \leq \kappa |y(\tau_*) - s|.$$

Therefore we get

$$\langle \eta, \dot{y}(\tau_*) \rangle = \langle \eta, \nu \rangle + \langle \eta, \dot{y}(\tau_*) - \nu \rangle \leq H_{F_i^\sharp}(s, -\eta) + \kappa |y(\tau_*) - s| \leq \kappa |y(\tau_*) - s|,$$

where the last inequality is obtained since $H_{F_i^\sharp}(s, -\eta) \leq 0$ by assumption (ii). Hence, we get

$$\begin{aligned}\dot{f}(\tau_*) = \lim_{\delta \to 0} \frac{d_{S_i}(y(\tau_* + \delta)) - d_{S_i}(y(\tau_*))}{\delta} &\leq \lim_{\delta \to 0} \frac{|y(\tau_* + \delta) - s| - |y(\tau_*) - s|}{\delta} \\ &= \langle \eta, \dot{y}(\tau_*) \rangle \leq \kappa |y(\tau_*) - s| = k f(\tau_*).\end{aligned}$$

■

Since f is Lipschitz, positive and $f(t) = 0$ ($y(t) = \alpha \in S_i$), then by using the Lemma above and Gronwall Lemma [9, Theorem 6.41], we get that $f \equiv 0$ on $[t, t+\varepsilon]$, which finishes the proof of step 1.
Notice that, every trajectory that lies entirely in $\cup_{i=1}^{n}\mathcal{M}_i$ also verifies Step 1 since it is a union of pairwise disjoint open sets.

Step 2. Let $\mathcal{M}$ be a union of subdomains such that $\cup_{i=1}^{n}\mathcal{M}_i \subseteq \mathcal{M}$ and denote by $\delta_{\mathcal{M}}$ the minimum dimension of the subdomains of $\mathcal{M}$. Let $\mathcal{M}_{k_0}$ be a subdomain such that $\mathcal{M}_{k_0} \subset \overline{\mathcal{M}} \setminus \mathcal{M}$ and its dimension is inferior or equal to $\delta_{\mathcal{M}}$. We show the following proposition:

Proposition 2.19.1. *If we have (ii) $\Longrightarrow$ (i) for every trajectory that lies entirely in $\mathcal{M}$ or lies entirely in $\mathcal{M}_{k_0}$, then (ii) $\Longrightarrow$ (i) is verified for every trajectory that lies in $\mathcal{M} \cup \mathcal{M}_{k_0}$.*

Proof. Let $y(.) \subseteq \mathcal{M} \cup \mathcal{M}_{k_0}$ be a trajectory of F on $[t, T]$ such that $y(t) \in S$. We define

$$J = \{\tau \in [t, T] \ : \ y(\tau) \notin \mathcal{M}_{k_0}\}.$$

The set J is open since $\mathcal{M}_{k_0}$ is of inferior dimension than $\mathcal{M}$, then it is a closed set relative to $\mathcal{M} \cup \mathcal{M}_{k_0}$ (equipped with inhereted topology from $\mathbb{R}^N$). Thus J can be written as a countable union of open intervals in the following way:

$$J = \bigcup_{i=1}^{\infty}(a_i, b_i),$$

such that the open sets (a_i, b_i) are pairwise disjoint and $a_i < b_i \leq a_{i+1}$, $\forall i \geq 1$. Notice that we necessarily have $y(a_i), y(b_i) \in \mathcal{M}_{k_0}$. Set $b_0 = t$. First, we prove that

$$\forall\, i \in \mathbb{N},\ y(b_i), y(a_{i+1}) \in S \text{ and } \quad y((a_{i+1}, b_{i+1})) \subset S.$$

We have $y(b_0) \in S$ by assumption. If $b_0 = a_1$ then we have $y(a_1) \in S$. If $b_0 < a_1$ then $y(b_0) \in S \cap \mathcal{M}_{k_0}$ and $y((b_0, a_1)) \subset \mathcal{M}_{k_0}$ almost everywhere. Hence, by assumption of the proposition, we have $y((b_0, a_1)) \subset S$ and therefore $y(a_1) \subset S$ since S is a closed set. Moreover, since $y(a_1) \in S \cap \overline{\mathcal{M}}$ and $y((a_1, b_1)) \subset \mathcal{M}$, then by assumption of the proposition, we have that $y((a_1, b_1)) \subset S$ and therefore $y(b_1) \in S$ since S is closed. By induction, following the same argument, we get that

$$\forall\, i \in \mathbb{N},\ y(b_i), y(a_{i+1}) \in S \text{ and } \quad y([a_{i+1}, b_{i+1}]) \subset S.$$

It remains to prove $y([t, T] \setminus J) \subset S$. If the set J was equal to a finite union of open intervals then the above argument would have been sufficient to prove that $y([t, T] \setminus J) \subset S$. However, this is not the case for all trajectories $y(.)$. The trajectories $y(.)$ can move in and out of $\mathcal{M}_{k_0}$ infinitely many times exhibiting the phenomenon known as the *Zeno effect*, or can reside in $\mathcal{M}_{k_0}$ for sets of time that have a strictly positive Lebesgue measure but are nowhere dense in $[t, T]$. To deal with this case,

we will approximate such trajectories $y(.)$ by ones that behave "nicely".
We fix $m \geq 1$ and set

$$J_m = \bigcup_{k=1}^{m} (a_k, b_k),$$

Which we can assume to satisfy

$$t = b_0 \leq a_1 < b_1 \leq a_2 < b_2 \leq ... \leq a_m < b_m \leq a_{m+1} := T.$$

We choose m large enough such that

$$\mathscr{L}(J \setminus J_m) < \frac{r}{2e^{\kappa T}||F||},$$

with κ being the Lipschitz constant of F_{k_0} and r is equal to

$$r := \inf_{\substack{w \in \overline{\mathcal{M}}_{k_0} \setminus \mathcal{M}_{k_0} \\ s \in [t,T]}} |y(s) - w|,$$

(notice that r is strictly positive and can be infinite). The choice of m is made in such a way to be able to apply the Filippov approximation theorem [10, Theorem 3.1.6] on manifolds (see [41, Remark 3.1]). We will approximate the arc $y([b_i, a_{i+1}])$, for some $i = 0, ..., m$, by trajectories that remain entirely in $\mathcal{M}_{k_0}$. By Filippov approximation theorem ([10, Theorem 3.1.6]) and [41, Proposition 3.2]), there exists $z_i(.)$ a trajectory of F_{k_0} on $[b_i, a_{i+1}]$ such that $z_i(b_i) = y(b_i) \in \mathcal{M}_{k_0} \cap S$, $z_i(.) \subset \mathcal{M}_{k_0}$ and

$$||y(.) - z_i(.)||_{L^\infty[b_i,a_{i+1}]} \leq e^{\kappa(a_{i+1}-b_i)}\rho_i \leq 2\, e^{\kappa T}\, ||F||\varepsilon_i,$$

where we denote $\varepsilon_i = \mathscr{L}(J \cap (b_i, a_{i+1}))$ and

$$\rho_i := \int_{b_i}^{a_{i+1}} d(\dot{y}(s), F_{k_0}(z_i(s)))ds \leq 2\, ||F||\varepsilon_i.$$

Since $\varepsilon_i \leq \mathscr{L}(J \setminus J_m)$, we get

$$||y(.) - z_i(.)||_{L^\infty[b_i,a_{i+1}]} \leq 2\, e^{\kappa T}\, ||F||\mathscr{L}(J \setminus J_m).$$

Furthermore, from the assumption of the proposition we have $z_i(.) \subset S$. Thus we get

$$d_S(y(.)) \leq ||y(.) - z_i(.)||_{L^\infty[b_i,a_{i+1}]} \leq 2\, e^{\kappa T}\, ||F||\mathscr{L}(J \setminus J_m), \qquad \forall m \geq 1.$$

By letting $m \to \infty$, we have $\mathscr{L}(J \setminus J_m) \to 0$. Therefore $y(.) \subset S$, which is the required result. ∎

Step 3. From the above Proposition, we deduce that $(ii) \Longrightarrow (i)$ by a simple finite induction argument starting from $\mathcal{M} = \cup_{i=1}^{n} \mathcal{M}_i$ and adding an interface $\mathcal{M}_{k_0} \subset \Lambda$, with $k_0 \in \{n+1, ..., n+l\}$, in such a way that decreases the dimension of $\mathcal{M}_{k_0}$ at each iteration.

To initiate the induction, take first $\mathcal{M} = \cup_{i=1}^{n} \mathcal{M}_i$ and $\mathcal{M}_{k_0}$, $k_0 \in \{n+1, ..., n+l\}$, with maximal dimension among the submanifolds that constitute the singular set. Then if $y(.)$ a trajectory of F such that $y(t) \in S \cap (\mathcal{M} \cup \mathcal{M}_{k_0})$ and $y([t,T])$ lies entirely either in $\mathcal{M}$ or $\mathcal{M}_{k_0}$, then Step 1 gives us that $y([t,T]) \subset S$. Furthermore, from the proposition above, we get that any trajectory $y(.)$ of F such that $y(t) \in S \cap (\mathcal{M} \cup \mathcal{M}_{k_0})$ and $y([t,T]) \subset \mathcal{M} \cup \mathcal{M}_{k_0}$, verifies $y([t,T]) \subset S$.

Now we prove the direct implication $(i) \implies (ii)$. For that, we use Lemma 2.14. See also [41, Proposition 5.1 and Lemma 5.2]. Suppose (S, F) is strongly invariant. Let $x \in S \cap \overline{\mathcal{M}}_i$, $\nu \in F_i^{\sharp}(x)$ and $\eta \in N_{S_i}^{p}(x)$ such that $|\eta|= 1$ be a proximal normal realised at $\sigma > 0$, from Definition 2.6.
Since $\nu \in F_i^{\sharp}(x)$, then by Lemma 2.14, there exists a C^1 trajectory $y(.)$ of $F_i^{\sharp}$, defined on some interval $[t, t+\varepsilon]$ with $\varepsilon > 0$ such that $y(t) = x$ and $\dot{y}(t) = \nu$ and $y(.) \subseteq \overline{\mathcal{M}}_i$. By the strong invariance hypothesis, we have $y(.) \subseteq S_i$. So we get

$$\langle \nu, \eta \rangle = \lim_{\tau \downarrow t} \Big\langle \frac{y(\tau) - x}{\tau - t}, \eta \Big\rangle \leq \lim_{\tau \downarrow t} \frac{1}{2\sigma(\tau - t)} |y(\tau) - x|^2 = 0.$$

By taking the supremum over $F_i^{\sharp}(x)$, we obtain the desired result. ∎

Chapter 3

Viscosity solutions of Hamilton Jacobi equations in proper CAT(0) spaces

Contents

3.1 Introduction

First order Hamilton Jacobi equations have been exstensively studied in the Euclidean space or more generally in infinite dimensional Banach spaces that enjoy the

Radon Nikodym property or have a smooth norm. A substential literature exists on the subject going back to several decades ago [1, 28, 29, 79]. More recently, there has been an increasing interest in studying first order Hamilton Jacobi equations posed in more general metric spaces. Typical examples include topological networks, the space of Borel probability measures, or more generally any generic metric space. This problem involves many challenging mathematical issues and has a wide range of applications in various fields such as data transmission, social network problems, traffic management problems, fluid mechanics, optimal control of multi-agent systems and mean field game problems.

Several new notions of viscosity were proposed for first order Hamilton Jacobi equations in metric spaces. Since a notion of a differential for real valued functions defined in a general metric space is not well defined, the Hamiltonians studied in this case depend on the differential of the unknown function only through its local Lipschitz constant, called the local slope. In [32, 80], the authors studied a class of Hamilton Jacobi equations of Eikonal type in a general metric space. The notion of viscosity used by the authors is defined via optimal control interpretations along absolutely continuous curves. This has the advantage to reduce the viscosity notion into a one dimensional problem and requires no structure on the space considered. In [81, 33, 34], the authors proposed a different notion of viscosity for a similar class of Hamilton Jacobi equations defined in a complete geodesic metric space using local slopes and suitable test functions. In [82], the authors provide a comparison between these notions of viscosity. In particular, they show that the notions coincide in the case of the Eikonal equation defined in a general geodesic space.

On the other hand, there is a growing interest in studying Hamilton Jacobi equations on a simpler structure in the form of a network. The latter is defined as a finite collection of isometric half-spaces glued together along their boundary. For example in the one dimensional case, a network is the result of gluing a finite number of half-lines along their origin. The subset where the gluing operation occurs is called the *junction*. On each branch of the network, one defines a Hamiltonian, the Hamiltonians are a priori independent from one another and a discontinuity occurs at the junction. Thanks to the smooth structure that each branch of the network possesses, one can define more general Hamilton Jacobi equations than the Eikonal type equations. The notion of viscosity solution is defined here using test functions that are continuous on the network and continuously differentiable on each branch. First, the special case of the Eikonal equation on networks has been considered in [20, 21]. Later in [24] the authors treated the case of convex Hamiltonians on each branch in a one dimensional network. In their work, an additional junction condition is considered, called the flux-limiter, in order to guarantee well-posedness of the problem. These results have been extended to the case of quasi-convex Hamiltonians in [22, 23] and the case of a higher dimensional network was treated in [83]. In [27] the authors studied the case of a one dimensional network with Hamiltonians that are not necessarily convex nor quasi-convex. They introduced a junction condition called the Kirchoff condition and proved well-posedness of the problem using the same notion of viscosity as in [22, 23]. Furthermore, they proved that the flux-limiter type condition at the junction is a special case of the Kirchoff-type condition. The book [43] offers a detailed discussion on the different junction conditions considered on networks and the notion of viscosity solution adopted in this space.

The current state of the art keeps many problems unsolved. Indeed, the techniques developed for the treatment of Hamilton Jacobi equations in a general metric space are restricted to a certain class of Hamilton Jacobi equations such as Eikonal equations. On the other hand, the setting of a network allows to study more general Hamilton Jacobi equations but the techniques used in this setting do not take advantage of the metric structure of the network and focus more on the differential structure that exists on each branch. Furthermore, extending these current results to a network where the branches have different Hausdorff dimensions is still a challenging question. The purpose of this chapter is to define a viscosity notion for first order Hamilton Jacobi equations in a class of metric spaces general enough that includes Euclidean spaces and networks. Furthermore, this viscosity notion should ideally coincide with the classical one developed in Euclidean spaces. Therefore, we focus our attention in this manuscript on developing a theory of first order viscosity notion in a subclass of metric spaces called CAT(0) spaces.

A metric space (X, d), is said to be a CAT(0) space[1] if, roughly speaking, it is a geodesic space and its geodesic triangles are "thinner" than the triangles of the Euclidean plane $\mathbb{R}^2$ (see Definition 3.1). This method of comparing geodesic triangles of a geodesic space with triangles from a model space, such as the Euclidean plane, was first introduced by Alexandrov [51, 52] to give a synthetic definition of curvature for geodesic spaces. In particular, CAT(0) spaces are spaces of curvature not greater than 0 in the sense of Alexandrov. Typical examples of CAT(0) spaces are Hilbert spaces, convex sets of Hilbert spaces, simply connected Riemannian manifolds with nonpositive sectional curvature, multi-dimensional networks and the space of Borel probability measures over the real line [85]. Although CAT(0) spaces do not possess any smooth structure, they carry a solid first order calculus similar to smooth manifolds with sectional curvature not greater than 0. For example, a notion of tangent cone T_xX is well defined at each point of X. The tangent cone is the metric counterpart of the tangent space in Riemannian geometry or the Bouligand tangent cone in convex analysis. Furthermore, a notion of differential is well defined for any function $u : X \to \mathbb{R}$ that is Lipschitz and DC. By DC functions we mean real valued functions that can be represented as a difference of two semiconvex functions. The exact definition of this class of functions is given in Definition 3.4. We refer to [86, 50, 52, 51, 53] for a more detailed discussion on the topic.

In this chapter, we propose to study first order Hamilton Jacobi equations in proper

[1]The acronyme "CAT" stands for the initials of the three mathematicians Cartan, Alexandrov and Toponogov. This notation was introduced by Gromov in 1987 [84, p.119].

CAT(0) spaces, i.e., CAT(0) spaces whose closed bounded sets are compact. More specifically, we consider the following stationary problems,

$$\begin{cases} H(u(x), x, D_x u) = 0, & \forall x \in \Omega, \\ u(x) = \ell(x), & \forall x \in \partial\Omega, \end{cases} \tag{3.1}$$

and the time dependent problems,

$$\begin{cases} \partial_t u + H(x, D_x u) = 0, \quad \forall\, (t, x) \in (0, +\infty) \times X, \\ u(0, x) = \ell(x), \quad x \in X, \end{cases} \tag{3.2}$$

where Ω is an open set, ℓ is a real valued continuous and bounded function on its domain, and u is a Lipschitz and DC function. The differential $D_x u : T_x X \to \mathbb{R}$ is defined in the tangent cone of X at a point x. The differential $D_x u$ is itself a Lipschitz, DC and positively homogeneous function (see Proposition 3.8.1). The Hamiltonian $H : \mathbb{R} \times X \times \mathrm{DC}_1(T_x X) \to \mathbb{R}$ is a real valued function. The set $\mathrm{DC}_1(T_x X)$ represents Lipschitz, positively homogeneous and DC functions on $T_x X$.

The viscosity notion we use here is different from what is currently present in the literature. We define the notion of viscosity using subsets of the class of Lipschitz and DC functions. More precisely, we test upper semicontinuous subsolutions with Lipschitz semiconvex functions and we test lower semicontinuous supersolutions with Lipschitz semiconcave functions. We prove comparison results that hold for any upper semicontinuous subsolution and any lower semicontinuous supersolution using the same techniques as in the classical theory of viscosity. In particular, we apply the variable doubling technique using the squared distance function in the same way as in [4, 3]. Comparison results guarantee uniqueness of the solution. Furthermore, we prove existence of the solution by virtue of Perron's method in the same way originally developped in [4, 3].

We give several examples showing the degree of generality of our setting. Namely, we show that the setting developed in this chapter coincides with classical setting when $X = \mathbb{R}^N$ by treating the examples of Hamilton Jacobi Bellman equations and Isaacs' equations defined in $\mathbb{R}^N$. Furthermore, we give several examples of Eikonal type equations and Eikonal type equations in the presence of an obstacle defined in proper CAT(0) spaces of the form:

- the proper CAT(0) space obtained by gluing together three half-lines of $\mathbb{R}^2$

$$\begin{cases} X_1 := [0, +\infty) e_1, \\ X_2 := [0, +\infty) e_2, \\ X_3 := [0, +\infty) e_3, \end{cases}$$

  along the origin point $A = \{0\}$;

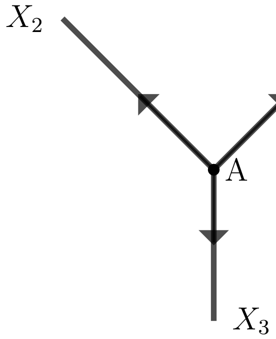


Figure 3.1: The space obtained by gluing X_1, X_2 and X_3 along A.

- the proper CAT(0) space obtained by gluing together the sets

$$\begin{cases} X_1 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_3 = 0\}, \\ X_2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_1 = x_2 = 0\}, \end{cases}$$

along the origin point $A = \{0\}$;

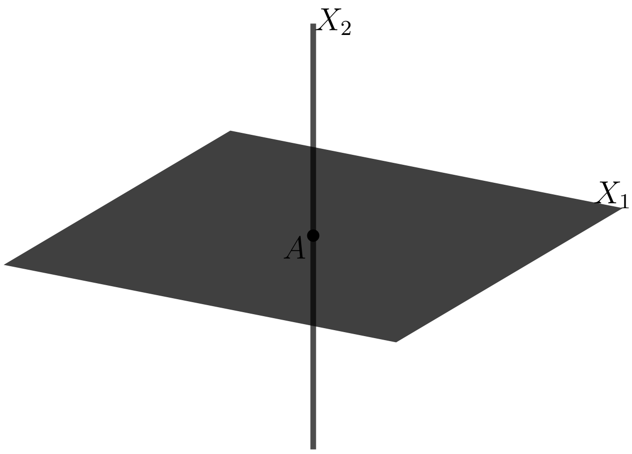


Figure 3.2: The space obtained by gluing X_1 and X_2 along A.

- the CAT(0) space obtained by gluing together the sets

$$\begin{cases} X_1 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_3 = 0\}, \\ X_2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_1 = x_2 = 0, \ x_3 \geq 0\}, \end{cases}$$

along the origin point $A = \{0\}$;

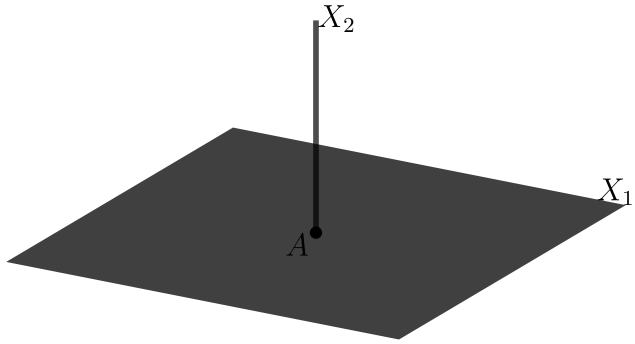


Figure 3.3: The space obtained by gluing X_1 and X_2 along A.

The gluing operation will be defined precisely in Section 3.2. The justification that the above spaces are proper CAT(0) spaces will also be given in Section 3.2. The examples of the Hamilton Jacobi equations we treat in the spaces above are of the following form:

$$\gamma u(x) + \sup_{\substack{v \in T_x X \\ |v|_x = 1}} \{-D_x u \centerdot v\} - b(x) = 0, \quad x \in X,$$

and

$$\min \left\{ \gamma u(x) + \sup_{\substack{v \in T_x X \\ |v|_x = 1}} \{-D_x u \centerdot v\} - b_1(x), \gamma u(x) - b_2(x) \right\} = 0, \quad x \in X,$$

where $\gamma > 0$ is a strictly positive constant and $b, b_1, b_2 : X \to \mathbb{R}$ are Lipschitz and bounded functions. For the time dependent case we treat the following Eikonal equation

$$\begin{cases} \partial_t u + \sup\limits_{\substack{v \in T_x X \\ |v|_x = 1}} \{-D_x u \centerdot v\} = 0, & (t, x) \in (0, +\infty) \times X, \\ u(0, x) = \ell(x). \end{cases}$$

where $\ell : X \to \mathbb{R}$ is a continuous and bounded function.

The chapter is organized the following way. In Section 3.2 we give the definition of CAT(0) spaces, we describe the gluing operation of a collection of CAT(0) spaces, we define the central notion of the tangent cone and we give the definition of DC functions and their differential in CAT(0) spaces. In Section 3.3 we define the notion of viscosity solution and the general form of the Hamiltonian we are going to work with. We show that we recover the main features of viscosity theory in this setting: the comparison principle and Perron's method. Finally, we give several examples showing the degree of generality of our setting. In particular, we treat classical examples of Hamilton Jacobi equations when the space $X = \mathbb{R}^N$ to demonstrate that our setting coincides with the classical one in $\mathbb{R}^N$ and we treat the case of Eikonal type equations defined in several structures as the ones presented above.

3.2 Calculus in CAT(0) spaces

Let us briefly recall some facts in metric geometry. Classical references are [86, 50, 52, 51, 53]. Let (X, d) be a metric space. For $x \in X$ and $r > 0$ we denote by $B(x, r)$ and $\overline{B}(x, r)$ the open and closed balls of center x and radius r respectively.

The metric space (X, d) is said to be *proper* if all of its closed bounded sets are compact sets.

Let $l > 0$. A metric space (X, d) is said to be a *geodesic space* if any two points $x, y \in X$ are connected by at least one *unit speed geodesic*, i.e. a map $\gamma : [0, l] \to X$ such that $\gamma_0 = x$, $\gamma_l = y$ and

$$d(\gamma_t, \gamma_s) = |t - s|, \quad \forall t, s \in [0, l].$$

In particular, we necessarily have $l = d(x, y)$. The image of γ is called the *geodesic segment* with endpoints x and y.

Let $I \subset \mathbb{R}$ be an interval. A map $\gamma : I \to X$ is said to be a *constant speed geodesic* if there exists a constant $D \geq 0$ such that

$$\forall s, t \in I, \quad d(\gamma_s, \gamma_t) = D|s - t|.$$

In what follows, we will refer to constant speed geodesics simply by *geodesics*.

A geodesic space (X, d) is said to be *geodesically extendible* if for every geodesic $\gamma : [a, b] \to X$ with $a < b \in \mathbb{R}$, there exists a geodesic $\tilde{\gamma} : (-\infty, +\infty) \to X$ such that

$$\tilde{\gamma}|_{[a,b]} = \gamma.$$

3.2.1 CAT(0) spaces

Let (X, d) be a geodesic space. In order to define the notion of curvature of X, we shall first introduce the notion of *geodesic triangles*. A geodesic triangle $\triangle(x, y, z)$ in a geodesic space is the result of three points $x, y, z \in X$, called the *vertices*, together with a choice of three corresponding geodesics, the *edges*, linking the vertices. A *comparison triangle* for the geodesic triangle $\triangle(x, y, z)$ is a geodesic triangle built in the Euclidean plane $(\mathbb{R}^2, d_{\mathbb{R}^2})$, denoted by $\bar{\triangle}(\bar{x}, \bar{y}, \bar{z})$, with $\bar{x}, \bar{y}, \bar{z} \in \mathbb{R}^2$, such that

$$d_{\mathbb{R}^2}(\bar{x}, \bar{y}) = d(x, y), \quad d_{\mathbb{R}^2}(\bar{y}, \bar{z}) = d(y, z), \quad d_{\mathbb{R}^2}(\bar{x}, \bar{z}) = d(x, z).$$

The choice of the comparison triangle is unique up to an isometry [50, Lemma I.2.14]. A point $a \in X$ is said to be between y and z provided that we have

$$d(y, a) + d(z, a) = d(y, z).$$

This means that the point a lies in a geodesic segment of y and z. The *comparison point* of a is the unique point $\bar{a} \in \mathbb{R}^2$, once the comparison triangle $\bar{\triangle}(\bar{x}, \bar{y}, \bar{z})$ is fixed, such that

$$d_{\mathbb{R}^2}(\bar{y}, \bar{a}) + d_{\mathbb{R}^2}(\bar{z}, \bar{a}) = d_{\mathbb{R}^2}(\bar{y}, \bar{z}).$$

Definition 3.1 (CAT(0) spaces)**.** A metric space (X, d) is called a CAT(0) space if it is a geodesic space and satisfies the following comparison triangle inequality: for any $x, y, z \in X$ and any point $a \in X$ between y and z, the comparison points $\bar{x}, \bar{y}, \bar{z}, \bar{a} \in \mathbb{R}^2$ satisfy

$$d(x, a) \leq d_{\mathbb{R}^2}(\bar{x}, \bar{a}). \tag{3.3}$$

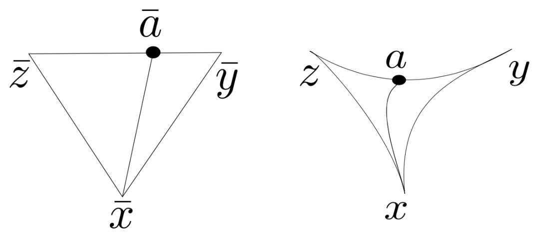


Figure 3.4: The comparison triangle on the left and the geodesic triangle on the right.

Remark 3.2.1. The comparison triangle inequality (3.3) is equivalent to the following one: for any $x, y, z \in X$ and any comparison points $\bar{x}, \bar{y}, \bar{z} \in \mathbb{R}^2$,

$$d^2(\gamma_t, x) \leq d^2_{\mathbb{R}^2}\Big((1-t)\bar{y} + t\bar{z}, \bar{x}\Big), \quad \forall t \in [0, 1], \tag{3.4}$$

where $\gamma : [0, 1] \to X$ is the geodesic joining $\gamma_0 = y$ and $\gamma_1 = z$. By expanding the right hand side of (3.4) using the elementary properties of the inner product in $\mathbb{R}^2$, it becomes

$$d^2(\gamma_t, x) \leq (1-t)d^2(\gamma_0, x) + td^2(\gamma_1, x) - t(1-t)d^2(\gamma_0, \gamma_1), \quad \forall t \in [0, 1]. \tag{3.5}$$

Inequality (3.5) can be used in an equivalent way as a definition of CAT(0) spaces. It can be understood as a synthetic inequality that quantifies the deficit of the curvature of X with respect to the Euclidean space $\mathbb{R}^2$, where inequality (3.5) is an equality. In other words, inequality (3.5) quantifies how much the triangle $\triangle(x, y, z)$ in X is thinner with respect to the triangle $\bar{\triangle}(\bar{x}, \bar{y}, \bar{z})$ in $\mathbb{R}^2$.

Example 3.1. Here are some examples of CAT(0) spaces [50, Example II.1.15].

- Euclidean spaces, Hilbert spaces (the only Banach spaces which are CAT(0)).
- Convex subsets of Hilbert spaces.
- Convex subsets of other CAT(0) spaces.
- The n-dimensional hyperbolic space, denoted $\mathbb{H}^n$. It is the unique simply connected, n-dimensional complete Riemannian manifold with a constant negative sectional curvature equal to -1.
- Simply connected Riemannian manifolds with sectional curvature not greater than 0.

- Metric $\mathbb{R}$-trees, i.e. any metric space T such that:
  - there exists a unique geodesic segment joining each pair of points $x, y \in T$; we denote it by $[x, y]$;
  - if $[x, y] \cap [y, z] = \{y\}$, then $[x, y] \cup [y, z] = [x, z]$.
- The 2-Wasserstein space over the real line, denoted $\mathcal{P}_2(\mathbb{R})$ [85, Proposition 4.1].

Following [50, Proposition II.1.4], an important result which is a consequence of Definition 3.1 is that in a CAT(0) space (X, d), any two points $x, y \in X$ are connected by a *unique* geodesic joining x and y.

Let (X, d) be a CAT(0) space. A subset $\mathcal{C} \subset X$ is said to be *convex* if for every $x, y \in \mathcal{C}$, the geodesic segment connecting x and y lies entirely in $\mathcal{C}$.
In the Euclidean plane $\mathbb{R}^2$, the open balls are convex. Hence, from Definition 3.1, it is straighforward to prove that the open balls of (X, d) are convex (see [50, Proposition II.1.4-(3)] for a detailed proof of this fact). Furthermore, any convex subset of X equipped with the distance d is also a CAT(0) space [50, Examples II.1.15].

Another useful result concerning CAT(0) spaces is that any product of two CAT(0) spaces is a CAT(0) space when equipped with the product distance, as the following lemma shows.

Lemma 3.2. *([50, Exercice II.1.16]). Let (Y, d_Y) and (Z, d_Z) be two* CAT(0) *spaces. Then the product space $(Y \times Z, d_{Y\times Z})$, equipped with the distance*

$$d^2_{Y\times Z}\Big((y_1, z_1), (y_2, z_2)\Big) := d^2_Y(y_1, y_2) + d^2_Z(z_1, z_2),$$

is a CAT(0) *space. Moreover, if Y and Z are proper spaces then the product space is also a proper space.*

3.2.2 Gluing constructions

In this section, we will discuss the most obvious way of gluing metric spaces, which is to attach them along isometric subsets. Furthermore, we will see that when the underlying metric spaces are CAT(0) spaces, and the isometric subsets are complete CAT(0) subspaces, then the resulting space by the gluing operation is a CAT(0) space. In this section, the set I will denote an arbitrary index set (countable or uncountable). The following definition can be found in [50, Definition I.5.23]

Definition 3.2. (Gluing operation). Let I be an index set. Let $(X_\lambda, d_\lambda)_{\lambda\in I}$ be a familly of metric spaces. Let $A_\lambda \subset X_\lambda$ be fixed closed subsets. Let A be a metric space and suppose that for each $\lambda \in I$, there exists an isometry $i_\lambda : A \to A_\lambda$. Let

$\cup_{\lambda\in I}X_\lambda$ be the disjoint union of the metric spaces X_λ, $\lambda \in I$. We define the space X as the quotient space of $\cup_I X_\lambda$ by the equivalence relation,

$$\forall x, y \in \cup_I X_\lambda,\ x\,\mathcal{R}\,y \ \Leftrightarrow\ \exists\, a \in A,\ \lambda, \lambda' \in I \ :\ x \in A_\lambda, y \in A_{\lambda'} \\ \text{and } \ i_\lambda^{-1}(\{x\}) = i_{\lambda'}^{-1}(\{y\}) = a,$$

where we identify each X_λ with its image in X. X is called *the glued space* along A and is denoted

$$X := \bigsqcup_A X_\lambda.$$

Some examples of glued spaces will be given below. The following theorem shows how to define a distance on the glued space X and summarizes its main properties.

Theorem 3.3. *([50, Lemma I.5.24]). Let* $X = \bigsqcup_A X_\lambda$*. Let* $x \in X_\lambda$ *and* $y \in X_{\lambda'}$*. we define the following function,*

$$d(x, y) := \begin{cases} d_\lambda(x, y) & \text{if } \lambda = \lambda', \\ \inf\limits_{a\in A} \{d_\lambda(x, i_\lambda(a)) + d_{\lambda'}(x, i_{\lambda'}(a)) & \text{if } \lambda \neq \lambda'. \end{cases}$$

We have

1. *d is a distance on X;*
2. *if I is finite and each (X_λ, d_λ) is proper, then (X, d) is proper;*
3. *if each space (X_λ, d_λ) is a geodesic space and A is proper, then (X, d) is a geodesic space.*

For $\mathrm{CAT}(0)$ spaces, we have a stronger result that we give in the next proposition.

Proposition 3.3.1. *([50, Theorem II.11.3] Gluing families of* $\mathrm{CAT}(0)$ *spaces). Let I be an index set. Let $(X_\lambda, d_\lambda)_{\lambda\in I}$ be a family of* $\mathrm{CAT}(0)$ *spaces. Let $A_\lambda \subset X_\lambda$ be closed subsets. Let A be a metric space and suppose that for all $\lambda \in I$, there exist isometries $i_\lambda : A \to A_\lambda$. Let $X = \sqcup_A X_\lambda$ be the resulting glued space along A.*

If A is a complete $\mathrm{CAT}(0)$ *space, then the glued space X is a* $\mathrm{CAT}(0)$ *space, endowed with the distance defined in Theorem 3.3.*

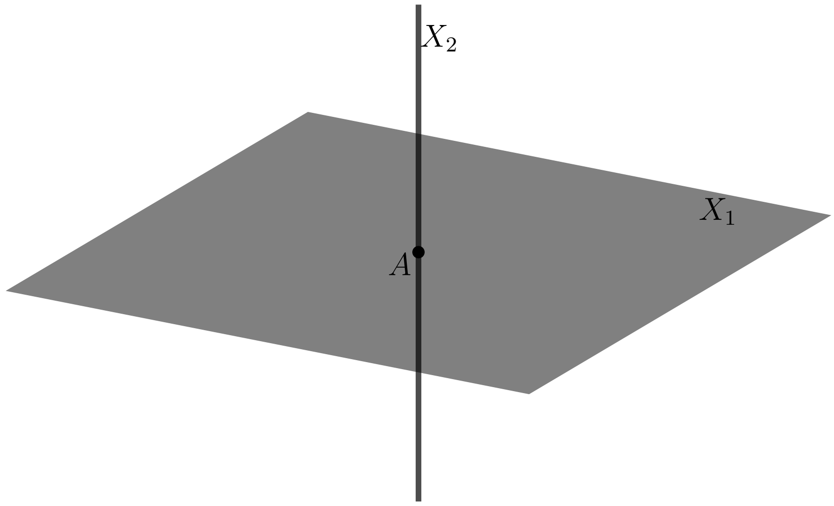
Example 3.4. Let X_1 and X_2 be the following two proper $\mathrm{CAT}(0)$ spaces:

$$\begin{cases} X_1 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ :\ x_3 = 0\}, \\ X_2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ :\ x_1 = x_2 = 0\}. \end{cases}$$

Let $A := \{0\}$. We consider the following glued space

$$X := X_1 \bigsqcup_A X_2,$$

along A. The resulting glued space X is a proper $\mathrm{CAT}(0)$ space.



Indeed, X_1 are X_2 are Euclidean spaces. Hence they are proper CAT(0) spaces when endowed with their Euclidean distances. Furthermore, A is a complete CAT(0) space since it is reduced to one point. Hence, according to Theorem 3.3 and Proposition 3.3.1, X is a proper CAT(0) space when endowed with its geodesic distance. The resulting distance is obtained thanks to Theorem 3.3 in the following way

$$\forall x, y \in X,\ d(x,y) := \begin{cases} |x-y|, \text{ if } \exists\, i \in \{1,2\}\ :\ x, y \in X_i, \\ |x|+|y|, \text{ otherwise}, \end{cases}$$

where $|.|$ denotes the Euclidean norm.

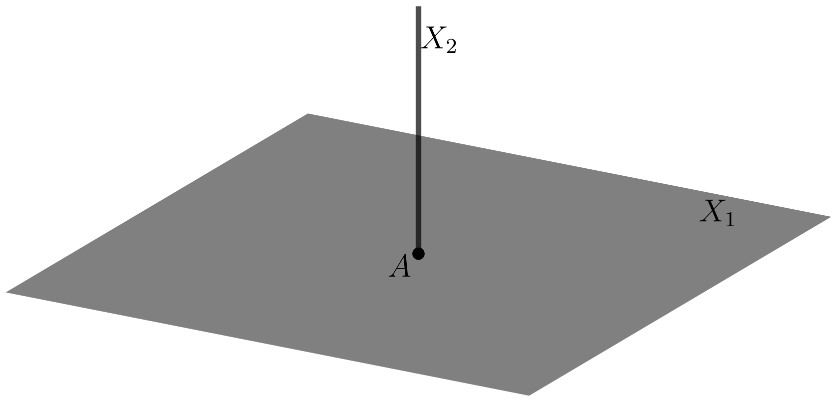
Example 3.5. Let X_1 and X_2 be the following two proper CAT(0) spaces:

$$\begin{cases} X_1 := \{(x_1, x_2, x_3) \in \mathbb{R}^3\ :\ x_3 = 0\}, \\ X_2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3\ :\ x_1 = x_2 = 0,\ x_3 \geq 0\}. \end{cases}$$

Let $A := \{0\}$. We consider the following glued space

$$X := X_1 \bigsqcup_A X_2,$$

along A. The resulting glued space X is a proper CAT(0) space.

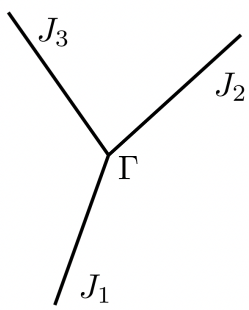


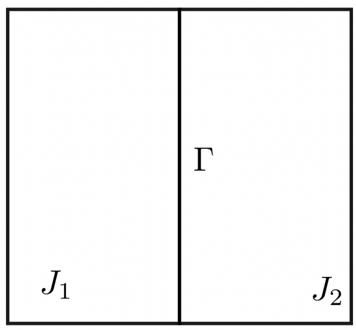
Indeed, X_1 are X_2 are closed convex subsets of Euclidean spaces. Hence they are proper CAT(0) spaces when endowed with their Euclidean distances. Furthermore, A is a complete CAT(0) space since it is reduced to one point. Hence, according to Theorem 3.3 and Proposition 3.3.1, X is a proper CAT(0) space when endowed with its geodesic distance. The resulting distance is obtained thanks to Theorem 3.3 in the following way

$$\forall x, y \in X,\ d(x,y) := \begin{cases} |x-y|, \text{ if } \exists\, i \in \{1,2\} \ :\ x, y \in X_i, \\ |x|+|y|, \text{ otherwise}, \end{cases}$$

where $|.|$ denotes the Euclidean norm.

Example 3.6. Here are more examples covered by this setting:





On the left, the space

$$J = \bigsqcup_{\Gamma} J_i$$

is the result of gluing three copies of the half-line $[0, +\infty)$ along the subset $\Gamma = \{0\}$. On the right, the space J is isometric to the Euclidean plane $\mathbb{R}^2$ obtained by gluing two copies of the half-plane $\{(x,y) \in \mathbb{R}^2 \ :\ x \geq 0\}$, along the subset $\Gamma = \{(x,y) \in \mathbb{R}^2 \ :\ x = 0\}$.

3.2.3 Tangent cone

In this section, we recall the notion of the tangent cone on geodesic spaces and give its main properties in the case of CAT(0) spaces. The tangent cone is a central notion in metric geometry, similar to the tangent space for differentiable manifolds or the Bouligand tangent cone in convex analysis. We refer to the bibliography mentioned at the beginning of this section for a more detailed discussion.

Let (X, d) be a geodesic space and let $x \in X$. We denote by $\mathrm{Geo}_x(X)$ the set of geodesics emanating from x and defined in some neighborhood of the form $[0, \varepsilon]$, with $\varepsilon > 0$. Let $\eta, \gamma \in \mathrm{Geo}_x(X)$. Then the following quantity

$$d_x(\eta, \gamma) := \limsup_{t \downarrow 0} \frac{d(\eta_t, \gamma_t)}{t} \tag{3.6}$$

is a pseudo-distance on the space $\mathrm{Geo}_x(X)$. Moreover, d_x defines an equivalence relation on $\mathrm{Geo}_x(X)$ in the following way:

$$\forall \eta, \gamma \in \mathrm{Geo}_x(X),\ \eta \sim \gamma \quad \text{if and only if} \quad d_x(\eta, \gamma) = 0.$$

The quotient space $\mathrm{Geo}_x(X)/\sim$ endowed with quotient distance, still denoted by d_x, is a metric space. The equivalence class of a geodesic $\gamma \in \mathrm{Geo}_x(X)$ under the equivalence relation $\sim$ is denoted by $\gamma_0' \in \mathrm{Geo}_x(X)/\sim$. It represents the initial *velocity* or *direction* of γ.

Definition 3.3 (Tangent cone)**.** Let (X, d) be a geodesic space and $x \in X$. The *tangent cone* at x is the metric space (T_xX, d_x), where T_xX is the abstract completion of $(\mathrm{Geo}_x(X)/\sim, d_x)$, i.e.

$$T_xX := \overline{\mathrm{Geo}_x(X)/\sim}^{d_x}.$$

We denote by $0_x \in T_xX$ the equivalence class of the geodesic with speed equal to 0 in $\mathrm{Geo}_x(X)/\sim$. It is called *the origin or the apex* of the tangent cone T_xX.

Example 3.7. If (X, d) is a simply connected manifold with sectional curvature not greater 0, then the tangent cone at a point $x \in X$ is isometric to the usual tangent space.

When (X, d) is a general geodesic space, the structure of the tangent cone at a point can be very wild and little can be said about it. However, when (X, d) is a CAT(0) space, then the tangent cone behaves nicely. This fact is exploited to built a first order calculus in (X, d). First, we have the following key result. If (X, d) is a CAT(0) space and $x \in X$, then the supremum limit in (3.6) is actually a limit. In fact, we have a stronger result. The tangent cone at a given point x of a CAT(0) space is a complete CAT(0) space when endowed with the distance d_x [50, Theorem II-3.19]. Furthermore, the tangent cone of a CAT(0) space has a structure resembling a Hilbert space. This is due to the fact that it is a complete CAT(0) space and it has a cone structure.
To make the latter statement clearer, first notice that for any $\lambda \geq 0$, the map sending the geodesics $(\gamma_t) : t \mapsto \gamma_t \in \mathrm{Geo}_x(X)$ to the geodesics $(\gamma_{\lambda t}) : t \mapsto \gamma_{\lambda t} \in \mathrm{Geo}_x(X)$ can be passed to the quotient $\mathrm{Geo}_x(X)/\sim$ and the resulting quotient map sending the equivalence class of (γ_t) to the equivalence class of $(\gamma_{\lambda t})$ is λ-Lipschitz on $\mathrm{Geo}_x(X)/\sim$. Indeed, for any two geodesics $(\gamma_t)_t$, $(\eta_t)_t$ that belong to $\mathrm{Geo}_x(X)$, we have

$$\lim_{t\downarrow 0} \frac{d(\gamma_{\lambda t}, \eta_{\lambda t})}{t} = \lambda \lim_{t\downarrow 0} \frac{d(\gamma_{\lambda t}, \eta_{\lambda t})}{\lambda t} = \lambda \lim_{s\downarrow 0} \frac{d(\gamma_s, \eta_s)}{s}.$$

Therefore, by passing to the quotient, the map is λ-Lipschitz from $\mathrm{Geo}_x(X)/\sim$ to itself. Hence it can be extended by continuity to T_xX and can be seen as the operation of *multiplication by a positive scalar*. We denote it the following way:

$$\forall\, v \in T_xX,\ \forall \lambda \geq 0, \quad \lambda v \in T_xX.$$

Thus T_xX has a structure of a cone. Moreover, for any $v, w \in T_xX$ and $\lambda \in \mathbb{R}^+$, we define the norm and the scalar product on T_xX the following way:

$$\text{Norm}: \quad |v|_x := d_x(v, 0_x), \tag{3.7a}$$

$$\text{Scalar product}: \quad \langle v, w\rangle_x := \frac{1}{2}(|v|_x^2 + |w|_x^2 - d_x^2(v, w)). \tag{3.7b}$$

Furthermore, we have the following results on the norm and scalar product.

Proposition 3.7.1. *([87, Proposition 2.11] Calculus on the tangent cone)***.** *Let (X, d) be a* CAT(0) *space, let $x \in X$ be a fixed point and T_xX be the tangent cone of X at x. Then the operations* (3.7a) *and* (3.7b) *are continuous in their variables. The operation* (3.7b) *is symmetric. Furthermore, we have*

$$|\lambda v|_x = \lambda |v|_x, \tag{3.8a}$$

$$\langle \lambda v, w\rangle_x = \langle v, \lambda w\rangle_x = \lambda \langle v, w\rangle_x, \tag{3.8b}$$

$$|\langle v, w\rangle_x| \leq |v|_x|w|_x \text{ and } \langle v, w\rangle_x = |v|_x|w|_x \text{ if and only if } |w|_x v = |v|_x w, \tag{3.8c}$$

for all $v, w \in T_xX$ and $\lambda \in \mathbb{R}^+$.

Since CAT(0) spaces are uniquely geodesic, meaning that any two points are connected by a unique unit speed geodesic, we introduce the following notation which is going to be useful throughout this chapter.

Notation 3.8. Let (X, d) be a CAT(0) space, and let $x, y \in X$. the unique unit speed geodesic connecting x and y is denoted by

$$t \mapsto G_t^{x,y}, \quad \forall\, t \in [0, d(x, y)].$$

Furthermore, we denote by

$$\uparrow_x^y := (G_0^{x,y})' \in T_xX$$

the direction of $G^{x,y}$ at x. The direction between x and y has a norm equal to 1, meaning that

$$|\uparrow_x^y|_x = 1.$$

3.2.4 DC calculus

In this section, we introduce the notion of real valued *directionally differentiable* functions in CAT(0) spaces. A special attention will be given to Lipschitz functions that are semiconvex or semiconcave since they are differentiable at every point according to this definition. These functions are going to serve us as test functions in the definition of viscosity notion in the next section.

Let (X, d) be a CAT(0) space and $x \in X$. Let $f : X \to \mathbb{R}$ be a function. We say that f has a *directional derivative* at x along the geodesic $\gamma : [0, \varepsilon] \to X$ emanating from x, with $\varepsilon > 0$, if the limit

$$\left.\frac{d}{dt}\right|_{t=0} f(\gamma_t) = \lim_{t \downarrow 0} \frac{f(\gamma_t) - f(\gamma_0)}{t}$$

exists and is finite.

Definition 3.4. Let $f : X \to \mathbb{R}$ be a function.

- We say that f is *semiconvex* if there exists $\lambda \in \mathbb{R}$ such that for every geodesic $\gamma : [0,1] \to X$ the following inequality holds:

$$f(\gamma_t) \le (1-t)f(\gamma_0) + tf(\gamma_1) - \frac{\lambda}{2}t(1-t)d^2(\gamma_0,\gamma_1), \tag{3.9}$$

  or equivalently, if the real-to-real function

$$t \mapsto f(\gamma_t) - \frac{\lambda}{2}d^2(\gamma_0,\gamma_1)t^2$$

  is convex. We also say that f is λ-*convex*. If $\lambda = 0$ then we simply say that f is convex.

- We say that f is *semiconcave* (or λ-*concave* for some $\lambda \in \mathbb{R}$) if and only if $-f$ is semiconvex (or $(-\lambda)$-convex).

- Finally, we say that f is a DC function if it can be represented as a difference of two semiconvex functions.

In particular, every semiconvex function is a DC function and every semiconcave function is also a DC function. Furthermore, we can define locally semiconvex and locally smiconcave functions as well.

Let $\Omega \subset X$ be an open subset. A function $f : \Omega \to \mathbb{R}$ is said to be *locally semiconvex* if for any point $x \in X$ there exists a neighborhood U_x of x such that for all geodesics $\gamma : [0,1] \to \Omega$ with endpoints in U_x we have that inequality (3.9) holds. Similarly, a function $f : \Omega \to \mathbb{R}$ is is said to be *locally semiconcave* if and only if $-f$ is locally semiconvex. Finally, a function $f : \Omega \to \mathbb{R}$ is said to be a *locally DC* function if it can be locally represented as a difference of two semiconvex functions.

Let $\Omega \subset X$ be an open subset and $x \in \Omega$. Let $f : \Omega \to \mathbb{R}$ be a locally Lipschitz and locally semiconvex function. Then the directional derivative of f along any geodesic emanating from x exists and is finite by [87, Proposition 2.16]. Furthermore, we define the *differential function* of f at x from its directional derivatives as the map $D_x f : (\mathrm{Geo}_x X/\sim) \to \mathbb{R}$ defined as

$$D_x f \centerdot \gamma_0' := \frac{d}{dt}\Big|_{t=0} f(\gamma_t) = \lim_{t\downarrow 0} \frac{f(\gamma_t) - f(\gamma_0)}{t}, \quad \forall \gamma \in \mathrm{Geo}_x X, \quad \gamma_0' \in \mathrm{Geo}_x X/\sim .$$

Notice that the above definition does not depend on the choice of the geodesic γ whose velocity is γ_0'. Moreover, the differential function is Lipschitz, convex and positively homogeneous. Thus it can be uniquely extended to the whole tangent cone $T_x X$ by density. These properties are collected in the next proposition.

Proposition 3.8.1. *([87, Proposition 2.16] Differential of semiconvex functions).* *Let $f : \Omega \to \mathbb{R}$ be a locally Lipschitz and locally semiconvex function around $x \in \Omega$. Then f is directionally differentiable at x and the differential function $D_x f : T_x X \to \mathbb{R}$ is Lipschitz, convex and positively homogeneous, i.e.,*

$$D_x f \centerdot (\lambda v) = \lambda D_x f \centerdot v, \quad \forall v \in T_x X \text{ and } \lambda \geq 0.$$

Similarly, if $f : \Omega \to \mathbb{R}$ is locally Lipschitz and locally semiconcave, then it is differentiable at any $x \in \Omega$ and its differential function is Lipschitz, concave, positively homogeneous and defined By

$$D_x f := -D_x(-f).$$

Finally, if $f : \Omega \to \mathbb{R}$ is a locally Lipschitz and locally DC function, then f is differentiable at any $x \in \Omega$ and its differential function is Lipschitz, DC and positively homogeneous.

We denote by $\mathrm{DC}_{\mathrm{lip}}(\Omega)$ the class of locally Lipschitz and locally DC functions on Ω. We also denote by $\mathrm{DC}_1(T_x X)$ the class of Lipschitz, DC and positively homogeneous functions on the tangent cone $T_x X$ at some point $x \in X$. Finally we denote by $\mathrm{DC}_1(TX)$ the set

$$\mathrm{DC}_1(TX) := \{(x, p_x) \in X \times \mathrm{DC}_1(T_x X)\},$$

which is the metric analogue of the cotangent bundle in this setting.

Next, we give several examples of locally Lipschitz and locally DC functions in CAT(0) spaces to demonstrate how abundant these functions are in this class of metric spaces. Moreover, we will give the explicit expression of their differential function at every point.

In the Euclidean plane $\mathbb{R}^2$, for $\bar{y} \in \mathbb{R}^2$ fixed, the Euclidean distance function $\bar{x} \mapsto d_{\mathbb{R}^2}(\bar{x}, \bar{y})$ is Lipschitz continuous and convex and the squared Euclidean distance function $\bar{x} \mapsto d^2_{\mathbb{R}^2}(\bar{x}, \bar{y})$, is locally Lipschitz continuous and 2-convex. It follows directly from Definition 3.1 that for any CAT(0) space (X, d) and $y \in X$ fixed, the distance function $x \mapsto d(x, y)$ is Lipschitz continuous and convex [50, Proposition II.2.2] and from Remark 3.2.1 that the squared distance function $x \mapsto d^2(x, y)$ is locally Lipschitz continuous and 2-convex. Furthermore, their differential functions are given explicitly in the next proposition, whose proof can be found in [87, Proposition 2.17].

Proposition 3.8.2. *Let (X, d) be a CAT(0) space. Let $y \in X$ be a fixed point. The following properties hold true.*

- *For all $x \in X$, we have*

$$\forall v \in T_x X, \quad D_x d(., y) \centerdot v = \begin{cases} -\langle v, \uparrow_x^y \rangle_x, & \text{if } x \neq y, \\ |v|_x, & \text{if } x = y, \end{cases}$$

  where $\uparrow_x^y$ is given in Notation 3.8.

- *For all $x \in X$, we have*

$$\forall v \in T_xX, \quad D_xd^2(.,y) \centerdot v = -2d(x,y)\langle v, \uparrow_x^y \rangle_x.$$

More generally, if (X,d) is a CAT(0) space and $\mathcal{C}$ is a complete convex subset of X, then the distance function to $\mathcal{C}$ is also Lipschitz and convex. We summarize the main properties of the distance function to a complete convex subset in the next proposition.

Proposition 3.8.3. *(Distance function to a convex set [50, Proposition II.2.4, Corollary II.2.5]). Let (X,d) be a CAT(0) space. Let $\mathcal{C}$ be a complete convex subset of X. Then the following holds:*

1. *for every $x \in X$, there exists a unique point $\pi(x)$ called the* projection *of x onto $\mathcal{C}$ such that*
$$d(x,\pi(x)) = d(x,\mathcal{C}) := \inf_{y \in \mathcal{C}} d(x,y);$$
2. *for all $x,y \in X$, we have $|d(x,\mathcal{C}) - d(y,\mathcal{C})| \leq d(x,y)$;*
3. *the function $x \mapsto d(x,\mathcal{C})$ is convex.*

Since the distance to a closed convex subset in a CAT(0) space is Lipschitz and convex, then according to Proposition 3.8.1, it is differentiable at every point. The next proposition is the first result of this chapter where we give the explicit expression of the differential of the distance function to a complete convex subset in a CAT(0) space.

Theorem 3.9 (Differential of the distance function to a complete convex set)**.** *Let (X,d) be a CAT(0) space. Let $\mathcal{C}$ be a complete convex subset of X. Then the following holds:*

$$\forall x \in X,\ \forall v \in T_xX, \quad D_xd(.,\mathcal{C}) \centerdot v = \begin{cases} -\langle \uparrow_x^{\pi(x)}, v \rangle_x, & \text{if } x \notin \mathcal{C}, \\ d_x(v, T_x\mathcal{C}), & \text{if } x \in \mathcal{C}, \end{cases}$$

where, $\pi(x)$ is the projection of x onto $\mathcal{C}$, $T_x\mathcal{C}$ is the tangent cone of $x \in \mathcal{C}$, when $(\mathcal{C},d)$ is seen as a complete CAT(0) space and

$$d_x(v, T_x\mathcal{C}) := \inf_{w \in T_x\mathcal{C}} d_x(v,w).$$

The tangent cone $T_x\mathcal{C}$ is a complete convex subset of the CAT(0) space (T_xX, d_x).

Proof. The proof is decomposed into two steps. Let $x \in X$.

$\boxed{\text{Step 1}}$. If $x \notin C$, then we have

$$\forall y \in X, \quad d(y,\mathcal{C}) \leq d(y,\pi(x)),$$

which implies that

$$\forall \gamma(.) \in \mathrm{Geo_x}, \quad \lim_{t \downarrow 0} \frac{d(\gamma_t, \mathcal{C}) - d(x, \mathcal{C})}{t} \le \lim_{t \downarrow 0} \frac{d(\gamma_t, \pi(x)) - d(x, \pi(x))}{t}.$$

The last inequality is equivalent to

$$\forall v \in \mathrm{Geo_x}/\sim, \quad D_x d(., \mathcal{C}) \centerdot v \le D_x d(., \pi(x)) \centerdot v.$$

By Lipschitz continuity of the differential functions, we get

$$\forall v \in T_x X, \quad D_x d(., \mathcal{C}) \centerdot v \le D_x d(., \pi(x)) \centerdot v,$$

and by Proposition 3.8.2 we have

$$\forall v \in T_x X, \quad D_x d(., \pi(x)) \centerdot v = -\langle \uparrow_x^{\pi(x)}, v \rangle_x.$$

Therefore, we get

$$\forall v \in T_x X, \quad D_x d(., \mathcal{C}) \centerdot v \le -\langle \uparrow_x^{\pi(x)}, v \rangle_x.$$

For the other inequality, let $v \in \mathrm{Geo_x}/\sim$ and let $\gamma : [0, r] \to X$ be a geodesic such that $\gamma'_0 = v$. First, by [50, Lemma II.3.20] we have

$$\lim_{s \to 0} d_x(\uparrow_x^{\pi(\gamma_s)}, \uparrow_x^{\pi(x)}) = 0.$$

Moreover, by Proposition 3.8.2 we have

$$D_x d(., \pi(x)) \centerdot v = -\langle v, \uparrow_x^{\pi(x)} \rangle_x.$$

Therefore, by the continuity of the scalar product asserted in Proposition 3.7.1 we get

$$D_x d(., \pi(x)) \centerdot v = -\langle v, \uparrow_x^{\pi(x)} \rangle_x = \lim_{s \downarrow 0} -\langle v, \uparrow_x^{\pi(\gamma_s)} \rangle_x.$$

Furthermore, we have

$$\begin{aligned} D_x d(., \pi(x)) \centerdot v = \lim_{s \downarrow 0} -\langle v, \uparrow_x^{\pi(\gamma_s)} \rangle_x &= \lim_{s \downarrow 0} \lim_{t \downarrow 0} \frac{d(\gamma_t, \pi(\gamma_s)) - d(x, \pi(\gamma_s))}{t} \\ &= \lim_{s \downarrow 0} \inf_{0 < t < r} \frac{d(\gamma_t, \pi(\gamma_s)) - d(x, \pi(\gamma_s))}{t}. \end{aligned}$$

The last equality is true since the real-to-real function

$$t \mapsto d(\gamma_t, \pi(\gamma_s))$$

is convex, so the incremental ratio

$$(0, r] \ni t \mapsto \frac{d(\gamma_t, \pi(\gamma_s)) - d(x, \pi(\gamma_s))}{t}$$

is non decreasing. The monotonicity property of the above incremental ratio also gives us

$$\lim_{s\downarrow 0}\inf_{0<t<r}\frac{d(\gamma_t,\pi(\gamma_s))-d(x,\pi(\gamma_s))}{t}\leq \lim_{s\downarrow 0}\inf_{s\leq t<r}\frac{d(\gamma_t,\pi(\gamma_s))-d(x,\pi(\gamma_s))}{t}$$
$$=\lim_{s\downarrow 0}\frac{d(\gamma_s,\pi(\gamma_s))-d(x,\pi(\gamma_s))}{s}.$$

Moreover we have

$$\forall s\in[0,r],\quad d(x,\pi(\gamma_s))\geq d(x,\mathcal{C}),\quad \text{and}\quad d(\gamma_s,\pi(\gamma_s))=d(\gamma_s,\mathcal{C}).$$

Therefore, we get

$$\begin{aligned}D_x d(.,\pi(x))\,\centerdot\, v &\leq \lim_{s\downarrow 0}\frac{d(\gamma_s,\pi(\gamma_s))-d(x,\pi(\gamma_s))}{s}\\ &\leq \lim_{s\downarrow 0}\frac{d(\gamma_s,\mathcal{C})-d(x,\mathcal{C})}{s}\\ &= D_x d(.,\mathcal{C})\,\centerdot\, v.\end{aligned}$$

This is true for any $v\in \mathrm{Geo_x}/\sim$. Lastly, by the Lipschitz continuity of the differentials we get

$$\forall v\in T_xX,\quad D_x d(.,\pi(x))\,\centerdot\, v\leq D_x d(.,\mathcal{C})\,\centerdot\, v,$$

which is the desired inequality.

$\boxed{\text{Step 2}}$. Let $x\in\mathcal{C}$. By the Lipschitz continuity of the differentials, it is enough to consider only geodesic directions in T_xX. Let $\gamma:[0,r]\to X$ be a geodesic such that $\gamma_0'=v\in\mathrm{Geo_x}/\sim$. Then we have

$$D_x d(.,\mathcal{C})\,\centerdot\, v=\lim_{t\downarrow 0}\frac{d(\gamma_t,\mathcal{C})}{t}=\inf_{0<t<r}\frac{d(\gamma_t,\mathcal{C})}{t},$$

where the last equality holds since the real-to-real function $t\mapsto d(\gamma_t,\mathcal{C})$ is convex. Consequently we have

$$D_x d(.,\mathcal{C})\,\centerdot\, v=\inf_{0<t<r}\frac{d(\gamma_t,\mathcal{C})}{t}=\inf_{0<t<r}\inf_{y\in\mathcal{C}}\frac{d(\gamma_t,y)}{t}.$$

On the other hand, we have

$$\begin{aligned}d_x(v,T_x\mathcal{C}) := \inf_{w\in T_x\mathcal{C}} d_x(v,w)=\inf_{\beta\in\mathrm{Geo}_x(\mathcal{C})}\lim_{t\downarrow 0}\frac{d(\gamma_t,\beta_t)}{t} &= \inf_{y\in\mathcal{C}}\lim_{t\downarrow 0}\frac{d(\gamma_t,y)}{t}\\ &=\inf_{y\in\mathcal{C}}\inf_{0<t<r}\frac{d(\gamma_t,y)}{t},\end{aligned}$$

where the last two equalities hold because

$$\{\beta_t\ :\ \beta:[0,r']\to X\in\mathrm{Geo}_x(\mathcal{C}),\ \text{for some } r'\geq 0,\ \text{and } t\in[0,r']\}=\{y\ :\ y\in\mathcal{C}\},$$

and the real-to-real function $t \mapsto d(\gamma_t, y)$ is convex.
Finally, notice that we have

$$\inf_{0<t<r} \inf_{y\in\mathcal{C}} \frac{d(\gamma_t, y)}{t} = \inf_{y\in\mathcal{C}} \inf_{0<t<r} \frac{d(\gamma_t, y)}{t} = \inf_{\substack{0<t<r \\ y\in\mathcal{C}}} \frac{d(\gamma_t, y)}{t}.$$

This ends the proof. ∎

3.3 Stationary Hamilton Jacobi equations in proper CAT(0) spaces

In this section, we study first order Hamilton Jacobi equations in proper CAT(0) spaces. We recall that a metric space is proper if its closed bounded sets are compact. We use subsets of Lipschitz DC functions as test functions to define the viscosity notion. More precisely, we use subsets of semiconvex functions to test subsolutions and subsets of semiconcave functions to test supersolutions. With this class of test functions, we will see that we can define a notion of viscosity for first order Hamilton Jacobi equations in proper CAT(0) spaces and recover the main features of the theory: the comparison principle and Perron's method. Throughout this section, (X, d) is a proper CAT(0) space.

First, we define the notion of viscosity used throughout this section. Let Ω be an open subset of X. We denote by $\overline{\Omega}$ its closure and we set $\partial\Omega := \overline{\Omega} \setminus \Omega$. Let

$$H : \mathbb{R} \times \mathrm{DC}_1(TX) \to \mathbb{R}$$

be a function called the *Hamiltonian* and $\ell : \partial\Omega \to \mathbb{R}$ be a bounded continuous function. We consider the following Hamilton Jacobi equation

$$\begin{cases} H(u(x), x, D_x u) = 0, & \forall x \in \Omega, \\ u(x) = \ell(x), & \forall x \in \partial\Omega, \end{cases} \tag{3.10}$$

where $u : \overline{\Omega} \to \mathbb{R}$ is a Lipschitz and DC function which is the unknown of equation (3.10). We give the following definition of *classical solutions* of equation (3.10).

Definition 3.5 (Classical solutions)**.** A Lipschitz and DC function $u : \overline{\Omega} \to \mathbb{R}$ is said to be a *classical solution* of (3.10) if for every $x \in \Omega$ we have

$$H(u(x), x, D_x u) = 0,$$

and $u = \ell$ on $\partial\Omega$.

Let $\mathcal{TEST}_-$ and $\mathcal{TEST}_+$ be two subsets of $\mathrm{DC}_{\mathrm{lip}}(\Omega)$. $\mathcal{TEST}_-$ and $\mathcal{TEST}_+$ will de given precisely later. We are now ready to define the notion of viscosity solutions. This definition is dependent upon the choice of $\mathcal{TEST}_-$ and $\mathcal{TEST}_+$.

Definition 3.6. (Viscosity solution).

- An upper semicontinuous function $u : \Omega \to \mathbb{R}$ is said to be a *viscosity subsolution* of (3.10) if, for any $\phi \in \mathcal{TEST}_-$ such that $u - \phi$ attains a local maximum at x, we have
$$H(u(x), x, D_x\phi) \leq 0.$$
- Similarly, a lower semicontinuous function $u : \Omega \to \mathbb{R}$ is said to be a *viscosity supersolution* of (3.10) if, for any $\phi \in \mathcal{TEST}_+$ such that $u - \phi$ attains a local minimum at x, we have
$$H(u(x), x, D_x\phi) \geq 0.$$
- A continuous function $u : \overline{\Omega} \to \mathbb{R}$ is said to be a viscosity solution of (3.10) if it is both a viscosity supersolution and a viscosity subsolution and satisfies the boundary condition
$$u = \ell, \quad \text{in } \partial\Omega.$$

3.3.1 Comparison principle

Let $H : \mathbb{R} \times \mathrm{DC}_1(TX) \to \mathbb{R}$ be a Hamiltonian and Ω be an open subset of X. We consider the following Hamilton Jacobi equation:

$$H(u(x), x, D_x u) = 0, \quad \forall\, x \in \Omega. \tag{3.11}$$

Let us give now the test functions we use for (3.11). We saw in the previous section that real valued locally Lipschitz and locally DC functions of X behave well in this setting. In particular, they are directionally differentiable at every point and the differential is Lipschitz, positively homogeneous and a DC function. Therefore, we will consider subsets of $\mathrm{DC}_{\mathrm{lip}}(\Omega)$ that verify the following properties given below.

Definition 3.7. (Test functions).
Let $\mathcal{TEST}_-$ be a subset of $\mathrm{DC}_{\mathrm{lip}}(\Omega)$ such that

- constant functions belong to $\mathcal{TEST}_-$;
- for all $\phi(.), \psi(.) \in \mathcal{TEST}_-$ and $a, b \geq 0$, $a\,\phi(.) + b\,\psi(.) \in \mathcal{TEST}_-$;
- let $y \in X$ be fixed. Then the function $x \mapsto d^2(x, y)$ belongs to $\mathcal{TEST}_-$.

Let $\mathcal{TEST}_+$ be a subset of $\mathrm{DC}_{\mathrm{lip}}(\Omega)$ such that

- $\mathcal{TEST}_+ = -\mathcal{TEST}_- := \{-\phi(.)\ : \phi(.) \in \mathcal{TEST}_-\}$.

Example 3.10. For example, one can take the following test functions

$$\mathcal{TEST}_- := \{\text{real valued locally Lipschitz and locally semiconvex functions}\}.$$

Thus we have

$$\mathcal{TEST}_+ = \{\text{real valued locally Lipschitz and locally semiconcave functions}\}.$$

We test subsolutions with $\mathcal{TEST}_-$ functions and supersolution with $\mathcal{TEST}_+$ functions. Next, we prove the comparison principle for the Hamilton Jacobi equation (3.11). We assume the following hypotheses on the Hamiltonian.

Hypothesis 3.1. The Hamiltonian H is such that there exists $K_{db} > 0$ such that for all $\alpha > 0$, for all $r \in \mathbb{R}$ and for all $x, y \in \Omega$, we have

$$H(r, x, D_x(-\alpha d^2(., y))) - H(r, y, D_y(\alpha d^2(x, .))) \leq K_{db} d(x, y)(1 + \alpha d(x, y)).$$

Hypothesis 3.2. The Hamiltonian H is such that there exists $\gamma > 0$ such that

$$\gamma(r - s) \leq H(r, x, p) - H(s, x, p) \quad \text{for all } r \geq s,\ x \in \Omega, \text{ and } p \in \mathrm{DC}_1(T_x X).$$

Now, we prove the following key lemma. It allows to use the variable doubling technique to prove comparison type results. It was first proven in [3, Proposition 3.7] in the particular case of Euclidean spaces. We prove it here for every metric space.

Lemma 3.11. *Let $\mathcal{O}$ be a subset of a metric space (Z, d_Z). Let $\Phi : \mathcal{O} \to \mathbb{R}$ be an upper semicontinuous function and $\Psi : \mathcal{O} \to \mathbb{R}$ be a lower semicontinuous function such that $\Psi \geq 0$, and*

$$M_{\alpha_n} = \sup_{z \in \mathcal{O}} \{\, \Phi(z) - \alpha_n \Psi(z) \,\},$$

with $(\alpha_n)_n \subset \mathbb{R}^+ \setminus \{0\}$ is an increasing sequence such that $\alpha_n \to +\infty$ as $n \to +\infty$. Suppose that $\lim_{\alpha_n \to +\infty} M_{\alpha_n}$ exists and

$$-\infty < \lim_{\alpha_n \to +\infty} M_{\alpha_n} < +\infty.$$

Let $z_{\alpha_n} \in \mathcal{O}$ be chosen such that

$$\lim_{\alpha_n \to +\infty} (M_{\alpha_n} - (\Phi(z_{\alpha_n}) - \alpha_n \Psi(z_{\alpha_n}))) = 0.$$

Then the following holds:

$$\begin{cases} (i) & \lim_{\alpha_n \to +\infty} \alpha_n \Psi(z_{\alpha_n}) = 0, \\ (ii) & \Psi(\hat{z}) = 0 \text{ and } \Phi(\hat{z}) = \sup_{\{\Psi(z)=0\}} \Phi(z) = \lim_{\alpha_n \to +\infty} M_{\alpha_n}, \\ & \text{whenever } \hat{z} \in \mathcal{O} \text{ is an accumulation point of } (z_{\alpha_n})_{\alpha_n}. \end{cases}$$

Proof. The proof is exactly the same as in [3, Proposition 3.7] even though it was asserted for Euclidean spaces. We give it hereafter for the sake of completeness. Let

$$\delta_{\alpha_n} = M_{\alpha_n} - (\Phi(z_{\alpha_n}) - \alpha_n\,\Psi(z_{\alpha_n})),$$

so that $\lim_{\alpha_n\to+\infty} \delta_{\alpha_n} = 0$. Since $\Psi \geq 0$, M_{α_n} decreases as α_n increases and $\lim_{\alpha_n\to+\infty} M_{\alpha_n}$ exists and is finite by assumption. Furthermore, we have:

$$M_{\frac{\alpha_n}{2}} \geq \Phi(z_{\alpha_n}) - \frac{\alpha_n}{2}\,\Psi(z_{\alpha_n}) = \Phi(z_{\alpha_n}) - \alpha_n\,\Psi(z_{\alpha_n}) + \frac{\alpha_n}{2}\,\Psi(z_{\alpha_n}) = M_{\alpha_n} - \delta_{\alpha_n} + \frac{\alpha_n}{2}\,\Psi(z_{\alpha_n}),$$

which implies that $\alpha_n\,\Psi(z_{\alpha_n}) \leq 2\,(\delta_{\alpha_n} + M_{\frac{\alpha_n}{2}} - M_{\alpha_n})$ and therefore

$$\lim_{\alpha_n\to+\infty} \alpha_n\,\Psi(z_{\alpha_n}) = 0.$$

Suppose now that there exists a subsequence of $(z_{\alpha_n})_{\alpha_n}$, not relabeled here, that converges to $\hat{z} \in \mathcal{O}$. Then $\lim_{\alpha_n\to+\infty} \Psi(z_{\alpha_n}) = 0$ and by lower semicontinuity and positivity of Ψ we also get $\Psi(\hat{z}) = 0$. Moreover, since

$$\Phi(z_{\alpha_n}) - \alpha_n\,\Psi(z_{\alpha_n}) = M_{\alpha_n} - \delta_{\alpha_n} \geq \sup_{\{\Psi(z)=0\}} \Phi(z) - \delta_{\alpha_n},$$

and Φ is upper semicontinuous, we get by letting $\alpha_n \to \infty$

$$\sup_{\{\Psi(z)=0\}} \Phi(z) \geq \Phi(\hat{z}) \geq \lim_{\alpha_n\to\infty} M_{\alpha_n} \geq \sup_{\{\Psi(z)=0\}} \Phi(z),$$

which forces equality everywhere. this ends the proof. ∎

In the next theorem, we prove the comparison principle on a bounded open subset of X. The proof is similar to the proof of the comparison principle in the classical theory of viscosity. The main difference here is that we use test functions that verify Definition 3.7.

Theorem 3.12 (Comparison on bounded domains)**.** *Assume H satisfies Hypotheses 3.1 and 3.2. Let Ω be an open* bounded *set of X and set $\partial\Omega = \overline{\Omega} \setminus \Omega$. Consider $u : \overline{\Omega} \to \mathbb{R}$ a bounded from above upper semicontinuous subsolution of* (3.11)*, and $v : \overline{\Omega} \to \mathbb{R}$ a bounded from below lower semicontinuous supersolution of* (3.11)*.*
Then $u \leq v$ in $\partial\Omega$ implies $u \leq v$ in $\overline{\Omega}$.

Proof. Let $M := \sup_{\overline{\Omega}}(u(x) - v(x))$. Assume by contradiction that $u \leq v$ in $\partial\Omega$ and $M > 0$.
For every $\alpha > 0$, define $\psi_\alpha : \overline{\Omega} \times \overline{\Omega} \to \mathbb{R}$ as

$$\psi_\alpha(x,y) = u(x) - v(y) - \frac{\alpha}{2} d^2(x,y), \quad \forall (x,y) \in \overline{\Omega} \times \overline{\Omega}.$$

Since u and $-v$ are bounded from above and $u-v$ is upper semicontinuous, the supremum $M_\alpha := \sup\limits_{\overline{\Omega}\times\overline{\Omega}} \psi_\alpha$ is reached. Let (x_α, y_α) be such that $M_\alpha = \psi_\alpha(x_\alpha, y_\alpha)$. We have

$$\lim_{\alpha\to+\infty}(M_\alpha - \psi_\alpha(x_\alpha, y_\alpha)) = 0, \quad \text{and} \quad -\infty < M \leq M_\alpha \leq \sup_{\overline{\Omega}}(u) + \sup_{\overline{\Omega}}(-v) < +\infty.$$

Since $\overline{\Omega}$ is closed and bounded, then it is compact by the assumption of X being proper. Hence we can take a subsequence $((x_{\alpha_n}, y_{\alpha_n}))_{\alpha_n}$ that converges as $\alpha_n \to +\infty$. We have

$$\lim_{\alpha_n\to+\infty}(M_{\alpha_n} - \psi_{\alpha_n}(x_{\alpha_n}, y_{\alpha_n})) = 0, \quad \text{and} \quad -\infty < \lim_{\alpha_n\to+\infty} M_{\alpha_n} < +\infty.$$

Therefore, we can apply Lemma 3.11 via the correspondences

$$Z = X\times X, \quad \mathcal{O} = \overline{\Omega}\times\overline{\Omega}, \quad \Phi(z) = u(x) - v(y), \quad \Psi(z) = \frac{1}{2}d^2(x,y),$$

and we get

$$\begin{cases} (i) & \lim\limits_{\alpha_n\to+\infty} \dfrac{\alpha_n}{2} d^2(x_{\alpha_n}, y_{\alpha_n}) = 0, \\ (ii) & \lim\limits_{\alpha_n\to+\infty} M_{\alpha_n} = M. \end{cases}$$

It follows that for α_n big enough we have $x_{\alpha_n}, y_{\alpha_n} \in \Omega$ since $u \leq v$ in $\partial\Omega$. Thus we get

$$H\Big(v(y_{\alpha_n}), y_{\alpha_n}, D_{y_{\alpha_n}}(-\frac{\alpha_n}{2}d^2(x_{\alpha_n}, .))\Big) \ \geq \ 0 \ \geq H\Big(u(x_{\alpha_n}), x_{\alpha_n}, D_{x_{\alpha_n}}(\frac{\alpha_n}{2}d^2(., y_{\alpha_n}))\Big). \tag{3.12}$$

Hence, using Hypotheses 3.1 and 3.2 and the above inequality, we get

$$\begin{aligned}
\gamma(u(x_{\alpha_n}) - v(y_{\alpha_n})) &\overset{3.2}{\leq} H\Big(u(x_{\alpha_n}), x_{\alpha_n}, D_{x_{\alpha_n}}(\frac{\alpha_n}{2}d^2(., y_{\alpha_n}))\Big) \\
&\quad - H\Big(v(y_{\alpha_n}), x_{\alpha_n}, D_{x_{\alpha_n}}(\frac{\alpha_n}{2}d^2(., y_{\alpha_n}))\Big) \\
&\overset{(3.12)}{\leq} H\Big(v(y_{\alpha_n}), y_{\alpha_n}, D_{y_{\alpha_n}}(-\frac{\alpha_n}{2}d^2(x_{\alpha_n}, .))\Big) \\
&\quad - H\Big(v(y_{\alpha_n}), x_{\alpha_n}, D_{x_{\alpha_n}}(\frac{\alpha_n}{2}d^2(., y_{\alpha_n}))\Big) \\
&\overset{3.1}{\leq} K_{db} d(x_{\alpha_n}, y_{\alpha_n}) \Big(1 + \frac{\alpha_n}{2} d(x_{\alpha_n}, y_{\alpha_n})\Big).
\end{aligned}$$

By letting $\alpha_n \to +\infty$, we get

$$\gamma M \leq 0,$$

a contradiction with $M > 0$. ■

If Ω is an unbounded open set of X, then we need the following additional hypothesis to prove the comparison principle.

Hypothesis 3.3. The Hamiltonian H is such that there exists $K_L > 0$ such that, for every $x \in \Omega$ and $r \in \mathbb{R}$, we have

$$\forall p_x, q_x \in \mathrm{DC}_1(T_xX), \quad \Big| H(r,x,p_x) - H(r,x,q_x) \Big| \leq K_L \sup_{|v|_x=1} |p_x \cdot v - q_x \cdot v|.$$

Remark 3.3.1. Note that the mapping

$$\mathrm{DC}_1(T_xX) \ni p_x \mapsto \sup_{|v|_x=1} |p_x \cdot v|$$

verifies all the axioms of a norm on $\mathrm{DC}_1(T_xX)$.

Remark 3.3.2. Hypothesis 3.3 asserts that the Hamiltonian H is Lipschitz continuous with respect to the variable p_x. When $X = \mathbb{R}^N$ and the test functions are continuously differentiable, then Hypothesis 3.3 is the same as the Lipschitz assumption on p_x usually required for the Hamiltonian in the classical theory of viscosity.

Theorem 3.13 (Comparison on unbounded domains)**.** *Assume H satisfies Hypotheses 3.1, 3.2 and 3.3. Let Ω be an open set of X and set $\partial\Omega = \overline{\Omega} \setminus \Omega$. Let $u : \overline{\Omega} \to \mathbb{R}$ be a bounded from above upper semicontinuous subsolution of (3.11), and $v : \overline{\Omega} \to \mathbb{R}$ a bounded from below lower semicontinuous supersolution of (3.11). Then*

$$u \leq v \text{ in } \partial\Omega \text{ implies } u \leq v \text{ in } \overline{\Omega}.$$

Proof. Let $M := \sup_{\overline{\Omega}}(u(x) - v(x))$. Assume by contradiction that $M > 0$ and $u \leq v$ in $\partial\Omega$.

Let

$$\varepsilon \in \left(0, \min\left\{M, \left(\frac{\gamma}{\gamma + 4K_L} M\right)^2, 1\right\}\right),$$

where γ and K_L are given in Hypotheses 3.2 and 3.3. Let $x_\varepsilon \in \Omega$ be such that

$$u(x_\varepsilon) - v(x_\varepsilon) \geq M - \varepsilon > 0.$$

For $\alpha > 0$, set

$$\psi_\alpha(x,y) = u(x) - v(y) - \Big(d^2(x,x_\varepsilon) + d^2(y,x_\varepsilon)\Big) - \frac{\alpha}{2} d^2(x,y), \quad \forall (x,y) \in \overline{\Omega} \times \overline{\Omega}.$$

It is clear that ψ_α is upper semicontinuous and bounded from above. Set $M_\alpha := \sup_{\overline{\Omega}\times\overline{\Omega}} \psi_\alpha$. We have

$$0 < \psi_\alpha(x_\varepsilon, x_\varepsilon)$$

and for all $x, y \in \overline{\Omega} \setminus \overline{B}\left(x_\varepsilon, \sqrt{|\sup_{\overline{\Omega}}(u)|+|\sup_{\overline{\Omega}}(-v)|}\right)$ we have

$$\psi_\alpha(x,y) \leq 0.$$

Hence, the supremum of ψ_α is reached on a compact set. Let (x_α, y_α) be such that M_α is reached. We have

$$\lim_{\alpha\to+\infty}(M_\alpha - \psi_\alpha(x_\alpha, y_\alpha)) = 0 \quad \text{and} \quad -\infty < M - \varepsilon \leq M_\alpha \leq \sup_{\overline{\Omega}}(u) + \sup_{\overline{\Omega}}(-v) < +\infty.$$

Since x_α, y_α are in a compact set, then we can take a subsequence $(x_{\alpha_n}, y_{\alpha_n})$ that converges as $\alpha_n \to +\infty$ and

$$\lim_{\alpha_n\to+\infty}(M_{\alpha_n} - \psi_{\alpha_n}(x_{\alpha_n}, y_{\alpha_n})) = 0, \quad \text{and} \quad -\infty < \lim_{\alpha_n\to+\infty} M_{\alpha_n} < +\infty.$$

Therefore, we can apply Lemma 3.11 via the correspondences

$$Z = X \times X, \quad \mathcal{O} = \overline{\Omega}\times\overline{\Omega}, \quad \Phi(z) = u(x) - v(y) - \Big(d^2(x, x_\varepsilon) + d^2(y, x_\varepsilon)\Big), \quad \Psi(z) = \frac{1}{2}d^2(x,y),$$

and we get

$$\begin{cases} (i) & \lim_{\alpha_n\to+\infty} \dfrac{\alpha_n}{2} d^2(x_{\alpha_n}, y_{\alpha_n}) = 0, \text{ and } x_{\alpha_n}, y_{\alpha_n} \to \hat{x} \in \overline{\Omega}, \\ (ii) & \lim_{\alpha_n\to+\infty} M_{\alpha_n} = \sup_{x\in\overline{\Omega}} u(x) - v(x) - 2d^2(x, x_\varepsilon) = u(\hat{x}) - v(\hat{x}) - 2d^2(\hat{x}, x_\varepsilon) > 0. \end{cases}$$

On the other hand, notice first that $\hat{x} \in \Omega$ since we have $u(\hat{x}) - v(\hat{x}) > 0$. It follows that for α_n big enough we have $x_{\alpha_n}, y_{\alpha_n} \in \Omega$ since $\hat{x} \in \Omega$. Furthermore, we have

$$M - \varepsilon \leq u(\hat{x}) - v(\hat{x}) - 2d^2(\hat{x}, x_\varepsilon) \implies 2d^2(\hat{x}, x_\varepsilon) \leq \varepsilon \implies d(\hat{x}, x_\varepsilon) \leq \sqrt{\varepsilon},$$

and

$$\begin{aligned} H\Big(v(y_{\alpha_n}), y_{\alpha_n}, D_{y_{\alpha_n}}(-\frac{\alpha_n}{2} d^2(x_{\alpha_n}, .) - d^2(., x_\varepsilon))\Big) \;\geq\; 0 \\ 0 \geq H\Big(u(x_{\alpha_n}), x_{\alpha_n}, D_{x_{\alpha_n}}(\frac{\alpha_n}{2} d^2(., y_{\alpha_n}) + d^2(., x_\varepsilon))\Big). \end{aligned} \tag{3.13}$$

Hence, it follows from Hypotheses 3.1, 3.2, 3.3 and the inequality above

$$\begin{aligned} \gamma(u(x_{\alpha_n}) - v(y_{\alpha_n})) &\overset{3.2}{\leq} H\Big(u(x_{\alpha_n}), x_{\alpha_n}, D_{x_{\alpha_n}}(\frac{\alpha_n}{2} d^2(y_{.,\alpha_n}) + d^2(., x_\varepsilon))\Big) - \\ &\quad H\Big(v(y_{\alpha_n}), x_{\alpha_n}, D_{x_{\alpha_n}}(\frac{\alpha_n}{2} d^2(., y_{\alpha_n}) + d^2(., x_\varepsilon))\Big) \\ &\overset{(3.13)}{\leq} H\Big(v(y_{\alpha_n}), y_{\alpha_n}, D_{y_{\alpha_n}}(-\frac{\alpha_n}{2} d^2(., x_{\alpha_n}) - d^2(., x_\varepsilon))\Big) - \\ &\quad H\Big(v(y_{\alpha_n}), x_{\alpha_n}, D_{x_{\alpha_n}}(\frac{\alpha_n}{2} d^2(., y_{\alpha_n}) + d^2(., x_\varepsilon))\Big) \\ &\leq K_{db} d(x_{\alpha_n}, y_{\alpha_n})(1 + \frac{\alpha_n}{2} d(x_{\alpha_n}, y_{\alpha_n})) + 2K_L(d(x_{\alpha_n}, x_\varepsilon) + d(x_{\alpha_n}, x_\varepsilon)), \end{aligned}$$

where the last inequality is obtained thanks to Proposition 3.8.2 and Hypotheses 3.1 and 3.3. Furthermore, notice that we have for all α_n

$$\gamma(M-\varepsilon) \leq \gamma(u(x_\varepsilon)-v(x_\varepsilon)) \leq \gamma(u(x_{\alpha_n})-v(y_{\alpha_n})).$$

Whence, by letting $\alpha_n \to +\infty$, we get

$$\gamma(M-\varepsilon) \leq 4K_L\sqrt{\varepsilon}$$

Moreover, we have

$$\gamma(M-\sqrt{\varepsilon}) \leq \gamma(M-\varepsilon),$$

since $0<\varepsilon<1$ by assumption. We get

$$\gamma(M-\sqrt{\varepsilon}) \leq \gamma(M-\varepsilon) \leq 4K_L\sqrt{\varepsilon} \Longrightarrow \sqrt{\varepsilon} \geq \frac{\gamma}{\gamma+4K_L}M.$$

This is a contradiction with $\sqrt{\varepsilon} < \dfrac{\gamma}{\gamma+4K_L}M$, which ends the proof. ■

3.3.2 Perron's method

Let Ω be an arbitrary open set of X. Let $H : \mathbb{R}\times \mathrm{DC}_1(TX) \to \mathbb{R}$ be a Hamiltonian and $\ell : \partial\Omega \to \mathbb{R}$ be a bounded and continuous function. We consider the following Hamilton Jacobi equation with Dirichlet boundary condition

$$\begin{cases} H(u(x), x, D_x u) = 0, & \forall x \in \Omega, \\ u(x) = \ell(x), & \forall x \in \partial\Omega. \end{cases} \tag{3.14}$$

We consider the following hypotheses on the Hamiltonian H.

Hypothesis 3.4. The Hamiltonian H is such that:

- $(i)-$ For every $\phi : \Omega \to \mathbb{R}$ such that $\phi \in \mathcal{TEST}_-$, the function

  $$(r,x) \mapsto H(r,x,D_x\phi)$$

  is lower semicontinuous;

- $(ii)-$ For every $\phi : \Omega \to \mathbb{R}$ such that $\phi \in \mathcal{TEST}_+$, the function

  $$(r,x) \mapsto H(r,x,D_x\phi)$$

  is upper semicontinuous.

Hypothesis 3.5. The Hamiltonian H is such that for every $\phi_1, \phi_2 \in \mathrm{DC}_1(TX)$, and every $(x,r) \in \Omega\times\mathbb{R}$, we have

$$\forall \eta \in T_xX, \quad D_x\phi_2 \centerdot \eta \leq D_x\phi_1 \centerdot \eta \implies H(r,x,D_x\phi_1) \leq H(r,x,D_x\phi_2).$$

Remark 3.3.3. Hypothesis 3.4 on the Hamiltonian depends on the choice of the test functions adopted in the definition of viscosity. It is a weaker assumption than the continuity assumption usually required for $H(r, ., .)$ when $X = \mathbb{R}^N$ and

$$\mathcal{TEST}_- = \mathcal{TEST}_+ = \{\text{Twice continuously differentiable functions}\}.$$

Indeed, when $X = \mathbb{R}^N$ and the test functions are twice continuously differentiable, Hypothesis 3.4 is automatically verified as a consequence of the continuity of the Hamiltonian and the regularity of the test functions.
Hypothesis 3.5 is needed in the case of general proper CAT(0) spaces in order to generalize Perron's method in this setting. More precisely, Hypothesis 3.5 gives us the following useful result given below.

Lemma 3.14. *Let $x_0 \in \Omega$ and $\phi : \Omega \to \mathbb{R}$ be a $\mathrm{DC}_{\mathrm{lip}}(\Omega)$ function. Assume that the Hamiltonian H verifies Hypothesis 3.5.*

- *If the inequality*
$$H(\phi(x_0), x_0, D_{x_0}\phi) \leq 0,$$
is verified at $x_0 \in \Omega$. Then ϕ is a viscosity subsolution at x_0 is the sense of Definition 3.6.

- *Similarly, if the inequality*
$$H(\phi(x_0), x_0, D_{x_0}\phi) \geq 0,$$
is verified at $x_0 \in \Omega$. Then ϕ is a viscosity supersolution at x_0 is the sense of Definition 3.6.

Proof. We will only prove the first part of the lemma. The other part is done in the exact same way.

Let $\phi_{test} \in \mathcal{TEST}_-$ such that $\phi - \phi_{test}$ attains a local maximum at x_0. Then in a small neighborhood V of x_0, we have

$$\forall y \in V, \quad \phi(y) - \phi(x_0) \leq \phi_{test}(y) - \phi_{test}(x_0).$$

The last inequality implies that for any geodesic $\beta : [0, r'] \to X$ emanating from x_0 we have

$$\lim_{t \downarrow 0} \frac{\phi(\beta(t)) - \phi(x_0)}{t} \leq \lim_{t \downarrow 0} \frac{\phi_{test}(\beta(t)) - \phi_{test}(x_0)}{t} \quad \Longleftrightarrow D_{x_0}\phi \centerdot \beta'_0 \leq D_{x_0}\phi_{test} \centerdot \beta'_0.$$

Consequently, by Proposition 3.8.1, the Lipschitz continuity of the differentials gives us

$$\forall \eta \in T_{x_0}X, \quad D_{x_0}\phi \centerdot \eta \leq D_{x_0}\phi_{test} \centerdot \eta.$$

Finally, by Hypothesis 3.5, we get

$$H(\phi(x_0), x_0, D_{x_0}\phi_{test}) \leq H(\phi(x_0), x_0, D_{x_0}\phi) \leq 0.$$

This ends the proof. ■

We will derive existence of the solution of (3.14) from the comparison result proven in Theorem 3.12. First, we define the *half-relaxed limits* of a sequence of functions.

Definition 3.8. (Half-relaxed limits).
Let $(u_\varepsilon)_{\varepsilon>0}$ be a family of uniformly locally bounded functions such that $u_\varepsilon : \Omega \to \mathbb{R}$. We define the following *half-relaxed limits* of the family $(u_\varepsilon)_\varepsilon$ as:

$$\limsup{}^* u_\varepsilon\,(x) = \limsup_{\substack{\varepsilon\to 0\\ \Omega\ni z\to x}} u_\varepsilon(z);$$

$$\liminf{}_* u_\varepsilon\,(x) = \liminf_{\substack{\varepsilon\to 0\\ \Omega\ni z\to x}} u_\varepsilon(z).$$

It is clear from the above definition that $\limsup^* u_\varepsilon$ is an upper semicontinuous function and that $\liminf_* u_\varepsilon$ is a lower semicontinuous function. Before getting to Perron's method, we need two key lemmas. They are classical results when $X = \mathbb{R}^N$ (see for example [4]).

Lemma 3.15. *Let $(v_\varepsilon)_{\varepsilon>0}$ be a family of uniformly locally bounded upper semicontinuous functions on Ω and $\bar{v} := \limsup^* v_\varepsilon$. Let $y \in \Omega$ be a strict local maximum point of $\bar{v}$ on Ω. Then there exists a subsequence $(v_{\varepsilon_n})_{\varepsilon_n}$ and a sequence $(y_{\varepsilon_n})_{\varepsilon_n}$ such that for all ε_n, y_{ε_n} is a local maximum point of v_{ε_n} in Ω, the sequence $(y_{\varepsilon_n})_{\varepsilon_n}$ converges to y and $v_{\varepsilon_n}(y_{\varepsilon_n})$ converges to $\bar{v}(y)$ as $\varepsilon_n \to 0$.*

Proof. Since y is a strict local maximum point of $\bar{v}$ on Ω, there exists $r > 0$ such that $\overline{B}(y,r) \subset \Omega$ and

$$\forall\, z \in \overline{B}(y,r) \setminus \{x\}, \quad \bar{v}(z) < \bar{v}(y).$$

On the other hand, $\overline{B}(y,r)$ is compact and v_ε is upper semicontinuous bounded from above on $\overline{B}(y,r)$, therefore for any $\varepsilon > 0$ there exists a maximum point y_ε of v_ε on $\overline{B}(y,r)$, i.e.,

$$\forall\, z \in \overline{B}(y,r), \quad v_\varepsilon(z) \le v_\varepsilon(y_\varepsilon).$$

Hence, by taking the limsup for $z \to y$ and $\varepsilon \to 0$, we get:

$$\bar{v}(y) \le \limsup_{\varepsilon}\; v_\varepsilon(y_\varepsilon).$$

Next we consider the right hand side of the last inequality. By extracting a subsequence of $(y_\varepsilon)_\varepsilon$, denoted by $(y_{\varepsilon_n})_{\varepsilon_n}$, we have $\limsup_{\varepsilon\to 0}\; v_\varepsilon(y_\varepsilon) = \lim_{\varepsilon_n\to 0}\; v_{\varepsilon_n}(y_{\varepsilon_n})$. Furthermore, since $\overline{B}(y,r)$ is compact, we may assume that $(y_{\varepsilon_n})_{\varepsilon_n}$ converges to some $\bar{y}$. But using again the definition of $\limsup^*$ at $\bar{y}$ we get,

$$\bar{v}(y) \le \limsup_{\varepsilon\to 0}\; v_\varepsilon(y_\varepsilon) = \lim_{\varepsilon_n\to 0}\; v_{\varepsilon_n}(y_{\varepsilon_n}) \le \bar{v}(\bar{y}).$$

Since y is a strict maximum point of $\bar{v}$, we get $\bar{y} = y$ and $v_{\varepsilon_n}(y_{\varepsilon_n}) \to \bar{v}(y)$. ∎

Lemma 3.16. *Assume H satisfies Hypothesis 3.4. Let $(u_\varepsilon)_{\varepsilon>0}$ be a family of uniformly locally bounded upper semicontinuous functions on Ω, and set $u = \limsup^* u_\varepsilon$. If for all $\varepsilon > 0$, u_ε is a subsolution of (3.14), then u is a subsolution of (3.14).*

Proof. Let $\phi \in \mathcal{TEST}_-$ be a function such that $u - \phi$ attains a local maximum at $x \in \Omega$. Then $\psi(.) = \phi(.) + d^2(., x)$ is also a $\mathcal{TEST}_-$ function such that $u - \psi$ attains a strict local maximum at x. Since ψ is continuous we have $u - \psi = \limsup^*(u_\varepsilon - \psi)$. Applying the lemma above, there exists a subsequence $(x_{\varepsilon_n})_{\varepsilon_n}$ such that $x_{\varepsilon_n} \to x$, $u_{x_{\varepsilon_n}}(x_{\varepsilon_n}) \to u(x)$ and $u_{\varepsilon_n} - \psi$ attains a local maximum at x_{ε_n}. So we get

$$H(u_{\varepsilon_n}(x_{\varepsilon_n}), x_{\varepsilon_n}, D_{x_{\varepsilon_n}}(\phi(.) + d^2(., x))) = H(u_{\varepsilon_n}(x_{\varepsilon_n}), x_{\varepsilon_n}, D_{x_{\varepsilon_n}}\psi) \leq 0.$$

On the other hand, Proposition 3.8.2 implies

$$D_x d^2(., x) = 0.$$

Hence, by Hypothesis 3.4 we get

$$H(u(x), x, D_x\phi) \leq \liminf_{\varepsilon_n \to 0} H(u_{\varepsilon_n}(x_{\varepsilon_n}), x_{\varepsilon_n}, D_{x_{\varepsilon_n}}\psi) \leq 0,$$

which is the required result. ■

In the next theorem, we derive existence of the solution from the comparison principle asserted in Theorem 3.13. The proof is similar to the one in the classical theory of viscosity. The difficulty here is the lack of continuity of the Hamiltonian. However, with the use of Hypotheses 3.4 and 3.5, we are able to recover the same result as in the classical setting.

Theorem 3.17 (Perron's method)**.** *Let Ω be an open set of X and set $\partial\Omega = \overline{\Omega} \setminus \Omega$. Suppose that the Hamiltonian H satisfies Hypotheses 3.1, 3.2, 3.3, 3.4 and 3.5. Assume that there exist $\underline{u} : \overline{\Omega} \to \mathbb{R}$ a bounded upper semicontinuous subsolution of (3.14) and $\overline{u} : \overline{\Omega} \to \mathbb{R}$ a bounded lower semicontinuous supersolution of (3.14). If*

$$\liminf{}_* \underline{u}(x) \geq \ell(x) \geq \limsup{}^* \overline{u}(x), \quad \forall x \in \partial\Omega,$$

then there exists a unique continuous and bounded viscosity solution of (3.14).

Proof. We define the set

$$\mathcal{S} = \{h : \overline{\Omega} \to \mathbb{R} \ : \ \underline{u} \leq h \leq \overline{u} \text{ and } h \text{ is a subsolution of } (3.14)\}.$$

The set $\mathcal{S}$ is nonempty since $\underline{u} \in \mathcal{S}$. For $x \in \overline{\Omega}$, we set

$$u(x) = \sup\{h(x),\ h \in \mathcal{S}\}.$$

We will show that $v := \limsup^* \{u\}$ is the viscosity solution of (3.14). First we show that v is a subsolution. Notice that v is obviously upper semicontinuous. Let

$\phi : \Omega \to \mathbb{R}$ be a $\mathcal{TEST}_-$ function such that $v - \phi$ attains a local maximum at $x_0 \in \Omega$. Without loss of generality, we can suppose that

$$v(x_0) = \phi(x_0).$$

By definition of v, there exists a sequence of points $x_j \to x_0$ and a sequence of functions $u_j \in \mathcal{S}$ such that

$$v(x_0) = \lim_{j\to\infty} u_j(x_j).$$

In particular $\limsup^*\{u_j\}(x_0) \geq v(x_0)$. On the other hand, by construction we have $v \geq \limsup^*\{u_j\}$. Therefore, at x_0 we have $\limsup^*\{u_j\}(x_0) = v(x_0)$. For the other points on a small enough neighborhood of x_0, we have

$$\phi \geq v \geq \limsup^*\{u_j\}.$$

Therefore, by using Lemma 3.16 on a small enough bounded open neighborhood of x_0, we get that $\limsup^*\{u_j\}$ is a subsolution at x_0, since Hypothesis 3.4 holds. Therefore, by definition we have

$$H(v(x_0), x_0, D_{x_0}\phi) \leq 0.$$

This shows that v is a subsolution at x_0. Now we show that $v_* := \liminf_* v$ is a supersolution. We argue by contradiction.
Suppose that there exists a point $x_0 \in \Omega$ and a function $\psi \in \mathcal{TEST}_+$ such that $v_* - \psi$ attains a local minimum at x_0, but

$$H(v_*(x_0), x_0, D_{x_0}\psi) < 0.$$

Without loss of generality, we can suppose that $\psi(x_0) = v_*(x_0)$. Thus we have

$$H(\psi(x_0), x_0, D_{x_0}\psi) < 0.$$

So, by Hypothesis 3.5, ψ is a strict viscosity subsolution of (3.14) at x_0. Indeed, let $\psi_{test} \in \mathcal{TEST}_-$ such that $\psi - \psi_{test}$ attains a local maximum at x_0. Then for all $y \in X$ in a small enough neighborhood of x_0, we have

$$\psi(y) - \psi(x_0) \leq \psi_{test}(y) - \psi_{test}(x_0).$$

This implies that for any geodesic $\beta : [0, r'] \to X$ emanating from x_0 we have

$$\lim_{t\downarrow 0} \frac{\psi(\beta(t)) - \psi(x_0)}{t} \leq \lim_{t\downarrow 0} \frac{\psi_{test}(\beta(t)) - \psi_{test}(x_0)}{t} \quad \iff D_{x_0}\psi \centerdot \beta_0' \leq D_{x_0}\psi_{test} \centerdot \beta_0'.$$

Consequently, by Proposition 3.8.1, the Lipschitz continuity of the differentials gives us

$$\forall \eta \in T_{x_0}X, \quad D_{x_0}\psi \centerdot \eta \leq D_{x_0}\psi_{test} \centerdot \eta.$$

Hence by Hypothesis 3.5, we get

$$H(\psi(x_0), x_0, D_{x_0}\psi_{test}) \leq H(\psi(x_0), x_0, D_{x_0}\psi) < 0.$$

Furthermore from Hypothesis 3.4, it is also a strict subsolution of (3.14) in a small enough neighborhood of x_0 by upper semicontinuity of the Hamiltonian. Indeed, the function

$$g : x \mapsto H(\psi(x), x, D_x\psi)$$

is upper semicontinuous. So the set $\{x \in \Omega \ : \ g(x) < 0\}$ is open. In particular, there exists a small enough neighborhood of x_0 such that for all $x \in X$ that belong to this neighborhood, we have

$$H(\psi(x), x, D_x\psi) < 0.$$

Theorefore, for any $\psi_{test} \in \mathcal{TEST}_-$ such that $\psi - \psi_{test}$ attains a local maximum at $x \in X$ belonging to a small enough neighborhood of x_0, we get

$$H(\psi(x), x, D_x\psi_{test}) \leq H(\psi(x), x, D_x\psi) < 0.$$

Moreover, for $\delta > 0$ small enough, $\tilde{\psi} = \psi + \delta$ is a subsolution on a small enough neighborhood of x_0 denoted by $B(x_0, r) \subset \Omega$, with $r > 0$, since the function

$$s \mapsto H(s, x, D_x\psi),$$

is upper semicontinuous also by Hypothesis 3.4.

We have $\tilde{\psi}(x_0) > v_*(x_0)$. This implies that there are points at every neighborhood of x_0 such that $\tilde{\psi}(x) > v(x)$. Let

$$w(x) := \begin{cases} \max\{v, \tilde{\psi}\}(x), & \text{if } x \in \overline{B}(x_0, \frac{r}{2}), \\ v(x), & \text{otherwise.} \end{cases}$$

By Lemma 3.16, w a subsolution of (3.14). Consequently, we have

$$v, w \in \mathcal{S}.$$

However, $w > v$ at some points, a contradiction.
Therefore, v_* is a viscosity supersolution of (3.14). Finally, observe that

$$\liminf{}_* \underline{u}(x) \leq \liminf{}_* v(x) \leq \limsup{}^* v(x) \leq \limsup{}^* \overline{u}(x), \quad \forall x \in \partial\Omega,$$

implies that $v(x) = \ell(x)$ on $\partial\Omega$. In the end, by Theorem 3.12, v is continuous, bounded and is the unique viscosity solution to equation (3.14). ■

3.3.3 Examples

We give hereafter some examples showing the degree of generality of this new setting. The first two examples are considered in the case where the state space is $\mathbb{R}^N$, with the change being that we use different sets of test functions from the classical theory. The third example treats the case of an Eikonal type equation in a general proper CAT(0) space which is geodesically extendible. The last example presents the case of an Eikonal type equation in the presence of an obstacle in a general proper CAT(0) space which is geodesically extendible. This last example shows that this new setting allows to treat nonconvex Hamiltonians as well.

Example 3.18. For $(X, d) = (\mathbb{R}^N, d_{\mathbb{R}^N})$, with $d_{\mathbb{R}^N}$ the Euclidean distance, consider the Hamiltonian

$$H(u(x), x, D_x u) := \gamma u(x) + \sup_{\alpha \in \mathcal{A}} \{ -D_x u \boldsymbol{.} f(x, \alpha) + b(x, \alpha) \}, \quad x \in \mathbb{R}^N,$$

where $\gamma > 0$, $\mathcal{A}$ is a compact metric space and $f : \mathbb{R}^N \times \mathcal{A} \to \mathbb{R}^N$ is a Lipschitz bounded function. The function $b : \mathbb{R}^N \times \mathcal{A} \to \mathbb{R}$ is a Lipschitz bounded function. We consider

$$\mathcal{TEST}_- = \{ \text{Locally Lipschitz and locally semiconvex functions of } \mathbb{R}^N \}.$$

$$\mathcal{TEST}_+ = \{ \text{Locally Lipshitz and locally semiconcave functions of } \mathbb{R}^N \}.$$

$\mathcal{TEST}_-$ and $\mathcal{TEST}_+$ satisfy all the requirements of Definition 3.7. Furthermore, it is straightforward to check that the Hamiltonian H satisfies Hypotheses 3.1, 3.2, 3.3 and 3.5.

It remains to prove that the Hamiltonian H satisfies Hypothesis 3.4. We start by Hypothesis 3.4-(i).

Let $\phi \in \mathcal{TEST}_-$. ϕ is a locally Lipschitz and semiconvex function. Hence it is Clarke regular [9, Definition 10.12]. Therefore, the function

$$(x, v) \mapsto D_x\phi \boldsymbol{.} v$$

is upper semicontinuous [9, Proposition 10.2]. Consequently, the function

$$(x, \alpha) \mapsto -D_x\phi \boldsymbol{.} f(x, \alpha)$$

is lower semicontinuous. Finally the function

$$x \mapsto \sup_{\alpha \in \mathcal{A}} \{-D_x\phi \boldsymbol{.}\ f(x, \alpha) + b(x, \alpha)\}$$

is lower semicontinuous as the pointwise supremum of a family of lower semicontinuous functions, which implies Hypothesis 3.4-(i).

Now we turn our attention to Hypothesis 3.4-(ii). Let $\phi \in \mathcal{TEST}_+$. ϕ is a locally

Lipschitz and locally semiconcave function. So $-\phi$ is a locally Lipschitz and locally semiconvex function. Hence it is Clarke regular, i.e. the function

$$(x, v) \mapsto -D_x\phi \centerdot v$$

is upper semicontinuous. Consequently, the function

$$(x, \alpha) \mapsto -D_x\phi \centerdot f(x, \alpha) + b(x, \alpha)$$

is also upper semicontinuous. Now, let $x \in \mathbb{R}^N$ and $(x_n)_n \subset \mathbb{R}^N$ be a sequence converging to x. Let $(\alpha_n)_n \subset \mathcal{A}$ be a sequence such that

$$\forall n \in \mathbb{N}, \quad \sup_{\alpha \in \mathcal{A}} \{-D_{x_n}\phi \centerdot f(x_n, \alpha) + b(x_n, \alpha)\} = -D_{x_n}\phi \centerdot f(x_n, \alpha_n) + b(x_n, \alpha_n).$$

Since $\mathcal{A}$ is a compact metric space, then we can assume, without loss of generality, that the sequence $(\alpha_n)_n$ converges to $\bar{\alpha} \in \mathcal{A}$. Finally, we have

$$\sup_{\alpha \in \mathcal{A}} \{-D_x\phi \centerdot f(x, \alpha) + b(x, \alpha)\} \geq -D_x\phi \centerdot f(x, \bar{\alpha}) + b(x, \bar{\alpha}) \geq$$
$$\limsup_{\substack{x_n \to x \\ \alpha_n \to \bar{\alpha}}} -D_{x_n}\phi \centerdot f(x_n, \alpha_n) + b(x_n, \alpha_n) \geq \limsup_{x_n \to x} \sup_{\alpha \in \mathcal{A}} \{-D_{x_n}\phi \centerdot f(x_n, \alpha) + b(x_n, \alpha)\}.$$

This implies Hypothesis 3.4-(ii).

On the other hand, $\underline{u}(x) = -C$, $\overline{u}(x) = C$ are a bounded upper semicontinuous subsolution and a bounded lower semicontinuous supersolution respectively for $C > 0$ big enough. Consequently, Theorem 3.17 applies and there exists a unique continuous and bounded viscosity solution to equation (3.14), with the Hamiltonian defined above.

Example 3.19. For $(X, d) = (\mathbb{R}^N, d_{\mathbb{R}^N})$, with $d_{\mathbb{R}^N}$ the Euclidean distance, consider the Hamiltonian

$$H(u(x), x, D_x u) := \gamma u(x) + \inf_{\alpha \in \mathcal{A}} \sup_{\beta \in \mathcal{B}} \{ -D_x u \centerdot f(x, \alpha, \beta) + b(x, \alpha, \beta) \}, \quad x \in \mathbb{R}^N,$$

where $\gamma > 0$, $\mathcal{A}, \mathcal{B}$ are compact metric spaces and $f : \mathbb{R}^N \times \mathcal{A} \times \mathcal{B} \to \mathbb{R}^N$ is a Lipschitz bounded function. The function $b : \mathbb{R}^N \times \mathcal{A} \times \mathcal{B} \to \mathbb{R}$ is a Lipschitz bounded function. We consider

$$\mathcal{TEST}_- = \{ \text{Locally Lipschitz and locally semiconvex functions of } \mathbb{R}^N \}.$$

$$\mathcal{TEST}_+ = \{ \text{Locally Lipshitz and locally semiconcave functions of } \mathbb{R}^N \}.$$

$\mathcal{TEST}_-$ and $\mathcal{TEST}_+$ satisfy all the requirements of Definition 3.7. Furthermore, it is straightforward to check that the Hamiltonian H satisfies Hypotheses 3.1, 3.2, 3.3 and 3.5.

It remains to prove that the Hamiltonian H satisfies Hypotheses 3.4. We start by

Hypothesis 3.4-(i).

Let $\phi \in \mathcal{TEST}_-$. ϕ be a locally Lipschitz and semiconvex function. Hence it is Clarke regular [9, Proposition 10.2 and Definition 10.12]. As we have seen from Example 3.18, we have

$$(x,\alpha) \mapsto \sup_{\beta\in\mathcal{B}} \{-D_x\phi \centerdot f(x,\alpha,\beta) + b(x,\alpha,\beta)\},$$

is lower semicontinuous. Now, let $(x_n)_n \subset \mathbb{R}^N$ be a sequence converging to $x \in \mathbb{R}^N$. Let $(\alpha_n)_n \subset \mathcal{A}$ be a sequence such that for all $n \in \mathbb{N}$ we have

$$\inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \{-D_{x_n}\phi \centerdot f(x_n,\alpha,\beta) + b(x_n,\alpha,\beta)\} = \sup_{\beta\in\mathcal{B}} -D_{x_n}\phi \centerdot f(x_n,\alpha_n,\beta) + b(x_n,\alpha_n,\beta).$$

Since $\mathcal{A}$ is a compact metric space, then we can assume, without loss of generality, that the sequence $(\alpha_n)_n$ converges to some $\bar{\alpha} \in \mathcal{A}$. Finally, we have

$$\begin{aligned}
\inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \{-D_x\phi \centerdot f(x,\alpha) + b(x,\alpha)\} &\leq \sup_{\beta\in\mathcal{B}} \{-D_x\phi \centerdot f(x,\bar{\alpha},\beta) + b(x,\bar{\alpha},\beta)\} \\
&\leq \liminf_{\substack{x_n\to x\\ \alpha_n\to\bar{\alpha}}} \sup_{\beta\in\mathcal{B}} \{-D_{x_n}\phi \centerdot f(x_n,\alpha_n,\beta) + b(x_n,\alpha_n,\beta)\} \\
&\leq \liminf_{x_n\to x} \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \{-D_{x_n}\phi \centerdot f(x_n,\alpha,\beta) + b(x_n,\alpha,\beta)\}.
\end{aligned}$$

This implies Hypothesis 3.4-(i).
Next, we turn our attention to Hypothesis 3.4-(ii). Let $\phi \in \mathcal{TEST}_+$. ϕ is a locally Lipschitz and semiconcave function. Therefore, from Example 3.18, we have

$$x \mapsto \sup_{\beta\in\mathcal{B}} \{-D_x\phi \centerdot f(x,\alpha,\beta) + b(x,\alpha,\beta)\},$$

is upper semicontinuous. Hence, the function

$$x \mapsto \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} -D_x\phi \centerdot f(x,\alpha,\beta),$$

is also upper semicontinuous since it is the pointwise infimum of a family of upper semicontinuous functions, which implies Hypothesis 3.4-(ii).

On the other hand, $\underline{u}(x) = -C$ and $\overline{u}(x) = C$ are a bounded upper semicontinuous subsolution and a bounded lower semicontinuous supersolution respectively for $C > 0$ big enough. Consequently, Theorem 3.17 applies and there exists a unique continuous viscosity solution to equation (3.14).

Remark 3.3.4. In the above two examples, the only change we made from the classical theory of viscosity, was to change the sets of test functions in $\mathbb{R}^N$. The interest here is limited, as we could also have chosen the test functions to be locally twice continuously differentiable functions for both the supersolution and the subsolution in the current setting as in the classical theory of viscosity. Indeed, a result due to Alexandrov [88] shows that locally twice continuously differentiable functions are locally DC functions. Actually they are locally both semiconvex and semiconcave functions. In other words, twice continuously differentiable functions constitute a subset of the intersection between the sets of locally semiconvex and semiconcave functions in $\mathbb{R}^N$. So the present setting and the classical theory of viscosity coincide in $\mathbb{R}^N$.

Example 3.20 (Eikonal type equation in proper geodesically extendible CAT(0) spaces)**.** Let (X, d) be a proper, *geodesically extendible* CAT(0) space. All spaces given in Examples, 3.4, 3.5 and 3.6 verify this condition. Consider the Hamiltonian

$$H(u(x), x, D_x u) := \gamma u(x) + \sup_{\substack{v \in T_x X \\ |v|_x = 1}} \{-D_x u \centerdot v\} - b(x), \quad x \in X,$$

where $\gamma > 0$ and $b : X \to \mathbb{R}$ is a Lipschitz function of constant $Lip(b)$ and bounded. We consider

$$\mathcal{TEST}_- = \{\text{Locally Lipschitz and locally semiconvex functions of } X\}.$$

$$\mathcal{TEST}_+ = \{\text{Locally Lipshitz and locally semiconcave functions of } X\}.$$

$\mathcal{TEST}_-$ satisfies all the requirements of Definition 3.7.
First, we prove that the above Hamiltonian verifies Hypothesis 3.1. Let $\alpha > 0$, $r \in \mathbb{R}$ and $x, y \in X$. We have by Proposition 3.8.2

$$\begin{aligned} &H(r, x, -D_x(\alpha d^2(., y))) - H(r, y, D_y(\alpha d^2(x, .))) = \\ &\qquad d(x,y)\Big(\sup_{\substack{v \in T_x X \\ |v|_x = 1}} \{-2\alpha\langle v, \uparrow_x^y\rangle_x\} - \sup_{\substack{v \in T_y X \\ |v|_y = 1}} \{2\alpha\langle v, \uparrow_y^x\rangle_y\}\Big) + b(y) - b(x). \end{aligned}$$

By inequality (3.8c), we have

$$\sup_{\substack{v \in T_x X \\ |v|_x = 1}} \{-2\alpha\langle v, \uparrow_x^y\rangle_x\} \le \sup_{\substack{v \in T_x X \\ |v|_x = 1}} \{2\alpha\, |v|_x \mid \uparrow_x^y|_x\} = 2\alpha,$$

and

$$\sup_{\substack{v \in T_y X \\ |v|_y = 1}} \{2\alpha\langle v, \uparrow_y^x\rangle_y\} = 2\alpha, \quad \text{reached at } v = \uparrow_y^x .$$

Hence we have

$$H(r, x, -D_x(\alpha d^2(., y))) - H(r, y, D_y(\alpha d^2(x, .))) \le b(y) - b(x) \le Lip(b) d(x, y),$$

which implies the result.
Hypotheses 3.2, 3.3 and 3.5 are straightforward. It remains to prove that the Hamiltonian verifies Hypotheses 3.4. We start by Hypothesis 3.4-(i). The proof is inspired

from [56, Lemma 1.3.4].
Let $\phi \in \mathcal{TEST}_-$. So ϕ is a locally Lipschitz and locally semiconvex function which implies that $\psi := -\phi$ is a locally Lipschitz and locally semiconcave function.

Let $x \in X$. Suppose that ψ is 2λ-concave for some $\lambda \in \mathbb{R}$ around x.

Let $\varepsilon > 0$ and let $y \in X$ near x such that

$$y \neq x, \qquad |\lambda| d(x,y) < \varepsilon, \quad \text{and} \quad \frac{\psi(y) - \psi(x)}{d(x,y)} \geq \sup_{\substack{v \in T_x X \\ |v|_x = 1}} \{D_x \psi \centerdot v\} - \varepsilon.$$

Let $(x_n)_n$ and $(y_n)_n$ be two sequences converging to x and y respectively. Let $[0, d(x_n, y_n)] \ni t \mapsto G_t^{x_n, y_n}$ be the unit speed geodesic connecting x_n and y_n. By definition, the 2λ-concavity of ψ implies that the real-to-real function

$$[0, d(x_n, y_n)] \ni t \mapsto \psi(G_t^{x_n, y_n}) - \lambda t^2$$

is concave. Therefore, the incremental ratio

$$(0, d(x_n, y_n)] \ni t \mapsto \frac{\psi(G_t^{x_n, y_n}) - \lambda t^2 - \psi(x_n)}{t}.$$

is non increasing. Hence, the 2λ-concavity of ψ gives

$$D_{x_n}\psi \centerdot \uparrow_{x_n}^{y_n} \geq \frac{\psi(y_n) - \psi(x_n) - \lambda d^2(x_n, y_n)}{d(x_n, y_n)} \geq \frac{\psi(y) - \psi(x)}{d(x,y)} - \varepsilon,$$

where the last inequality is obtained when n is large enough. Hence we get

$$\sup_{\substack{v \in T_{x_n} X \\ |v|_{x_n} = 1}} \{D_{x_n} \psi \centerdot v\} \geq D_{x_n}\psi \centerdot \uparrow_{x_n}^{y_n} \geq \sup_{\substack{v \in T_x X \\ |v|_x = 1}} \{D_x \psi \centerdot v\} - 2\varepsilon.$$

By taking the infimum limit in the left hand side of the last inequality, we get

$$\liminf_{\substack{x_n \to x \in X \\ r_n \to r \in \mathbb{R}}} \gamma r_n + \sup_{\substack{v \in T_{x_n} X \\ |v|_{x_n} = 1}} \{-D_{x_n} \phi \centerdot v\} - b(x_n) \geq \gamma r + \sup_{\substack{v \in T_x X \\ |v|_x = 1}} \{-D_x \phi \centerdot v\} - b(x) - 2\varepsilon.$$

Since ε is arbitrary, we get that Hypothesis 3.4-(i) is verified.
Now we prove Hypothesis 3.4-(ii). Let $\phi \in \mathcal{TEST}_+$. So ϕ is a locally Lipschitz and locally semiconcave function which implies that $\psi := -\phi$ is a locally Lipschitz and locally semiconvex function.

Let $x \in X$. Suppose that ψ is 2λ-convex around x for some $\lambda \in \mathbb{R}$. Let $0 < \varepsilon < M$ be two stricly positive constants, $(x_n)_n \subset X$ be a sequence converging to x and $(y_n)_n \subset X$ be a sequence such that

$$d(x, y_n) \leq M, \quad \text{and} \quad \forall n \in \mathbb{N},\ D_{x_n}\psi \centerdot \uparrow_{x_n}^{y_n} \geq \sup_{\substack{v \in T_{x_n} X \\ |v|_{x_n} = 1}} \{D_{x_n} \psi \centerdot v\} - \varepsilon.$$

Since (X,d) is geodesically extendible, we can also always choose y_n such that

$$d(x,y_n) \geq \varepsilon, \quad \forall n \in \mathbb{N}.$$

Since $(y_n)_n \subset \overline{B}(x,M)$, then we can suppose, without loss of generality, that it converges to some $y \in X$. Moreover $y \neq x$ since we have $d(x,y_n) \geq \varepsilon$ for all $n \in \mathbb{N}$. Let $t \mapsto G_t^{x_n,y_n}$ be the unique unit speed geodesic between x_n and y_n.
By the extendibility property of X, we can extend all geodesics G^{x_n,y_n} to be defined in the same interval $[0,K]$, with $K > M$ large enough. Then by Arzela–Ascoli theorem [52, Theorem 2.5.14] there exists a converging subsequence (not relabeled here) of the sequence of curves $(G^{x_n,y_n})_n$. Moreover, by [52, Proposition 2.5.17], the limit curve is $G^{x,y}$; the unit speed geodesic starting from x and passing through y, and defined in $[0,K]$.
Furthermore, the 2λ-convexity of ψ gives

$$\begin{aligned}
\sup_{\substack{v\in T_{x_n}X\\ |v|_{x_n}=1}} \{D_{x_n}\psi \centerdot v\} - \varepsilon \leq D_{x_n}\psi \centerdot \uparrow_{x_n}^{y_n} &= D_{x_n}\psi \centerdot \uparrow_{x_n}^{G_t^{x_n,y_n}} \\
&\leq \frac{\psi(G_t^{x_n,y_n}) - \psi(x_n) - \lambda d^2(x_n, G_t^{x_n,y_n})}{d(x_n, G_t^{x_n,y_n})} \\
&\leq \frac{\psi(G_t^{x,y}) - \psi(x)}{d(x, G_t^{x,y})} + \varepsilon,
\end{aligned}$$

where the last inequality holds when taking t small enough and n big enough. Hence we get

$$\begin{aligned}
\limsup_{x_n\to x} \sup_{\substack{v\in T_{x_n}X\\ |v|_{x_n}=1}} \{D_{x_n}\psi \centerdot v\} \leq \lim_{t\downarrow 0} \frac{\psi(G_t^{x,y}) - \psi(x)}{d(x, G_t^{x,y})} + 2\varepsilon &= \lim_{t\downarrow 0} \frac{\psi(G_t^{x,y}) - \psi(x)}{t} + 2\varepsilon \\
&\leq \sup_{\substack{v\in T_xX\\ |v|_x=1}} \{D_x\psi \centerdot v\} + 2\varepsilon.
\end{aligned}$$

Finally we get

$$\limsup_{\substack{x_n\to x\in X\\ r_n\to r\in\mathbb{R}}} \gamma r_n + \sup_{\substack{v\in T_{x_n}X\\ |v|_{x_n}=1}} \{D_{x_n}\psi \centerdot v\} - b(x_n) \leq \gamma r + \sup_{\substack{v\in T_xX\\ |v|_x=1}} \{D_x\psi \centerdot v\} - b(x) + 2\varepsilon,$$

wihch is equivalent to

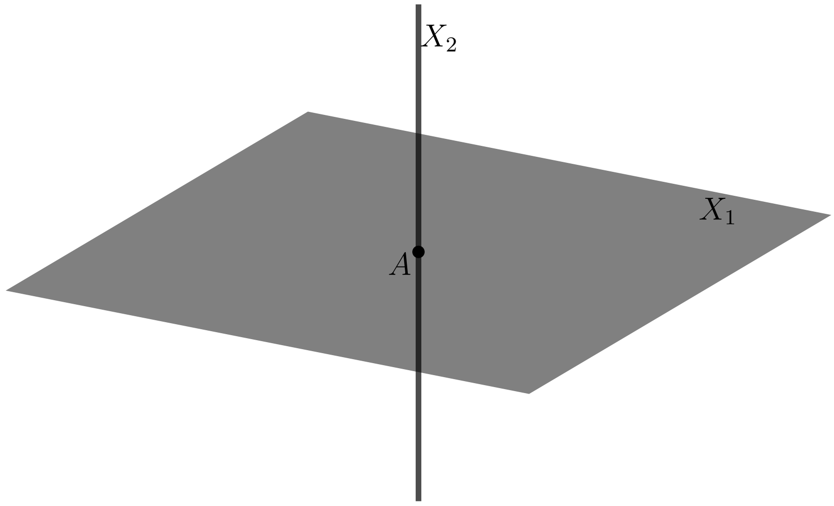
$$\limsup_{\substack{x_n\to x\in X\\ r_n\to r\in\mathbb{R}}} \gamma r_n + \sup_{\substack{v\in T_{x_n}X\\ |v|_{x_n}=1}} \{-D_{x_n}\phi \centerdot v\} - b(x_n) \leq \gamma r + \sup_{\substack{v\in T_xX\\ |v|_x=1}} \{-D_x\phi \centerdot v\} - b(x) + 2\varepsilon.$$

Since ε is arbitrary, we get the result.
In summary, the Hamiltonian H verifies all Hypotheses 3.1, 3.2, 3.3, 3.4 and 3.5. Furthermore, the functions $\underline{u}(x) = -C$, $\overline{u}(x) = C$ are bounded upper semicontinuous subsolution and bounded lower semicontinuous supersolution respectively for

$C > 0$ big enough. Hence by Theorem 3.17, there exists a unique bounded and continuous viscosity solution to the Hamilton Jacobi equation

$$H(u(x), x, D_x u) = 0, \quad \forall x \in X.$$

Remark 3.3.5. Let us take the proper, geodesically extendible CAT(0) space given in Example 3.4:



where X_1 and X_2 are the following two proper CAT(0) spaces:

$$\begin{cases} X_1 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_3 = 0\}, \\ X_2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_1 = x_2 = 0\}, \end{cases}$$

and $A := \{0\}$. The glued space

$$X := X_1 \bigsqcup_A X_2,$$

along A is a proper, geodesically extendible CAT(0) space when endowed with the following distance:

$$\forall x, y \in X, \ d(x, y) := \begin{cases} |x - y|, \text{ if } \exists\, i \in \{1, 2\} \ : \ x, y \in X_i, \\ |x| + |y|, \text{ otherwise}, \end{cases}$$

where $|.|$ is the Euclidean norm on $\mathbb{R}^3$. The tangent cone at a point $x \in X$ is:

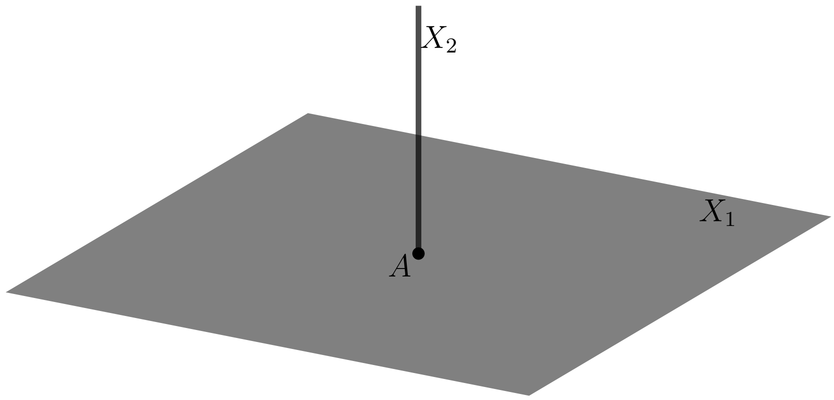
$$T_x X = \begin{cases} X_1 & \text{if } x \in X_1 \setminus A, \\ X_2 & \text{if } x \in X_2 \setminus A, \\ X & \text{if } x \in A. \end{cases}$$

Let $\gamma > 0$ and $b : X \to \mathbb{R}$ be a Lipschitz bounded function. The equation studied in Example 3.20 has the following expression:

$$\begin{cases} \gamma u(x) + \sup\limits_{\substack{v\in X_1\\|v|=1}} \{-D_x u \centerdot v\} - b(x) = 0, & \text{if } x \in X_1 \setminus A, \\ \gamma u(x) + \sup\limits_{\substack{v\in X_2\\|v|=1}} \{-D_x u \centerdot v\} - b(x) = 0, & \text{if } x \in X_2 \setminus A, \\ \gamma u(x) + \sup\limits_{\substack{v\in X\\|v|=1}} \{-D_x u \centerdot v\} - b(x) = 0, & \text{if } x \in A. \end{cases}$$

From Example 3.20, the above equation admits a unique continuous and bounded viscosity solution.

Remark 3.3.6. Let us take the proper, geodesically extendible CAT(0) space given in Example 3.5:



where X_1 and X_2 are the following two proper CAT(0) spaces:

$$\begin{cases} X_1 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_3 = 0\}, \\ X_2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_1 = x_2 = 0, \ x_3 \geq 0\}, \end{cases}$$

and $A := \{0\}$. The glued space

$$X := X_1 \bigsqcup_A X_2,$$

along A is a proper, geodesically extendible CAT(0) space when endowed with the following distance:

$$\forall x, y \in X, \ d(x,y) := \begin{cases} |x-y|, \text{ if } \exists\, i \in \{1,2\} \ : \ x, y \in X_i, \\ |x|+|y|, \text{ otherwise.} \end{cases}$$

The tangent cone at a point $x \in X$ is:

$$T_x X = \begin{cases} X_1 & \text{if } x \in X_1 \setminus A, \\ \mathbb{R}e_3 & \text{if } x \in X_2 \setminus A, \\ X & \text{if } x \in A. \end{cases}$$

Let $\gamma > 0$ and $b : X \to \mathbb{R}$ be a Lipschitz bounded function. The equation studied in Example 3.20 has the following expression:

$$\begin{cases} \gamma u(x) + \sup\limits_{\substack{v \in X_1 \\ |v|=1}} \{-D_x u \centerdot v\} - b(x) = 0, & \text{if } x \in X_1 \setminus A, \\ \gamma u(x) + \sup\limits_{\substack{v \in \mathbb{R}e_3 \\ |v|=1}} \{-D_x u \centerdot v\} - b(x) = 0, & \text{if } x \in X_2 \setminus A, \\ \gamma u(x) + \sup\limits_{\substack{v \in X \\ |v|=1}} \{-D_x u \centerdot v\} - b(x) = 0, & \text{if } x \in A. \end{cases}$$

From Example 3.20, the above equation admits a unique continuous and bounded viscosity solution.

Example 3.21 (Nonconvex Hamiltonian)**.** Let (X, d) be a proper, geodesically extendible CAT(0) space. Consider the Hamiltonian

$$H(u(x), x, D_x u) := \min\Big\{\gamma u(x) + \sup_{\substack{v \in T_x X \\ |v|_x = 1}} \{-D_x u \centerdot v\} - b_1(x), \gamma u(x) - b_2(x)\Big\}, \quad x \in X,$$

where $\gamma > 0$ and $b_1 : X \to \mathbb{R}$ and $b_2 : X \to \mathbb{R}$ are Lipschitz and bounded functions. From Example 3.20, the Hamiltonian

$$H_1(u(x), x, D_x u) := \gamma u(x) + \sup_{\substack{v \in T_x X \\ |v|_x = 1}} \{-D_x u \centerdot v\} - b_1(x),$$

verifies all Hypotheses 3.1, 3.2, 3.3, 3.4 and 3.5. Furthermore, it is easy to check that the Hamiltonian

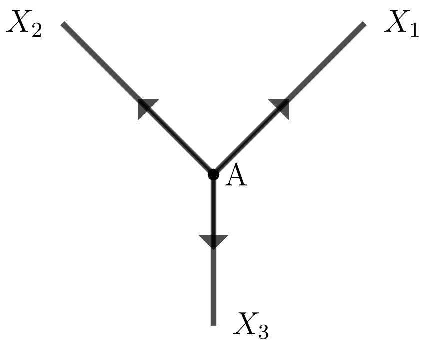
$$H_2(u(x), x, D_x u) := \gamma u(x) - b_2(x),$$

also verifies all Hypotheses 3.1, 3.2, 3.3, 3.4 and 3.5. Consequently, the Hamitlonian H verifies all the mentioned Hypotheses as well since it is in the form of a minimum of two Hamiltonians that verify the same Hypotheses. Furthermore, the functions $\underline{u}(x) = -C$, $\overline{u}(x) = C$ are bounded upper semicontinuous subsolution and bounded lower semicontinuous supersolution respectively for $C > 0$ big enough. Hence, by Theorem 3.17, there exists a unique bounded and continuous viscosity solution to the equation

$$H(u(x), x, D_x u) = 0, \quad \forall x \in X.$$

Remark 3.3.7. Let e_1, e_2 and e_3 be three unit vectors of $\mathbb{R}^2$. Let us take the proper, geodesically extendible CAT(0) obtained by gluing together three half-lines, denoted by X_1, X_2 and X_3 along the origin point $A = \{0\}$:

$$\begin{cases} X_1 := [0, +\infty)e_1, \\ X_2 := [0, +\infty)e_2, \\ X_3 := [0, +\infty)e_3. \end{cases}$$



The glued space

$$X := \bigsqcup_A X_i,$$

along A is a proper, geodesically extendible CAT(0) space when endowed with the following distance:

$$\forall x, y \in X,\ d(x,y) := \begin{cases} |x-y|, \text{ if } \exists\, i \in \{1,2,3\} \ : \ x, y \in X_i, \\ |x|+|y|, \text{ otherwise.} \end{cases}$$

The tangent cone at a point $x \in X$ is:

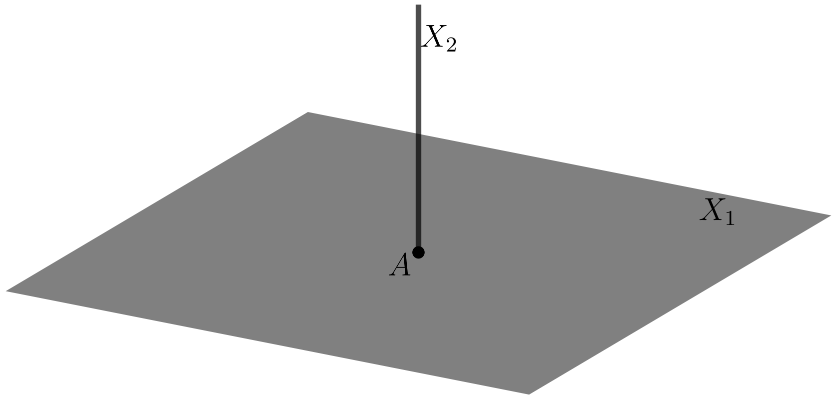
$$T_x X = \begin{cases} \mathbb{R}e_i & \text{if } x \in X_i \setminus A, \text{ and } \ = 1, 2, 3, \\ X & \text{if } x \in A. \end{cases}$$

Let $\gamma > 0$ and $b_1 : X \to \mathbb{R}$ and $b_2 : X \to \mathbb{R}$ be Lipschitz and bounded functions. The equation studied in Example 3.21 has the following expression:

$$\begin{cases} \min\Big\{\gamma u(x) + \sup\limits_{\substack{v \in \mathbb{R}e_i \\ |v|=1}} \{-D_x u \centerdot v\} - b_1(x), \gamma u(x) - b_2(x)\Big\} = 0, & \text{if } x \in X_i \setminus A, \\ \min\Big\{\gamma u(x) + \sup\limits_{\substack{v \in X \\ |v|=1}} \{-D_x u \centerdot v\} - b_1(x), \gamma u(x) - b_2(x)\Big\} = 0, & \text{if } x \in A. \end{cases}$$

From Example 3.21, the above equation admits a unique continuous and bounded viscosity solution.

Remark 3.3.8. Let us take the proper, geodesically extendible CAT(0) space given in Example 3.5:



where X_1 and X_2 are the following two proper CAT(0) spaces:

$$\begin{cases} X_1 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_3 = 0\}, \\ X_2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_1 = x_2 = 0, \ x_3 \geq 0\}, \end{cases}$$

and $A := \{0\}$. The glued space

$$X := X_1 \bigsqcup_A X_2,$$

along A is a proper, geodesically extendible CAT(0) space when endowed with the following distance:

$$\forall x, y \in X, \ d(x,y) := \begin{cases} |x-y|, \text{ if } \exists\, i \in \{1,2\} \ : \ x, y \in X_i, \\ |x|+|y|, \text{ otherwise.} \end{cases}$$

The tangent cone at a point $x \in X$ is:

$$T_x X = \begin{cases} X_1 & \text{if } x \in X_1 \setminus A, \\ \mathbb{R}e_3 & \text{if } x \in X_2 \setminus A, \\ X & \text{if } x \in A. \end{cases}$$

Let $\gamma > 0$ and $b_1 : X \to \mathbb{R}$ and $b_2 : X \to \mathbb{R}$ be Lipschitz and bounded functions. The equation studied in Example 3.21 has the following expression:

$$\begin{cases} \min\Big\{\gamma u(x) + \sup\limits_{\substack{v \in X_1 \\ |v|=1}} \{-D_x u \centerdot v\} - b_1(x), \gamma u(x) - b_2(x)\Big\} = 0, & \text{if } x \in X_1 \setminus A, \\ \min\Big\{\gamma u(x) + \sup\limits_{\substack{v \in \mathbb{R}e_3 \\ |v|=1}} \{-D_x u \centerdot v\} - b_1(x), \gamma u(x) - b_2(x)\Big\} = 0, & \text{if } x \in X_2 \setminus A, \\ \min\Big\{\gamma u(x) + \sup\limits_{\substack{v \in X \\ |v|=1}} \{-D_x u \centerdot v\} - b_1(x), \gamma u(x) - b_2(x)\Big\} = 0, & \text{if } x \in A. \end{cases}$$

From Example 3.21, the above equation admits a unique continuous and bounded viscosity solution.

3.4 Time dependent Hamilton Jacobi equations in proper CAT(0) spaces

In this section we discuss time dependent Hamilton Jacobi equations in a proper CAT(0) space (X, d). First, notice that by means of Lemma 3.2, the product space $[0, +\infty) \times X$ is also a CAT(0) space. One could consider the time variable as being one of the state variables and use the setting developped for the stationary case. Although it is possible to do it, we choose to treat the time dependent case separately, as it has its own specificities. Let $H : \mathrm{DC}_1(TX) \to \mathbb{R}$ be a Hamiltonian and $\ell : X \to \mathbb{R}$ be a bounded and continuous function. We consider the following Hamilton Jacobi equation:

$$\begin{cases} \partial_t u + H(x, D_x u) = 0, & \forall\,(t,x) \in (0, +\infty) \times X, \\ u(0,x) = \ell(x), & x \in X, \end{cases} \tag{3.15}$$

where $u : [0, +\infty) \times X$ is a Lipschitz and DC function. The term $\partial_t u$ is the usual right derivative with respect to time, i.e.

$$\partial_t u = \lim_{r \downarrow 0} \frac{u(t + r, x) - u(t, x)}{r}.$$

Let $\mathrm{C}^2((0, +\infty))$ be the space of twice continuously differentiable functions of $(0, +\infty)$. We will take test functions that are $\mathrm{C}^2((0, +\infty))$ with respect to the time variable and in the class of DC functions with respect to the space variable.

Definition 3.9. Let $\mathcal{TEST}_-$ and $\mathcal{TEST}_+$ be two subsets of $\mathrm{DC}_{\mathrm{lip}}((0, \infty) \times X)$ such that:

$$\mathcal{TEST}_- := \{(t, x) \mapsto \phi_1(t) + \phi_2(x) \ : \ \phi_1 \text{ is } \mathrm{C}^2((0, +\infty)) \text{ and } \phi_2 \text{ is locally Lipschitz and locally semiconvex}\},$$

and

$$\mathcal{TEST}_+ := \{(t, x) \mapsto \phi_1(t) + \phi_2(x) \ : \ \phi_1 \text{ is } \mathrm{C}^2((0, +\infty)) \text{ and } \phi_2 \text{ is locally Lipschitz and locally semiconcave}\}.$$

3.4.1 Comparison principle

Next, we prove a comparison result in the time dependent case. Since the Hamiltonian in (3.15) does not depend on $u(x)$, there is no need to assume Hypothesis 3.2 on the Hamiltonian. By assuming only Hypotheses 3.1 and 3.3, we can prove the comparison principle for the time dependent case, as the following theorem shows.

Theorem 3.22. *Assume H satisfies Hypotheses 3.1 and 3.3. Let $u : [0,+\infty) \times X \to \mathbb{R}$ be a bounded from above upper semicontinuous subsolution of (3.15), and $v : [0,+\infty) \times X \to \mathbb{R}$ a bounded from below lower semicontinuous supersolution of equation (3.15). Then it holds:*

$$\sup_{[0,+\infty)\times X}(u-v)_+ \leq \sup_{\{0\}\times X}(u-v)_+,$$

where $(r)_+ = \max(r,0)$.

Proof. Without loss of generality, we can suppose that $\sup\limits_{\{0\}\times X}(u-v)_+ = 0$. Let $M := \sup\limits_{[0,+\infty)\times X}(u(t,x)-v(t,x))$. It suffices to prove that $M \leq 0$.
Assume by contradiction that $M > 0$. Let $\lambda > 0$ sufficiently small so that

$$\sup_{[0,+\infty)\times X}(u(t,x)-v(t,x)-\lambda t) > 0.$$

Let $(t_0,x_0) \in [0,+\infty) \times X)$ be such that

$$u(t_0,x_0)-v(t_0,x_0)-\lambda t_0 > 0.$$

Let $\varepsilon \in (0,M)$. For every $\alpha > 0$, define $\psi_\alpha : [0,+\infty)^2 \times X^2 \to \mathbb{R}$ as

$$\psi_\alpha(t,s,x,y) = u(t,x)-v(s,y)-\frac{\lambda}{2}(t+s)-\frac{\varepsilon}{2}\Big(d^2(x,x_0)+d^2(y,x_0)\Big)-\frac{\alpha}{2}|t-s|^2-\frac{\alpha}{2}d^2(x,y).$$

It is clear that ψ_α is upper semicontinuous and bounded from above. We denote by $M_\alpha := \sup\psi_\alpha$, where the supremum is taken over $[0,+\infty)^2 \times X^2$. Furthermore, we have

$$0 < \psi_\alpha(t_0,t_0,x_0,x_0),$$

and for all $x, y \notin B\left(x_0, \sqrt{2\frac{|\sup(u)|+|\sup(-v)|}{\varepsilon}}\right)$ or $t,s \geq 2\frac{|\sup(u)|+|\sup(-v)|}{\lambda}$ we have

$$\psi_\alpha(t,s,x,y) \leq 0.$$

Hence, the supremum of ψ_α is reached in a compact set independent of α. Let $(t_\alpha,s_\alpha,x_\alpha,y_\alpha)$ be such that $M_\alpha = \psi_\alpha(t_\alpha,s_\alpha,x_\alpha,y_\alpha)$. We have

$$\lim_{\alpha\to+\infty}(M_\alpha - \psi_\alpha(t_\alpha,s_\alpha,x_\alpha,y_\alpha)) = 0$$

and

$$-\infty < u(t_0,x_0)-v(t_0,x_0)-\lambda t_0 \leq M_\alpha \leq \sup_{[0,+\infty)\times X}(u) + \sup_{[0,+\infty)\times X}(-v) < +\infty.$$

Since $(t_\alpha,s_\alpha,x_\alpha,y_\alpha)$ is in a compact set, we take a subsequence $(t_{\alpha_n},s_{\alpha_n},x_{\alpha_n},y_{\alpha_n})$ such that $(t_{\alpha_n},s_{\alpha_n},x_{\alpha_n},y_{\alpha_n})$ converges as $\alpha_n \to +\infty$ and

$$\lim_{\alpha_n\to+\infty}(M_{\alpha_n}-\psi_{\alpha_n}(t_{\alpha_n},s_{\alpha_n},x_{\alpha_n},y_{\alpha_n}))=0,\quad\text{and}\quad -\infty<\lim_{\alpha_n\to+\infty}M_{\alpha_n}<+\infty.$$

Therefore, we can apply Lemma 3.11 via the correspondences

$$Z=\mathcal{O}=[0,+\infty)^2\times X^2,\quad \Phi(z)=u(t,x)-v(s,y)-\frac{\lambda}{2}(t+s)-\frac{\varepsilon}{2}\Big(d^2(x,x_0)+d^2(y,x_0)\Big),$$

$$\Psi(z)=\frac{1}{2}|t-s|^2+\frac{1}{2}d^2(x,y),$$

and we get

$$\begin{cases}(i)\quad \displaystyle\lim_{\alpha_n\to+\infty}\frac{\alpha_n}{2}d^2(x_{\alpha_n},y_{\alpha_n})+\frac{\alpha_n}{2}|t_{\alpha_n}-s_{\alpha_n}|^2=0,\\ \qquad\text{and } x_{\alpha_n},y_{\alpha_n}\to\hat{x}\in X,\ t_{\alpha_n},s_{\alpha_n}\to\hat{t}\in[0,+\infty),\\ (ii)\quad \displaystyle\lim_{\alpha_n\to+\infty}M_{\alpha_n}\geq u(t_0,x_0)-v(t_0,x_0)-\lambda t_0>0.\end{cases}$$

Moreover, we have

$$0<u(t_0,x_0)-v(t_0,x_0)-\lambda t_0\leq u(\hat{t},\hat{x})-v(\hat{t},\hat{x})-\lambda\hat{t}-\varepsilon d^2(\hat{x},x_0).\tag{3.16}$$

This implies

$$\varepsilon d^2(\hat{x},x_0)\leq M\quad\Longrightarrow\quad \varepsilon d(\hat{x},x_0)\leq\sqrt{M\,\varepsilon}.$$

On the other hand, notice that $\hat{t}\neq 0$ since $u(\hat{t},\hat{x})-v(\hat{t},\hat{x})>0$. It follows that for α_n big enough we have $t_{\alpha_n}\neq 0$. Furthermore, we have

$$\begin{aligned}-\frac{\lambda}{2}+\alpha_n(t_{\alpha_n}-s_{\alpha_n})+H\Big(y_{\alpha_n},D_{y_{\alpha_n}}(-\frac{\alpha_n}{2}d^2(x_{\alpha_n},.)-\frac{\varepsilon}{2}d^2(.,x_\varepsilon))\Big)\ \geq\ 0\ \geq\\ \frac{\lambda}{2}+\alpha_n(t_{\alpha_n}-s_{\alpha_n})+H\Big(x_{\alpha_n},D_{x_{\alpha_n}}(\frac{\alpha_n}{2}d^2(.,y_{\alpha_n})+\frac{\varepsilon}{2}d^2(.,x_\varepsilon))\Big).\end{aligned}\tag{3.17}$$

Hence, it follows from Hypotheses 3.1 and 3.3 and the inequality above

$$\begin{aligned}\lambda &\overset{(3.17)}{\leq} H\Big(y_{\alpha_n},-D_{y_{\alpha_n}}(\frac{\alpha_n}{2}d^2(x_{\alpha_n},.)+\frac{\varepsilon}{2}d^2(.,x_\varepsilon))\Big)-H\Big(x_{\alpha_n},D_{x_{\alpha_n}}(\frac{\alpha_n}{2}d^2(.,y_{\alpha_n})+\frac{\varepsilon}{2}d^2(.,x_\varepsilon))\Big)\\ &\overset{3.1,\,3.3}{\leq} K_{db}d(x_{\alpha_n},y_{\alpha_n})(1+\frac{\alpha_n}{2}d(x_{\alpha_n},y_{\alpha_n}))+K_L\varepsilon(d(y_{\alpha_n},x_0)+d(x_{\alpha_n},x_0)).\end{aligned}$$

By letting $\alpha_n\to+\infty$, we get

$$\lambda\leq 2K_L\sqrt{M\,\varepsilon}.$$

The last inequality is valid for any $0<\varepsilon<M$, a contradiction. ■

3.4.2 Perron's method

Next, we prove Perron's method on the product space $[0,+\infty) \times X$. First, notice that if we assume that the Hamiltonian H verifies Hypothesis 3.4, then the same hypothesis is verified by the full Hamiltonian

$$\partial_t u + H(x, D_x u),$$

in the product space $[0,+\infty) \times X$ since it is a CAT(0) space and the test functions chosen in Definition 3.9 are locally continuously differentiable with respect to the time variable.

Theorem 3.23. *Let $\Omega = (0,+\infty) \times X$ and set $\partial\Omega = \{0\} \times X$. Assume H satisfies Hypotheses 3.1 and 3.3. Suppose that there exist $\underline{u} : \overline{\Omega} \to \mathbb{R}$ a locally bounded and bounded from above upper semicontinuous subsolution of (3.15) and $\overline{u} : \overline{\Omega} \to \mathbb{R}$ a locally bounded and bounded from below lower semicontinuous supersolution of (3.15) such that*

$$\liminf{}_* \underline{u}(t,x) \geq \ell(x) \geq \limsup{}^* \overline{u}(t,x), \quad \forall (t,x) \in \partial\Omega.$$

Then there exists a unique continuous viscosity solution of (3.15).

Proof. We define the set

$$\mathcal{S} = \{h : \overline{\Omega} \to \mathbb{R} \ : \ \underline{u} \leq h \leq \overline{u} \text{ and } h \text{ is a subsolution of (3.15)}\}.$$

The set $\mathcal{S}$ is nonempty since $\underline{u} \in \mathcal{S}$. For $x \in \overline{\Omega}$, we set

$$u(t,x) = \sup\{h(t,x), \ h \in \mathcal{S}\}.$$

We will show that $v := \limsup^* \{u\}$ is the viscosity solution of (3.15). First we show that v is a subsolution. Notice that v is obviously upper semicontinuous. Let $\phi : \Omega \to \mathbb{R}$ be a $\mathcal{TEST}_-$ function such that $v - \phi$ attains a local maximum at $(t_0, x_0) \in \Omega$. Without loss of generality, we can suppose that

$$v(t_0, x_0) = \phi(t_0, x_0).$$

By definition of v, there exists a sequence of points $(t_j, x_j) \to x_0$ and a sequence of functions $u_j \in \mathcal{S}$ such that

$$v(t_0, x_0) = \lim_{j\to\infty} u_j(t_j, x_j).$$

In particular $\limsup^*\{u_j\}(t_0, x_0) \geq v(t_0, x_0)$. On the other hand, by construction we have $v \geq \limsup^*\{u_j\}$. Therefore, at (t_0, x_0) we have $\limsup^*\{u_j\}(t_0, x_0) = v(t_0, x_0)$. For the other points on a small enough neighborhood of (t_0, x_0), we have

$$\phi \geq v \geq \limsup{}^*\{u_j\}.$$

Therefore, by using Lemma 3.16 on a small enough bounded open neighborhood of (t_0, x_0), we get that $\limsup^*\{u_j\}$ is a subsolution at (t_0, x_0), since Hypothesis 3.4 holds. Therefore, by definition we have

$$\partial_{t_0}\phi + H(x_0, D_{x_0}\phi) \leq 0.$$

This shows that v is a subsolution at (t_0, x_0). Now we show that $v_* := \liminf_* v$ is a supersolution. We argue by contradiction.
Suppose that there exists a point $(t_0, x_0) \in \Omega$ and a function $\psi \in \mathcal{TEST}_+$ such that $v_* - \psi$ attains a local minimum at (t_0, x_0), but

$$\partial_{t_0}\psi + H(x_0, D_{x_0}\psi) < 0.$$

By Hypothesis 3.4 and the continuity of $t \mapsto \partial_t\psi$, we get that the function

$$(t, x) \mapsto \partial_t\psi + H(x, D_x\psi)$$

is upper semicontinuous. Hence, on a small open neighborhood of (t_0, x_0), denoted by $(t_0 - r, t_0 + r) \times B(x_0, r)$ we have

$$\forall (t, x) \in (t_0 - r, t_0 + r) \times B(x_0, r), \quad \partial_t\psi + H(D_x\psi) < 0.$$

Furthermore, by Hypothesis 3.5, ψ is a strict viscosity subsolution of (3.15) on $(t_0 - r, t_0 + r) \times B(x_0, r)$.
Indeed, let $\psi_- \in \mathcal{TEST}_-$ such that $\psi - \psi_-$ attains a locall maximum at $(t, x) \in (t_0 - r, t_0 + r) \times B(x_0, r)$. Notice that since $\psi \in \mathcal{TEST}_+$ and $\psi_- \in \mathcal{TEST}_-$, then they are of the from

$$\psi(t, x) = f(t) + g(x), \quad \psi_-(t, x) = f_-(t) + g_-(x),$$

where f and f_- are twice continuously differentiable functions, and g and g_- are locally Lipschitz and locally semiconcave and semiconvex respectively. we have for all (s, y) in a small neighborhood of (t, x):

$$f(t) + g(x) - (f_-(t) + g_-(x)) \geq f(s) + g(y) - (f_-(s) + g_-(y)).$$

It follows from the above inequality that we have for all s in a small neighborhood of t:

$$f(t) - f_-(t) \geq f(s) - f_-(s),$$

and for all y in a small neighborhood of x we get

$$g(x) - g_-(x) \geq g(y) - g_-(y).$$

This implies that

$$\partial_t f = \partial_t f_- \quad \text{and} \quad \partial_x g \leq \partial_x g_-.$$

Therefore, by Hypothesis 3.5 we get

$$\partial_t \psi_- + H(x, D_x \psi_-) \leq \partial_t \psi + H(x, D_x \psi) < 0.$$

On the other hand, without loss of generality, we can suppose that $\psi(t_0, x_0) = v_*(t_0, x_0)$.

Moreover for $\delta > 0$, $\tilde{\psi} = \psi + \delta$ is also a viscosity subsolution on a small enough neighborhood $(t_0 - r, t_0 + r) \times B(x_0, r)$. We have $\tilde{\psi}(t_0, x_0) > v_*(t_0, x_0)$. This implies that there are points at every neighborhood of (t_0, x_0) such that $\tilde{\psi}(t, x) > v(t, x)$. Let

$$w(t,x) := \begin{cases} \max\{v, \tilde{\psi}\}(t,x), & \text{if } (t,x) \in (t_0 - \frac{r}{2}, t_0 + \frac{r}{2}) \times \overline{B}(x_0, \frac{r}{2}), \\ v(t,x), & \text{otherwise.} \end{cases}$$

By Lemma 3.16, w a subsolution of (3.15). Consequently, we have

$$v, w \in \mathcal{S}.$$

However, $w > v$ at some points, a contradiction. Therefore, v_* is a viscosity supersolution of (3.15).

Finally, observe that

$$\liminf{}_* \underline{u}(t,x) \leq \liminf{}_* v(x) \leq \limsup{}^* v(t,x) \leq \limsup{}^* \overline{u}(t,x), \quad \forall (t,x) \in \partial\Omega,$$

implies that $v(t,x) = \ell(x)$ on $\partial\Omega$. In the end, by Theorem 3.22, v is continuous, bounded and is the unique viscosity solution to equation (3.15). ■

Remark 3.4.1. A sufficient condition to guarantee the existence of a locally bounded and bounded from above upper semicontinuous subsolution $\underline{u}$ and a locally bounded and bounded from below lower semicontinuous supersolution $\overline{u}$ of (3.15), that verify all the conditions of Theorem 3.23, is to suppose the following condition on the Hamiltonian H.

Hypothesis 3.6. The Hamiltonian H is such that

$$X \ni x \mapsto H(x, 0_{\mathrm{DC}_1(T_x X)}) \text{ is bounded.}$$

Indeed if Hypothesis 3.6 holds, let

$$C := \sup_{x \in X} |H(x, 0_{\mathrm{DC}_1(T_x X)})|, \quad \text{and} \quad M := \sup_{x \in X} |\ell(x)|.$$

Then for $(t,x) \in [0, +\infty) \times X$, the following two functions

$$\underline{u}(t,x) := M - Ct, \quad \text{and} \quad \overline{u}(t,x) := -M + Ct$$

are respectively a locally bounded and bounded from above upper semicontinuous subsolution and a locally bounded and bounded from below lower semicontinuous supersolution of (3.15) on $(0, +\infty) \times X$ and

$$M = \liminf{}_* \underline{u}(t,x) \geq \ell(x) \geq \limsup{}^* \overline{u}(t,x) = -M, \quad \forall (t,x) \in \{0\} \times X.$$

3.4.3 Examples

Example 3.24 (Time dependent Eikonal equation in proper geodesically extendible CAT(0) spaces)**.** Let (X, d) be a proper, geodesically extendible CAT(0) space. All spaces given in Examples, 3.4, 3.5 and 3.6 verify this condition. Consider the Hamiltonian

$$H(x, D_x u) := \sup_{\substack{v \in T_x X \\ |v|_x = 1}} \{-D_x u \centerdot v\}, \quad x \in X.$$

We consider the Hamilton Jacobi equation (3.15) with the Hamiltonian H defined above:

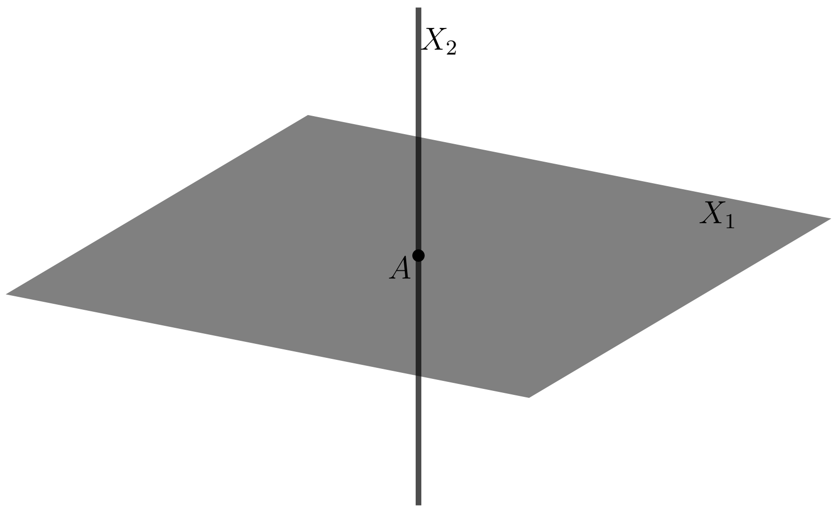
$$\begin{cases} \partial_t u + \sup\limits_{\substack{v \in T_x X \\ |v|_x = 1}} \{-D_x u \centerdot v\} = 0, & (t, x) \in (0, +\infty) \times X, \\ u(0, x) = \ell(x), & \text{if } x \in X. \end{cases}$$

This equation is the *time dependent Eikonal equation*. We consider the test functions given in Definition 3.9 for the viscosity notion. From Example 3.20, we know that the Hamiltonian H verifies Hypotheses 3.1 and 3.3. Hence we can apply Theorem 3.22 for any bounded from above upper semicontinuous subsolution and any bounded from below lower semicontinuous supersolution. Furthermore, from Example 3.20, we know that the Hamiltonian verifies Hypotheses 3.4 and 3.5. Finally, we have

$$\forall x \in X, \quad H(x, 0_{\mathrm{DC}_1(T_x X)}) = 0.$$

Consequently, Hypothesis 3.6 is also verified by the Hamiltonian. In conclusion, Theorem 3.23 applies and there exists a unique bounded and continuous viscosity solution to equation (3.15) with the Hamiltonian H defined above.

Remark 3.4.2. Let us take the proper, geodesically extendible CAT(0) space given in Example 3.4:



where X_1 and X_2 are the following two proper CAT(0) spaces:

$$\begin{cases} X_1 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_3 = 0\}, \\ X_2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_1 = x_2 = 0\}, \end{cases}$$

and $A := \{0\}$. The glued space

$$X := X_1 \bigsqcup_A X_2,$$

along A is a proper, geodesically extendible CAT(0) space when endowed with the following distance:

$$\forall x, y \in X,\ d(x,y) := \begin{cases} |x-y|, \text{ if } \exists\, i \in \{1,2\} \ : \ x, y \in X_i, \\ |x|+|y|, \text{ otherwise}, \end{cases}$$

where $|.|$ is the Euclidean norm on $\mathbb{R}^3$. The tangent cone at a point $x \in X$ is:

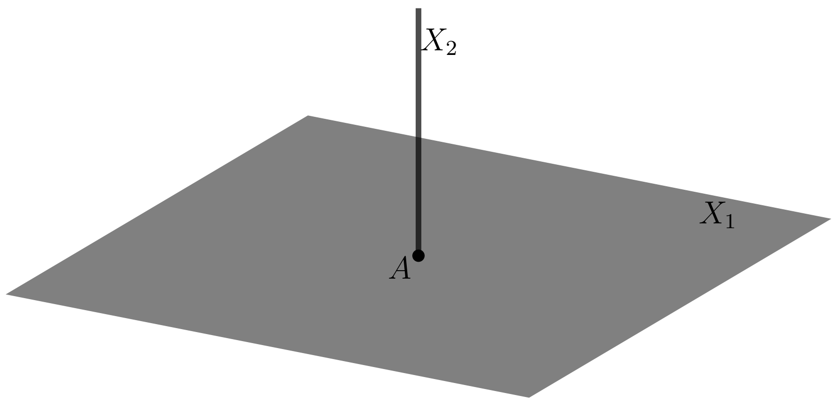
$$T_x X = \begin{cases} X_1 & \text{if } x \in X_1 \setminus A, \\ X_2 & \text{if } x \in X_2 \setminus A, \\ X & \text{if } x \in A. \end{cases}$$

Let $\ell : X \to \mathbb{R}$ be a continuous bounded function. The equation studied in Example 3.24 has the following expression:

$$\begin{cases} \partial_t u + \sup\limits_{\substack{v \in X_1 \\ |v|=1}} \{-D_x u \centerdot v\} = 0, & \text{if } (t,x) \in (0,+\infty) \times X_1 \setminus A, \\ \partial_t u + \sup\limits_{\substack{v \in X_2 \\ |v|=1}} \{-D_x u \centerdot v\} = 0, & \text{if } (t,x) \in (0,+\infty) \times X_2 \setminus A, \\ \partial_t u + \sup\limits_{\substack{v \in X \\ |v|=1}} \{-D_x u \centerdot v\} = 0, & \text{if } (t,x) \in (0,+\infty) \times A, \\ u(0,x) = \ell(x). \end{cases}$$

From Example 3.24, the above equation admits a unique continuous and bounded viscosity solution.

Remark 3.4.3. Let us take the proper, geodesically extendible CAT(0) space given in Example 3.5:



where X_1 and X_2 are the following two proper CAT(0) spaces:

$$\begin{cases} X_1 := \{(x_1,x_2,x_3) \in \mathbb{R}^3 \ : \ x_3 = 0\}, \\ X_2 := \{(x_1,x_2,x_3) \in \mathbb{R}^3 \ : \ x_1 = x_2 = 0, \ x_3 \geq 0\}, \end{cases}$$

and $A := \{0\}$. The glued space

$$X := X_1 \bigsqcup_A X_2,$$

along A is a proper, geodesically extendible CAT(0) space when endowed with the following distance:

$$\forall x,y \in X, \ d(x,y) := \begin{cases} |x-y|, \text{ if } \exists\, i \in \{1,2\} \ : \ x,y \in X_i, \\ |x|+|y|, \text{ otherwise.} \end{cases}$$

The tangent cone at a point $x \in X$ is:

$$T_xX = \begin{cases} X_1 & \text{if } x \in X_1 \setminus A, \\ \mathbb{R}e_3 & \text{if } x \in X_2 \setminus A, \\ X & \text{if } x \in A. \end{cases}$$

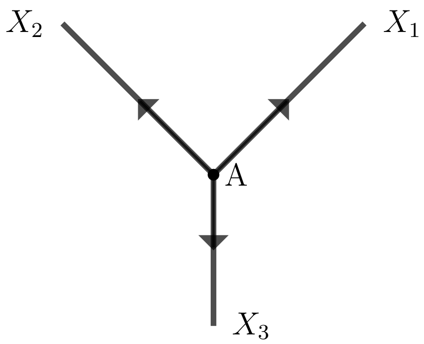
Let $\ell : X \to \mathbb{R}$ be a continuous bounded function. The equation studied in Example 3.24 has the following expression:

$$\begin{cases} \partial_t u + \sup\limits_{\substack{v \in X_1 \\ |v|=1}} \{-D_x u \,.\, v\} = 0, & \text{if } (t,x) \in (0,+\infty) \times X_1 \setminus A, \\ \partial_t u + \sup\limits_{\substack{v \in \mathbb{R}e_3 \\ |v|=1}} \{-D_x u \,.\, v\} = 0, & \text{if } (t,x) \in (0,+\infty) \times X_2 \setminus A, \\ \partial_t u + \sup\limits_{\substack{v \in X \\ |v|=1}} \{-D_x u \,.\, v\} = 0, & \text{if } (t,x) \in (0,+\infty) \times A, \\ u(0,x) = \ell(x), \quad \text{if } x \in X. \end{cases}$$

From Example 3.24, the above equation admits a unique continuous and bounded viscosity solution.

Remark 3.4.4. Let e_1, e_2 and e_3 be three unit vectors of $\mathbb{R}^2$. Let us take the proper, geodesically extendible CAT(0) obtained by gluing together three half-lines, denoted by X_1, X_2 and X_3 along the origin point $A = \{0\}$:

$$\begin{cases} X_1 := [0,+\infty)e_1, \\ X_2 := [0,+\infty)e_2, \\ X_3 := [0,+\infty)e_3. \end{cases}$$



The glued space

$$X := \bigsqcup_A X_i,$$

along A is a proper, geodesically extendible CAT(0) space when endowed with the following distance:

$$\forall x, y \in X,\ d(x,y) := \begin{cases} |x-y|, \text{ if } \exists\, i \in \{1,2,3\} \ : \ x, y \in X_i, \\ |x|+|y|, \text{ otherwise.} \end{cases}$$

The tangent cone at a point $x \in X$ is:

$$T_x X = \begin{cases} \mathbb{R}e_i & \text{if } x \in X_i \setminus A, \text{ and } \ = 1,2,3, \\ X & \text{if } x \in A. \end{cases}$$

Let $\ell : X \to \mathbb{R}$ be a continuous bounded function. The equation studied in Example 3.24 has the following expression:

$$\begin{cases} \partial_t u + \sup\limits_{\substack{v \in \mathbb{R}e_i \\ |v|=1}} \{-D_x u \,.\, v\} = 0, & \text{if } (t,x) \in (0,+\infty) \times X_i \setminus A, \\ \partial_t u + \sup\limits_{\substack{v \in X \\ |v|=1}} \{-D_x u \,.\, v\} = 0, & \text{if } (t,x) \in (0,+\infty) \times A, \\ u(0,x) = \ell(x), \quad \text{if } x \in X. \end{cases}$$

From Example 3.24, the above equation admits a unique continuous and bounded viscosity solution.

Chapter 4

Deterministic optimal control problem in Riemannian manifolds under probability knowledge of the initial condition

Contents

4.1 Introduction

The study of optimal control problems and viscosity theory in Wasserstein spaces has gained more and more momentum in the last decade, due to its potential real-world applications in modeling multi-agent systems. The potential real-world applications include crowd dynamics modeling [89], opinion formation process modeling [90], herd analysis [91], autonomous multi-vehicle navigation [92] and modeling uncertainties on the initial state of a deterministic controlled system [36, 37]. These problems look into the evolution of a large number of agents, considered to be indistinguishable from one another, subject to local and nonlocal interactions that depend on the density of the distribution of all agents.

A suitable way to model these problems is through a macroscopic approach, where we consider the collection of all agents that belong to a state space denoted X (typically the Euclidean space or a Riemannian manifold), to be a density that evolves through time. If we assume further that the total number of all agents remains constant at all time, then we can normalize the density and assume that the total mass of the system is equal to 1 at all time. Therefore, the evolution of the system, seen as a probability density in the space of Borel probability measures over X and denoted $\mathcal{P}(X)$, is described by a curve $t \mapsto \mu_t \in \mathcal{P}(X)$, where μ_t is the probability density of the system at time $t \geq 0$. The conservation of the mass of the system at all time $t \geq 0$ is described by the *continuity equation*

$$\partial_t \mu_t + div(w_t(.)\mu_t) = 0, \tag{4.1}$$

where $w_t(.)$ is a time dependent vector field and the equation is understood in the sense of distributions.

In this chapter, we propose to study a deterministic controlled system on a compact Riemannian manifold M with imperfect information on the initial state of the system, i.e. the controller only knows the initial condition through a Borel probability measure $\mu_0 \in \mathcal{P}(M)$, along which the initial state is distributed. This could be regarded as a multi-agent system where the nonlocal interations between the agents are not considered. More precisely, let $T > 0$ and consider the following controlled system

$$\begin{cases} \dot{Y}(t) = f(Y(t), u(t)), \quad t \in [t_0, T], \\ Y(t_0) = x_0, \quad u(t) \in U, \end{cases} \tag{4.2}$$

where $f : M \times U \to TM$ is the dynamics, assumed to be Lipschitz with respect to the first variable and continuous with respect to the second variable, $x_0 \in M$ and $t_0 \in [0, T]$. The set U is the set of admissible control values which is assumed to be a compact subset of some metric space. The control function $u(.) \in U$ is a Borel measurable function $u : [t_0, T] \to U$. To emphasize the dependence of trajectories of the controlled system (4.2) on the control function $u(.)$, the initial time t_0 and the initial position x_0, we denote them by

$$t \mapsto Y_t^{t_0, x_0, u}.$$

The main feature of this problem is that the initial state x_0 is not perfectly known, but rather distributed along the probability measure μ_0. The evolution curves of the initial uncertainty, denoted $t \mapsto \mu_t^{t_0, \mu_0, u}$, are obtained by the pushforward of μ_0

with the flow at time t of the controlled equation (4.2). Therefore, the curves are of the form

$$\begin{cases} \mu_t^{t_0,\mu_0,u} = Y_t^{t_0,.,u} \sharp \mu_0, \quad t \in [t_0, T], \quad \text{and} \quad x \mapsto Y_t^{t_0,x,u} \text{ is the flow of (4.2)}, \\ \mu_{t_0}^{t_0,\mu_0,u} = \mu_0. \end{cases}$$

Furthermore, notice that since $f(., u(t))$ is Lipschitz continuous and bounded, then it is a known fact that the evolution trajectory $t \mapsto \mu_t^{t_0,\mu_0,u}$, of the uncertainty μ_0, is the unique solution to the continuity equation

$$\begin{cases} \partial_t \mu_t^{t_0,\mu_0,u} + div(f(., u(t))\mu_t^{t_0,\mu_0,u}) = 0, \; t \in [t_0, T], \\ \mu_{t_0}^{t_0,\mu_0,u} = \mu_0, \end{cases}$$

in the distributional sense [60, 93]. The controller aims at minimizing the following final cost:

$$L(\mu) = \int \ell(y) d\mu(y),$$

where $\ell : M \to \mathbb{R}$ is a Lipschitz function. An immediate consequence of this assumption is that the function $L : \mathcal{P}(M) \to \mathbb{R}$ inherits the Lipschitz property from ℓ as well. The quantity $L(\mu_T^{t_0,\mu_0,u})$ represents the expectation of the deterministic final cost with respect to the measure $\mu_T^{t_0,\mu_0,u}$. To this optimal control problem, we associate the following value function

$$\vartheta(t_0, \mu_0) = \inf_{u(.) \in U} L(\mu_T^{t_0,\mu_0,u}).$$

In the literature, a similar problem was studied by various authors in the space of Borel probability measures over the Euclidean space $\mathbb{R}^N$. In particular, it was addressed in [37] a differential game problem with uncertainties on the initial condition and in [36] a Mayer optimal control problem with uncertainties on the initial condition. We stress on the difference between the set of trajectories $t \mapsto \mu_t$ considered here and the set of trajectories considered in [36]. Indeed, in the latter case, the set of the evolution curves $t \mapsto \mu_t$ of the initial uncertainty appears to be larger. The trajectories $t \mapsto \mu_t$ are solutions to the following continuity equation

$$\begin{cases} \partial_t \mu_t + div(w_t \mu_t) = 0, \; t \in [t_0, T], \\ \mu_{t_0} = \mu_0, \end{cases}$$

where $w_t(.)$ is a vector field such that

$$w_t(.) \in \{ f(., u) \; : \; u \in U \},$$

which gives rise to trajectories $t \mapsto \mu_t$ that may not be obtained by the pushforward of the initial uncertainty μ_0 by the flow at time t of the controlled equation (4.2).

In this manuscript, we want to study the evolution of the lack of information on the initial condition in (4.2), modeled by a Borel probability measure μ_0. Hence, we only consider trajectories $t \mapsto \mu_t$ that are obtained by the pushforward of the initial uncertainty μ_0 with the flow at time t of the controlled equation (4.2). Finally, we mention that more general controlled systems on the space of Borel probability measures over the Euclidean space are studied in [38, 94, 95, 96], where the nonlocal interactions between the agents are taken into account.

The first main goal of this chapter is to study the properties and the regularity of the value function. In particular, we show that the value function is Lipschitz continuous with respect to both its variables and that it verifies the dynamic programming principle. The second main goal of this chapter is to characterize the value function as a unique viscosity solution to a Hamilton Jacobi Bellman equation (HJB in short) defined on the Wasserstein space $\mathcal{P}(M)$. Ideally, the HJB equation should have the following form:

$$\begin{cases} \partial_t v + H(\mu, D_\mu v) = 0, \quad (t, \mu) \in [0, T) \times \mathcal{P}(M), \\ v(T, \mu) = L(\mu). \end{cases} \tag{4.3}$$

The Hamiltonian H and the derivative with respect to the measure variable $D_\mu v$ are to be defined in a suitable way. Furthermore, we want to define a viscosity notion for time dependent Hamilton Jacobi equations in $\mathcal{P}(M)$ so that the we can prove a comparison principle that holds for any bounded upper semicontinuous subsolution and any bounded lower semicontinuous supersolution. To give a precise definition of all these notions in $\mathcal{P}(M)$, we will rely on the pseudo-Riemannian structure of $\mathcal{P}(M)$, presented in Section 4.3.

Concerning viscosity theory in Wasserstein spaces, several notions have been introduced in the literature to study first order Hamilton Jacobi equations in the Wasserstein space over the Euclidean space $\mathcal{P}_2(\mathbb{R}^N)$. One approach relies on introducing a generalization of sub/super differentials to the space $\mathcal{P}_2(\mathbb{R}^N)$ [36, 38, 37]. Another approach is to define the notion of viscosity, in an extrinsic way, by "lifting" the Hamilton Jacobi equation to a Hilbert space, then one uses the viscosity theory developed in Banach spaces, developed in [28, 29]. All these approaches only give a comparison princple that holds only for any uniformly continuous subsolution and any uniformly continuous supersolution. In this chapter, we use a different approach. We aim at transposing the viscosity theory techniques that are used in the classical theory ([3]) to the space of Borel probability measures $\mathcal{P}(M)$. In particular, we define a suitable notion of viscosity using a class of real-valued functions that admit *directional derivatives* at all points $\mu \in \mathcal{P}(M)$. We then prove a comparison principle that holds for any bounded upper semicontinuous subsolution and any bounded lower semicontinuous supersolution. Finally, we prove that the value function is the unique viscosity solution to the above HJB equation by using the dynamic programming principle verified by the value function.

The chapter is structured as follows. In Section 4.2, we formulate the Mayer problem in the space of probability measures and we give the main properties of the value functions. In Section 4.3, we recall some results of optimal transport theory and the geometry of $\mathcal{P}(M)$. In particular, we give a characterization of the geodesics in $\mathcal{P}(M)$, we describe the pseudo-Riemannian structure of $\mathcal{P}(M)$ and we give a definition of real-valued directionally differentiable functions in $\mathcal{P}(M)$. Section 4.4 is devoted to the study of a suitable HJB equation that characterizes the value function. In particular, we define the Hamiltonian we are going to work with, then we define a notion of viscosity using the class of functions that are directionally differentiable, we prove a comparison principle that holds for any bounded upper semicontinuous subsolution and any bounded lower semicontinuous supersolution and we prove that the value function is the viscosity solution of the HJB equation via the dynamic programming principle.

4.2 Setting of the problem

Throughout this manuscript, $(M, \langle ., . \rangle)$ is a finite dimensional, compact and connected Riemannian manifold without boundary. We denote by $|\,.\,|$ the associated norm on the tangent bundle TM, and by $d(.,.)$ its Riemannian distance on M. The metric space (M, d), is a complete and compact space and its topology is equivalent to the topology of the differentiable manifold M. The tangent bundle TM is itself a complete Riemannian manifold when endowed with the Sasaki metric [97]. We denote by $d_{TM}(.,.)$ its Riemannian distance on TM associated to the Sasaki metric (see Appendix 4.5.2).

We denote by $\mathcal{P}(M)$ the set of Borel probability measures over M and $\mathcal{P}_2(M)$ the set of Borel probability measures with bounded second moment

$$\mathcal{P}_2(M) := \{\, \mu \in \mathcal{P}(M) \ : \ \int d^2(x, x_0) d\mu(x) < \infty, \quad \forall\, x_0 \in M \,\}.$$

Actually, since M is compact, we have $\mathcal{P}_2(M) = \mathcal{P}(M)$ but we will keep using the notation $\mathcal{P}_2(M)$. Recall that for any two topological spaces X and Z, any Borel probability measure μ on X and any Borel function $g : X \to Z$, the pushforward measure $g\sharp\mu$ on Z is defined by

$$g\sharp\mu(A) = \mu(g^{-1}(A)) \quad \forall A \subset Z, \text{ a Borel set},$$

or equivalently,

$$\int h \; dg\sharp\mu = \int h \circ g \; d\mu, \quad \forall\, h : Z \to \mathbb{R}, \text{ Borel measurable and bounded.}$$

We define the Wasserstein distance $d_W(.,.)$ over $\mathcal{P}_2(M)$ by

$$d_W(\mu, \nu) := \sqrt{\inf \Big\{ \int d^2(x, y) d\gamma(x, y) \Big\}},$$

where the infimum is taken over all Borel probability measures of $M \times M$ that have marginals μ and ν, i.e.

$$\gamma \in \mathcal{P}(M\times M) \ : \ \gamma(A\times M) = \mu(A) \text{ and } \gamma(M\times B) = \nu(B) \quad \forall A, B, \text{ Borel sets of } M.$$

Such Borel probability measures γ are called admissible plans of μ and ν and the set of such plans is denoted $Adm(\mu, \nu)$. It is well known that d_W verifies all the axioms of a distance and that the infimum is always reached [98, Theorem 1.5]. The admissible plans where the minimum is achieved are called optimal transport plans and the set of such plans is denoted $Opt(\mu, \nu) \subset Adm(\mu, \nu)$.

Let $T > 0$ and U be a compact subset of a metric space. Consider the controlled system, defined for $T > t_0 \geq 0$ and $x_0 \in M$, as

$$\begin{cases} \dot{Y}(t) = f(Y(t), u(t)), & \text{for almost every } t \in [t_0, T], \\ Y(t_0) = x_0, u(t) \in U, \end{cases} \tag{4.4}$$

where $f : M \times U \to TM$ satisfies the following assumptions:

$$\textbf{(H)} \begin{cases} f : M \times U \to TM \text{ is continuous and Lipschitz with respect to the state, i.e.} \\ \qquad \exists\, k > 0 \ : \quad d_{TM}(f(x, u), f(y, u)) \leq k\, d(x, y), \quad \forall\, u \in U, \ (x, y) \in M \times M. \end{cases}$$

$(\boldsymbol{H_{co}})$ For all $x \in M$, the set of functions $\{x \mapsto f(x, u) \ : \ u \in U\}$ is convex.

.

Remark 4.2.1. Since M and U are compact, then the vector field f is bounded. Furthermore, the Lipschitz assumption on $f(., u)$ in Hypothesis **(H)** is equivalent to the following: there exists $k' > 0$ such that for all $u \in U$, $x, y \in M$ and every smooth curve $\alpha : [0, 1] \to M$ joining x and y, we have

$$|\tau^{\alpha}_{x,y}(f(x, u)) - f(y, u)| \leq k' \,\text{length}(\alpha),$$

with $\tau^{\alpha}_{x,y}$ is the parallel transport of $f(x, u)$ along the curve α and length(α) is the Riemannian length of the curve α (see [99, Lemma II.A.2.4]). We set

$$Lip(f) := \max(k, k').$$

Remark 4.2.2. Hypotheses **(H)** and $(\boldsymbol{H_{co}})$ are verified in particular by control-affine systems of the following form. Let $m \geq 1$ and let U be a compact convex subset of $\mathbb{R}^m$. Let $(f_0, ..., f_m)$ be an $(m + 1)$-tuple of Lipschitz vector fields on M. Then the dynamics

$$f(x, u) = f_0(x) + \sum_{i=1}^{m} u_i f_i(x), \quad \forall (x, u) \in M \times U,$$

verifies **(H)** and $(\boldsymbol{H_{co}})$.

We define the set of open-loop controls by

$$\mathcal{U} := \{u : [0,T] \to U \ : \ u(.) \text{ is measurable}\}.$$

Under the assumption **(H)**, classical results of ordinary differential equations hold. In particular, for any control $u(.) \in \mathcal{U}$ and $x_0 \in M$, there exists a unique Lipschitz trajectory $t \mapsto Y_t^{t_0,x_0,u}$ defined on all $[t_0,T]$ that is a solution to the controlled sytem (4.4). Moreover, we have the following estimates.

Proposition 4.0.1. *There exist $C_1, C_2 > 0$ positive constants such that for all $x_0, z_0 \in M$, for all $t_0 \in [0,T]$, and $t \mapsto Y_t^{t_0,x_0,u}$, $t \mapsto Y_t^{t_0,z_0,u}$ be solutions of (4.4), it holds:*

$$\forall t \in [t_0,T], \quad d(Y_t^{t_0,x_0,u}, Y_t^{t_0,z_0,u}) \le C_1\, d(x_0,z_0),$$
$$d(Y_t^{t_0,x_0,u}, x_0) \le C_2\, |t-t_0|, \quad t \in [t_0,T].$$

Proof. (Sketch). Since M is compact, then all the statements are local in nature. The global result is obtained by compactness of M and of $[t_0,T]$. First, by using Nash embedding theorem, M can be embedded isometrically into a Euclidean space $(\mathbb{R}^N, ||.||)$, with $N > 0$ big enough. Let $x_0 \in M$ and V be a small enough open neighborhood of x_0. Then for $z_0 \in V$ we can apply the usual theory in $\mathbb{R}^N$ and get

$$||Y_t^{t_0,x_0,u} - Y_t^{t_0,z_0,u}|| \le e^{Lip(f)T}\, ||x_0 - z_0||.$$

Then by using the fact that the Euclidean distance is equivalent to the Riemannian distance in V, we get the result. The second assertion can be established with similar arguments, by taking t small enough so that $Y_t^{t_0,x_0,u} \in V$. ■

The control problem aims at minimizing the final cost

$$\int \ell(Y_T^{t_0,x_0,u})\, d\mu_0(x_0),$$

over all trajectories that are solutions of the dynamics (4.4) with the initial condition at time t_0 is $x_0 \in M$, distributed along the measure $\mu_0 \in \mathcal{P}_2(M)$. We consider the following assumption:

$$\textbf{(}\boldsymbol{H_\ell}\textbf{)} \quad \ell : M \to \mathbb{R} \text{ is Lipschitz continuous with constant } Lip(\ell).$$

When μ_0 is equal to the Dirac mass δ_{x_0}, the resulting system corresponds to the classical case without uncertainties on the initial condition. This problem is thoroughly studied in the literature (see for example [9, 76]). When μ_0 is any probability measure of $\mathcal{P}_2(M)$, it is better to see this problem as an optimal control problem defined in the space of Borel probability measures $\mathcal{P}_2(M)$. First we rewrite the final cost the following way

$$\int \ell(Y_T^{t_0,x_0,u})\, d\mu_0(x_0) = \int \ell(y)\, dY_T^{t_0,\cdot,u}\sharp\mu_0\,(y),$$

and we minimize this cost over the set of trajectories $t \mapsto \mu_t^{t_0,\mu_0,u}$ of the space $\mathcal{P}_2(M)$ that verify

$$\begin{cases} \mu_t^{t_0,\mu_0,u} = Y_t^{t_0,.,u}\sharp\mu_0, & t \in [t_0,T], \text{ and } x \mapsto Y_t^{t_0,x,u} \text{ is the flow of (4.4)}, \\ \mu_{t_0}^{t_0,\mu_0,u} = \mu_0. \end{cases}$$

Hence if we set

$$\forall\, \mu \in \mathcal{P}_2(M), \quad L(\mu) = \int \ell(y) d\mu(y),$$

then the final cost becomes

$$\int \ell(y)\, dY_T^{t_0,.,u}\sharp\mu_0\,(y) = L(\mu_T^{t_0,\mu_0,u}).$$

For any $u(.) \in \mathcal{U}$, the map $x \mapsto f(x,u(t))$ is Lipschitz continuous and bounded independently of t. Hence, the curve $t \mapsto \mu_t^{t_0,\mu_0,u}$ is the unique continuous solution of the continuity equation (see [60, 93])

$$\begin{cases} \partial_t \mu_t^{t_0,\mu_0,u} + div(f(.,u(t))\mu_t^{t_0,\mu_0,u}) = 0, & t \in [t_0,T], \\ \mu_{t_0}^{t_0,\mu_0,u} = \mu_0, \end{cases}$$

in the sense of distributions, i.e.

$$\begin{cases} \displaystyle\int_{t_0}^T \int_M (\partial_t\phi(t,x) + \langle \nabla_x\phi(t,x), f(x,u(t))\rangle) d\mu_t^{t_0,\mu_0,u}(x)dt = 0, & \forall \phi \in C_c^\infty([t_0,T]\times M), \\ \mu_{t_0}^{t_0,\mu_0,u} = \mu_0, \end{cases}$$

where $C_c^\infty([t_0,T]\times M)$ is the class of smooth functions of $[t_0,T]\times M$ with compact support. Therefore, the above optimal control problem can be rewritten as

$$\begin{cases} \displaystyle\inf_{u(.)\in\mathcal{U}} L(\mu_T^{t_0,\mu_0,u}), \\ \text{such that } \begin{cases} \partial_t \mu_t^{t_0,\mu_0,u} + div(f(.,u(t))\mu_t^{t_0,\mu_0,u}) = 0, & t \in [t_0,T], \\ \mu_{t_0}^{t_0,\mu_0,u} = \mu_0, \end{cases} \end{cases} \tag{4.5}$$

and the infimum is reached, since the set of trajectories of (4.4) is compact in the topology of uniform convergence under Hypotheses $\boldsymbol{(H)}$ and $\boldsymbol{(H_{co})}$ [76, Theorem 1, pp 60]. The associated *value function* to the above optimal control problem is defined as

$$\vartheta(t_0,\mu_0) := \inf_{u(.)\in\mathcal{U}} L(\mu_T^{t_0,\mu_0,u}) = \inf_{u(.)\in\mathcal{U}} \int \ell(y)\, d\mu_T^{t_0,\mu_0,u}(y).$$

Under Hypotheses $\boldsymbol{(H)}$, $\boldsymbol{(H_\ell)}$ and $\boldsymbol{(H_{co})}$, we can already prove two properties of the value function.

Theorem 4.1 (Dynamic programming principle)**.** *Let $\mu \in \mathcal{P}_2(M)$, $t \in [0,T]$ and $h \in [t, T-t]$. Assume **(H)**, **(H_ℓ)** and **(H_{co})**. Then it holds*

$$\vartheta(t,\mu) = \inf_{u(.)\in\mathcal{U}} \vartheta(t+h, \mu_{t+h}^{t,\mu,u}).$$

Proof. Let $u_0(.) \in \mathcal{U}$ be such that

$$\vartheta(t,\mu) = \int \ell(Y_T^{t,x,u_0})d\mu(x) = \int \ell \, d\mu_T^{t,\mu,u_0}.$$

We have

$$\begin{aligned}
\vartheta(t,\mu) = \int \ell(Y_T^{t,x,u_0})d\mu(x) &= \int \ell(Y_T^{t+h,Y_{t+h}^{t,x,u_0},u_0})d\mu(x) \\
&= \int \ell(Y_T^{t+h,x,u_0})d(Y_{t+h}^{t,.,u_0}\sharp\mu)(x) \\
&= \int \ell(Y_T^{t+h,x,u_0})d\mu_{t+h}^{t,\mu,u_0}(x) \\
&\geq \inf_{u(.)\in\mathcal{U}} \int \ell(Y_T^{t+h,x,u})d\mu_{t+h}^{t,\mu,u_0}(x) \\
&= \inf_{u(.)\in\mathcal{U}} \int \ell \, d\mu_T^{t+h,\mu_{t+h}^{t,\mu,u_0},u} \\
&= \vartheta(t+h, \mu_{t+h}^{t,\mu,u_0}) \geq \inf_{v(.)\in\mathcal{U}} \vartheta(t+h,\mu_{t+h}^{t,\mu,v}).
\end{aligned}$$

It remains to prove the other inequality. Let $u : [t,T] \to U$ and $u_{opt} : [t+h,T] \to U$ be such that

$$\int \ell \, d\mu_T^{t+h,\mu_{t+h}^{t,\mu,u},u_{opt}} = \vartheta(t+h,\mu_{t+h}^{t,\mu,u})$$

Let $u^* : [0,T] \to U$ be the control function defined by

$$u^*(s) = \begin{cases} u(s), & \text{if } s \in [t,t+h], \\ u_{opt}(s), & \text{if } s \in [t+h,T]. \end{cases}$$

Thus we get

$$\begin{aligned}
\vartheta(t,\mu) \leq \int \ell \, d\mu_T^{t,\mu,u^*} &= \int \ell \, d\mu_T^{t+h,\mu_{t+h}^{t,\mu,u},u_{opt}} \\
&= \vartheta(t+h,\mu_{t+h}^{t,\mu,u}).
\end{aligned}$$

By taking the infimum over $u(.) \in \mathcal{U}$ we get the result. ■

Proposition 4.1.1. *Assume **(H)**, **(H_ℓ)** and **(H_{co})**. Then, the value function ϑ is Lipschitz continuous on $[0,T] \times \mathcal{P}_2(M)$. In particular, ϑ is bounded.*

Proof. Let $t \in [0,T]$, $\mu, \sigma \in \mathcal{P}_2(M)$. There exists a trajectory $s \mapsto Y_s^{t,x,u}$ such that

$$\int \ell(Y_T^{t,x,u})d\sigma(x) = \vartheta(t,\sigma).$$

Hence, we have

$$\vartheta(t,\mu) - \vartheta(t,\sigma) \leq \int \ell(Y_T^{t,x,u})d\mu(x) - \int \ell(Y_T^{t,x,u})d\sigma(x).$$

Let $\gamma \in Opt(\mu,\sigma)$. Then we get

$$\begin{aligned}\int \ell(Y_T^{t,x,u})d\mu(x) - \int \ell(Y_T^{t,x,u})d\sigma(x) &= \int \Big(\ell(Y_T^{t,x,u}) - \ell(Y_T^{t,y,u})\Big)d\gamma(x,y)\\ &\leq Lip(\ell)C_1 \int d(x,y)d\gamma(x,y)\\ &\leq Lip(\ell)C_1\sqrt{\int d^2(x,y)d\gamma(x,y)}\\ &= Lip(\ell)C_1 d_W(\mu,\sigma),\end{aligned}$$

where $C_1 > 0$ is defined in Proposition 4.0.1. Thus we get

$$\vartheta(t,\mu) - \vartheta(t,\sigma) \leq Lip(\ell)C_1 d_W(\mu,\sigma).$$

We can exchange the roles of σ and μ to get the exact same inequality. Therefore, we get the Lipschitz continuity with respect to the state variable. To prove Lipschitz continuity with respect to time, let $t, s \in [0,T]$. We assume, without loss of generality, that $0 \leq t < s \leq T$. By Theorem 4.1, there exists a trajectory $r \mapsto Y_r^{t,x,u}$ such that

$$\vartheta(t,\sigma) = \vartheta(s, Y_s^{t,\cdot,u}\sharp\sigma).$$

We have

$$\begin{aligned}|\vartheta(s,\sigma) - \vartheta(t,\sigma)| &= |\vartheta(s,\sigma) - \vartheta(s, Y_s^{t,\cdot,u}\sharp\sigma)|\\ &\leq Lip(\ell)C_1 d_W(\sigma, Y_s^{t,\cdot,u}\sharp\sigma)\\ &\leq Lip(\ell)C_1\sqrt{\int d^2(x, Y_s^{t,x,u})d\sigma(x)}\\ &\leq Lip(\ell)C_1C_2|t-s|,\end{aligned}$$

where $C_1, C_2 > 0$ are defined in Proposition 4.0.1. Thus ϑ is Lipschitz continuous with respect to the time variable, and the proof is completed. ■

In the classical theory of viscosity, the value function is the unique viscosity solution of the Hamilton Jacobi Bellman equation [4]. The goal of the next two sections is to show that the value function, in this setting, is also a viscosity solution to a Hamilton Jacobi Bellman equation of the form

$$\begin{cases} \partial_t v + H(\mu, D_\mu v) = 0, & (t,\mu) \in [0,T) \times \mathcal{P}_2(M), \\ v(T,\mu) = L(\mu). \end{cases}$$

In order to define the Hamiltonian and the notation $D_\mu v$ rigorously, we will analyse the geometric structure of Wasserstein spaces in the next section.

4.3 Wasserstein space over compact Riemannian Manifolds

The first subsection aims to give some geometric and topological properties of the Wasserstein space and to give a characterization of the geodesics in the Wasserstein space. In the second subsection we describe the pseudo-Riemannian structure that the Wasserstein space enjoys. In particular, we shed some light on where this structure behaves “nicely” and where it degenerates. Finally, in the last subsection we give the definition of directionally differentiable functions in the Wasserstein space. All these tools are going to be necessary to give a precise definition of the *Hamiltonian* and *viscosity notion* for Hamilton Jacobi equations in $\mathcal{P}_2(M)$.

4.3.1 Geometric and topological properties of Wasserstein space

The Wasserstein space $(\mathcal{P}_2(M), d_W)$ inherits many geometric and topological properties from the base space (M,d). Indeed, since (M,d) is a Polish space (because it is a complete and separable metric space), then $(\mathcal{P}_2(M), d_W)$ is a Polish space. Also, since M is compact, then $(\mathcal{P}_2(M), d_W)$ is also compact ([35, Chapter 6]).
Next, we recall the definition of geodesic spaces. Let (X, d_X) be a metric space. A curve $\alpha : [0,1] \to X$ is called a *constant speed geodesic* if

$$d_X(\alpha_t, \alpha_s) = |t-s| d_X(\alpha_0, \alpha_1), \quad \forall t, s \in [0,1].$$

The metric space (X, d_X) is said to be a *geodesic space* if any two points of X are connected by at least one constant speed geodesic. In what follows, we intend by ‘geodesic’, a constant speed geodesic. Note that the metric spaces (M,d) and (TM, d_{TM}) are geodesic spaces. Furthermore, the Wasserstein space $(\mathcal{P}_2(M), d_W)$ inherits this property from (M,d) and is also a geodesic space (see [60] or [35]).

We denote by $\mathcal{P}(TM)$ the set of Borel probability measures over TM. We define the Wasserstein space over (TM, d_{TM}) by

$$\mathcal{P}_2(TM) = \{ \eta \in \mathcal{P}(TM) \ : \ \int d^2_{TM}\Big((x,v),(x_0,v_0)\Big) d\eta(x,v) < \infty, \quad \forall\, (x_0,v_0) \in TM \} \tag{4.6}$$

endowed with its corresponding Wasserstein distance. It is sufficient that the condition

$$\int d_{TM}^2\Big((x,v),(x_0,v_0)\Big)d\eta(x,v) < \infty$$

in (4.6) is verified for only one point $(x_0,v_0) \in TM$. Thus if we take $(x_0, 0_{x_0}) \in TM$, then from the definition of d_{TM} (see Appendix 4.5.2), this condition is equivalent to

$$\int |v|^2 d\eta(x,v) < \infty.$$

For $\mu \in \mathcal{P}_2(M)$, we denote by $\mathcal{P}_2(TM)_\mu \subset \mathcal{P}_2(TM)$, the set of measures γ such that $\pi^M \sharp \gamma = \mu$, where $\pi^M : TM \to M$ is the canonical projection onto M. This set is equivalent to the set of measures $\gamma \in \mathcal{P}(TM)$ such that

$$\pi^M \sharp \gamma = \mu, \quad \text{and} \quad \int |v|^2 d\gamma(x,v) < \infty.$$

Let $\exp : TM \to M$ be the exponential map of $(M, \langle ., . \rangle)$. The exponential $\mathbf{exp}_\mu(\gamma)$ of a measure $\gamma \in \mathcal{P}_2(TM)_\mu$ is defined by

$$\mathbf{exp}_\mu(\gamma) := \exp \sharp \gamma \in \mathcal{P}_2(M).$$

We define the map $\mathbf{exp}_\mu^{-1} : \mathcal{P}_2(M) \to \mathcal{P}_2(TM)_\mu$ by

$$\mathbf{exp}_\mu^{-1}(\nu) := \{\, \gamma \in \mathcal{P}_2(TM)_\mu \ : \ \mathbf{exp}_\mu(\gamma) = \nu \text{ and } \int |v|^2 d\gamma(x,v) = (d_W(\mu,\nu))^2 \,\},$$

or in other words, the set of measures $\gamma \in \mathcal{P}_2(TM)$ such that $(\pi^M, \exp)\sharp\gamma$ is an optimal plan from μ to ν and

$$\int |v|^2 d\gamma(x,v) = (d_W(\mu,\nu))^2.$$

We introduce the following notation

$$\Delta_t(x,v) = (x, tv), \quad \forall t \in \mathbb{R},\ (x,v) \in TM \quad \text{and} \quad t \centerdot \gamma = \Delta_t \sharp \gamma, \quad \forall t \in \mathbb{R},\ \gamma \in \mathcal{P}_2(TM).$$

Lemma 4.2. *([68, Theorem 1.11]) Let $\mu, \nu \in \mathcal{P}_2(M)$. A curve $(\mu_t)_{t\in[0,1]} \subset \mathcal{P}_2(M)$ is a geodesic connecting μ to ν if and only if there exists a measure $\gamma \in \mathbf{exp}_\mu^{-1}(\nu)$ such that*

$$\mu_t := (\exp \circ \Delta_t)\sharp\gamma = \mathbf{exp}_\mu(t \centerdot \gamma), \quad \forall\, t \in [0,1]. \tag{4.7}$$

The measure γ uniquely defines the geodesic $(\mu_t)_{t\in[0,1]}$.

Remark 4.3.1. The map $\mathbf{exp}_\mu^{-1}$ is not really an inverse map to $\mathbf{exp}_\mu$ since only the measures $\gamma \in \mathcal{P}_2(TM)_\mu$ such that $(\pi^M, \exp)\sharp\gamma$ are optimal plans between μ and ν and

$$\int |v|^2 d\gamma(x,v) = (d_W(\mu,\nu))^2$$

are considered. While this might seem confusing, the map $\mathbf{exp}_\mu^{-1}$ is defined this way so that for all $\gamma \in \mathbf{exp}_\mu^{-1}(\nu)$, the curve $t \mapsto \mathbf{exp}_\mu(t \cdot \gamma)$ is a geodesic connecting μ and ν.

From Lemma 4.2, we get the following result about geodesics emanating from any $\mu \in \mathcal{P}_2(M)$.

Proposition 4.2.1. *([68, Proposition 1.12]) Let $\mu \in \mathcal{P}_2(M)$ and let $(\mu_t)_t$ be a geodesic emanating from μ and defined in some interval $[0, \varepsilon]$, with $\varepsilon > 0$. Then there exists a unique measure $\gamma \in \mathcal{P}_2(TM)_\mu$ such that*

$$\mu_t = \mathbf{exp}_\mu(t \cdot \gamma), \quad t \in [0, \varepsilon].$$

To summarize, we have seen in this section that the Wasserstein space $(\mathcal{P}_2(M), d_W)$ is a compact geodesic space and each geodesic starting from $\mu \in \mathcal{P}_2(M)$ can be characterized by a measure $\gamma \in \mathcal{P}_2(TM)_\mu$ as shown in Proposition 4.2.1. We stress on the fact that *not* all the curves of the form

$$t \mapsto \mathbf{exp}_\mu(t \cdot \gamma), \quad t \in [0, \varepsilon],\ \gamma \in \mathcal{P}_2(TM)_\mu$$

are geodesics but all the geodesics are of this form.

4.3.2 The space of gradient vector fields and the tangent cone in $\mathcal{P}_2(M)$

The space $(\mathcal{P}_2(M), d_W)$ has a pseudo-Riemannian structure. This has been first pointed out by Otto in [65] and was justified rigorously by Ambrosio-Gigli-Savaré in the case of Wasserstein spaces over the Euclidean space $\mathcal{P}_2(\mathbb{R}^N)$ [60, Chapter 8]. We give hereafter the construction in $\mathcal{P}_2(M)$ following [100]. Let $L^2(\mu, TM)$ be the space of squared integrable vector fields with respect to $\mu \in \mathcal{P}_2(M)$, i.e. vector fields $w : M \to TM$ such that

$$||w||^2_{L^2(\mu,TM)} := \int_M \langle w(x), w(x) \rangle d\mu(x) < +\infty.$$

The *space of gradient vector fields* at μ is the following Hilbert space:

$$SpGr_\mu(\mathcal{P}_2(M)) := \overline{\{\nabla\phi \ : \ \phi \in C_c^\infty(M)\}}^{L^2(\mu,TM)}.$$

Here $C_c^\infty(M)$ is the space of smooth functions of M with compact support. We denote by

$$\pi^\mu : L^2(\mu, TM) \to SpGr_\mu(\mathcal{P}_2(M))$$

the orthogonal projection map onto the space of gradient vector fields. The space of gradient vector fields is linked to the continuity equation in the following way. Let

$I \subset \mathbb{R}$ be an open interval and $(\mu_t)_{t\in I}$ be a Lipschitz curve in $\mathcal{P}_2(M)$. Then, following [100, Proposition 2.5], there exists a family of vector fields $w : I \times M \to TM$, $(t,x) \mapsto w_t(x) \in T_xM$, such that

$$\int_r^s ||w_t||^2_{L^2(\mu_t,TM)}\, dt < \infty, \quad \forall r,s \in I\ : r < s,$$

and the continuity equation

$$\frac{d}{dt}\mu_t + div(w_t\mu_t) = 0, \tag{4.8}$$

is satisfied in the distributional sense, i.e.

$$\int_I \int_M (\partial_t\phi(t,x) + \langle \nabla_x\phi(t,x), w_t(x)\rangle) d\mu_t(x)dt = 0, \quad \forall \phi \in C_c^\infty(I \times M).$$

The family of vector fields w is not unique in general. Indeed, notice that if the family of vector fields w defined above verifies the continuity equation (4.8) in the sense of distributions, then the family of vector fields $(t,x) \mapsto \pi^{\mu_t} \circ w_t(x)$ also verifies the continuity equation (4.8) in the sense of distributions. However, it can be shown [100, Propositions 2.4 and 2.5] that $(t,x) \mapsto \pi^{\mu_t} \circ w_t(x)$ is the unique family of vector fields that verifies (4.8) with minimal $L^2(\mu_t,TM)$ norm for almost all $t \in I$ and

$$\pi^{\mu_t} \circ w_t(.) \in SpGr_{\mu_t}(\mathcal{P}_2(M)), \quad \text{for almost all } t \in I.$$

One can think of $\pi^{\mu_t} \circ w_t(.)$ as the *velocity vector field* at time t for the curve $(\mu_t)_{t\in I}$. The construction of the space of gradient vector fields is analytical and it has the advantage to retain the link between Lipschitz curves of $\mathcal{P}_2(M)$ and the continuity equation (4.8). We will use this point of view to justify the expression of the Hamiltonian associated to the optimal control problem (4.5).

On the other hand, there is another point of view that also justifies rigorously the pseudo-Riemannian structure of $(\mathcal{P}_2(M), d_W)$ which consists in using tools of *metric geometry*. In short, for sufficiently well-behaved metric spaces, one can define a *tangent cone* at every point of the space. The tangent cone is the metric counterpart of the tangent space for Riemannian manifolds. For $(\mathcal{P}_2(M), d_W)$, it was shown in [68, 58] that the notion of tangent cone is well-defined at every point. We give hereafter the definition of the tangent cone in $(\mathcal{P}_2(M), d_W)$ following [68]. First we define the space of directions at a point. Let $\mu \in \mathcal{P}_2(M)$. The space of directions at μ is the set of "initial velocities" of geodesics emanating from μ:

$$Dir_\mu := \left\{ \gamma \in \mathcal{P}_2(TM)_\mu\ :\ t \mapsto \mathbf{exp}_\mu(t\mathbf{.}\gamma) \text{ is a geodesic defined in some interval } [0,\varepsilon] \right\}.$$

This definition is a direct consequence of Proposition 4.2.1. In this point of view, The measures γ are seen as the "initial velocities" of their corresponding geodesics starting from μ in analogy with Riemannian geometry. Next, we are going to define the *tangent cone* at μ. Following [68, Section 3], we define the following distance W_μ on Dir_μ. For all $\gamma, \eta \in Dir_\mu$, the limit

$$W_\mu(\gamma, \eta) := \lim_{t \downarrow 0} \frac{d_W\Big(\mathbf{exp}_\mu(t \centerdot \gamma), \mathbf{exp}_\mu(t \centerdot \eta)\Big)}{t}.$$

exists and defines a distance on Dir_μ [68, Corollary 5.6].

Definition 4.1. (Tangent cone). Let $\mu \in \mathcal{P}_2(M)$. The tangent cone $T_\mu \mathcal{P}_2(M)$ is the following set

$$\begin{aligned} T_\mu \mathcal{P}_2(M) &:= \overline{Dir_\mu}^{W_\mu} \\ &= \overline{\Big\{\gamma \in \mathcal{P}_2(TM)_\mu \ : \ t \mapsto \mathbf{exp}_\mu(t \centerdot \gamma) \text{ is a geodesic defined in some } [0, \varepsilon]\Big\}}^{W_\mu} \end{aligned}$$

with the closure taken with respect to the distance W_μ. By closure we intend the abstract completion of Dir_μ with respect to W_μ. One can see clearly the structure of a cone in $T_\mu \mathcal{P}_2(M)$ since we have

$$\forall \gamma \in T_\mu \mathcal{P}_2(M), \ \forall \lambda \in \mathbb{R}^+, \quad \lambda \centerdot \gamma \in T_\mu \mathcal{P}_2(M).$$

The above definition of W_μ is necessary for the tangent cone to be well-defined for reasons we will not develop here. This construction is not specific to Wasserstein spaces. It is valid for a large class of metric spaces [51]. In the case of Wasserstein spaces, the interested reader can check [68, 58] for more details. The important idea to retain here is that the tangent cone is always defined as the completion of the space of directions with respect to this distance.

Next, we highlight the connexion between the tangent cone and the space of gradient vector fields following [68, Section 6]. For any $\gamma \in \mathcal{P}_2(TM)_\mu$, we define its *barycentric projection* the following way

$$\mathcal{B}_\mu : \mathcal{P}_2(TM)_\mu \to L^2(\mu, TM) \ : \quad \mathcal{B}_\mu(\gamma)(x) = \int v d\gamma_x(v),$$

where $\{\gamma_x\}_{x \in M}$ is the disintegration of γ with respect to the projection π^M (see appendix 4.5.1). The barycentric projection verifies the following equality:

$$\forall g(.) \in L^2(\mu, TM), \quad \forall x \in M, \quad \mathcal{B}_\mu(g \sharp \mu)(x) = g(x).$$

Furthermore, the barycentric projection is characterized by the following equality

$$\begin{aligned} \int \langle w(x), v \rangle d\gamma(x, v) &= \int \langle w(x), \int v d\gamma_x(v) \rangle \, d\pi^M \sharp \gamma(x) \\ &= \int \langle w(x), \mathcal{B}_\mu(\gamma)(x) \rangle \, d\mu(x), \quad \forall w \in L^2(\mu, TM). \end{aligned} \tag{4.9}$$

Following [68, Corollary 6.4 and Proposition 6.3], the barycentric projection and the pushforward map link the tangent cone with the space of gradient vector fields in the following way:

$$\begin{aligned} SpGr_\mu(\mathcal{P}_2(M)) &= \{g(.) \in L^2(\mu, TM) \ : \ g\sharp\mu \in T_\mu\mathcal{P}_2(M)\}, \\ &= \{\, \mathcal{B}_\mu(\gamma)(x) \ : \ \gamma \in T_\mu\mathcal{P}_2(M)\}. \end{aligned} \tag{4.10}$$

Consequently, the space of gradient vector fields can be seen as a subset of the tangent cone since we trivially get from (4.10)

$$\forall\mu \in \mathcal{P}_2(M), \quad w(.) \in SpGr_\mu(\mathcal{P}_2(M)) \Longleftrightarrow w\sharp\mu \in T_\mu\mathcal{P}_2(M) \text{ and } w(.) \in L^2(\mu, TM).$$

More generally, given $g(.) \in L^2(\mu, TM)$, the measure $\pi^\mu \circ g\sharp\mu$ belongs to the tangent cone, i.e.

$$\pi^\mu \circ g\sharp\mu \in T_\mu\mathcal{P}_2(M), \quad \text{since} \quad \pi^\mu \circ g(.) \in SpGr_\mu(\mathcal{P}_2(M)).$$

A natural question that raises itself is when the tangent cone and the space of gradient vector fields are equal (up to an isometry). This question was answered by Gigli in [68]. It was shown that the two sets are equal at some $\mu \in \mathcal{P}_2(M)$ if and only if μ is a “regular measure”, meaning that it gives zero measure to any hypersurface of M which, locally, is the graph of the difference of two convex functions [68, Corollary 6.6]. A regular measure μ is characterized by the following property: for any $\sigma \in \mathcal{P}_2(M)$, there exists a unique optimal transport plan between μ and σ and it is induced by a map, i.e. there exists a Borel measurable map $T : M \to M$ such that $(Id, T)\sharp\mu$ is the optimal transport plan between μ and σ. This is a refinement of Brenier-McCann's result [101] in which the same property was proven to be true for the case where μ is absolutely continuous with respect to the Riemannian volume form. Intuitively, it means that when μ is a regular measure, the Riemannian structure on $\mathcal{P}_2(M)$ behaves nicely, since the tangent cone is equal to the space of gradient vector fields, so it is a Hilbert space, in contrast with when μ is not a regular measure where the structure of the tangent cone degenerates. This distinction is important for us because we want to build a robust viscosity notion for first order Hamilton Jacobi equations that will allow us to treat them in all $\mathcal{P}_2(M)$. Therefore, we will use the tangent cone to define directionally differentiable functions since the tangent cone encodes all the information about initial velocities of geodesics starting from μ. We will then use directionally differentiable functions to define the viscosity notion.

4.3.3 Semiconvex/Semiconcave/DC functions

In $\mathcal{P}_2(M)$, real-valued Lipschitz semiconvex or semiconcave functions admit directional derivatives at every point. These functions are going to serve us as test functions in the definition of viscosity notion. Moreover, the squared Wasserstein distance function $d_W^2(.,\sigma)$ (for some $\sigma \in \mathcal{P}_2(M)$ fixed) is Lipschitz and semiconcave. An explicit formula will be given for its directional derivatives at every point.

Let $F : \mathcal{P}_2(M) \to \mathbb{R}$ be a function and $\mu \in \mathcal{P}_2(M)$. We say that F has a *directional derivative* at μ along a geodesic $\alpha : [0,\varepsilon] \to \mathcal{P}_2(M)$ emanating from μ, with $\varepsilon > 0$, if the limit

$$\frac{d}{dt}\Big|_{t=0} F(\alpha_t) = \lim_{t\downarrow 0} \frac{F(\alpha_t) - F(\alpha_0)}{t}$$

exists and is finite (notice that the continuity of F is not required in this definition). A particular class of functions that admit directional derivatives are Lipschitz functions that can be represented as a difference of semiconvex functions. We refer to them as Lipschitz and *DC functions*. We define them hereafter.

Definition 4.2. Let $F : \mathcal{P}_2(M) \to \mathbb{R}$ be a function.

- We say that F is semiconcave if there exists $\lambda \in \mathbb{R}$ such that for every geodesic $\alpha : [0,1] \to \mathcal{P}_2(M)$ the following inequality holds

$$F(\alpha_t) \geq (1-t)F(\alpha_0) + tF(\alpha_1) - \frac{\lambda}{2} t(1-t) d_W^2(\alpha_0, \alpha_1).$$

- Similarly, we say that F is semiconvex if and only if $-F$ is semiconcave.

- Finally, we say that F is a DC function if it can be represented as a difference of two semiconvex functions.

In particular, every semiconvex function is a DC function and every semiconcave function is also a DC function.
Let $\mu \in \mathcal{P}_2(M)$ and $F : \mathcal{P}_2(M) \to \mathbb{R}$ be a Lipschitz and semiconcave function. The directional derivative of F at μ along a geodesic α emanating from μ

$$\frac{d}{dt}\Big|_{t=0} F(\alpha_t) = \lim_{t\downarrow 0} \frac{F(\alpha_t) - F(\alpha_0)}{t},$$

exists and is finite by [51, Proposition 6.14]. Furthermore, by Proposition 4.2.1, every geodesic α emanating from μ is of the following form

$$\alpha_t = \mathbf{exp}_\mu(t \centerdot \gamma), \quad \text{for some } \gamma \in Dir_\mu \text{ and } t \in [0,\varepsilon].$$

So we define the *differential function* of F on Dir_μ, denoted by $D_\mu F(\mu)$ in the following way:

$$\forall \gamma \in Dir_\mu, \quad D_\mu F(\mu)(\gamma) := \lim_{t\downarrow 0} \frac{F(\mathbf{exp}_\mu(t \centerdot \gamma)) - F(\mu)}{t}.$$

Moreover, following [51, Proposition 6.14] the differential function

$$\gamma \mapsto D_\mu F(\mu)(\gamma)$$

is Lipschitz and positively homogeneous in (Dir_μ, W_μ), with a Lipschitz constant inferior or equal to the Lipschitz constant of F. Consequently, we can extend the *differential function* $D_\mu F(\mu)(.)$ to be defined on the whole tangent cone $(T_\mu \mathcal{P}_2(M), W_\mu)$ by density.
Similarly, if $F : \mathcal{P}_2(M) \to \mathbb{R}$ is Lipschitz and semiconvex, then it is directionally differentiable and its differential function is Lipschitz and positively homogeneous and is defined by

$$D_\mu F(\mu)(.) = -D_\mu(-F)(\mu)(.).$$

Finally, if $F : \mathcal{P}_2(M) \to \mathbb{R}$ is a Lipschitz and DC function then it is directionally differentiable and its differential is Lipschitz and positively homogeneous.
For $\mu \in \mathcal{P}_2(M)$, we denote by $\mathcal{C}_\mu(\mathcal{P}_2(M))$ the class of Lipschitz and positively homogeneous functions of $T_\mu \mathcal{P}_2(M)$ and we set

$$\mathcal{C}(\mathcal{P}_2(M)) := \bigcup_{\mu \in \mathcal{P}_2(M)} \{\mu\} \times \mathcal{C}_\mu(\mathcal{P}_2(M)),$$

to be the metric analogue of the *cotangent bundle* in $\mathcal{P}_2(M)$. Next, we give an explicit expression of the directional derivatives of the squared Wasserstein distance. The next result shows that the squared Wasserstein distance is a *semiconcave* function.

Proposition 4.2.2. *([68, Proposition 4.1]). Let $\sigma \in \mathcal{P}_2(M)$ be fixed. Then the squared Wasserstein distance*

$$\mathcal{P}_2(M) \ni \nu \mapsto d_W^2(\nu, \sigma)$$

is a Lipschitz and semiconvave function.

In particular, the squared Wasserstein distance function is directionally differentiable. In fact, a much more general result holds: if $F(.) = d^2(., \sigma)$, then the limit

$$\lim_{t \downarrow 0} \frac{F(\mathbf{exp}_\mu(t \centerdot \gamma)) - F(\mu)}{t}$$

exists for all curves of the form

$$t \mapsto \mathbf{exp}_\mu(t \centerdot \gamma), \quad \text{for some } \gamma \in \mathcal{P}_2(TM)_\mu \text{ and } t \in [0, \varepsilon]. \tag{4.11}$$

even though they are not geodesics. We will only give a weaker version of the expression of the above limit, when

$$\gamma = g\sharp\mu, \quad g(.) \in L^2(\mu, TM).$$

The general result can be found in [68, Theorem 4.2].

Proposition 4.2.3. *(Derivative of the squared Wasserstein distance) Let $\mu, \sigma \in \mathcal{P}_2(M)$, and $g(.) \in L^2(\mu, TM)$. Let $\gamma = g\sharp\mu \in \mathcal{P}_2(TM)_\mu$. Let $F : \mathcal{P}_2(M) \to \mathbb{R}$ be the function*

$$\forall \nu \in \mathcal{P}_2(M), \quad F(\nu) = d_W^2(\nu, \sigma).$$

Then it holds

$$\lim_{t\downarrow 0}\frac{F(\mathbf{exp}_\mu(t\cdot\gamma))-F(\mu)}{t} = -2\sup_{\zeta\in\mathbf{exp}_\mu^{-1}(\sigma)}\int\langle g(x),v\rangle d\zeta(x,v). \tag{4.12}$$

We stress on the fact that equality (4.12) holds for all curves of the form (4.11) even though they are not geodesics. Next, we show the following result concerning the differential of the squared Wasserstein distance, which is a consequence of Proposition 4.2.3 and the properties of the barycentric projection.

Theorem 4.3. *Let $\mu,\sigma\in\mathcal{P}_2(M)$, and $g(.)\in L^2(\mu,TM)$. Let $\gamma=\pi^\mu\circ g\sharp\mu\in T_\mu\mathcal{P}_2(M)$. Let $F:\mathcal{P}_2(M)\to\mathbb{R}$ be the function*

$$\forall\nu\in\mathcal{P}_2(M),\quad F(\nu)=d_W^2(\nu,\sigma).$$

Then it holds

$$\begin{aligned}D_\mu F(\mu)(\gamma)=\lim_{t\downarrow 0}\frac{F(\mathbf{exp}_\mu(t\cdot\gamma))-F(\mu)}{t} &= -2\sup_{\zeta\in\mathbf{exp}_\mu^{-1}(\sigma)}\int\langle\pi^\mu\circ g(x),v\rangle d\zeta(x,v)\\ &= -2\sup_{\zeta\in\mathbf{exp}_\mu^{-1}(\sigma)}\int\langle g(x),v\rangle d\zeta(x,v).\end{aligned}$$

Proof. First, we show that

$$D_\mu F(\mu)(\gamma)=\lim_{t\downarrow 0}\frac{F(\mathbf{exp}_\mu(t\cdot\gamma))-F(\mu)}{t}.$$

Since $\gamma\in T_\mu\mathcal{P}_2(M)$ and $D_\mu F(\mu)(.)$ is Lipschitz, then there exists a sequence of measures $(\gamma_n)_n\subset Dir_\mu$ such that $W_\mu(\gamma_n,\gamma)\to 0$ and $D_\mu F(\mu)(\gamma_n)\to D_\mu F(\mu)(\gamma)$ as n tends to infinity. We have

$$\begin{aligned}\left|D_\mu F(\mu)(\gamma)-\lim_{t\downarrow 0}\frac{F(\mathbf{exp}_\mu(t\cdot\gamma))-F(\mu)}{t}\right| &= \left|\lim_{n\to\infty}D_\mu F(\mu)(\gamma_n)-\lim_{t\downarrow 0}\frac{F(\mathbf{exp}_\mu(t\cdot\gamma))-F(\mu)}{t}\right|\\ &= \left|\lim_{n\to\infty}\lim_{t\downarrow 0}\frac{F(\mathbf{exp}_\mu(t\cdot\gamma_n))-F(\mathbf{exp}_\mu(t\cdot\gamma))}{t}\right|\\ &\le Lip(F)\lim_{n\to\infty}\lim_{t\downarrow 0}\frac{d_W\Big(\mathbf{exp}_\mu(t\cdot\gamma_n),\mathbf{exp}_\mu(t\cdot\gamma)\Big)}{t}\\ &= Lip(F)\lim_{n\to\infty}W_\mu(\gamma_n,\gamma)=0,\end{aligned}$$

where $Lip(F)$ is the Lipschitz constant of F. This implies the result. Furthermore, Proposition 4.2.3 gives us

$$\lim_{t\downarrow 0}\frac{F(\mathbf{exp}_\mu(t\cdot\gamma))-F(\mu)}{t} = -2\sup_{\zeta\in\mathbf{exp}_\mu^{-1}(\sigma)}\int\langle\pi^\mu\circ g(x),v\rangle d\zeta(x,v).$$

It remains to prove the last equality. First, notice that from Lemma 4.2 we trivially get

$$\mathbf{exp}_\mu^{-1}(\sigma) \subset T_\mu \mathcal{P}_2(M).$$

Furthermore, we know from equality (4.10) that if $\zeta \in T_\mu \mathcal{P}_2(M)$, then $\mathcal{B}_\mu(\zeta) \in SpGr_\mu(\mathcal{P}_2(M))$. Hence, from (4.9) and (4.10) we deduce that for any $\zeta \in \mathbf{exp}_\mu^{-1}(\sigma)$ we have

$$\begin{aligned}\int \langle g(x), v\rangle d\zeta(x,v) &= \int \langle g(x), \int v d\zeta_x(v)\rangle d\mu(x)\\ &= \int \langle g(x), \mathcal{B}_\mu(\zeta)(x)\rangle d\mu(x)\\ &= \int \langle \pi^\mu \circ g(x), \mathcal{B}_\mu(\zeta)(x)\rangle d\mu(x)\\ &= \int \langle \pi^\mu \circ g(x), v\rangle d\zeta(x,v),\end{aligned}$$

which implies the last equality. ■

4.4 Time dependent Hamilton Jacobi Bellman equation in $\mathcal{P}_2(M)$

We have defined all the elements we need to give a precise definition of the Hamiltonian and the viscosity notion. In this section, we prove that the value function is the unique viscosity solution to a Hamilton Jacobi Bellman equation. First, we give a justification for the Hamiltonian we are going to work with, based on Otto's point of view of the pseudo-Riemaniann structure of $\mathcal{P}_2(M)$. We recall from Section 4.2 that the value function ϑ is equal to

$$\forall (t_0,\mu_0) \in [0,T]\times\mathcal{P}_2(M),\ \vartheta(t_0,\mu_0) = \begin{cases} \inf\limits_{u\in\mathcal{U}} L(\mu_T^{t_0,\mu_0,u}) \\ \text{such that} \begin{cases} \partial_t \mu_t^{t_0,\mu_0,u} + div(f(.,u(t))\mu_t^{t_0,\mu_0,u}) = 0, \\ \mu_{t_0}^{t_0,\mu_0,u} = \mu_0, \quad t\in[t_0,T], \end{cases} \end{cases}$$

and the continuity equation

$$\begin{cases} \partial_t \mu_t^{t_0,\mu_0,u} + div(f(x,u(t))\mu_t^{t_0,\mu_0,u}) = 0, \quad t\in[t_0,T], \\ \mu_{t_0}^{t_0,\mu_0,u} = \mu_0, \end{cases}$$

is understood in the sense of distributions, i.e.

$$\begin{cases} \displaystyle\int_{t_0}^T \int_M (\partial_t \phi(t,x) + \langle \nabla_x \phi(t,x), f(x,u(t))\rangle) d\mu_t^{t_0,\mu_0,u}(x) dt = 0, \quad \forall \phi \in C_c^\infty([t_0,T]\times M), \\ \mu_{t_0}^{t_0,\mu_0,u} = \mu_0 \end{cases}$$

Moreover, by the discussion made at the beginning of Section 4.3.2, every trajectory $t \mapsto \mu_t^{t_0,\mu_0,u}$ is also a solution to the continuity equation

$$\begin{cases} \partial_t \mu_t^{\mu_0,u} + div(\pi^{\mu_t} \circ f(x,u(t))\mu_t^{t_0,\mu_0,u}) = 0, \quad t \in [t_0,T], \\ \mu_{t_0}^{t_0,\mu_0,u} = \mu_0, \end{cases}$$

in the distributional sense. Hence, the quantity

$$\pi^{\mu_t} \circ f(.,u(t))\sharp\mu_t^{t_0,\mu_0,u} \in T_{\mu_t^{t_0,\mu_0,u}}\mathcal{P}_2(M)$$

can be seen as the *velocity* at time t of the trajectories $t \mapsto \mu_t^{t_0,\mu_0,u}$. This heuristic argument motivates us to consider the following Hamiltonian $H : \mathcal{C}(\mathcal{P}_2(M)) \to \mathbb{R}$ defined by

$$\forall (\mu, p_\mu) \in \mathcal{C}(\mathcal{P}_2(M)), \quad H(\mu,p_\mu) = \inf_{u \in U} p_\mu\Big(\pi^\mu \circ f(.,u)\sharp\mu\Big). \tag{4.13}$$

The definition of the Hamiltonian here resembles the one we usually encounter when Hamilton Jacobi equations are studied on a differentiable manifold. The only difference here is that since $(\mathcal{P}_2(M), d_W)$ is a metric space, the Hamiltonian is defined on the metric cotangent bundle. We consider the following Hamilton Jacobi equation

$$\begin{cases} \partial_t v + H(\mu, D_\mu v) = 0, \quad (t,\mu) \in [0,T) \times \mathcal{P}_2(M), \\ v(T,\mu) = L(\mu) = \displaystyle\int \ell d\mu. \end{cases} \tag{4.14}$$

We will take test functions that are twice continuously differentiable with respect to the time variable and in the class of DC functions with respect to the measure variable in order to define the notions of viscosity supersolution and viscosity subsolution.

Definition 4.3. (Test functions).
Let $\mathcal{TEST}_1$ be the set defined as:

$$\mathcal{TEST}_1 := \{(t,\mu) \mapsto \psi(t) + a\, d_W^2(\mu,\sigma) \ : \ a \in \mathbb{R}^+, \ \sigma \in \mathcal{P}_2(M), \ \psi(.) \in C^2([0,T],\mathbb{R})\}.$$

We set $\mathcal{TEST}_2 = -\mathcal{TEST}_1$, so we have

$$\mathcal{TEST}_2 = \{(t,\mu) \mapsto \psi(t) + a\, d_W^2(\mu,\sigma) \ : \ a \in \mathbb{R}^-, \ \sigma \in \mathcal{P}_2(M), \ \psi(.) \in C^2([0,T],\mathbb{R})\}.$$

Definition 4.4. (Viscosity solutions).

- We say that a function $v : [0,T) \times \mathcal{P}_2(M) \to \mathbb{R}$ satisfies the inequality

$$\partial_t v + H(\mu, D_\mu v) \geq 0,$$

at $(t, \mu) \in [0, T) \times \mathcal{P}_2(M)$ in the viscosity sense if v is upper semicontinuous and for all $\mathcal{TEST}_1$ functions $\phi : [0, T] \times \mathcal{P}_2(M) \to \mathbb{R}$ such that $v - \phi$ attains a maximum at (t, μ), we have

$$\partial_t \phi + H(\mu, D_\mu \phi) \geq 0.$$

A function v satisfying $\partial_t v + H(\mu, D_\mu v) \geq 0$ on $[0, T) \times \mathcal{P}_2(M)$ in the viscosity sense is called a *viscosity subsolution* of (4.14).

- Similarly, we say that a function $v : [0, T) \times \mathcal{P}_2(M) \to \mathbb{R}$ satisfies the inequality

$$\partial_t v + H(\mu, D_\mu v) \leq 0,$$

at $(t, \mu) \in [0, T) \times \mathcal{P}_2(M)$ in the viscosity sense if v is lower semicontinuous and for all $\mathcal{TEST}_2$ functions $\phi : [0, T] \times \mathcal{P}_2(M) \to \mathbb{R}$ such that $v - \phi$ attains a minimum at (t, μ), then

$$\partial_t \phi + H(\mu, D_\mu \phi) \leq 0.$$

A function v satisfying $\partial_t v + H(\mu, D_\mu v) \leq 0$ on $[0, T) \times \mathcal{P}_2(M)$ in the viscosity sense is called a *viscosity supersolution* of (4.14).

- We say that a continuous function $v : [0, T] \times \mathcal{P}_2(M) \to \mathbb{R}$ is a *viscosity solution* of (4.14) if it is both a supersolution and a subsolution on $[0, T) \times \mathcal{P}_2(M)$ and verifies

$$v(T, \mu) = L(\mu).$$

Discussion on the notion of viscosity

The notion of viscosity was introduced in [1] to prove well-posedness of Hamilton Jacobi equations in the Euclidean space $\mathbb{R}^N$, where the test functions used were continuously differentiable functions for both the supersolution and subsolution. In this case, one can identify the differential of the test functions with its gradient and the Hamiltonian is assumed to be a continuous mapping from $\mathbb{R}^N \times \mathbb{R}^N$ to $\mathbb{R}$. Shortly after, the notion was extended to any Banach space, denoted V, which possesses the Radon-Nikodym property [28, 29]. The test functions used in this setting were Fréchet differentiable functions for both the supersolution and subsolution. The Fréchet differential belongs to the dual space of V, denoted V^*, and the Hamiltonian is assumed to be a continuous mapping from $V \times V^*$ to $\mathbb{R}$. The notion of viscosity can also be extended to Riemannian manifolds using continuously differentiable functions and the Hamiltonian is assumed to be a continuous mapping from the cotangent bundle to $\mathbb{R}$ [102].
The common features between these definitions are that the state spaces considered in all these examples possess a structure rich enough so that one can assume continuity of the Hamiltonian with respect to both its variables in the topology of the product of the space and its dual space in the case of Banach spaces or the topology

of the cotangent bundle in the case of differentiable manifolds. Furthermore, the Fréchet differentiability/continuous differentiability can be defined in these spaces and Fréchet differentiable/continuously differentiable functions exist "in abundance" in order to use them as test functions. In particular, the squared distance function of the state spaces considered is differentiable. This function is particularly important in viscosity theory because it is used to apply the *variable doubling* technique to obtain the comparison results that guarantee uniqueness of the viscosity solution. Furthermore, one can derive existence of the solution from the comparison results using Perron's method (see for example [3, 4]).

In $(\mathcal{P}_2(M), d_W)$ this approach seems to be less straightforward. On the one hand, $(\mathcal{P}_2(M), d_W)$ is a metric space that does not have any bundle structure that can be exploited to assume continuity of the Hamiltonian on an interesting topology (indeed, the metric cotangent bundle defined in Section 4.3.3 can be endowed with the disjoint union distance which is not very useful). On the other hand, the notion of Fréchet differentiability/continuous differentiability is not well-defined in this space, due to the fact that the structure of the tangent cone at $\mu \in \mathcal{P}_2(M)$ degenerates whenever μ is not a regular measure.

The most known approach to circumvent these difficulties in Wasserstein spaces is through the so-called Lions differentiability [63]. The idea is the following: given a real-valued function of $\mathcal{P}_2(M)$, one considers its "lift" to the space of squared integrable random variables of a probability space equipped with an atomless probability measure (for example, a closed ball of M equipped with the normalized volume form). The lifted function depends on the random variables only through their law in $\mathcal{P}_2(M)$. One then defines Lions differentiable functions in $\mathcal{P}_2(M)$ as the set of functions such that their lift is Fréchet differentiable in the space of squared integrable random variables. This approach was studied in detail in [40] for the space $\mathcal{P}_2(\mathbb{R}^N)$. However, the functions that verify this notion of differentiability are not "abundant" in Wasserstein spaces. For example, the squared Wasserstein distance is not differentiable according to this definition. In fact, it was shown in [103] that the squared Wasserstein distance is differentiable according to this notion at some $\mu \in \mathcal{P}_2(\mathbb{R}^N)$ if and only if μ is a regular measure. This result is not surprising since the pseudo-Riemannian structure degenerates whenever μ is not regular. This presents a major issue for studying Hamilton Jacobi equations in Wasserstein spaces since we can no longer extend viscosity-type techniques (variable doubling, Perron's method) in this setting.

Two possible approaches can be considered to solve this problem. The first approach would consist in restricting the treatment of Hamilton Jacobi equations to the set of regular measures. The difficulty using this method is that the set of regular measures is not locally compact and not geodesically convex as it was shown in [104]. The second approach would be to relax the Lions differentiability condition and look for test functions that would still be differentiable in a suitable sense and exist "in abundance" in $\mathcal{P}_2(M)$. The latter approach is the one adopted in this manuscript.

The notion of differentiability that is most suitable in $\mathcal{P}_2(M)$ is directional differentiability, presented in Section 4.3.3. Indeed, all Lipschitz and DC functions are directionally differentiable at every point. The class of DC functions includes the squared distance function and most of the known functionals in $\mathcal{P}_2(M)$ (the internal energy functional, the potential energy functional, the interaction energy functional, the entropy functional...) [35, Chapter 15]. The test functions chosen in Definition 4.3 constitute a subset of the class of DC functions. More precisely, we choose a subset of semiconcave functions to test subsolutions and a subset of semiconvex functions to test supersolutions. However, this approach comes with a major difficulty which is that the Hamiltonian (4.13) is not continuous in this setting. This is the most delicate part to deal with. Luckily for us, for Hamiltonians of type (4.13), and the test functions chosen in Definition 4.3, we have enough information to guarantee well-posedness of the Hamilton Jacobi equation (4.14) without any further assumptions on the regularity of the Hamiltonian.

Next, we prove a comparison principle that holds for any bounded upper semicontinuous subsolution and any bounded lower semicontinuous supersolution. First, we need two key results.

Proposition 4.3.1. *For all $\sigma, \mu \in \mathcal{P}_2(M)$ and $a > 0$, we have:*

$$H(\mu, a\, D_\mu(d_W^2(\mu,\sigma))) - H(\sigma, -a\, D_\sigma(d_W^2(\mu,\sigma))) \leq 2a\, Lip(f)\, d_W^2(\mu,\sigma).$$

Proof. For any $(x,v) \in TM$, let $\tau_{x,\exp_x(v)}$ be the parallel transport from x to $\exp_x(v)$ along the curve $[0,1] \ni t \to \exp_x(tv)$ (see Appendix 4.5.2). First, since the parallel transport $\tau_{x,\exp_x(v)}$ preserves the Riemannian metric, we have

$$\forall\, (x,v) \in TM,\ \langle f(x,u), v\rangle = \langle \tau_{x,\exp_x(v)}(f(x,u)), \tau_{x,\exp_x(v)}(v)\rangle \text{ and } |\tau_{x,\exp_x(v)}(v)| = |v|.$$

Furthermore, since $f(.,u)$ is Lipschitz, then by Remark 4.2.1 we have

$$\forall\, x \in M,\ \forall v \in T_xM, \quad |\tau_{x,\exp_x(v)}(f(x,u)) - f(\exp_x(v),u)| \leq Lip(f)\, |v|.$$

Thus we get for every $(x,v) \in TM$

$$\begin{aligned}\langle \tau_{x,\exp_x(v)}(f(x,u)), -\tau_{x,\exp_x(v)}(v)\rangle &\leq \langle f(\exp_x(v),u), -\tau_{x,\exp_x(v)}(v)\rangle + Lip(f)|v||\tau_{x,\exp_x(v)}(v)| \\ &= \langle f(\exp_x(v),u), -\tau_{x,\exp_x(v)}(v)\rangle + Lip(f)|v|^2\end{aligned}$$

Let $\sigma, \mu \in \mathcal{P}_2(M)$, $a > 0$ and $\zeta \in \mathbf{exp}_\mu^{-1}(\sigma)$. Then we have

$$\begin{aligned}-\int \langle f(x,u), v\rangle d\zeta(x,v) &= -\int \langle \tau_{x,\exp_x(v)}(f(x,u)), \tau_{x,\exp_x(v)}(v)\rangle d\zeta(x,v) \\ &\leq \int \langle f(\exp_x(v),u), -\tau_{x,\exp_x(v)}(v)\rangle d\zeta(x,v) + Lip(f)\int |v|^2 d\zeta(x,v) \\ &= \int \langle f(\exp_x(v),u), -\tau_{x,\exp_x(v)}(v)\rangle d\zeta(x,v) + Lip(f) d_W^2(\sigma,\mu),\end{aligned}$$

where the last equality holds since $\zeta \in \mathbf{exp}_\mu^{-1}(\sigma)$.
Let $\beta : TM \to TM$ defined for every $(x, v) \in TM$ by

$$\beta(x, v) = (\exp_x(v), -\tau_{x,\exp_x(v)}(v)).$$

Then it comes

$$\begin{aligned}-\int \langle f(x, u), v\rangle d\zeta(x, v) &\leq \int \langle f(\exp_x(v), u), -\tau_{x,\exp_x(v)}(v)\rangle d\zeta(x, v) + Lip(f)d_W^2(\sigma, \mu)\\ &= \int \langle f(x, u), v\rangle d\beta\sharp\zeta(x, v) + Lip(f)d_W^2(\sigma, \mu).\end{aligned}$$

Set $\widetilde{\zeta} = \beta\sharp\zeta$. Notice that we have

$$\pi^M\sharp\widetilde{\zeta} = \exp\sharp\zeta = \sigma, \quad \exp\sharp\widetilde{\zeta} = \pi^M\sharp\zeta = \mu, \quad \int |v|^2 d\widetilde{\zeta}(x, v) = d_W(\mu, \sigma)^2,$$

since

$$\forall\, (x, v) \in TM, \quad \pi^M \circ \beta(x, v) = \exp_x(v), \quad \exp\circ\beta(x, v) = x,$$

and

$$\int |v|^2 d\widetilde{\zeta}(x, v) = \int | - \tau_{x,\exp_x(v)}(v)|^2 d\zeta(x, v) = \int |v|^2 d\zeta(x, v) = d_W(\mu, \sigma)^2.$$

Thus $\widetilde{\zeta} \in \mathbf{exp}_\sigma^{-1}(\mu)$, and therefore it follows from Theorem 4.3 that

$$\begin{aligned}D_\mu d_W^2(\mu, \sigma)\Big(\pi^\mu \circ f(., u)\sharp\mu\Big) &\leq -2\int \langle f(x, u), v\rangle d\zeta(x, v)\\ &\leq 2\int \langle f(x, u), v\rangle d\widetilde{\zeta}(x, v) + 2Lip(f)d_W^2(\sigma, \mu)\\ &\leq -D_\sigma d_W^2(\mu, \sigma)\Big(\pi^\mu \circ f(., u)\sharp\sigma\Big) + 2Lip(f)d_W^2(\sigma, \mu).\end{aligned}$$

By multiplying by a and taking the infimum over $u \in U$, we get the desired result. ∎

Remark 4.4.1. The above result is of fundamental importance to prove the comparison principle. Indeed, it will allow us to use the variable doubling technique without assuming any extra-regularity on the Hamiltonian. Furthermore, the proof can also be adapted if for example the base space is the Euclidean space $\mathbb{R}^N$, rather than the compact manifold M. The reason is that the squared Wasserstein distance in $\mathcal{P}_2(\mathbb{R}^N)$ is a semiconvave function and its directional derivatives have an expression similar to (4.12) (see [60, Theorem 7.3.2 and Proposition 7.3.6]).

Proposition 4.3.2. *Let $\mathcal{O}$ be a subset of a metric space (X, d_X), $\Phi : \mathcal{O} \to \mathbb{R}$ be upper semicontinuous, $\Psi : \mathcal{O} \to \mathbb{R}$ be lower semicontinuous such that $\Psi \geq 0$, and*

$$\Gamma_a = \sup_{\mathcal{O}} \{\, \Phi(x) - a\,\Psi(x)\,\},$$

with $a > 0$. Suppose $-\infty < \lim_{a\to+\infty} M_a < +\infty$ and let $x_a \in \mathcal{O}$ be chosen such that

$$\lim_{a\to+\infty} (\Gamma_a - (\Phi(x_a) - a\,\Psi(x_a))) = 0.$$

Then the following holds:

$$\begin{cases} (i) & \lim_{a\to+\infty} a\,\Psi(x_a) = 0, \\ (ii) & \Psi(\hat{x}) = 0 \text{ and } \lim_{a\to+\infty} \Gamma_a = \Phi(\hat{x}) = \sup_{\{\Psi(x)=0\}} \Phi(x), \\ & \text{whenever } \hat{x} \in \mathcal{O} \text{ is a limit of } (x_a)_a, \text{ as } a \to +\infty. \end{cases}$$

Proof. The proof is exactly the same as in [3, Proposition 3.7], even though it was asserted only for Euclidean spaces. We give here below the proof for the sake of completeness. Let

$$\varepsilon_a = \Gamma_a - (\Phi(x_a) - a\,\Psi(x_a)),$$

so that $\lim_{a\to\infty} \varepsilon_a = 0$. Since $\Psi > 0$, Γ_a decreases as a increases and $\lim_{a\to+\infty} \Gamma_a$ exists and is finite by assumption. Furthermore, we have:

$$\Gamma_{\frac{a}{2}} \geq \Phi(x_a) - \frac{a}{2}\,\Psi(x_a) \geq \Phi(x_a) - a\,\Psi(x_a) + \frac{a}{2}\,\Psi(x_a) = \Gamma_a - \varepsilon_a + \frac{a}{2}\,\Psi(x_a),$$

which implies that $a\,\Psi(x_a) \leq 2\,(\varepsilon_a + \Gamma_{\frac{a}{2}} - \Gamma_a)$, hence $\lim_{a\to+\infty} a\,\Psi(x_a) = 0$.
Suppose now $a_n \to +\infty$ and $x_{a_n} \to \hat{x} \in \mathcal{O}$. Then $\lim_{a_n\to+\infty} \Psi(x_{a_n}) = 0$ and by lower semicontinuity $\Psi(\hat{x}) = 0$. Moreover, since

$$\Phi(x_{a_n}) - a_n\,\Psi(x_{a_n}) = \Gamma_{a_n} - \varepsilon_{a_n} \geq \sup_{\{\Psi(x)=0\}} \Phi(x) - \varepsilon_{a_n},$$

and Φ is upper semicontinuous, the result holds. ■

Remark 4.4.2. Proposition 4.3.2 is a very general statement. It only requires assumptions on the topology of the considered space. Furthermore, this result holds for non locally compact metric spaces.

Theorem 4.4 (Comparison principle)**.** *Assume **(H)** and **(H_ℓ)**. Let $v, w : [0,T] \times \mathcal{P}_2(M) \to \mathbb{R}$ be respectively a bounded upper semicontinuous subsolution and a bounded lower semicontinuous supersolution on $[0,T] \times \mathcal{P}_2(M)$. Then we have:*

$$\sup_{[0,T]\times\mathcal{P}_2(M)} (v - w)_+ \leq \sup_{\{T\}\times\mathcal{P}_2(M)} (v - w)_+,$$

where $(r)_+ = \max(r, 0)$.

Proof. Let $\Gamma := \sup_{[0,T]\times\mathcal{P}_2(M)} (v-w)$. First, by replacing v by $v - \sup_{\{T\}\times\mathcal{P}_2(M)} (v-w)_+$, which is still a subsolution, it suffices to prove that $\Gamma \leq 0$.
By contradiction, suppose that $\Gamma > 0$. Let $0 < \alpha \leq \Gamma$ and let

$$v^\alpha(t,\mu) = v(t,\mu) + \alpha(t-T).$$

v^α is still a subsolution of (4.14). Furthermore, we take α small enough so that

$$\Gamma^\alpha := \sup_{[0,T]\times\mathcal{P}_2(M)} (v^\alpha - w) > 0.$$

We construct test functions the following way:

$$\psi_a(t,s,\mu,\sigma) = v^\alpha(t,\mu) - w(s,\sigma) - \frac{a}{2}(d_W^2(\mu,\sigma)+|t-s|^2).$$

Since v, w are bounded, $v-w$ is upper semicontinuous and $[0,T]\times\mathcal{P}_2(M)$ is compact, then $\Gamma_a = \sup\psi_a$ is reached.
Let (t_a,s_a,μ_a,σ_a) be such that Γ_a is reached. Without loss of generality, we can suppose that (t_a,s_a,μ_a,σ_a) converges and $\lim_{a\to+\infty}\Gamma_a$ exists as $a\to+\infty$ (take a subsequence if necessary). We have

$$\lim_{a\to+\infty}(\Gamma_a - \psi_a(t_a,s_a,\mu_a,\sigma_a)) = 0 \quad \text{and} \quad -\infty < \lim_{a\to+\infty}\Gamma_a < +\infty.$$

Therefore, we can apply Proposition 4.3.2 via the correspondences

$$X = \mathcal{O} = [0,T]^2\times(\mathcal{P}_2(M))^2, \quad \Phi(x) = v^\alpha(t,\mu)-w(s,\sigma), \quad \Psi(x) = \frac{1}{2}\left(d_W^2(\mu,\sigma)+|t-s|^2\right),$$

and we get

$$\begin{cases} (i) & \lim_{a\to+\infty}\frac{a}{2}(d_W^2(\mu_a,\sigma_a)+|t_a-s_a|^2) = 0, \\ & \mu_a,\sigma_a \to \hat\mu \in \mathcal{P}_2(M),\ t_a,s_a \to \hat t \in [0,T],\ \text{as } a\to\infty, \\ (ii) & \lim_{a\to+\infty}\Gamma_a = \psi_a(\hat t,\hat t,\hat\mu,\hat\mu) = \Gamma^\alpha. \end{cases}$$

Hence, when a is big enough, we have $t_a, s_a \notin \{T\}$ since $v^\alpha(\hat t,\hat\mu) - w(\hat t,\hat\mu) > 0$. Then since v is a subsolution and w is a supersolution, we get

$$-\alpha + a(t_a-s_a) + H(\mu_a, \frac{a}{2}D_\mu d_W^2(\mu_a,\sigma_a)) \geq 0 \geq a(t_a-s_a) + H(\sigma_a, -\frac{a}{2}D_\sigma d_W^2(\mu_a,\sigma_a)).$$

Thus we get from Proposition 4.3.1

$$0 \leq -\alpha + H(\mu_a, \frac{a}{2}D_\mu d_W^2(\mu_a,\sigma_a)) - H(\sigma_a, -\frac{a}{2}D_\sigma d_W^2(\mu_a,\sigma_a)) \leq -\alpha + aLip(f)(d_W(\mu_a,\sigma_a))^2.$$

By letting a tend to infinity, we get $\alpha \leq 0$, a contradiction. ■

A similar comparison result was obtained in [37, 38] for similar Hamilton Jacobi equations defined in the Wasserstein space over the Euclidean space. However, it holds only for uniformly continuous subsolutions and supersolutions. Here, with the new definition of viscosity, the comparison principle holds for equation (4.14) for any bounded upper semicontinuous subsolution and bounded lower semicontinuous supersolution. Before proving existence of the solution for equation (4.14), we need the following proposition.

Proposition 4.4.1. *Let $t \mapsto Y_t^{t_0,x_0,u}$ be a trajectory of (4.2). Let $\mu, \sigma \in \mathcal{P}_2(M)$. Then, there exists a subsequence, $(t_n)_n \downarrow t_0$ and a vector field $b(.) \in L^2(\mu, TM)$, such that*

$$b(.) \in \overline{co}\{f(.,u) \ : \ u \in U\},$$

where $\overline{co}$ stands for the closed convex hull of the set of functions $f(.,u)$ with $u \in U$, and verifies

$$\lim_{t_n \downarrow t_0} \frac{\Big(d_W(Y_{t_n}^{t_0,x_0,u}\sharp\mu, \sigma)\Big)^2 - \Big(d_W(\mu,\sigma)\Big)^2}{t_n - t_0} = \lim_{t_n \downarrow t_0} \frac{\Big(d_W\Big(\mathbf{exp}_\mu\Big((t_n - t_0) \centerdot b\sharp\mu\Big), \sigma\Big)\Big)^2 - \Big(d_W(\mu,\sigma)\Big)^2}{t_n - t_0}$$

Proof. First, notice that if such a vector field $b(.) \in L^2(\mu, TM)$ exists, then we have

$$\left|\frac{\Big(d_W(Y_{t_n}^{t_0,x_0,u}\sharp\mu, \sigma)\Big)^2 - \Big(d_W\Big(\mathbf{exp}_\mu\Big((t_n - t_0) \centerdot b\sharp\mu\Big), \sigma\Big)\Big)^2}{t_n - t_0}\right| \leq$$
$$\frac{d_W\Big(Y_{t_n}^{t_0,.,u}\sharp\mu, \mathbf{exp}_\mu((t_n - t_0) \centerdot b)\sharp\mu\Big)}{t_n - t_0}\Big(d_W\Big(\mathbf{exp}_\mu\Big((t_n - t_0) \centerdot b\sharp\mu\Big), \sigma\Big) + d_W\Big(Y_{t_n}^{t_0,x_0,u}\sharp\mu, \sigma\Big)\Big).$$

Hence it suffices to prove that

$$\lim_{t_n \downarrow t_0} \frac{d_W\Big(Y_{t_n}^{t_0,.,u}\sharp\mu, \mathbf{exp}_\mu((t_n - t_0) \centerdot b\sharp\mu)\Big)}{t_n - t_0} = 0.$$

By Nash embedding theorem, we can assume that M is isometrically embedded into a Euclidean space $(\mathbb{R}^N, ||.||)$ with $N > 0$ big enough. We have

$$Y_t^{t_0,x_0,u} = x_0 + \int_{t_0}^t f(Y_s^{t_0,x_0,u}, u(s))ds,$$

and the quantity

$$x_0 \mapsto \frac{1}{t - t_0} \int_{t_0}^t f(Y_s^{t_0,x_0,u}, u(s))ds$$

is uniformly bounded independently of t and x_0. Let $(t_n)_n \downarrow t_0$, and let $b_n(.)$ be the sequence of functions defined as

$$\forall x_0 \in M, \quad b_n(x_0) := \frac{1}{t_n - t_0} \int_{t_0}^{t_n} f(Y_s^{t_0,x_0,u}, u(s))ds.$$

The sequence $(b_n(.))_n$ is uniformly bounded. Furthermore, it is equiLipschitz. Indeed we have

$$\begin{aligned}\forall x_0, y_0 \in M, \quad ||b_n(x_0) - b_n(y_0)|| &\leq \frac{1}{t_n - t_0} \int_{t_0}^{t_n} ||f(Y_s^{t_0,x_0,u}, u(s)) - f(Y_s^{t_0,y_0,u}, u(s))||ds \\ &\leq C_1 Lip(f) d(x_0, y_0),\end{aligned}$$

where C_1 is the constant from Proposition 4.0.1. Hence, by Arzelà–Ascoli Theorem, there exists a subsequence of $(t_n)_n$ (not relabelled here) and a function $b(.)$ such that

$$\forall x_0 \in M, \quad b_n(x_0) \to b(x_0), \quad \text{as } n \text{ tends to infinity.}$$

Moreover, $b(.) \in L^2(\mu, TM)$ since it is the pointwise limit of measurable and uniformly bounded functions. On the other hand, there exists $(\varepsilon_n) \downarrow 0$ such that

$$b_n(.) \in \overline{co}\Big(\bigcup_{|t_0 - s| \leq \varepsilon_n} \{f(Y_s^{t_0,.,u}, u(s))\}\Big).$$

Hence

$$b(.) \in \overline{co}\{f(., u) \ : \ u \in U\}.$$

Consider the curve $t \mapsto \exp_{x_0}((t - t_0)b(x_0))$. For any $x_0 \in M$, we have

$$|| \exp_{x_0}((t - t_0)b(x_0)) - (x_0 + (t - t_0)b(x_0))|| = o(|t - t_0|),$$

since the two curves are smooth and have the same position and velocity at t_0. Then, we get

$$\lim_{t_n \downarrow t_0} \frac{1}{t_n - t_0} \Big|\Big| Y_{t_n}^{t_0,x_0,u} - \exp_{x_0}((t_n - t_0)b(x_0)) \Big|\Big| = \lim_{t_n \downarrow t_0} \Big|\Big| \frac{1}{t_n - t_0} \int_{t_0}^{t_n} f(Y_s^{t_0,x_0,u}, u(s))ds - b(x_0) \Big|\Big| = 0.$$

On the other hand, since Nash embedding is biLipschitz, we get

$$\lim_{t_n \downarrow t_0} \frac{1}{t_n - t_0} d\Big(Y_{t_n}^{t_0,x_0,u}, \exp_{x_0}((t_n - t_0)b(x_0))\Big) = \lim_{t_n \downarrow t_0} \frac{1}{t_n - t_0} \Big|\Big| Y_{t_n}^{t_0,x_0,u} - \exp_{x_0}((t_n - t_0)b(x_0)) \Big|\Big| = 0.$$

Thus we obtain

$$\lim_{t_n \downarrow t_0} \frac{1}{(t_n - t_0)^2} d_W^2\Big(Y_{t_n}^{t_0,.,u} \sharp\mu, \mathbf{exp}_\mu((t_n - t_0) \centerdot b\sharp\mu)\Big) \leq \lim_{t_n \downarrow t_0} \frac{1}{(t_n - t_0)^2} \int d^2\Big(Y_{t_n}^{t_0,x_0,u}, \exp_{x_0}((t_n - t_0)b(x_0))\Big) d\mu(x_0) = 0,$$

by dominated convergence, which implies the result. ■

Theorem 4.5. *Assume* $(\boldsymbol{H})$, $(\boldsymbol{H_\ell})$ *and* $(\boldsymbol{H_{co}})$. *Then the value function* ϑ *is the unique continuous viscosity solution to* (4.14).

Proof. First we prove that ϑ is a supersolution. Let $\phi \in \mathcal{TEST}_2$, such that $\vartheta - \phi$ attains a minimum at $(t_0, \mu_0) \in [0,T) \times \mathcal{P}_2(M)$.
So there exists, $(a,\sigma) \in \mathbb{R}^- \times \mathcal{P}_2(M)$ and $\psi(.) \in C^2([0,T],\mathbb{R})$ such that

$$\phi(t,\mu) = \psi(t) + a\, d_W^2(\mu,\sigma),$$

and

$$\forall (t,\mu) \in [0,T) \times \mathcal{P}_2(M),\ \phi(t,\mu) - \phi(t_0,\mu_0) \le \vartheta(t,\mu) - \vartheta(t_0,\mu_0).$$

Let $t \mapsto Y_t^{t_0,x,u}$ be a trajectory of (4.2) such that $\vartheta(t_0,\mu_0) = \vartheta(t_0+h, Y_{t_0+h}^{t_0,.,u}\sharp\mu)$. So we get for all $h \in [0, T-t_0)$,

$$\phi(t_0+h, Y_{t_0+h}^{t_0,.,u}\sharp\mu_0) - \phi(t_0,\mu_0) \le \vartheta(t_0+h, Y_{t_0+h}^{t_0,.,u}\sharp\mu_0) - \vartheta(t_0,\mu_0) \le 0.$$

Thus along a subsequence $(h_n)_n \to 0$, by dividing by h_n and letting h_n tend to 0, we get by Proposition 4.4.1 and Theorem 4.3,

$$\begin{aligned}\partial_t\phi + \inf_{u\in U} D_\mu\phi\Big(\pi^\mu \circ f(.,u)\sharp\mu_0\Big) &= \partial_t\phi + \inf_{b(.)\in\overline{co}\{f(.,u)\}} D_\mu\phi\Big(\pi^\mu \circ b\sharp\mu_0\Big)\\ &\le \partial_t\phi + D_\mu\phi\Big(\pi^\mu\circ b\sharp\mu_0\Big)\\ &\le 0,\end{aligned}$$

where the first equality is obtained by Hypothesis $(\boldsymbol{H_{co}})$.

To prove that ϑ is a subsolution, let $\phi \in \mathcal{TEST}_1$, such that $\vartheta - \phi$ attains a maximum at $(t_0,\mu_0) \in [0,T) \times \mathcal{P}_2(M)$. So there exists $(a,\sigma) \in \mathbb{R}^+ \times \mathcal{P}_2(M)$ and $\psi(.) \in C^2([0,T],\mathbb{R})$ such that

$$\phi(t,\mu) = \psi(t) + a\, d_W^2(\mu,\sigma),$$

and

$$\forall (t,\mu) \in [0,T) \times \mathcal{P}_2(M),\ \phi(t,\mu) - \phi(t_0,\mu_0) \ge \vartheta(t,\mu) - \vartheta(t_0,\mu_0).$$

Let $t \mapsto Y_t^{t_0,x,u}$ be a trajectory that is a solution to the controlled system (4.4) with constant control $u \in U$. So we get for all $h \in [0, T-t_0)$,

$$\phi(t_0+h, Y_{t_0+h}^{t_0,.,u}\sharp\mu_0) - \phi(t_0,\mu_0) \ge \vartheta(t_0+h, Y_{t_0+h}^{t_0,.,u}\sharp\mu_0) - \vartheta(t_0,\mu_0) \ge 0.$$

On the other hand, by the same reasoning as in Proposition 4.4.1, we get

$$\lim_{h\downarrow 0} \frac{\Big(d_W(Y_{t_0+h}^{t_0,x_0,u}\sharp\mu,\sigma)\Big)^2 - \Big(d_W(\mu,\sigma)\Big)^2}{h} = \lim_{h\downarrow 0} \frac{\Big(d_W\Big(\mathbf{exp}_\mu\Big(h \centerdot (f(.,u)\sharp\mu)\Big),\sigma\Big)\Big)^2 - \Big(d_W(\mu,\sigma)\Big)^2}{h}.$$

Therefore, by dividing by h and letting h tend to 0, we get by Theorem 4.3

$$\partial_t\phi + D_\mu\phi\Big(\pi^\mu \circ f(.,u)\sharp\mu_0\Big) \ge 0.$$

By taking the infimum over $u \in U$, we get the result.

Finally, the final condition of (4.14) is trivially verified by ϑ. Hence, the value function ϑ is a continuous bounded solution to (4.14) and it is unique by Theorem 4.4. ■

4.5 Appendices

4.5.1 Disintegration theorem

We recall here the disintegration theorem. For more details, we refer to [60, Theorem 5.3.1].

Theorem 4.6. *Let X, Y be two Polish spaces (i.e. complete and separable metric spaces), $\mu \in \mathcal{P}(X)$, let $r : X \to Y$ be a Borel map and let $\nu = r\sharp\mu \in \mathcal{P}(Y)$. Then, there exists a $\nu-$a.e. uniquely determined Borel family of probability measures $\{\mu_y\}_{y \in Y} \subset \mathcal{P}(X)$ such that:*

$$\mu_y(X \setminus r^{-1}(y)) = 0, \quad \textit{for } \nu\textit{−a.e. } y \in Y,$$

and

$$\int_X f(x)\, d\mu(x) = \int_Y \left(\int_{r^{-1}(y)} f(x)\, d\mu_y(x) \right) d\nu(y), \quad \textit{for every Borel map } f : X \to [0, +\infty].$$

4.5.2 Riemannian manifolds

We recall some standard notions of Riemannian geometry. Classical references are for example [105, 106]. We consider a connected differentiable manifold M with empty boundary endowed with a Riemannian metric $\langle ., . \rangle$ and we assume that $(M, \langle ., . \rangle)$ is a complete Riemaniann manifold. Let $d(., .)$ be the Riemannian distance on $(M, \langle ., . \rangle)$. The metric space (M, d) is a complete space and its topology is equivalent to the topology of the manifold M. For any $x \in M$, we denote by T_xM the tangent space of M at x, by $TM := \cup_{x \in M} \{x\} \times T_xM$ the tangent bundle and by $\pi^M : TM \to M$ the canonical projection. Let ∇ be the Levi-Civita connection associated to $(M, \langle ., . \rangle)$. A vector field $V : M \to TM$ is a mapping such that

$$\pi^M \circ V(x) = x, \quad \forall x \in M.$$

Let $\alpha : [a, b] \to M$ be a smooth curve. The connection ∇ induces a linear isometry between $T_{\alpha(a)}M$ and $T_{\alpha(t)}M$, for all $t \in [a, b]$. More precisely, for all $v \in T_{\alpha(a)}$, there exists a unique vector field V along α, satisfying

$$\nabla_{\dot\alpha(t)} V(\alpha(t)) = 0, \quad \forall t \in [a, b], \text{ and } V(\alpha(a)) = v.$$

The resulting isometry, called the *parallel transport* along α from $\alpha(a)$ to $\alpha(b)$, and denoted by $\tau^{\alpha}_{\alpha(a),\alpha(b)}$ is defined by

$$\tau^{\alpha}_{\alpha(a),\alpha(b)}(v) = V(\alpha(b)), \quad \forall\, v \in T_{\alpha(a)}M.$$

There holds that $\tau^{\alpha}_{\alpha(b_1),\alpha(b_2)} \circ \tau^{\alpha}_{\alpha(a),\alpha(b_1)} = \tau^{\alpha}_{\alpha(a),\alpha(b_2)}$ and $(\tau^{\alpha}_{\alpha(a),\alpha(b)})^{-1} = \tau^{\alpha}_{\alpha(b),\alpha(a)}$. For convenience, we will drop the superscript α, whenever it is clear from the context which curve α is used.

Let $\exp : TM \to M$ be the exponential map. For every $x \in M$, the function exp maps straight lines of T_xM, $x \in M$, passing through $0_x \in T_xM$ to geodesics of M passing through x. Since $(M, \langle .,. \rangle)$ is supposed to be complete, it is a consequence of Hopf-Rinow theorem, that the exponential map is defined on all the tangent bundle. However it may not be a diffeomorphism.

The tangent bundle TM is itself a complete Riemannian manifold when endowed with the Sasaki metric [97]. The Riemannian distance d_{TM} on TM associated with the Sasaki metric is defined by

$$\forall (u,v) \in TM \times TM, \quad d^2_{TM}(u,v) := \inf \{\, (\text{length}(\alpha))^2 + |\tau^{\alpha}_{\pi^M(u),\pi^M(v)}(u) - v|^2 \,\},$$

where the infimum is taken over all smooth curves $\alpha : [0,1] \to M$ connecting $\pi^M(u)$ and $\pi^M(v)$ and its length is defined by

$$\text{length}(\alpha) := \int_0^1 \sqrt{\langle \dot\alpha(t), \dot\alpha(t) \rangle}\, dt = \int_0^1 |\dot\alpha(t)|\, dt,$$

where $|.|$ is the norm associated to the Riemannian metric $\langle .,. \rangle$ on the tangent bundle TM.

Chapter 5

Conclusions and future directions

Let us conclude this work by summarizing the main contributions that have been displayed in this thesis and by mentioning some future directions that we think are interesting to investigate.

Conclusions

In Chapter 2, we studied a Hamilton Jacobi Bellman equation associated to a stratified domain Mayer optimal control problem. We introduced a viscosity notion adapted to the stratification where test functions are continuously differentiable on the closure of each subdomain of the stratification. With this notion of viscosity, we proved that the value function is the unique viscosity solution of the discontinuous Hamilton Jacobi Bellman equation. The proof is based on the dynamic programming principle verified by the value function and on the invariance properties proved in this discontinuous setting. Furthermore, the invariance properties give us a comparison principle valid for any upper semicontinuous subsolution and any lower semicontinuous supersolution. Moreover, we proved some stability results in the presence of perturbations on the Hamiltonians on each domain. Finally, the comparison principle combined with the new notion of viscosity allowed us to prove a general convergence result for monotone numerical schemes approximating the Hamilton Jacobi Bellman equation in this setting, which generalizes the classical result due to Barles and Souganidis [47].

The invariance properties were proven by considering the *essential dynamics* of the stratified setting first introduced by Barnard and Wolenski [41]. Barnard and Wolenski along with the work of Rao and Zidani [18] provided us with the intuition to solve this problem. The proof of the strong invariance property given in [41] needed further investigation, which was the limiting factor in [18] to prove a strong comparison principle similar to the one we provided here. We improved on the work of Barnard and Wolenski by assuming further that the essential dynamics are lower semicontinuous, an assumption that was not present in their paper. Furthermore, we introduced a different notion of viscosity than the one considered by Rao and

Zidani which turned out to be more adequate in our setting. Other key assumptions that play an important part in this setting are proximal smoothness and relative wedgeness hypotheses on the closure of the domains. These hypotheses, though they are necessary for technical reasons, can be sharpened more. We will discuss these hypotheses through the framework of Chapter 3

Chapter 3 of this work was devoted to develop a first order viscosity theory on proper CAT(0) spaces. CAT(0) spaces are geodesic spaces with upper curvature bound equal to 0 in the sense of Alexandrov. CAT(0) spaces enjoy a rigid structure that allows for a first order calculus to be possible on them. In particular, a notion of tangent cone is well defined at every point and it has a structure resembling that of a Hilbert space. Furthermore, a notion of differential exists for Lipschitz and DC functions. We defined the notion of viscosity using test functions that are Lipschitz and DC. More precisely, we test subsolutions with Lipschitz semiconvex functions and we test supersolutions with Lipschitz semiconcave functions. Particular cases of Lipschitz and DC functions are the squared distance function and the distance function to a closed convex subset. Under mild assumptions on the Hamiltonian considered, we proved the comparison principle using the variable doubling technique in the exact same manner as in the classical theory of viscosity. Furthermore, we derived existence of the viscosity solution using Perron's method in a similar way as in the classical theory. Finally, we showed through several examples that this framework not only extends to a wide class of metric spaces useful in several real-world applications, but it also applies to the Euclidean space and manifolds of nonpositive sectional curvature.

These results are, to our point of view, are the tip of an iceberg consisting of a theory of viscosity in geodesic spaces of one curvature bound in the sense of Alexandrov. In this chapter, we only gave some elements of this idea for spaces with upper bound curvature equal to 0 in the sense of Alexandrov. Moreover, in the last chapter, we treated a particular case of a geodesic space with the flavor of lower curvature bound in the sense of Alexandrov, which is the Wasserstein space over a compact Riemannian manifold.

In Chapter 4, we studied well posedness of a Hamilton Jacobi Bellman equation coming from an optimal control problem defined in the Wasserstein space over a compact Riemannian manifold. We started by exploiting the formal Riemannian structure that the Wasserstein space enjoys. In particular, the continuity equation and the space of gradient vector fields allow to define the optimal control problem on the Wasserstein space. Furthemore, we explained the intricacies between the formal Riemannian structure and the the "2-uniform structure" of the Wasserstein space, which can be seen as a generalization of the lower curvature bound in the sense of Alexandrov. In particular, we explained that the tangent cone in the sense of metric geometry is well defined and the space of gradient vector fields is always a subset of the tangent cone. Furthermore, we explained that metric geometry tools clarify where the formal Riemannian structure is a good approximation of the Wasserstein space and where it degenerates. We explained that this distinction is important since we wanted to define a notion of viscosity that is robust enough so that viscosity-type techniques could be transposed to Wasserstein spaces.

We defined the value function associated to the optimal control problem and we proved that it is Lipschitz continuous under standard assumptions on the dynamics of the system. Furthermore, we proved that the value function verifies a dynamic programming principle which allows us to assert that it is a viscosity solution in a suitable viscosity sense. Finally, we defined the notion of viscosity using Lipschitz and DC function. In particular, we test subsolutions with functions that are Lipschitz and semiconcave and we test supersolutions with Lipschitz and semiconvex functions. The uniqueness of the solution is obtained thanks to a comparison principle that we proved to hold for any upper semicontinuous subsolution and any lower smeicontinuous supersolution. We believe that the example treated in this chapter, is a simple example for the notion of viscosity that we introduced and it can be applied to a more general class of Hamilton Jacobi equations posed in Wasserstein spaces.

Future directions

Let us mention here some open questions that can be seen as a natural continuation of the work done in this thesis.

- **Extension of the viscosity notion to CAT(κ) spaces.** The proofs given in Chapter 3 for viscosity theory on proper CAT(0) spaces are all local in nature. Hence, they could be extended, in a rather straightforward manner, to any proper CAT(κ) space for any $\kappa \in \mathbb{R}$. Interesting examples of proper CAT(κ) spaces are ramified spaces where Camilli, Marchi and Schieborn studied the Eikonal equation in them [21] and proximally smooth subsets of Euclidean spaces or smooth manifolds with curvature bounded from above, as proven by Lytchak in [107]. Furthermore, in Chapter 2, we assumed proximal smoothness for the closure of the domains of the stratification. It would be interesting to investigate whether the results obtained in Chapter 2 could be otained using purely the viscosity notion introduced in the setting of Chapter 3. Moreover, we conjecture that viscosity techniques would not need the relative wedgeness assumption. Indeed, although it was necessary for technical reasons related to nonsmooth analysis tools, we could not find an example of a proximally smooth set that does not automatically verify the relative wedgeness assumption. The relative wedgeness assumption seems to compensate for the lack of a general optimal control theory in CAT(κ) spaces.

- **An optimal control theory in CAT(κ) spaces.** One of the most challenging questions that has not been addressed in these pages concerns the optimal control interpretation of the examples of Hamilton Jacobi equations given in Chapter 3. A first step in this direction would be to develop a theory of controlled gradient flows in CAT(0) spaces. A recent preprint by Conforti, Kraaij and Tonon [108] investigates a general comparison principle for Hamilton Jacobi equations associated to a linearly controlled gradient flow system posed in a general metric space. The author conjectures that since CAT(κ) spaces enjoy more structure than a general metric space, one might be able to recover stronger results.

- **Extension of the viscosity notion in $\mathcal{P}(M)$ to more general Hamiltonians.** In Chapter 4, we introduced a notion of viscosity in the Wasserstein space over a compact Riemannian manifold to study well posedness of a Hamilton Jacobi Bellman equation associated to a simple model of a multi-agent optimal control problem where the non local interactions between the agents are not considered. The extension to more general Hamilton Jacobi equations coming from multi-agent optimal control problems is the natural next step of the work presented here. Furthermore, an interesting question would be to compare the various notions of viscosity that exist in the literature with the current one proposed in this manuscript. Moreover, it would be interesting to apply the notion given here to the first order master equation coming from mean field games and to compare it to the results obtained by Cardaliaguet, Delarue, Lasry and Lions in [39].

Bibliography

[1] M. G. Crandall and P.-L. Lions, "Viscosity solutions of Hamilton-Jacobi equations," *Transactions of the American mathematical society*, vol. 277, no. 1, pp. 1–42, 1983.

[2] ——, "Condition d'unicité pour les solutions généralisées des équations de Hamilton-Jacobi du premier ordre," *C. R. Acad. Sci., Paris, Sér. I*, vol. 292, pp. 183–186, 1981.

[3] M. G. Crandall, H. Ishii, and P.-L. Lions, "User's guide to viscosity solutions of second order partial differential equations," *Bulletin of the American mathematical society*, vol. 27, no. 1, pp. 1–67, 1992.

[4] G. Barles, "Solutions de viscosité des équations de Hamilton-Jacobi," *Collection SMAI*, 1994.

[5] W. Fleming and H. Soner, "Controlled markov processes and viscosity solutions springer verlag, new york," *Search in*, 1993.

[6] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations.* Springer Science & Business Media, 2008.

[7] H. Ishii, "Hamilton-jacobi equations with discontinuous hamiltonians on arbitrary open sets," *Bull. Fac. Sci. Eng. Chuo Univ*, vol. 28, no. 28, p. 1985, 1985.

[8] E. Barron and R. Jensen, "Semicontinuous viscosity solutions for hamilton–jacobi equations with convex hamiltonians," *Communications in Partial Differential Equations*, vol. 15, no. 12, pp. 293–309, 1990.

[9] F. Clarke, *Functional analysis, calculus of variations and optimal control.* Springer Science & Business Media, 2013, vol. 264.

[10] F. H. Clarke, *Optimization and nonsmooth analysis.* Siam, 1990, vol. 5.

[11] J.-P. Aubin and H. Frankowska, *Set-valued analysis.* Springer Science & Business Media, 2009.

[12] R. B. Vinter, *Optimal control.* Springer, 2010.

[13] H. Frankowska, "Lower semicontinuous solutions of Hamilton–Jacobi–Bellman equations," *SIAM Journal on Control and Optimization*, vol. 31, no. 1, pp. 257–272, 1993.

[14] A. Bressan and Y. Hong, "Optimal control problems on stratified domains," *NHM*, vol. 2, no. 2, pp. 313–331, 2007.

[15] G. Barles, A. Briani, and E. Chasseigne, "A Bellman approach for two-domains

optimal control problems in $\mathbb{R}^n$," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 19, no. 3, pp. 710–739, 2013.

[16] ——, "A Bellman approach for regional optimal control problems in $\mathbb{R}^n$," *SIAM Journal on Control and Optimization*, vol. 52, no. 3, pp. 1712–1744, 2014.

[17] G. Barles and E. Chasseigne, "(Almost) everything you always wanted to know about deterministic control problems in stratified domains," *arXiv preprint arXiv:1412.7556*, 2014.

[18] Z. Rao and H. Zidani, "Hamilton–Jacobi–Bellman equations on multi-domains," in *Control and Optimization with PDE Constraints*. Springer, 2013, pp. 93–116.

[19] Z. Rao, A. Siconolfi, and H. Zidani, "Transmission conditions on interfaces for Hamilton–Jacobi–Bellman equations," *Journal of Differential Equations*, vol. 257, no. 11, pp. 3978–4014, 2014.

[20] D. Schieborn, "Viscosity solutions of Hamilton-Jacobi equations of eikonal type on ramified spaces," Ph.D. dissertation, Citeseer, 2006.

[21] F. Camilli, D. Schieborn, and C. Marchi, "Eikonal equations on ramified spaces," *Interfaces and Free Boundaries*, vol. 15, no. 1, pp. 121–140, 2013.

[22] C. Imbert and R. Monneau, "Flux-limited solutions for quasi-convex Hamilton-Jacobi equations on networks," *Annales scientifiques de l'ENS*, vol. 50, no. 2, 2017.

[23] ——, "The vertex test function for Hamilton-Jacobi equations on networks," Jun. 2013, 52 pages. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00832545

[24] C. Imbert, R. Monneau, and H. Zidani, "A Hamilton-Jacobi approach to junction problems and application to traffic flows," *ESAIM: Control, Optimisation, and Calculus of Variations.*, vol. 19(01), pp. 129–166, 2013.

[25] Y. Achdou, F. Camilli, A. Cutrì, and N. Tchou, "Hamilton–Jacobi equations constrained on networks," *Nonlinear Differential Equations and Applications NoDEA*, vol. 20, no. 3, pp. 413–445, 2013.

[26] Y. Achdou and N. Tchou, "Hamilton-Jacobi equations on networks as limits of singularly perturbed problems in optimal control: dimension reduction," *Communications in Partial Differential Equations*, vol. 40, no. 4, pp. 652–693, 2015.

[27] P.-L. Lions and P. Souganidis, "Well posedness for multi-dimensional junction problems with Kirchoff-type conditions," *Rendiconti Lincei - Matematica e Applicazioni*, vol. 28, 04 2017.

[28] M. G. Crandall and P.-L. Lions, "Hamilton-Jacobi equations in infinite dimensions I. Uniqueness of viscosity solutions," *Journal of functional analysis*, vol. 62, no. 3, pp. 379–396, 1985.

[29] ——, "Hamilton-Jacobi equations in infinite dimensions. II. Existence of viscosity solutions," *Journal of Functional Analysis*, vol. 65, no. 3, pp. 368–405, 1986.

[30] ——, "Hamilton-Jacobi equations in infinite dimensions, III," *Journal of functional analysis*, vol. 68, no. 2, pp. 214–247, 1986.

[31] ——, "Viscosity solutions of Hamilton-Jacobi equations in infinite dimensions. IV. hamiltonians with unbounded linear terms," *Journal of Functional Analysis*, vol. 90, no. 2, pp. 237–283, 1990.

[32] Y. Giga, N. Hamamuki, and A. Nakayasu, "Eikonal equations in metric spaces," *Transactions of the American Mathematical Society*, vol. 367, no. 1, pp. 49–66, 2015.

[33] W. Gangbo and A. Święch, "Metric viscosity solutions of hamilton–jacobi equations depending on local slopes," *Calculus of Variations and Partial Differential Equations*, vol. 54, no. 1, pp. 1183–1218, 2015.

[34] L. Ambrosio and J. Feng, "On a class of first order Hamilton–Jacobi equations in metric spaces," *Journal of Differential Equations*, vol. 256, no. 7, pp. 2194–2245, 2014.

[35] C. Villani, *Optimal transport: old and new.* Springer, 2009, vol. 338.

[36] A. Marigonda and M. Quincampoix, "Mayer control problem with probabilistic uncertainty on initial positions," *Journal of Differential Equations*, vol. 264, no. 5, pp. 3212–3252, 2018.

[37] P. Cardaliaguet and M. Quincampoix, "Deterministic differential games under probability knowledge of initial condition," *International Game Theory Review*, vol. 10, no. 01, pp. 1–16, 2008.

[38] C. Jimenez, A. Marigonda, and M. Quincampoix, "Optimal control of multi-agent systems in the Wasserstein space," *Calculus of Variations and Partial Differential Equations*, vol. 59, no. 2, pp. 1–45, 2020.

[39] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions, *The Master Equation and the Convergence Problem in Mean Field Games.* Princeton: Princeton University Press, 2019. [Online]. Available: https://doi.org/10.1515/9780691193717

[40] W. Gangbo and A. Tudorascu, "On differentiability in the Wasserstein space and well-posedness for Hamilton–Jacobi equations," *Journal de Mathématiques Pures et Appliquées*, vol. 125, pp. 119–174, 2019.

[41] R. C. Barnard and P. R. Wolenski, "Flow invariance on stratified domains," *Set-Valued and Variational Analysis*, vol. 21, no. 2, pp. 377–403, 2013.

[42] F. Camilli and A. Siconolfi, "Hamilton-jacobi equations with measurable dependence on the state variable," *Advances in Differential Equations*, vol. 8, no. 6, pp. 733–768, 2003.

[43] G. Barles and E. Chasseigne, "An illustrated guide of the modern approches of Hamilton-Jacobi equations and control problems with discontinuities," *arXiv preprint arXiv:1812.09197*, 2018.

[44] J. Guerand and M. Koumaiha, "Error estimates for a finite difference scheme associated with Hamilton–Jacobi equations on a junction," *Numerische Mathematik*, vol. 142, no. 3, pp. 525–575, 2019.

[45] E. Carlini, A. Festa, and N. Forcadel, "A Semi-Lagrangian scheme

for Hamilton–Jacobi–Bellman equations on networks," *SIAM Journal on Numerical Analysis*, vol. 58, no. 6, pp. 3165–3196, 2020. [Online]. Available: https://doi.org/10.1137/19M1260931

[46] P. S. Morfe, "Convergence & rates for Hamilton–Jacobi equations with kirchoff junction conditions," *Nonlinear Differential Equations and Applications NoDEA*, vol. 27, no. 1, pp. 1–69, 2020.

[47] G. Barles and P. E. Souganidis, "Convergence of approximation schemes for fully nonlinear second order equations," *Asymptotic analysis*, vol. 4, no. 3, pp. 271–283, 1991.

[48] A. D. Alexandrov, "Uber eine verallgemeinerung der Riemannscen geometrie," *Schriftenreiche der Institut fur Mathematik*, vol. 1, pp. 33–84, 1957.

[49] W. Ballmann, "Manifolds of non positive curvature," in *Arbeitstagung Bonn 1984*. Springer, 1985, pp. 261–268.

[50] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*. Springer Science & Business Media, 2013, vol. 319.

[51] S. Alexander, V. Kapovitch, and A. Petrunin, "Alexandrov geometry: foundations," *arXiv preprint arXiv:1903.08539*, 2022.

[52] D. Burago, I. D. Burago, Y. Burago, S. Ivanov, S. V. Ivanov, and S. A. Ivanov, *A course in metric geometry*. American Mathematical Soc., 2001, vol. 33.

[53] M. Bacák, "Convex analysis and optimization in Hadamard spaces," in *Convex Analysis and Optimization in Hadamard Spaces*. de Gruyter, 2014.

[54] G. Perel'man, "DC structure on Alexandrov space," *Preprint*, 1994.

[55] L. Ambrosio and J. Bertrand, "DC calculus," *Mathematische Zeitschrift*, vol. 288, no. 3, pp. 1037–1080, 2018.

[56] A. Petrunin, "Semiconcave functions in alexandrov's geometry," *Surveys in Differential Geometry*, vol. 11, 03 2013.

[57] A. Lytchak, "Open map theorem for metric spaces," *St. Petersburg Mathematical Journal*, vol. 17, no. 3, pp. 477–491, 2006.

[58] S.-I. Ohta, "Gradient flows on Wasserstein spaces over compact Alexandrov spaces," *American journal of mathematics*, vol. 131, no. 2, pp. 475–516, 2009.

[59] S.-I. Ohta and M. Pálfia, "Gradient flows and a trotter-kato formula of semiconvex functions on CAT(1)-spaces," *American Journal of Mathematics*, vol. 139, no. 4, pp. 937–965, 2017.

[60] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.

[61] U. Mayer, "Gradient flows on nonpositively curved metric spaces and harmonic maps," *Communications in Analysis and Geometry*, vol. 6, pp. 199–253, 1998.

[62] C. Jimenez, A. Marigonda, and M. Quincampoix, "Dynamical systems and Hamilton-Jacobi-Bellman equations on the Wasserstein space and their L^2 representations," 2022, cvgmt preprint. [Online]. Available: http://cvgmt.sns.it/paper/5584/

[63] P.-L. Lions, "Cours au Collège de France," *Available at www.college-de-france.fr*, 2007.

[64] R. J. McCann, *A convexity theory for interacting gases and equilibrium crystals.* Princeton University, 1994.
[65] F. Otto, "The geometry of dissipative evolution equations: the porous medium equation," 2001.
[66] J. Lott and C. Villani, "Ricci curvature for metric-measure spaces via optimal transport," *Annals of Mathematics*, pp. 903–991, 2009.
[67] K.-T. Sturm, "On the geometry of metric measure spaces," *Acta mathematica*, vol. 196, no. 1, pp. 65–131, 2006.
[68] N. Gigli, "On the inverse implication of Brenier-McCann theorems and the structure of $(\mathcal{P}_2(M), W_2)$," *Methods and Applications of Analysis*, vol. 18, no. 2, pp. 127–158, 2011.
[69] H. Ishii, "A boundary value problem of the dirichlet type for Hamilton-Jacobi equations," *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, vol. 16, no. 1, pp. 105–135, 1989.
[70] G. Barles and E. Chasseigne, "(almost) everything you always wanted to know about deterministic control problems in stratified domains," *Networks and Heterogeneous Media*, vol. 10, 12 2014.
[71] D. Ghilli, Z. Rao, and H. Zidani, "Junction conditions for finite horizon optimal control problems on multi-domains with continuous and discontinuous solutions," *ESAIM: Control, Optimisation and Calculus of Variations*, 07 2017.
[72] F. Camilli and C. Marchi, "A comparison among various notions of viscosity solution for Hamilton–Jacobi equations on networks," *Journal of Mathematical Analysis and applications*, vol. 407, no. 1, pp. 112–118, 2013.
[73] F. H. Clarke, R. Stern, and P. Wolenski, "Proximal smoothness and the lower-C^2 property," *J. Convex Anal*, vol. 2, no. 1-2, pp. 117–144, 1995.
[74] A. F. Filippov, "Differential equations with discontinuous right-hand side," *Amer. Math. Soc. Trans.*, vol. 42, pp. 199–231, 1964.
[75] L. Evans and A. M. Society, *Partial Differential Equations*, ser. Graduate studies in mathematics. American Mathematical Society, 1998. [Online]. Available: https://books.google.fr/books?id=5Pv4LVB_m8AC
[76] J.-P. Aubin and A. Cellina, *Differential inclusions: set-valued maps and viability theory.* Springer Science & Business Media, 2012, vol. 264.
[77] F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth analysis and control theory.* Springer Science & Business Media, 2008, vol. 178.
[78] A. Lechicki and A. Spakowski, "A note on intersection of lower semicontinuous multifunctions," *Proceedings of the American Mathematical Society*, vol. 95, no. 1, pp. 119–122, 1985.
[79] V. Barbu and G. Da Prato, "Hamilton-jacobi equations in Hilbert spaces; variational and semigroup approach," *Annali di Matematica pura ed applicata*, vol. 142, no. 1, pp. 303–349, 1985.
[80] A. Nakayasu, "On metric viscosity solutions for Hamilton-Jacobi equations of evolution type," *arXiv: Analysis of PDEs*, 2014.

[81] W. Gangbo and A. Święch, "Optimal transport and large number of particles," *Discrete & Continuous Dynamical Systems*, vol. 34, no. 4, p. 1397, 2014.

[82] L. Qing, S. Nageswari, and Z. Xiaodan, "Equivalence of solutions of Eikonal equation in metric spaces," *Journal of Differential Equations*, vol. 272, pp. 979–1014, 2021.

[83] C. Imbert and R. Monneau, "Quasi-convex Hamilton-Jacobi equations posed on junctions: the multi-dimensional case," *arXiv: Analysis of PDEs*, 2014.

[84] M. Gromov, "Hyperbolic groups," in *Essays in group theory.* Springer, 1987, pp. 75–263.

[85] B. Kloeckner, "A geometric study of wasserstein spaces: Euclidean spaces," *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, vol. 9, no. 2, pp. 297–323, 2010.

[86] Y. Burago, M. Gromov, and G. Perel'man, "A.D. Alexandrov spaces with curvature bounded below," *Russian mathematical surveys*, vol. 47, no. 2, pp. 1–58, 1992.

[87] S. Di Marino, N. Gigli, E. Pasqualetto, and E. Soultanis, "Infinitesimal hilbertianity of locally CAT(κ)-spaces," *The Journal of Geometric Analysis*, vol. 31, no. 8, pp. 7621–7685, 2021. [Online]. Available: https://doi.org/10.1007/s12220-020-00543-7

[88] A. Aleksandrov, "The existence almost everywhere of the second differential of a convex function and some associated properties of surfaces," *Ucenye Zapiski Leningr. Gos. Univ. Ser. Mat*, vol. 37, pp. 3–35, 1939.

[89] E. Cristiani, B. Piccoli, and A. Tosin, *Multiscale modeling of pedestrian dynamics.* Springer, 2014, vol. 12.

[90] N. Bellomo, M. Herrero, and A. Tosin, "On the dynamics of social conflicts: Looking for the black swan," *Kinetic and Related Models*, vol. 6, pp. 459–479, 02 2013.

[91] C. M. Topaz and A. L. Bertozzi, "Swarming patterns in a two-dimensional kinematic model for biological groups," *SIAM Journal on Applied Mathematics*, vol. 65, no. 1, pp. 152–174, 2004.

[92] W. Ren and R. W. Beard, *Distributed consensus in multi-vehicle cooperative control.* Springer, 2008, vol. 27, no. 2.

[93] P. Bernard, "Young measures, superposition and transport," *Indiana University mathematics journal*, pp. 247–275, 2008.

[94] B. Bonnet and H. Frankowska, "Differential inclusions in Wasserstein spaces: The Cauchy-Lipschitz framework," *Journal of Differential Equations*, vol. 271, pp. 594–637, 2021.

[95] Z. Badreddine and H. Frankowska, "Solutions to Hamilton–Jacobi equation on a Wasserstein space," *Calculus of Variations and Partial Differential Equations*, vol. 61, no. 1, pp. 1–41, 2022.

[96] B. Piccoli, "Measure differential inclusions," in *2018 IEEE Conference on Decision and Control (CDC).* IEEE, 2018, pp. 1323–1328.

[97] S. Sasaki, "On the differential geometry of tangent bundles of Riemannian

manifolds II," *Tohoku Mathematical Journal, Second Series*, vol. 14, no. 2, pp. 146–155, 1962.

[98] L. Ambrosio and N. Gigli, "A user's guide to optimal transport," in *Modelling and optimisation of flows on networks*. Springer, 2013, pp. 1–155.

[99] W. P. Thurston, S. Patterson *et al.*, *Fundamentals of Hyperbolic Manifolds: Selected Expositions*. Cambridge University Press, 2006, vol. 328.

[100] M. Erbar, "The heat equation on manifolds as a gradient flow in the Wasserstein space," in *Annales de l'IHP Probabilités et statistiques*, vol. 46, no. 1, 2010, pp. 1–23.

[101] R. J. McCann, "Polar factorization of maps on Riemannian manifolds," *Geometric & Functional Analysis GAFA*, vol. 11, no. 3, pp. 589–608, 2001.

[102] D. Azagra, J. Ferrera, and B. Sanz, "Viscosity solutions to second order partial differential equations on riemannian manifolds," *Journal of Differential Equations*, vol. 245, no. 2, pp. 307–336, 2008.

[103] A. Alfonsi and B. Jourdain, "Squared quadratic Wasserstein distance: optimal couplings and Lions differentiability," *ESAIM: Probability and Statistics*, vol. 24, pp. 703–717, 2020.

[104] N. Juillet, "On displacement interpolation of measures involved in Brenier's theorem," *Proceedings of the American Mathematical Society*, vol. 139, no. 10, pp. 3623–3632, 2011.

[105] P. Petersen, *Riemannian geometry*. Springer, 2006, vol. 171.

[106] M. P. Do Carmo, *Riemannian geometry*. Springer Science & Business Media, 2013.

[107] A. Lytchak, "Almost convex subsets," *Geometriae Dedicata*, vol. 115, no. 1, pp. 201–218, 2005.

[108] G. Conforti, R. Kraaij, and D. Tonon, "Hamilton–Jacobi equations for controlled gradient flows: the comparison principle," *arXiv preprint arXiv:2111.13258*, 2021.

ECOLE DOCTORALE DE MATHEMATIQUES HADAMARD

Title: Viscosity theory of first order Hamilton Jacobi equations in some metric spaces.

Key words: Optimal control, Hamilton Jacobi equations, Discontinuous Hamiltonians, CAT(0) spaces, Networks, Multi-agent systems, Wasserstein spaces, Semiconvex functions, Semiconcave functions.

Abstract: In the first part of this thesis, we study a discontinuous first order Hamilton Jacobi Bellman equation defined on a stratification of $\mathbb{R}^N$. The latter is a finite and disjoint union of smooth submanifolds of $\mathbb{R}^N$ called the *subdomains* of $\mathbb{R}^N$. On each subdomain, a continuous Hamiltonian is defined on it, However the global Hamiltonian in $\mathbb{R}^N$ presents discontinuities once one goes from one subdomain to the other. We use nonsmooth analysis techniques to prove that the value function is the unique viscosity solution to the discontinuous Hamilton Jacobi Bellman equation in this setting. Moreover, we prove some stability results in the presence of perturbations on the discontinuous Hamiltonian. Finally, by virtue of the comparison principle, we prove a general convergence result of monotone numerical schemes approximating this problem.

The second part of this thesis is concerned with defining a novel notion of viscosity for first order Hamilton Jacobi equations defined in proper CAT(0) spaces. We exploit the additional structure that these spaces enjoy to study stationary and time-dependent first order Hamilton-Jacobi equations in them. In particular, we want to recover the main features of viscosity theory: *the comparison principle and Perron's method.* We define the notion of viscosity using test functions that are Lipschitz and can be represented as a difference of two semiconvex function. We show that this new notion of viscosity coincides with the classical one in $\mathbb{R}^N$ by studying some classical examples of Hamilton Jacobi equations. Furthermore, we prove existence and uniqueness of the solution of Eikonal type equations posed in more general CAT(0) spaces.

In the third part of this thesis, we study a Mayer optimal control problem on the space of Borel probability measures over a compact Riemannian manifold M. We define the notion of viscosity in this space in a similar manner as in the previous part by taking test functions that are Lipschitz and can be written as a difference of two semiconvex functions. With this choice of test functions, we extend the notion of viscosity to Hamilton Jacobi Bellman equations in Wasserstein spaces and we establish that the value function is the unique viscosity solution of a Hamilton Jacobi Bellman equation in the Wasserstein space over M.

Titre : Théorie de viscosité des équations de Hamilton Jacobi du premier ordre sur certains espaces métriques.

Mots clés : Commande optimale, Équations de Hamilton Jacobi, Hamiltoniens discontinus, Espaces CAT(0), Networks, Systèmes multi-agents, Espaces de Wasserstein, Fonctions semiconvexes, fonctions semiconcaves.

Résumé : La première partie de cette thèse est consacrée à l'étude d'une équation de Hamilton Jacobi Bellman discontinue, définie sur une stratification de $\mathbb{R}^N$. Cette dernière est le résultat d'une union d'une collection finie de sous-variétés lisses et disjointes de $\mathbb{R}^N$, que l'on nomme *les sous-domaines*. Sur chaque sous-domaine, un Hamiltonien continu y est défini. Cependant, le Hamiltonien global sur $\mathbb{R}^N$ présente des discontinuités lorsque l'on passe d'un sous-domaine à l'autre. On utilise les techniques de l'analyse non lisse pour montrer que la fonction valeur est l'unique solution de viscosité de l'équation de Hamilton Jacobi Bellman définie dans ce chapitre. De plus, on prouve quelques résultats de stabilité en présence de perturbations sur le Hamiltonien discontinu. Finalement, en vertu du principe de comparaison, on montre un résultat de convergence général pour les schémas numériques monotones qui approchent ce problème.

La deuxième partie de cette thèse est consacrée au dévelopement d'une nouvelle notion de viscosité pour les équations de Hamilton Jacobi du premier ordre définies sur les espaces CAT(0) propres. On exploite la strucutre de ces espaces pour étudier les equations de Hamilton Jacobi du premier ordre stationnaires et dépendantes du temps. En particulier, le but du chapitre est de retrouver les principaux résultats de la théorie de la viscosité : *le principe de comparaison et la méthode de Perron.* On définit la notion de viscosité en utilisant des fonctions test qui sont Lipschitz et qui peuvent être représentées comme une différence de deux fonctions semiconvexes. On montre que cette notion de viscosité coïncide avec la notion classique dévelopée sur $\mathbb{R}^N$ en étudiant quelques exemples d'équations classiques. De surcroît, on prouve l'existence et l'unicité de la solution de certaines équations du type Eikonal posées sur des espaces CAT(0) plus généraux.

La troisième partie de la thèse se focalise sur l'étude d'un problème de commande optimale de Mayer sur l'espace des mesures Boréliennes de probabilité sur une variété compacte M. On définit la notion de viscosité sur cet espaces de la même manière que dans la deuxième partie de la thèse en considérant des fonctions test qui sont Lipschitz et qui peuvent être représentées par une différence de deux fonctions semiconvexes. Avec ce choix de fonctions test, on étend la notion de viscosité aux équations de Hamilton Jacobi Bellman définies sur l'espace de Wasserstein et on établit que la fonction valeur associée au problème de commande optimale et l'unique solution de viscosité sur l'espace de Wasserstein sur M.

Institut Polytechnique de Paris
91120 Palaiseau, France