HAL
open science

Minimal-time mean field games

Guilherme Mazanti, Filippo Santambrogio

To cite this version:

Guilherme Mazanti, Filippo Santambrogio. Minimal-time mean field games. Mathematical Models and Methods in Applied Sciences, 2019, 29 (8), pp.1413–1464. 10.1142/S0218202519500258 . hal-01768363v2

HAL Id: hal-01768363

https://hal.science/hal-01768363v2

Submitted on 7 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MINIMAL-TIME MEAN FIELD GAMES

GUILHERME MAZANTI AND FILIPPO SANTAMBROGIO

ABSTRACT. This paper considers a mean field game model inspired by crowd motion where agents want to leave a given bounded domain through a part of its boundary in minimal time. Each agent is free to move in any direction, but their maximal speed is bounded in terms of the average density of agents around their position in order to take into account congestion phenomena.

After a preliminary study of the corresponding minimal-time optimal control problem, we formulate the mean field game in a Lagrangian setting and prove existence of Lagrangian equilibria using a fixed point strategy. We provide a further study of equilibria under the assumption that agents may leave the domain through the whole boundary, in which case equilibria are described through a system of a continuity equation on the distribution of agents coupled with a Hamilton–Jacobi equation on the value function of the optimal control problem solved by each agent. This is possible thanks to the semiconcavity of the value function, which follows from some further regularity properties of optimal trajectories obtained through Pontryagin Maximum Principle. Simulations illustrate the behavior of equilibria in some particular situations.

CONTENTS

1. Introduction 2
2. Notations and preliminary definitions 3
3. The MFG model 5
4. Preliminary study of the minimal-time optimal control problem 8
4.1. Elementary properties of the value function 8
4.2. Hamilton–Jacobi equation 11
4.3. Consequences of Pontryagin Maximum Principle 13
4.4. Normalized gradient and the optimal control 15
5. Existence of equilibria for minimal-time MFG 19
5.1. Lipschitz continuity of the value function 19
5.2. Measurable selection of optimal trajectories 20
5.3. Fixed point formulation for equilibria 21
6. The MFG system 23
7. Examples 26
7.1. Mean field game on a segment 26
7.2. A Braess-type paradox 29
7.3. Mean field game on a disk 32
References 36
Index of notations 38

2010 *Mathematics Subject Classification.* 91A13, 49N70, 49K15, 35Q91.

Key words and phrases. Optimal control, Nash equilibrium, time-dependent eikonal equation, congestion games, Pontryagin Maximum Principle, MFG system.

This work was partially supported by a public grant as part of the "Investissement d'avenir" project, reference ANR-11-LABX-0056-LMH, LabEx LMH, PGMO project VarPDEMFG, and by the ANR (Agence Nationale de la Recherche) project ANR-16-CE40-0015-01. The first author was also partially supported by the Hadamard Mathematics LabEx (LMH) through the grant number ANR-11-LABX-0056-LMH in the "Investissement d'avenir" project.

1. Introduction

Introduced around 2006 by Jean-Michel Lasry and Pierre-Louis Lions [58–60] and independently by Peter E. Caines, Minyi Huang, and Roland P. Malhamé [51–53], mean field games (written simply MFGs in this paper for short) are differential games with a continuum of players, assumed to be rational, indistinguishable, individually neglectable, and influenced only by the average behavior of other players through a mean-field type interaction. Their original purpose was to provide approximations of Nash equilibria of games with a large number of symmetric players, with motivations from economics [58–60] and engineering [51–53]. In this paper, we use the words "player" and "agent" interchangeably to refer to those taking part in a game.

Since their introduction, mean field games have attracted much research effort and several works have investigated subjects such as approximation results (how games with a large number of symmetric players converge to MFGs [25, 56]), numerical approximations [1, 2, 31, 44], games with large time horizon [22, 27], variational mean field games [8, 28, 63, 66], games on graphs or networks [16, 17, 41, 45], or the characterization of equilibria using the master equation [9, 26, 32]. We refer to [23, 42, 46] for more details and further references on mean field games.

The goal of this paper is to study a simple mean field game model for crowd motion. The mathematical modeling and analysis of crowd motion has been the subject of a very large number of works from many different perspectives, motivated not only by understanding but also by controlling and optimizing crowd behavior [38, 39, 48, 49, 61, 62, 64]. Some mean field game models inspired by crowd motion have been considered in the literature, such as in [8, 15, 28, 57]. Most of these models, as well as most mean field games model in general, consider that the movement of agents takes place in a fixed time interval, and that each agent is free to choose their speed, which is only penalized in the cost. The goal of this paper is to propose and study a novel model where the final time for the movement of an agent is free, and is actually the agent's minimization criterion, and an agent's maximal speed is constrained in terms of the average distribution of agents around their position. This is motivated by the fact that, in some crowd motion situations, an agent may not be able to move faster by simply paying some additional cost, since the congestion provoked by other agents may work as a physical barrier for the agent to increase their speed. In some sense the closest MFG model to ours is the one in [3], where congestion effects are modeled via a cost which is multiplicative in the speed and the density (and a constraint could be obtained in the limit); yet, also the model in [3] is set on a fixed time horizon and cannot catch the limit case where speed and density are related by a constraint. As a result, the model that we propose is novel and deserves the preliminary study that we develop in the present paper.

The model we consider in this paper is related to Hughes' model for crowd motion [54, 55]. As the model we propose here, Hughes' model also considers agents who aim at leaving in minimal time a bounded domain under a congestion-dependent constraint on their speeds. In the present paper, for modeling and existence purposes, we consider both the case where the agents exit through all the boundary and the case where the exit is only a part of it, while some results characterizing equilibria in terms of a system of PDEs require regularity which cannot be obtained when the exit is only trhough a part of the boundary.

The main difference between our model and Hughes' is that, in the latter, at each time, an agent moves in the optimal direction to the boundary assuming that the distribution of agents remains constant, whereas in our model, agents take into account the future evolution of the distribution of agents in the computation of their optimal trajectories. This accounts for the time derivative in the Hamilton–Jacobi equation from (6.1), which is the main difference between (6.1) and the equations describing the motion of agents in Hughes' model. The knowledge of the future necessary for solving the minimization problem in our model can be interpreted, similarly to other mean field games, as an anticipation of future behavior of other agents based on past experiences in the same situation. In particular, a mean field game model such as ours should be suitable for modeling the behavior of a crowd walking on the streets or subway tunnels, but not for panic situations such as emergency evacuations.

We formulate the notion of equilibrium of a mean field game in this paper in a Lagrangian setting. Contrarily to the classical approach for mean field games consisting on defining an equilibrium in terms of a time-varying measure m_t describing the distribution of agents at time t, the Lagrangian approach relies instead on describing the motion of agents by a measure on the set of all possible trajectories. This is a classical approach in optimal transport problems (see, e.g., [5,69,72]), which has been used for instance in [14] to study incompressible flows, in [29] for Wardrop equilibria in traffic flow, or in [10] for branched transport problems. The Lagrangian approach has also been used for defining equilibria of mean field games, for instance in [8,18,24,28].

Our main results are (i) Theorem 5.1, stating the existence of an equilibrium for the mean field game model we propose in this paper, and (ii) Theorem 6.1, stating that equilibria satisfy a continuity equation on the time-dependent measure representing the distribution of agents coupled with a Hamilton–Jacobi equation on the value function of the time-minimization problem solved by each agent. The proof of Theorem 5.1 relies on a reformulation of the notion of equilibrium (provided in Definition 3.3) in terms of a fixed point and on Kakutani fixed point theorem. The Hamilton–Jacobi equation from Theorem 6.1 can be obtained by standard methods on optimal control, but the continuity equation relies on further properties of the value function, and in particular its semiconcavity, obtained thanks to a detailed study of optimal trajectories.

The notion of equilibrium used in this paper is standard in non-atomic congestion games, and a typical example of these games can be observed in Wardrop equilibria. This notion of equilibrium, introduced in [73] for vehicular traffic flow models, has many connections with MFGs and in particular with minimal-time MFGs, and has been studied in several other works [29, 30, 40, 47]. In Wardrop equilibria the players are the vehicles and their goal is to minimize their traveling time; a major difference with respect to the notion used in this paper is that Wardrop equilibria are usually considered in the stationary case where the flow of vehicles is constant, whereas MFGs explicitly address the case of a time-dependent flow of agents.

The paper is organized as follows. Section 2 provides the main notations used in this paper and recalls some classical results and definitions. Section 3 presents the mean field game model that we consider, together with an associated optimal control problem, presenting and discussing the main assumptions that are used in this paper. The associated optimal control problem is studied in Section 4, its main results being used in Section 5 to obtain existence of an equilibrium to our mean field game and in Section 6 to prove that the distribution of agents and the value function of the optimal control problem solved by each agent satisfy a system of partial differential equations. Examples and numerical simulations are provided in Section 7.

2. Notations and preliminary definitions

Let us set the main notations used in this paper. The usual Euclidean norm in $\mathbb{R}^d$ is denoted by $|\cdot|$, and we write $\mathbb{S}^{d-1}$ for the usual unit Euclidean sphere in $\mathbb{R}^d$, i.e., $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d \mid |x| = 1\}$. Given two sets A, B, the notation $f : A \rightrightarrows B$ indicates that f is a set-valued map from A to B, i.e., f maps a point $x \in A$ to a subset $f(x) \subset B$.

For a metric space X, $x \in X$, and $r \in \mathbb{R}_+$, the open and closed balls of center x and radius r are denoted respectively by $B_X(x,r)$ and $\overline{B}_X(x,r)$, these notations being simplified respectively to $B_d(x,r)$ and $\overline{B}_d(x,r)$ when $X = \mathbb{R}^d$, with a further simplification to B_d and $\overline{B}_d$ if $x = 0$ and $r = 1$. Given two metric spaces X and Y, we denote by $\mathcal{C}(X,Y)$ the set of continuous functions from X to Y and by $\mathrm{Lip}(X,Y)$ the set of those which are Lipschitz continuous. When $X \subset \mathbb{R}^d$, we will also make use of the space $\mathcal{C}^{1,1}(X,\mathbb{R})$.

Given a complete, separable, and bounded metric space X with metric $\mathbf{d}$, let $\mathcal{C}_X = \mathcal{C}(\mathbb{R}_+, X)$ be endowed with the topology of uniform convergence on compact sets ($\mathcal{C}_X$ is a Polish space for this distance, see, e.g., [12, Corollary 3, page X.9; Corollary, page X.20; and Corollary, page X.25]). Whenever needed, we endow $\mathcal{C}_X$ with the complete distance

$$\mathbf{d}_{\mathcal{C}_X}(\gamma_1, \gamma_2) = \sum_{n=1}^{\infty} \frac{1}{2^n} \sup_{t \in [0,n]} \mathbf{d}(\gamma_1(t), \gamma_2(t)).$$

For $c \geq 0$, we denote by $\mathrm{Lip}_c(X)$ the subset of $\mathcal{C}_X$ containing all c-Lipschitz continuous functions. Recall that, if X is compact, then, thanks to Arzelà–Ascoli Theorem [12, Corollary 3, page X.19], $\mathrm{Lip}_c(X)$ is compact. For $t \in \mathbb{R}_+$, we denote by $e_t : \mathcal{C}_X \to X$ the evaluation map $e_t(\gamma) = \gamma(t)$.

For a given topological space X, let $\mathcal{P}(X)$ denote the set of all Borel probability measures on X. If X, endowed with a metric $\mathbf{d}$, is a complete, separable, and bounded metric space, we endow $\mathcal{P}(X)$ with the usual Wasserstein distance W_1 defined by its dual formulation (see, e.g., [5, Chapter 7] and [69, Chapter 5]):

$$W_1(\mu,\nu) = \sup\left\{ \int_X \phi(x)\,\mathrm{d}(\mu-\nu)(x) \;\middle|\; \phi \in \mathrm{Lip}(X,\mathbb{R}) \text{ is 1-Lipschitz} \right\}. \tag{2.1}$$

Let X be a metric space with metric $\mathbf{d}$, $a,b \in \mathbb{R}$ with $a < b$, and $\gamma : (a,b) \to X$. The *metric derivative* of γ at a point $t \in (a,b)$ is defined by

$$|\dot{\gamma}|(t) = \lim_{s\to t} \frac{\mathbf{d}(\gamma(s),\gamma(t))}{|s-t|}$$

whenever this limit exists. Recall that, if γ is absolutely continuous, then $|\dot{\gamma}|(t)$ exists for almost every $t \in (a,b)$ (see, e.g., [5, Theorem 1.1.2]).

We shall also need in this paper some classical tools from non-smooth analysis, which we briefly recall now, following the presentation from [35, Chapter 2].

Definition 2.1. Let $O \subset \mathbb{R}^d$ be open, $\phi : O \to \mathbb{R}$ be Lipschitz continuous, $x \in O$, and $v \in \mathbb{R}^d$. We define *generalized directional derivative* $\phi^\circ(x,v)$ of ϕ at x in the direction v by

$$\phi^\circ(x,v) = \limsup_{\substack{y\to x\\ h\searrow 0}} \frac{\phi(y+hv)-\phi(y)}{h},$$

and the *generalized gradient* $\partial^{\mathrm{C}}\phi(x)$ of ϕ at x by

$$\partial^{\mathrm{C}}\phi(x) = \{\xi \in \mathbb{R}^d \mid \phi^\circ(x,v) \geq \xi\cdot v \text{ for all } v \in \mathbb{R}^d\}. \tag{2.2}$$

The generalized gradient is also sometimes referred to as *Clarke's gradient* in the literature, which explains the notation ∂^{C} used here. We shall also need the notion of partial generalized gradients.

Definition 2.2. Let $O_1 \subset \mathbb{R}^{d_1}$, $O_2 \subset \mathbb{R}^{d_2}$ be open sets, $\phi : O_1 \times O_2 \to \mathbb{R}$ be Lipschitz continuous, and $(x_1,x_2) \in O_1 \times O_2$. The *partial generalized gradient* $\partial_1^{\mathrm{C}}\phi(x_1,x_2) = \partial_{x_1}^{\mathrm{C}}\phi(x_1,x_2)$ is defined as the generalized gradient of the function $x_1 \mapsto \phi(x_1,x_2)$. The partial generalized gradient $\partial_2^{\mathrm{C}}\phi = \partial_{x_2}^{\mathrm{C}}\phi$ is defined similarly.

Recall that, in general, there is no link between $\partial^{\mathrm{C}}\phi(x_1,x_2)$ and $\partial_1^{\mathrm{C}}\phi(x_1,x_2) \times \partial_2^{\mathrm{C}}\phi(x_1,x_2)$ (see, e.g., [35, Example 2.5.2]).

Definition 2.3. Let $C \subset \mathbb{R}^d$ be non-empty and $\mathbf{d}_C : \mathbb{R}^d \to \mathbb{R}_+$ be the distance to C, defined by $\mathbf{d}_C(x) = \inf_{y\in C}|x-y|$. Let $x \in C$. We define the *tangent cone* $T_C(x)$ to C at x by

$$T_C(x) = \{v \in \mathbb{R}^d \mid \mathbf{d}_C^\circ(x,v) = 0\},$$

and the *normal cone* $N_C(x)$ to C at x by

$$N_C(x) = \{\xi \in \mathbb{R}^d \mid \xi\cdot v \leq 0 \text{ for all } v \in T_C(x)\}.$$

We refer to [34, Chapter 10] and [35, Chapter 2] for standard properties of generalized directional derivatives, generalized gradients, and tangent and normal cones.

3. THE MFG MODEL

We provide in this section a mathematical description of the mean field game model considered in this paper. Let $(X,\mathbf{d})$ be a complete and separable metric space, $\Gamma \subset X$ be non-empty and closed, and $K : \mathcal{P}(X) \times X \to \mathbb{R}_+$. We consider the following mean field game, denoted by $\mathrm{MFG}(X,\Gamma,K)$. Agents evolve in X, their distribution at time $t \in \mathbb{R}_+$ being given by a probability measure $m_t \in \mathcal{P}(X)$. The goal of each agent is to reach the exit Γ in minimal time, and, in order to model congestion, we assume the speed of an agent at a position x in time t to be bounded by $K(m_t,x)$.

Notice that, for a given agent, their choice of trajectory γ depends on the distribution of all agents m_t, since the speed of γ, i.e. its metric derivative $|\dot{\gamma}|$, should not exceed $K(m_t,x)$. On the other hand, the distribution of the agents m_t itself depends on how agents choose their trajectories γ. We are interested here in the equilibrium situations, meaning that, starting from a time evolution of the distribution of agents $m : \mathbb{R}_+ \to \mathcal{P}(X)$, the trajectories γ chosen by agents induce an evolution of the initial distribution of agents m_0 that is precisely given by m.

In order to provide a more mathematically precise definition of this mean field game and the notion of equilibrium, we first introduce an optimal control problem where agents evolving in X want to reach Γ in minimal time, their speed being bounded by some time- and state-dependent function $k : \mathbb{R}_+ \times X \to \mathbb{R}_+$. Here k will not depend on the density of the agents, it will be considered as given. This optimal control problem is denoted in the sequel by $\mathrm{OCP}(X,\Gamma,k)$.

Definition 3.1 ($\mathrm{OCP}(X,\Gamma,k)$)**.** Let $(X,\mathbf{d})$ be a complete and separable metric space, $\Gamma \subset X$ be non-empty and closed, and $k : \mathbb{R}_+ \times X \to \mathbb{R}_+$ be continuous.

(a) A curve $\gamma \in \mathrm{Lip}(\mathbb{R}_+,X)$ is said to be *k-admissible* for $\mathrm{OCP}(X,\Gamma,k)$ if its metric derivative $|\dot{\gamma}|$ satisfies $|\dot{\gamma}|(t) \leq k(t,\gamma(t))$ for almost every $t \in \mathbb{R}_+$. The set of all k-admissible curves is denoted by $\mathrm{Adm}(k)$.
(b) Let $t_0 \in \mathbb{R}_+$. The *first exit time* after t_0 of a curve $\gamma \in \mathrm{Lip}(\mathbb{R}_+,X)$ is the number $\tau(t_0,\gamma) \in [t_0,+\infty]$ defined by

$$\tau(t_0,\gamma) = \inf\{t \geq 0 \mid \gamma(t+t_0) \in \Gamma\}.$$

(c) Let $t_0 \in \mathbb{R}_+$ and $x_0 \in X$. A curve $\gamma \in \mathrm{Lip}(\mathbb{R}_+,X)$ is said to be a *time-optimal curve* or *time-optimal trajectory* (or simply *optimal curve* or *optimal trajectory*) for (k,t_0,x_0) if $\gamma \in \mathrm{Adm}(k)$, $\gamma(t) = x_0$ for every $t \in [0,t_0]$, $\tau(t_0,\gamma) < +\infty$, $\gamma(t) = \gamma(t_0+\tau(t_0,\gamma)) \in \Gamma$ for every $t \in [t_0+\tau(t_0,\gamma), +\infty)$, and

$$\tau(t_0,\gamma) = \min_{\substack{\beta \in \mathrm{Adm}(k) \\ \beta(t_0)=x_0}} \tau(t_0,\beta). \tag{3.1}$$

The set of all optimal curves for (k,t_0,x_0) is denoted by $\mathrm{Opt}(k,t_0,x_0)$.

Remark 3.2. If $X \subset \mathbb{R}^d$, the metric derivative of a curve $\gamma \in \mathrm{Lip}(\mathbb{R}_+,X)$ coincides with the norm of the usual derivative when it exists (cf. e.g. [5, Remark 1.1.3]), and one obtains that a curve γ is k-admissible if and only if there exists a measurable function $u : \mathbb{R}_+ \to \overline{B}_d$ such that

$$\dot{\gamma}(t) = k(t,\gamma(t))u(t). \tag{3.2}$$

System (3.2) can be seen as a control system, where $\gamma(t) \in X \subset \mathbb{R}^d$ is the state and $u(t) \in \overline{B}_d$ is the control input. This point of view allows one to formulate (3.1) as an optimal control problem, justifying the terminology used in this paper. Optimal control techniques are a key point in the study of equilibria for $\mathrm{MFG}(X,\Gamma,K)$ carried out in the sequel.

The relation between the optimal control problem $\mathrm{OCP}(X,\Gamma,k)$ and the mean field game MFG (X,Γ,K) is that, given $K : \mathcal{P}(X) \times X \to \mathbb{R}_+$, players from the mean field game $\mathrm{MFG}(X,\Gamma,K)$ solve the optimal control problem $\mathrm{OCP}(X,\Gamma,k)$ with $k : \mathbb{R}_+ \times X \to \mathbb{R}_+$ given by $k(t,x) = K(m_t,x)$, where m_t is the distribution of players at time t. Using this relation, one can provide the definition of an equilibrium for $\mathrm{MFG}(X,\Gamma,K)$.

Definition 3.3 (Equilibrium of MFG(X,Γ,K))**.** Let $(X,\mathbf{d})$ be a complete and separable metric space, $\Gamma\subset X$ be non-empty and closed, and $K:\mathcal{P}(X)\times X\to\mathbb{R}_+$ be continuous. Let $m_0\in\mathcal{P}(X)$. A measure $Q\in\mathcal{P}(\mathcal{C}_X)$ is said to be a *Lagrangian equilibrium* (or simply *equilibrium*) of MFG(X,Γ,K) with initial condition m_0 if $e_{0\#}Q=m_0$ and Q-almost every $\gamma\in\mathcal{C}_X$ is an optimal curve for $(k,0,\gamma(0))$, where $k:\mathbb{R}_+\times X\to\mathbb{R}_+$ is defined by $k(t,x)=K(e_{t\#}Q,x)$.

In order to simplify the notations in the sequel, given $K:\mathcal{P}(X)\times X\to\mathbb{R}_+$ and $m:\mathbb{R}_+\to\mathcal{P}(X)$, we define $k:\mathbb{R}_+\times X\to\mathbb{R}_+$ by $k(t,x)=K(m_t,x)$ for $(t,x)\in\mathbb{R}_+\times X$ and say that $\gamma\in\mathrm{Lip}(\mathbb{R}_+,X)$ is m-admissible for MFG(X,Γ,K) if it is k-admissible for OCP(X,Γ,k), and denote $\mathrm{Adm}(k)$ simply by $\mathrm{Adm}(m)$. We also say that, given $(t_0,x_0)\in\mathbb{R}_+\times X$, $\gamma\in\mathrm{Lip}(\mathbb{R}_+,X)$ is an optimal trajectory for (m,t_0,x_0) if it is an optimal trajectory for (k,t_0,x_0) for the optimal control problem OCP(X,Γ,k), and denote the set of optimal trajectories for (m,t_0,x_0) by $\mathrm{Opt}(m,t_0,x_0)$. Given $Q\in\mathcal{P}(\mathcal{C}_X)$, we also consider the time-dependent measure $\mu^Q:\mathbb{R}_+\to\mathcal{P}(X)$ given by $\mu^Q_t=e_{t\#}Q$, and denote $\mathrm{Adm}(\mu^Q)$ and $\mathrm{Opt}(\mu^Q,t_0,x_0)$ simply by $\mathrm{Adm}(Q)$ and $\mathrm{Opt}(Q,t_0,x_0)$, respectively.

Remark 3.4. Since players do not necessarily arrive at the target set Γ all at the same time, the behavior of players who have not yet arrived may be influenced by the players who already arrived at Γ since K is not necessarily local, i.e., $K(m_t,x)$ may depend on the measure m_t on points other than only x. On the other hand, after arriving at Γ, players are no longer submitted to the time minimization criterion, and thus their trajectory might in principle be arbitrary after their arrival time. In order to avoid ambiguity, we have decided to fix the behavior of players who already arrived at Γ by saying that a trajectory is optimal only when it remains at its arrival position after its arrival time. Similarly, we assume that optimal trajectories starting at a time $t_0>0$ remain constant on the interval $[0,t_0]$.

The study of OCP(X,Γ,k) and MFG(X,Γ,K) carried out in this paper requires some assumptions on X, Γ, K and k. For simplicity, we state all such assumptions here, and refer to them wherever needed in the sequel.

Hypotheses 3.5.

(a) The metric space $(X,\mathbf{d})$ is compact and $\Gamma\subset X$ is non-empty and closed.
(b) The function $k:\mathbb{R}_+\times X\to\mathbb{R}_+$ is Lipschitz continuous and there exist $K_{\min},K_{\max}\in\mathbb{R}_+^*$ such that, for all $(t,x)\in\mathbb{R}_+\times X$, one has $k(t,x)\in[K_{\min},K_{\max}]$.
(c) There exists $D>0$ such that, for every $x,y\in X$, there exist $T\in[0,D\mathbf{d}(x,y)]$ and $\gamma\in\mathrm{Lip}([0,T],X)$ such that $\gamma(0)=x$, $\gamma(T)=y$, and $|\dot\gamma|(t)=1$ for almost every $t\in[0,T]$.
(d) One has $X=\overline{\Omega}$ for $\Omega\subset\mathbb{R}^d$ open, bounded, and non-empty.
(e) (d) holds and $\Gamma=\partial\Omega$.
(f) (d) holds and $k\in\mathrm{Lip}(\mathbb{R}\times\mathbb{R}^d,\mathbb{R}_+)$.
(g) (d) holds and $k\in\mathcal{C}^{1,1}(\mathbb{R}\times\mathbb{R}^d,\mathbb{R}_+)$.
(h) (d) holds and Ω satisfies the *uniform exterior sphere property*: there exists $r>0$ and $C\subset\mathbb{R}^d\setminus\Omega$ such that
$$\mathbb{R}^d\setminus\Omega=\bigcup_{x\in C}\overline{B}_d(x,r).$$
(i) The function $K:\mathcal{P}(X)\times X\to\mathbb{R}_+$ is Lipschitz continuous and there exist $K_{\min},K_{\max}\in\mathbb{R}_+^*$ such that, for all $(\mu,x)\in\mathcal{P}(X)\times X$, one has $K(\mu,x)\in[K_{\min},K_{\max}]$.
(j) (d) holds and $K:\mathcal{P}(\overline{\Omega})\times\mathbb{R}^d$ is given by $K=g\circ E$, where $g\in\mathcal{C}^{1,1}(\mathbb{R}_+,\mathbb{R}_+^*)$ and $E:\mathcal{P}(\overline{\Omega})\times\mathbb{R}^d\to\mathbb{R}_+$ is given by
$$E(\mu,x)=\int_{\overline{\Omega}}\chi(x-y)\eta(y)\,\mathrm{d}\mu(y),\tag{3.3}$$
with $\chi\in\mathcal{C}^{1,1}(\mathbb{R}^d,\mathbb{R}_+)$, $\eta\in\mathcal{C}^{1,1}(\mathbb{R}^d,\mathbb{R}_+)$, χ and η bounded, and $\eta(x)=0$ and $\nabla\eta(x)=0$ for every $x\in\partial\Omega$.

Hypotheses 3.5(a)–(c) are the standard assumptions needed in all the results from Section 4 for studying OCP(X,Γ,k). Note that (c) provides a relation between the distance $\mathbf{d}(x,y)$ and the

length of curves from x to y in X, stating that the former is, up to a constant, an upper bound on the latter. This assumption is satisfied when the geodesic metric induced by $\mathbf{d}$ is equivalent to $\mathbf{d}$ itself. In particular, it is satisfied if X is a length space. Moreover, (c) implies that X is path-connected.

One uses Hypotheses 3.5(d)–(h) in Section 4 to obtain more properties of $\mathrm{OCP}(X,\Gamma,k)$, such as the facts that the value function satisfies a Hamilton–Jacobi equation and is semiconcave, and that optimal trajectories can be described via the Pontryagin Maximum Principle. Assumption (d) allows one to work on a subset of an Euclidean space, making it easier to give a meaning to the Hamilton–Jacobi equation, for instance.

Assumption (e) states that the goal of an agent is to leave the domain Ω through any part of its boundary. This simplifying assumption is first used when applying Pontryagin Maximum Principle to $\mathrm{OCP}(X,\Gamma,k)$, and, even though it is not strictly needed at this point, as remarked in the beginning of Section 4.3, it is important in Section 4.4 to deduce the semiconcavity of the value function in Proposition 4.18. Indeed, without assumption (e), one must take into account in $\mathrm{OCP}(X,\Gamma,k)$ the state constraint $\gamma(t)\in\overline{\Omega}$, and value functions of optimal control problems with state constraints may fail to be semiconcave (see, e.g., [19, Example 4.4]).

Notice that (e) is not needed to obtain existence of equilibria for $\mathrm{MFG}(X,\Gamma,K)$ in Section 5, nor for establishing the Hamilton–Jacobi equation on the value function φ of $\mathrm{OCP}(X,\Gamma,k)$ in Section 4.2. However, to prove that m satisfies the continuity equation from (6.1) and hence complete the description of equilibria by the MFG system in Section 6, one uses the characterization of optimal trajectories of $\mathrm{OCP}(X,\Gamma,k)$ in terms of the normalized gradient of φ carried out in Section 4.4, the semiconcavity of φ being a key ingredient in the proof of the main result of that section, Theorem 4.20.

Assumption (f) is not restrictive with respect to (b), stating that one considers k to be extended to the whole space $\mathbb{R}\times\mathbb{R}^d$ in a Lipschitz manner, and is included in the list of hypotheses only for simplifying the statements of the results. A classical technique to obtain Lipschitz extensions of Lipschitz continuous functions is by inf-convolution (see, e.g., [50]). Assumptions (g) and (h) are important to obtain the semiconcavity of the value function of $\mathrm{OCP}(X,\Gamma,k)$ in Section 4.4.

Concerning Hypotheses 3.5(i) and (j), they are used in Sections 5 and 6 to study $\mathrm{MFG}(X,\Gamma,K)$. Assumption (i) is the counterpart of (b), with (b) being obtained from (i) when k is given by $k(t,x)=K(e_{t\#}Q,x)$ (see Corollary 5.3). Assumption (j), which is reasonable for modeling reasons (the function g is typically required to be non-increasing, but this is not crucial for the mathematical results we present), will be important in order for this k to satisfy also (g) when Q is an equilibrium (see Proposition 6.2). Notice that (j) implies (i); more precisely, we have the following result.

Proposition 3.6. *Let $\Omega\subset\mathbb{R}^d$ be open, bounded, and non-empty, $X=\overline{\Omega}$, $g\in\mathrm{Lip}(\mathbb{R}_+,\mathbb{R}_+^*)$, $\chi\in\mathrm{Lip}(\mathbb{R}^d,\mathbb{R}_+)$, $\eta\in\mathrm{Lip}(\overline{\Omega},\mathbb{R}_+)$, and define $K:\mathcal{P}(\overline{\Omega})\times\overline{\Omega}\to\mathbb{R}_+$ by*

$$K(\mu,x)=g\left[\int_{\overline{\Omega}}\chi(x-y)\eta(y)\,\mathrm{d}\mu(y)\right].$$

Then K satisfies Hypothesis 3.5(i).

Proof. Let $E:\mathcal{P}(\overline{\Omega})\times\overline{\Omega}\to\mathbb{R}_+$ be defined by (3.3). Let $M_1>0$ be a Lipschitz constant for χ and η and $M_2=\max\{\max_{x\in\overline{\Omega}}\eta(x),\max_{(x,y)\in\overline{\Omega}^2}\chi(x-y)\}$. Then, for every $y\in\overline{\Omega}$, the function $z\mapsto\chi(y-z)\eta(z)$ is $2M_1M_2$-Lipschitz continuous on $\overline{\Omega}$ and, in particular, $z\mapsto\frac{\chi(y-z)\eta(z)}{2M_1M_2}$ is 1-Lipschitz continuous on $\overline{\Omega}$.

Let us first show that $K\in\mathrm{Lip}(\mathcal{P}(\overline{\Omega})\times\overline{\Omega},\mathbb{R}_+)$. Take $\mu,\nu\in\mathcal{P}(\overline{\Omega})$ and $x,y\in\overline{\Omega}$. Then

$$\begin{aligned}
&E(\mu,x)-E(\nu,y)\\
&=\int_{\overline{\Omega}}\chi(x-z)\eta(z)\,\mathrm{d}\mu(z)-\int_{\overline{\Omega}}\chi(y-z)\eta(z)\,\mathrm{d}\nu(z)\\
&=\int_{\overline{\Omega}}[\chi(x-z)-\chi(y-z)]\,\eta(z)\,\mathrm{d}\mu(z)+2M_1M_2\int_{\overline{\Omega}}\frac{\chi(y-z)\eta(z)}{2M_1M_2}\,\mathrm{d}(\mu-\nu)(z)
\end{aligned}$$

$$\leq M_1 M_2 |x-y| + 2M_1 M_2 W_1(\mu,\nu),$$

where we use (2.1). A similar computation starting from $E(\nu,y) - E(\mu,x)$ completes the proof of the fact that $E \in \mathrm{Lip}(\mathcal{P}(\overline{\Omega}) \times \overline{\Omega}, \mathbb{R}_+)$, and then $K = g \circ E \in \mathrm{Lip}(\mathcal{P}(\overline{\Omega}) \times \overline{\Omega}, \mathbb{R}_+)$ since $g \in \mathrm{Lip}(\mathbb{R}_+, \mathbb{R}_+^*)$.

Finally, it follows from the definition of E that $E(\mu,x) \leq M_2^2$ for all $(\mu,x) \in \mathcal{P}(\overline{\Omega}) \times \overline{\Omega}$, and thus

$$K_{\min} = \inf_{(\mu,x)\in\mathcal{P}(\overline{\Omega})\times\overline{\Omega}} g \circ E(\mu,x) \geq \min_{x\in[0,M_2^2]} g(x) > 0,$$
$$K_{\max} = \sup_{(\mu,x)\in\mathcal{P}(\overline{\Omega})\times\overline{\Omega}} g \circ E(\mu,x) \leq \max_{x\in[0,M_2^2]} g(x) < +\infty,$$

as required. □

4. Preliminary study of the minimal-time optimal control problem

In this section, we provide several properties for the optimal control problem $\mathrm{OCP}(X,\Gamma,k)$. Minimal-time optimal control problems are a classical subject in the optimal control literature (see, e.g., [21, 35, 65]). Most works consider the case of an autonomous control system with smooth dynamics in the Euclidean space or a smooth manifold, dealing with the non-autonomous case by a classical state-augmentation technique. Since the study of $\mathrm{MFG}(X,\Gamma,K)$ requires some properties of $\mathrm{OCP}(X,\Gamma,k)$ in less regular cases (for instance when k is only Lipschitz continuous), we provide here a detailed presentation of the properties of $\mathrm{OCP}(X,\Gamma,k)$ in order to highlight which hypotheses are required for each result.

The proofs of classical properties for $\mathrm{OCP}(X,\Gamma,k)$ are provided when the particular structure of $\mathrm{OCP}(X,\Gamma,k)$ allows for simplifications with respect to classical proofs in the literature, and omitted otherwise. Several properties presented here are new and rely on the particular structure of $\mathrm{OCP}(X,\Gamma,k)$. This is the case, for instance, of the lower bound on the time variation of the value function (Proposition 4.5 and Lemma 4.8) and the characterization of optimal trajectories in terms of the normalized gradient presented in Section 4.4.

We start in Section 4.1 by studying the value function φ corresponding to $\mathrm{OCP}(X,\Gamma,k)$. We then specialize to the case of an optimal control problem on $\overline{\Omega} \subset \mathbb{R}^d$ and establish a Hamilton–Jacobi equation for φ in Section 4.2, before proving further properties of optimal trajectories obtained from Pontryagin Maximum Principle in Section 4.3. We conclude this section by a characterization of the optimal control in terms of a normalized gradient in Section 4.4, which also considers the continuity of the normalized gradient.

4.1. Elementary properties of the value function.

As stated in Remark 3.2, at least in the case where X is a subset of a normed vector space, admissible trajectories for some $k : \mathbb{R}_+ \times X \to \mathbb{R}_+$ can be seen as trajectories of the control system (3.2) and the minimization problem (3.1), as an optimal control problem. Inspired by this interpretation, we make use of an usual strategy in optimal control, namely that of considering the value function associated with the optimal control problem. We start by recalling the classical definition of the value function for $\mathrm{OCP}(X,\Gamma,k)$.

Definition 4.1. Let (X,Γ,k) be as in Definition 3.1. The *value function* of $\mathrm{OCP}(X,\Gamma,k)$ is the function $\varphi : \mathbb{R}_+ \times X \to \mathbb{R}_+ \cup \{+\infty\}$ defined by

$$\varphi(t,x) = \inf_{\substack{\gamma\in\mathrm{Adm}(k)\\ \gamma(t)=x}} \tau(t,\gamma). \tag{4.1}$$

The goal of this section is to provide some properties of φ, in particular its Lipschitz continuity and the fact that it satisfies a Hamilton–Jacobi equation. We gather in the next proposition some elementary properties of φ.

Proposition 4.2. *Consider the optimal control problem* $\mathrm{OCP}(X,\Gamma,k)$ *and assume that Hypotheses 3.5(a)–(c) hold.*

(a) *There exists $M > 0$ such that, for every $(t,x) \in \mathbb{R}_+ \times X$, one has $\varphi(t,x) \leq M$.*
(b) *For every $(t,x) \in \mathbb{R}_+ \times X$, there exists an optimal trajectory $\gamma \in \mathrm{Lip}(\mathbb{R}_+, X)$ for (k,t,x). In particular, the infimum in* (4.1) *is attained at γ.*

The proof of Proposition 4.2 follows from standard techniques: the upper bound M in (a) can be taken as $M = \frac{D\,\mathrm{diam}(X)}{K_{\min}}$, where $\mathrm{diam}(X) = \sup_{x,y\in X} \mathbf{d}(x,y) < +\infty$ is the diameter of X, and the existence of an optimal trajectory proven using compactness of minimizing sequences.

We now turn to the proof of Lipschitz continuity of φ. We first show, in Proposition 4.3, that φ is Lipschitz continuous with respect to t, uniformly with respect to x, before completing the proof of Lipschitz continuity of φ in Proposition 4.4.

Proposition 4.3. *Consider the optimal control problem* $\mathrm{OCP}(X,\Gamma,k)$ *and assume that Hypotheses 3.5(a)–(c) hold. Then the map $t \mapsto \varphi(t,x)$ is Lipschitz continuous, uniformly with respect to $x \in X$.*

Proof. Let $t_1, t_2 \in \mathbb{R}_+$ and assume, with no loss of generality, that $t_1 \leq t_2$. We prove the result by showing that one has both

$$\varphi(t_1,x) - \varphi(t_2,x) \leq t_2 - t_1 \tag{4.2}$$

and

$$\varphi(t_2,x) - \varphi(t_1,x) \leq C(t_2 - t_1) \tag{4.3}$$

for some constant $C > 0$ independent of t_1, t_2, and x.

Let us first prove (4.2). Let $\gamma_2 \in \mathrm{Opt}(k,t_2,x)$. Define $\gamma_1 \in \mathrm{Lip}(\mathbb{R}_+, X)$ by

$$\gamma_1(s) = \begin{cases} x, & \text{if } s \leq t_2, \\ \gamma_2(s), & \text{if } s > t_2. \end{cases}$$

Then $\gamma_1(t) = x$ for $t \in [0,t_1]$, $\gamma_1(t) = \gamma_2(t) = \gamma_2(t_2 + \varphi(t_2,x)) \in \Gamma$ for $t > t_2 + \varphi(t_2,x)$, $|\dot\gamma_1|(s) = 0$ for $s < t_2$, and $|\dot\gamma_1|(s) \leq k(s,\gamma_1(s))$ for almost every $s > t_2$, which proves that $\gamma_1 \in \mathrm{Adm}(k)$ and $\varphi(t_1,x) \leq \tau(t_1,\gamma_1) \leq t_2 - t_1 + \varphi(t_2,x)$, yielding (4.2).

We now turn to the proof of (4.3). Let $M > 0$ be as in Proposition 4.2(a) and take $\gamma_1 \in \mathrm{Opt}(k,t_1,x)$. Notice that, thanks to Hypothesis 3.5(b), the map

$$\mathbb{R}_+ \times \mathbb{R}_+ \ni (t,s) \mapsto \frac{k(t,\gamma_1(s))}{k(s,\gamma_1(s))}$$

is lower bounded by $\frac{K_{\min}}{K_{\max}}$, upper bounded by $\frac{K_{\max}}{K_{\min}}$, and globally Lipschitz continuous. Let $M_0 > 0$ be a Lipschitz constant for this map. Let $\phi : [t_2, +\infty) \to \mathbb{R}$ be the unique function satisfying

$$\begin{cases} \dot\phi(t) = \dfrac{k(t,\gamma_1(\phi(t)))}{k(\phi(t),\gamma_1(\phi(t)))} \\ \phi(t_2) = t_1. \end{cases}$$

Notice that ϕ is strictly increasing and maps $[t_2,+\infty)$ onto $[t_1,+\infty)$, its inverse ϕ^{-1} being defined on $[t_1,+\infty)$. Let $\gamma_2 : \mathbb{R}_+ \to X$ be given by

$$\gamma_2(s) = \begin{cases} x, & \text{if } s \leq t_2, \\ \gamma_1(\phi(s)), & \text{if } s > t_2. \end{cases}$$

Then $\gamma_2 \in \mathrm{Lip}(\mathbb{R}_+, X)$, $\gamma_2(t) = x$ for every $t \in [0,t_2]$, $\gamma_2(t) = \gamma_2(\phi^{-1}(t_1 + \varphi(t_1,x))) \in \Gamma$ for every $t \geq \phi^{-1}(t_1 + \varphi(t_1,x))$, $|\dot\gamma_2|(s) = 0$ whenever $0 < s < t_2$ or $s > \phi^{-1}(t_1 + \varphi(t_1,x))$, and, for almost every $s \geq t_2$, one has $|\dot\gamma_2|(s) = \left|\dot\phi(s)\right| |\dot\gamma_1|(\phi(s)) \leq k(s,\gamma_2(s))$. Hence $\gamma_2 \in \mathrm{Adm}(k)$ and $t_2 + \varphi(t_2,x) \leq \phi^{-1}(t_1 + \varphi(t_1,x))$, i.e., $\phi(t_2 + \varphi(t_2,x)) \leq t_1 + \varphi(t_1,x)$. Yet, for every $\sigma \geq 0$,

$$\phi(t_2+\sigma) - (t_2+\sigma) = t_1 - t_2 + \int_{t_2}^{t_2+\sigma} \left[\frac{k(s,\gamma_1(\phi(s)))}{k(\phi(s),\gamma_1(\phi(s)))} - 1 \right] \mathrm{d}s,$$

and thus

$$|\phi(t_2+\sigma)-(t_2+\sigma)| \le |t_1-t_2| + M_0 \int_{t_2}^{t_2+\sigma} |\phi(s)-s|\,\mathrm{d}s,$$

which yields, by Gronwall's inequality, that $|\phi(t_2+\sigma)-(t_2+\sigma)| \le |t_1-t_2|\,e^{M_0\sigma}$. Then

$$|\phi(t_2+\varphi(t_2,x))-(t_2+\varphi(t_2,x))| \le |t_1-t_2|\,e^{MM_0},$$

which proves that

$$t_1+\varphi(t_1,x) \ge \phi(t_2+\varphi(t_2,x)) \ge t_2+\varphi(t_2,x) - e^{MM_0}(t_2-t_1),$$

and thus

$$\varphi(t_2,x)-\varphi(t_1,x) \le (e^{MM_0}-1)(t_2-t_1),$$

which concludes the proof of (4.3). □

Proposition 4.4. *Consider the optimal control problem* $\mathrm{OCP}(X,\Gamma,k)$ *and assume that Hypotheses 3.5(a)–(c) hold. Then* φ *is Lipschitz continuous on* $\mathbb{R}_+\times X$.

Proof. Let D be as in Hypothesis 3.5(c), $x_1,x_2\in X$, and $t_1,t_2\in\mathbb{R}_+$. According to Hypothesis 3.5(c), there exist $T_0\in[0,D\mathbf{d}(x_1,x_2)]$ and $\gamma_0\in\mathrm{Lip}([0,T_0],X)$ such that $\gamma_0(0)=x_1$, $\gamma_0(T_0)=x_2$, and $|\dot\gamma_0|(s)=1$ for almost every $s\in[0,T_0]$. Denote by $M>0$ the Lipschitz constant of the map $t\mapsto\varphi(t,x)$, which is independent of $x\in X$ according to Proposition 4.3.

Set $T=\frac{T_0}{K_{\min}}$ and $\sigma_2=\max\{t_2,T\}$. Let $\gamma_2\in\mathrm{Opt}(k,\sigma_2,x_2)$ and define $\gamma_1\in\mathrm{Lip}(\mathbb{R}_+,X)$ by

$$\gamma_1(t)=\begin{cases} x_1, & \text{if } t\le\sigma_2-T,\\ \gamma_0(K_{\min}(t+T-\sigma_2)), & \text{if } \sigma_2-T<t\le\sigma_2,\\ \gamma_2(t), & \text{if } t>\sigma_2.\end{cases}$$

Then $\gamma_1(t)=x_1$ for every $t\in[0,\sigma_2-T]$, $\gamma_1(t)=\gamma_1(\sigma_2+\varphi(\sigma_2,x_2))\in\Gamma$ for every $t\ge\sigma_2+\varphi(\sigma_2,x_2)$, and $|\dot\gamma_1|(t)\le k(t,\gamma_1(t))$ for almost every $t\ge0$, which proves that $\gamma_1\in\mathrm{Adm}(k)$ and $\varphi(\sigma_2-T,x_1)\le T+\varphi(\sigma_2,x_2)$. Hence, using Proposition 4.3, one obtains that

$$\begin{aligned}
&\varphi(t_1,x_1)-\varphi(t_2,x_2)\\
&\le M|t_1-\sigma_2+T|+T+M|\sigma_2-t_2|\\
&\le M|t_1-t_2|+M|t_2-\sigma_2+T|+M|\sigma_2-t_2|+T = M|t_1-t_2|+(M+1)T\\
&= M|t_1-t_2|+\frac{M+1}{K_{\min}}T_0 \le M|t_1-t_2|+D\frac{M+1}{K_{\min}}\mathbf{d}(x_1,x_2).
\end{aligned}$$

One can bound $\varphi(t_2,x_2)-\varphi(t_1,x_1)$ in exactly the same manner by exchanging the roles of (t_1,x_1) and (t_2,x_2) and replacing γ_0 by $\gamma_0(T_0-\cdot)$. □

Another important property of the value function is presented in the next proposition, and is useful for providing a lower bound on the time derivative of the value function when it exists.

Proposition 4.5. *Consider the optimal control problem* $\mathrm{OCP}(X,\Gamma,k)$ *and assume that Hypotheses 3.5(a)–(c) hold. There exists* $c>0$ *such that, for every* $x\in X$ *and* $t_1,t_2\in\mathbb{R}_+$ *with* $t_1\ne t_2$*, one has*

$$\frac{\varphi(t_2,x)-\varphi(t_1,x)}{t_2-t_1}\ge c-1. \tag{4.4}$$

In particular, if $t_1<t_2$*, then* $t_1+\varphi(t_1,x)<t_2+\varphi(t_2,x)$.

The last statement of the proposition means that, if two optimal trajectories start at the same point x on different times, the one which started sooner will arrive first at Γ.

Proof. It suffices to prove (4.4) in the case $t_2>t_1$, the other case being obtained by exchanging the role of t_1 and t_2. We then assume from now on that $t_2>t_1$. Notice also that, if (4.4) holds, then, for $t_2>t_1$, one has $\varphi(t_2,x)-\varphi(t_1,x)\ge(c-1)(t_2-t_1)>-(t_2-t_1)$, which yields $t_1+\varphi(t_1,x)<t_2+\varphi(t_2,x)$.

Let $\gamma_2 \in \mathrm{Opt}(k,t_2,x)$ and $\phi : [t_1,+\infty) \to \mathbb{R}$ be the unique function satisfying

$$\begin{cases} \dot{\phi}(t) = \dfrac{k(t,\gamma_2(\phi(t)))}{k(\phi(t),\gamma_2(\phi(t)))} \\ \phi(t_1) = t_2. \end{cases}$$

Notice that ϕ is strictly increasing and maps $[t_1,+\infty)$ onto $[t_2,+\infty)$, its inverse ϕ^{-1} being defined on $[t_2,+\infty)$. Moreover, since $\psi(t) = t$ is a solution of $\dot{\psi}(t) = \frac{k(t,\gamma_2(\psi(t)))}{k(\psi(t),\gamma_2(\psi(t)))}$ and $\phi(t_1) = t_2 > t_1 = \psi(t_1)$, one obtains that $\phi(t) > \psi(t) = t$ for every $t \in \mathbb{R}_+$.

Let $M > 0$ be as in Proposition 4.2(a) and $M_0 > 0$ be a Lipschitz constant for k. Notice that, for every $t \in \mathbb{R}_+$, one has

$$\dot{\phi}(t) - 1 = \frac{k(t,\gamma_2(\phi(t))) - k(\phi(t),\gamma_2(\phi(t)))}{k(\phi(t),\gamma_2(\phi(t)))}.$$

Fix $t_0 \in [t_1,t_1+M]$. Then, for every $t \in [t_1,t_0]$,

$$\phi(t_0) - t_0 = \phi(t) - t + \int_t^{t_0} \frac{k(s,\gamma_2(\phi(s))) - k(\phi(s),\gamma_2(\phi(s)))}{k(\phi(s),\gamma_2(\phi(s)))}\,\mathrm{d}s,$$

which yields

$$|\phi(t) - t| \le |\phi(t_0) - t_0| + \frac{M_0}{K_{\min}} \int_t^{t_0} |\phi(s) - s|\,\mathrm{d}s,$$

and thus, by Gronwall's inequality, for every $t \in [t_1,t_0]$,

$$|\phi(t) - t| \le |\phi(t_0) - t_0|\, e^{\frac{M_0}{K_{\min}}(t_0 - t)}.$$

In particular, using the fact that $\phi(t) > t$ for every $t \in \mathbb{R}_+$, one obtains from the above, setting $t_0 = t_1 + \varphi(t_1,x) \in [t_1,t_1+M]$ and $t = t_1$, that

$$\phi(t_1 + \varphi(t_1,x)) - (t_1 + \varphi(t_1,x)) \ge c(t_2 - t_1), \tag{4.5}$$

where $c = e^{-\frac{MM_0}{K_{\min}}}$.

Let $\gamma_1 \in \mathrm{Lip}(\mathbb{R}_+,X)$ be defined by $\gamma_1(t) = \gamma_2(\phi(t))$ for $t > t_1$ and $\gamma_1(t) = x$ for $t \le t_1$. Then, for almost every $t \in [t_1,+\infty)$, one has $|\dot{\gamma}_1|(t) = \left|\dot{\phi}(t)\right| |\dot{\gamma}_2|(\phi(t)) \le k(t,\gamma_1(t))$, which proves that $\gamma_1 \in \mathrm{Adm}(k)$. Moreover, $\gamma_1(t_1) = x$ and $\gamma_1(\phi^{-1}(t_2 + \varphi(t_2,x))) \in \Gamma$, which proves that the first exit time after t_1 of γ_1 satisfies $\tau(t_1,\gamma_1) \le \phi^{-1}(t_2 + \varphi(t_2,x))$, and hence $\varphi(t_1,x) \le \phi^{-1}(t_2 + \varphi(t_2,x)) - t_1$. Thus $\phi(t_1 + \varphi(t_1,x)) \le t_2 + \varphi(t_2,x)$, and, using (4.5), we get that $t_1 + \varphi(t_1,x) + c(t_2 - t_1) \le t_2 + \varphi(t_2,x)$, yielding (4.4). □

We gather in the next result two further properties of the value function and of optimal trajectories; the proof is straightforward and thus omitted here.

Proposition 4.6. *Consider the optimal control problem* $\mathrm{OCP}(X,\Gamma,k)$ *and assume that Hypotheses 3.5(a)–(c) hold. Let* $(t_0,x_0) \in \mathbb{R}_+ \times X$ *and* $\gamma \in \mathrm{Opt}(k,t_0,x_0)$.

(a) *For every* $h \in [0,\varphi(t_0,x_0)]$, *one has* $\varphi(t_0+h,\gamma(t_0+h)) + h = \varphi(t_0,x_0)$.

(b) *Let* $t_1 \in [t_0,t_0+\varphi(t_0,x_0)]$ *and consider the trajectory* $\widetilde{\gamma} : \mathbb{R}_+ \to X$ *defined by* $\widetilde{\gamma}(t) = \gamma(t)$ *for* $t > t_1$ *and* $\widetilde{\gamma}(t) = \gamma(t_1)$ *for* $t \le t_1$. *Then* $\widetilde{\gamma} \in \mathrm{Opt}(k,t_1,\gamma(t_1))$.

4.2. **Hamilton–Jacobi equation.** In this section, we collect further results on the value function φ under the additional assumption that Hypothesis 3.5(d) holds. Before establishing the main result of this section, Theorem 4.9, which provides a Hamilton–Jacobi equation for φ, we recall the definition of superdifferential of a function and obtain a lower bound on the time component of any vector on the superdifferential of φ as consequence of Proposition 4.5.

Definition 4.7. Let $A \subset \mathbb{R}^d$, $w : A \to \mathbb{R}$, and $x \in A$. The *superdifferential* of w at x is the set $D^+w(x)$ defined by

$$D^+w(x) = \left\{ p \in \mathbb{R}^d \;\middle|\; \limsup_{y \to x} \frac{w(y) - w(x) - p \cdot (y-x)}{|y-x|} \le 0 \right\}.$$

Lemma 4.8. *Consider the optimal control problem* $\mathrm{OCP}(X,\Gamma,k)$ *and assume that Hypotheses 3.5(a)–(d) hold. There exists* $c>0$ *such that, for every* $(t_0,x_0)\in\mathbb{R}_+^*\times\overline{\Omega}$*, and* $(p_0,p_1)\in D^+\varphi(t_0,x_0)$*, one has* $p_0\geq c-1$.

Proof. Let $c>0$ be as in Proposition 4.5. Since $(p_0,p_1)\in D^+\varphi(t_0,x_0)$, one has

$$\limsup_{\substack{t\to t_0\\ x\to x_0}}\frac{\varphi(t,x)-\varphi(t_0,x_0)-p_0(t-t_0)-p_1\cdot(x-x_0)}{\left(|t-t_0|^2+|x-x_0|^2\right)^{1/2}}\leq 0. \tag{4.6}$$

For $h>0$ small, one has, using Proposition 4.5,

$$\sup_{\substack{t\in(t_0-h,t_0+h)\\ x\in B(x_0,h)\\ (t,x)\neq(t_0,x_0)}}\frac{\varphi(t,x)-\varphi(t_0,x_0)-p_0(t-t_0)-p_1\cdot(x-x_0)}{\left(|t-t_0|^2+|x-x_0|^2\right)^{1/2}}$$
$$\geq\sup_{t\in(t_0,t_0+h)}\frac{\varphi(t,x_0)-\varphi(t_0,x_0)-p_0(t-t_0)}{t-t_0}\geq c-1-p_0,$$

and thus, taking the limit as $h\searrow 0$, one obtains from (4.6) that $c-1-p_0\leq 0$. □

Our next result provides a Hamilton–Jacobi equation for φ. Its proof is based on classical techniques on optimal control and is omitted here (see, e.g., [7, Chapter IV, Proposition 2.3]).

Theorem 4.9. *Consider the optimal control problem* $\mathrm{OCP}(X,\Gamma,k)$*, assume that Hypotheses 3.5(a)–(d) hold, and let* Ω *be as in Hypothesis 3.5(d) and* φ *be the value function from Definition 4.1. Consider the Hamilton–Jacobi equation on* $\mathbb{R}_+\times\overline{\Omega}$

$$-\partial_t\varphi(t,x)+|\nabla_x\varphi(t,x)|\,k(t,x)-1=0. \tag{4.7}$$

Then φ *is a viscosity subsolution of* (4.7) *on* $\mathbb{R}_+\times\Omega$*, a viscosity supersolution of* (4.7) *on* $\mathbb{R}_+\times(\overline{\Omega}\setminus\Gamma)$*, and satisfies* $\varphi(t,x)=0$ *for* $(t,x)\in\mathbb{R}_+\times\Gamma$.

Proposition 4.5 yields a lower bound on the time derivative of φ, which can be used to obtain information on the gradient of φ and on the optimal control of optimal trajectories thanks to the Hamilton–Jacobi equation (4.7).

Corollary 4.10. *Consider the optimal control problem* $\mathrm{OCP}(X,\Gamma,k)$*, assume that Hypotheses 3.5 (a)–(d) hold, and let* Ω *be as in Hypothesis 3.5(d) and* $(t_0,x_0)\in\mathbb{R}_+\times\overline{\Omega}$.

(a) *If* $x_0\notin\Gamma$ *and* φ *is differentiable at* (t_0,x_0)*, then* $\partial_t\varphi(t,x)>-1$ *and* $\nabla_x\varphi(t,x)\neq 0$.
(b) *Let* $\gamma\in\mathrm{Opt}(k,t_0,x_0)$ *and assume that* $t\in[t_0,t_0+\varphi(t_0,x_0))$ *is such that* γ *is differentiable at* t*,* $\gamma(t)\in\Omega$*, and* φ *is differentiable at* $(t,\gamma(t))$*. Then*

$$\dot\gamma(t)=-k(t,\gamma(t))\frac{\nabla_x\varphi(t,\gamma(t))}{|\nabla_x\varphi(t,\gamma(t))|}. \tag{4.8}$$

Proof. To prove (a), notice first that, for $(t_0,x_0)\in\mathbb{R}_+\times(\overline{\Omega}\setminus\Gamma)$ at which φ is differentiable, it follows from (4.4) that $\partial_t\varphi(t_0,x_0)\geq c-1>-1$. Since φ is a viscosity supersolution of (4.7) on $\mathbb{R}_+\times(\overline{\Omega}\setminus\Gamma)$, then

$$-\partial_t\varphi(t_0,x_0)+|\nabla_x\varphi(t_0,x_0)|\,k(t_0,x_0)-1\geq 0$$

(see, e.g., [37, Corollary I.6]). Since $\partial_t\varphi(t_0,x_0)>-1$ and $k(t_0,x_0)\geq K_{\min}>0$, one obtains that $|\nabla_x\varphi(t_0,x_0)|>0$, i.e., $\nabla_x\varphi(t_0,x_0)\neq 0$.

In order to prove (b), notice that, since $\gamma\in\mathrm{Adm}(k)$ and γ is differentiable at t, we can write $\dot\gamma(t)=k(t,\gamma(t))u$ for a certain $u\in\overline{B}_d$. Applying Proposition 4.6(a), one obtains, for small h,

$$\varphi(t+h,\gamma(t+h))+h=\varphi(t,\gamma(t)).$$

Differentiating with respect to h at $h=0$ yields

$$\partial_t\varphi(t,\gamma(t))+\nabla_x\varphi(t,\gamma(t))\cdot\dot\gamma(t)+1=0.$$

On the other hand, since φ is differentiable at $(t,\gamma(t))$, (4.7) holds pointwisely at $(t,\gamma(t))$ (see, e.g., [37, Corollary I.6]). Then, comparing the two expressions, we obtain

$$\nabla_x\varphi(t,\gamma(t))\cdot\dot\gamma(t)+|\nabla_x\varphi(t,\gamma(t))|\,k(t,\gamma(t))=0.$$

Using $\dot\gamma(t)=k(t,\gamma(t))u$ and $k(t,\gamma(t))>0$, we get

$$\nabla_x\varphi(t,\gamma(t))\cdot u=-\,|\nabla_x\varphi(t,\gamma(t))|\,.$$

Since $u\in\overline{B}_d$ and, by (a), $\nabla_x\varphi(t,\gamma(t))\neq 0$, this implies that

$$u=-\frac{\nabla_x\varphi(t,\gamma(t))}{|\nabla_x\varphi(t,\gamma(t))|},$$

and thus (4.8) holds. □

Notice that, since φ is Lipschitz continuous, one obtains as a consequence of Corollary 4.10(a) that $\partial_t\varphi(t,x)>-1$ and $\nabla_x\varphi(t,x)\neq 0$ for almost every $(t,x)\in\mathbb{R}_+\times\Omega$.

Definition 4.11. Consider the optimal control problem $\mathrm{OCP}(X,\Gamma,k)$ and assume that Hypothesis 3.5(d) holds. Given $(t_0,x_0)\in\mathbb{R}_+\times\overline{\Omega}$ and $\gamma\in\mathrm{Opt}(k,t_0,x_0)$, we say that a measurable function $u:\mathbb{R}_+\to\overline{B}_d$ is an *optimal control* associated with γ if $\dot\gamma(t)=k(t,\gamma(t))u(t)$ for almost every $t\in\mathbb{R}_+$.

In terms of Definition 4.11, Corollary 4.10(b) states that any optimal control u associated with an optimal trajectory γ satisfies $u(t)=-\frac{\nabla_x\varphi(t,\gamma(t))}{|\nabla_x\varphi(t,\gamma(t))|}$ whenever γ is differentiable at $t\in[t_0,t_0+\varphi(t_0,x_0))$, $\gamma(t)\in\Omega$, and φ is differentiable at $(t,\gamma(t))$. Even though φ and γ are both Lipschitz continuous, and hence differentiable almost everywhere, φ may be nowhere differentiable along a given optimal trajectory when $d\geq 2$, and thus Corollary 4.10(b) is not sufficient to characterize the optimal control for every optimal trajectory.

4.3. **Consequences of Pontryagin Maximum Principle.** As a first step towards providing a characterization of the optimal control associated with an optimal trajectory $\gamma\in\mathrm{Opt}(k,t_0,x_0)$, we apply Pontryagin Maximum Principle to $\mathrm{OCP}(X,\Gamma,k)$ to obtain a relation between the optimal control and the costate from Pontryagin Principle and deduce a differential inclusion for the optimal control. To do so, we will assume the more restrictive Hypothesis 3.5(e) on the target set Γ. Notice that, under Hypothesis 3.5(d), $\mathrm{OCP}(\overline{\Omega},\Gamma,k)$ is an optimal control problem with the state constraint $\gamma(t)\in\overline{\Omega}$ for every $t\in\mathbb{R}_+$, but such a state constraint becomes redundant when one assumes that $\Gamma=\partial\Omega$, since optimal trajectories starting at $\overline{\Omega}$ stop as soon as they reach the target set $\partial\Omega$, meaning that they will automatically always remain in $\overline{\Omega}$.

Even though versions of Pontryagin Maximum Principle for non-smooth dynamics and state constraints are available [35, Theorem 5.2.3], as well as techniques for adapting the unconstrained maximum principle to the constrained case [20], which could be used to study $\mathrm{OCP}(\overline{\Omega},\Gamma,k)$ without Hypothesis 3.5(e), we prefer to state its conclusions under Hypothesis 3.5(e) for simplicity since this assumption will be needed in the sequel in Section 4.4. However, notice the need for non-smooth statements of the Pontryagin Maximum Principle (which involve differential inclusions), the reason being that we are not assuming Hypothesis 3.5(g).

In the next result, $\pi_t:\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}$ and $\pi_x:\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}^d$ denote the canonical projections onto the factors of the product $\mathbb{R}\times\mathbb{R}^d$.

Proposition 4.12. *Consider the optimal control problem* $\mathrm{OCP}(X,\Gamma,k)$*, assume that Hypotheses 3.5(a)–(f) hold, and let* Ω *be as in Hypothesis 3.5(d). Let* $(t_0,x_0)\in\mathbb{R}_+\times\overline{\Omega}$*,* $\gamma\in\mathrm{Opt}(k,t_0,x_0)$*,* $T=\varphi(t_0,x_0)$*, and* $u:\mathbb{R}_+\to\overline{B}_d$ *be a measurable optimal control associated with* γ*. Then there exist* $\lambda\in\{0,1\}$ *and absolutely continuous functions* $p:[t_0,t_0+T]\to\mathbb{R}^d$ *and* $h:[t_0,t_0+T]\to\mathbb{R}$ *such that*

(a) *For almost every* $t\in[t_0,t_0+T]$*,*

$$\begin{cases}\dot h(t)\in|p(t)|\,\pi_t\partial^{\mathrm{C}}k(t,\gamma(t)),\\ \dot p(t)\in-\,|p(t)|\,\pi_x\partial^{\mathrm{C}}k(t,\gamma(t)).\end{cases}\tag{4.9}$$

(b) *One has*

$$u(t) = \frac{p(t)}{|p(t)|}$$

almost everywhere on $\{t \in [t_0, t_0+T] \mid p(t) \neq 0\}$.

(c) *For almost every* $t \in [t_0, t_0+T]$, *one has* $h(t) = |p(t)|\, k(t, \gamma(t)) - \lambda$, *and* $h(t_0+T) = 0$.

(d) $-p(t_0+T) \in N_{\partial\Omega}(\gamma(t_0+T))$.

(e) $\lambda + \max_{t\in[t_0,t_0+T]} |p(t)| > 0$.

Proposition 4.12 can be obtained from [35, Theorem 5.2.3] using a classical technique of state augmentation to transform (3.2) into an autonomous control system on the augmented state variable $z = (t, \gamma)$.

Notice that Proposition 4.12(b) characterizes the optimal control in terms of the costate p whenever the costate is non-zero. Our next result states that this happens everywhere on $[t_0, t_0+T]$.

Lemma 4.13. *Consider the optimal control problem* $\mathrm{OCP}(X, \Gamma, k)$, *assume that Hypotheses 3.5(a)–(f) hold, and let* Ω, (t_0, x_0), γ, T, λ, p, *and* h *be as in the statement of Proposition 4.12. Then* $p(t) \neq 0$ *for every* $t \in [t_0, t_0+T]$.

Proof. Let M be a Lipschitz constant for k and $\beta : [t_0, t_0+T] \to \mathbb{R}^d$ be a measurable function such that $\beta(t) \in \pi_x \partial^{\mathrm{C}} k(t, \gamma(t))$ and $\dot{p}(t) = -|p(t)|\, \beta(t)$ for almost every $t \in [t_0, t_0+T]$. Then $|\beta(t)| \leq M$ for almost every $t \in [t_0, t_0+T]$ (see, e.g., [35, Proposition 2.1.2]) and thus, for every $t, t_1 \in [t_0, t_0+T]$,

$$|p(t)| \leq |p(t_1)| + M \int_{\min\{t_1,t\}}^{\max\{t_1,t\}} |p(s)|\, \mathrm{d}s.$$

Hence, by Gronwall's inequality, for every $t, t_1 \in [t_0, t_0+T]$,

$$|p(t)| \leq |p(t_1)|\, e^{M|t-t_1|}.$$

One then concludes that, if there exists $t_1 \in [t_0, t_0+T]$ such that $p(t_1) = 0$, then $p(t) = 0$ for every $t \in [t_0, t_0+T]$. Thus, by Proposition 4.12(c), one has $h(t) = -\lambda$ for every $t \in [t_0, t_0+T]$, and, since $h(t_0+T) = 0$, it follows that $\lambda = 0$, contradicting Proposition 4.12(e). Thus $p(t) \neq 0$ for every $t \in [t_0, t_0+T]$. □

Combining Lemma 4.13 with the differential inclusion for p from (4.9), one obtains the following differential inclusion for u.

Corollary 4.14. *Consider the optimal control problem* $\mathrm{OCP}(X, \Gamma, k)$, *assume that Hypotheses 3.5(a)–(f) hold, and let* Ω, (t_0, x_0), γ, T, *and* u *be as in the statement of Proposition 4.12. Then* $\gamma \in \mathcal{C}^{1,1}([t_0, t_0+T], \mathbb{R}^d)$, $u \in \mathrm{Lip}([t_0, t_0+T], \mathbb{S}^{d-1})$, *and* (γ, u) *solves the system*

$$\begin{cases} \dot{\gamma}(t) = k(t, \gamma(t))u(t), \\ \dot{u}(t) \in -\mathrm{Pr}^{\perp}_{u(t)}\, \pi_x \partial^{\mathrm{C}} k(t, \gamma(t)), \end{cases} \tag{4.10}$$

where, for $x \in \mathbb{S}^{d-1}$, $\mathrm{Pr}^{\perp}_x : \mathbb{R}^d \to T_x\mathbb{S}^{d-1}$ *is the projection of* v *onto the tangent space to* x *of* $\mathbb{S}^{d-1}$, *defined for* $v \in \mathbb{R}^d$ *by* $\mathrm{Pr}^{\perp}_x v = v - (x \cdot v)x$.

Proof. Let p be as in the statement of Proposition 4.12. Thanks to Lemma 4.13, it follows from Proposition 4.12(b) that $u(t) = \frac{p(t)}{|p(t)|}$, and in particular u is absolutely continuous and takes values in $\mathbb{S}^{d-1}$ for every $t \in [t_0, t_0+T]$. Let $\beta : [t_0, t_0+T] \to \mathbb{R}^d$ be a measurable function such that $\beta(t) \in \pi_x \partial^{\mathrm{C}} k(t, \gamma(t))$ and $\dot{p}(t) = -|p(t)|\, \beta(t)$ for almost every $t \in [t_0, t_0+T]$. Then, for almost every $t \in [t_0, t_0+T]$,

$$\dot{u}(t) = \frac{\dot{p}(t)\, |p(t)| - \frac{p(t)\cdot\dot{p}(t)}{|p(t)|} p(t)}{|p(t)|^2} = -\beta(t) + [u(t) \cdot \beta(t)]\, u(t),$$

which yields the differential inclusion for u in (4.10). Since k is Lipschitz continuous, one obtains that $|\dot{u}(t)|$ is bounded for almost every $t \in [t_0, t_0+T]$, and thus u is Lipschitz continuous. It follows

from the differential equation on γ in (4.10) that $\dot\gamma$ is Lipschitz continuous on $[t_0, t_0+T]$, and hence $\gamma \in \mathcal{C}^{1,1}([t_0, t_0+T], \mathbb{R}^d)$. □

Corollary 4.14 states that every optimal trajectory satisfies, together with its optimal control, (4.10). However, given $(t_0, x_0) \in \mathbb{R}_+ \times \overline{\Omega}$, $u_0 \in \mathbb{S}^{d-1}$, a solution (γ, u) of (4.10) with initial condition $\gamma(t_0) = x_0$ and $u(t_0) = u_0$ may not yield an optimal trajectory. In order to understand when a solution of (4.10) is an optimal trajectory, we introduce the following definition.

Definition 4.15. Let $(t_0, x_0) \in \mathbb{R}_+ \times \Omega$. We define the set $\mathcal{U}_{(t_0,x_0)}$ of *optimal directions* at (t_0, x_0) as the set of all $u_0 \in \mathbb{S}^{d-1}$ such that there exists a solution (γ, u) of (4.10) with $\gamma(t_0) = x_0$ and $u(t_0) = u_0$ satisfying $\gamma \in \mathrm{Opt}(k, t_0, x_0)$.

Thanks to Proposition 4.2 and Corollary 4.14, $\mathcal{U}_{(t_0,x_0)}$ is non-empty. Our next result shows that, along an optimal trajectory γ, $\mathcal{U}_{(t,\gamma(t))}$ is a singleton, except possibly at its initial and final points.

Proposition 4.16. *Consider the optimal control problem* $\mathrm{OCP}(X, \Gamma, k)$ *and assume that Hypotheses 3.5(a)–(f) hold. Let* $(t_0, x_0) \in \mathbb{R}_+ \times \Omega$, $\gamma \in \mathrm{Opt}(k, t_0, x_0)$. *Then, for every* $t \in (t_0, t_0 + \varphi(t_0, x_0))$, $\mathcal{U}_{(t,\gamma(t))}$ *contains exactly one element.*

Proof. In order to simplify the notations, set $T = \varphi(t_0, x_0)$, fix $t_1 \in (t_0, t_0+T)$ and let $x_1 = \gamma(t_1)$. Let u be an optimal control associated with γ. By Corollary 4.14, $u \in \mathrm{Lip}([t_0, t_0+T], \mathbb{S}^{d-1})$. Thanks to Proposition 4.6(b), $u(t_1) \in \mathcal{U}_{(t_1,x_1)}$; suppose to have another element $u_1 \in \mathcal{U}_{(t_1,x_1)}$: we will prove that $u_1 = u(t_1)$.

Since $u_1 \in \mathcal{U}_{(t_1,x_1)}$, there exists a solution $(\widetilde\gamma, \widetilde u)$ of (4.10) with $\widetilde\gamma(t_1) = x_1$, $\widetilde u(t_1) = u_1$, and $\widetilde\gamma \in \mathrm{Opt}(k, t_1, x_1)$. Let $(\widehat\gamma, \widehat u)$ be defined on $\mathbb{R}_+$ by

$$\widehat\gamma(t) = \begin{cases} \gamma(t), & \text{if } t < t_1, \\ \widetilde\gamma(t), & \text{if } t \geq t_1, \end{cases} \qquad \widehat u(t) = \begin{cases} u(t), & \text{if } t < t_1, \\ \widetilde u(t), & \text{if } t \geq t_1. \end{cases}$$

We can see that $\widehat\gamma \in \mathrm{Opt}(k, t_0, x_0)$ and $\widehat u$ is an optimal control associated with $\widehat\gamma$. Hence, by Corollary 4.14, $\widehat\gamma \in \mathcal{C}^{1,1}([t_0, t_0+T], \mathbb{R}^d)$ and $\widehat u \in \mathrm{Lip}([t_0, t_0+T], \mathbb{S}^{d-1})$. In particular, $\widehat u$ is continuous, which proves that $u(t_1) = \widetilde u(t_1) = u_1$ and that $u(t_1)$ is the unique element of $\mathcal{U}_{(t_1,x_1)}$. □

4.4. **Normalized gradient and the optimal control.** In this section, we are interested in the additional properties one gets when assuming Hypothesis 3.5(g), the main goal being to characterize the optimal control u in a similar way to Corollary 4.10(b) but for all optimal trajectories and all times. A first result obtained from the extra regularity from Hypothesis 3.5(g) is the following immediate consequence of Corollary 4.14.

Corollary 4.17. *Consider the optimal control problem* $\mathrm{OCP}(X, \Gamma, k)$, *assume that Hypotheses 3.5(a)–(g) hold, and let* Ω, (t_0, x_0), γ, T, *and* u *be as in the statement of Proposition 4.12. Then* $\gamma \in \mathcal{C}^{2,1}([t_0, t_0+T], \mathbb{R}^d)$, $u \in \mathcal{C}^{1,1}([t_0, t_0+T], \mathbb{S}^{d-1})$, *and* (γ, u) *solves the system*

$$\begin{cases} \dot\gamma(t) = k(t, \gamma(t))u(t), \\ \dot u(t) = -\mathrm{Pr}^{\perp}_{u(t)} \nabla_x k(t, \gamma(t)). \end{cases} \tag{4.11}$$

Thanks to the extra regularity assumption from Hypothesis 3.5(g) and to Hypothesis 3.5(h), one may also use classical results in optimal control to conclude that φ is a semiconcave function.

Proposition 4.18. *Consider the optimal control problem* $\mathrm{OCP}(X, \Gamma, k)$ *and assume that Hypotheses 3.5(a)–(h) hold. Then* φ *is semiconcave in* $\mathbb{R}_+ \times \Omega$.

Proof. This is a classical fact and we will use [21, Theorem 8.2.7]. This Theorem is stated on the whole time-space $\mathbb{R} \times \mathbb{R}^d$, and requires an autonomous control system.

Hence, we will set $\mathcal{K} = \mathbb{R}^d \setminus \Omega$: $\mathcal{K}$ is closed and, due to the uniform exterior sphere property of Ω, $\mathcal{K}$ satisfies a uniform interior sphere property. Consider the optimal control problem of reaching $\mathcal{K}$ in minimal time with dynamics (3.2), starting at time $t \in \mathbb{R}$ from a point $x \in \mathbb{R}^d$, and

denote by $\widetilde{\varphi}$ its value function (the only difference from $\mathrm{OCP}(\overline{\Omega},\partial\Omega,k)$ is that we allow negative initial times). Of course, $\widetilde{\varphi}$ and φ coincide on $\mathbb{R}_+\times\overline{\Omega}$.

We rewrite (3.2) as an autonomous control system on the variable $z=(t,\gamma)$ and apply [21, Theorem 8.2.7] to the corresponding autonomous optimal control problem with target set $\mathbb{R}\times\mathcal{K}$ to conclude that $\widetilde{\varphi}$ is locally semiconcave on $\mathbb{R}\times\Omega$. Notice that, since one has an upper bound on the exit time for all optimal trajectories thanks to Proposition 4.2(a) (and its immediate generalization to the optimal control problem for $\widetilde{\varphi}$), it follows from the proof of [21, Theorem 8.2.7] that $\widetilde{\varphi}$ is semiconcave on $\mathbb{R}\times\Omega$. Hence, φ is semiconcave on $\mathbb{R}_+\times\Omega$. □

Definition 4.19. Let $A\subset\mathbb{R}^d$, $B\subset\mathbb{R}$.

(a) Let $w:A\to\mathbb{R}$ and $x\in A$. We say that $p\in\mathbb{R}^d$ is a *reachable gradient* of w at x if there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in $A\setminus\{x\}$ such that, for every $n\in\mathbb{N}$, w is differentiable at x_n, and $x_n\to x$ and $Dw(x_n)\to p$ as $n\to\infty$. The set of all reachable gradients of w at x is denoted by $D^*w(x)$.

(b) Let $w:B\times A\to\mathbb{R}$ and $(t,x)\in B\times A$. We say that $\omega\in\mathbb{S}^{d-1}$ is the *normalized gradient* of w with respect to x at (t,x) if $D^+w(t,x)$ is non-empty and, for every $(p_0,p_1)\in D^+w(t,x)$, one has $p_1\neq 0$ and $\frac{p_1}{|p_1|}=\omega$. In this case, ω is denoted by $\widehat{\nabla w}(t,x)$.

Recall that, for a semiconcave function w defined on an open set $A\subset\mathbb{R}^d$, for every $x\in A$, one has $D^+w(x)=\mathrm{co}\,D^*w(x)=\partial^{\mathrm{C}}w(x)$ and these sets are non-empty (see, e.g., [21, Proposition 3.3.4 and Theorem 3.3.6]). Here, $\mathrm{co}\,B$ denotes the convex hull of B.

Theorem 4.20. *Consider the optimal control problem* $\mathrm{OCP}(X,\Gamma,k)$ *and assume that Hypotheses 3.5(a)–(h) hold. Let* $(t_0,x_0)\in\mathbb{R}_+^*\times\Omega$. *Then* φ *admits a normalized gradient at* (t_0,x_0) *if and only if* $\mathcal{U}_{(t_0,x_0)}$ *contains only one element. In this case,* $\mathcal{U}_{(t_0,x_0)}=\{-\widehat{\nabla\varphi}(t_0,x_0)\}$.

Proof. Notice first that, thanks to Proposition 4.18, φ is semiconcave in $\mathbb{R}_+\times\Omega$. Assume that φ admits a normalized gradient at (t_0,x_0) and take $u_0\in\mathcal{U}_{(t_0,x_0)}$. Let (γ,u) be a solution of (4.10) with $\gamma(t_0)=x_0$, $u(t_0)=u_0$, and $\gamma\in\mathrm{Opt}(k,t_0,x_0)$. By Proposition 4.6(a), for every $h\in[0,\varphi(t_0,x_0)]$, one has

$$\varphi(t_0+h,\gamma(t_0+h))+h=\varphi(t_0,x_0).$$

Let $\psi:[0,\varphi(t_0,x_0)]\to\mathbb{R}_+\times\overline{\Omega}$ be given by $\psi(h)=(t_0+h,\gamma(t_0+h))$. Then $\varphi\circ\psi(h)+h=\varphi(t_0,x_0)$. Since $h\mapsto\varphi(t_0,x_0)-h$ is differentiable on $[0,\varphi(t_0,x_0)]$, with derivative equal to -1, one obtains that $\varphi\circ\psi$ is differentiable on $[0,\varphi(t_0,x_0)]$, with $(\varphi\circ\psi)'(h)=-1$ for every $h\in[0,\varphi(t_0,x_0)]$. Moreover, $\psi\in\mathcal{C}^1([0,\varphi(t_0,x_0)],\mathbb{R}_+\times\overline{\Omega})$, with $\psi'(h)=(1,\dot\gamma(t_0+h))$. Hence, by the chain rule for generalized gradients (see, e.g., [35, Theorem 2.3.10]), one has, for every $h\in(0,\varphi(t_0,x_0))$,

$$\{-1\}=\partial^{\mathrm{C}}(\varphi\circ\psi)(h)\subset\partial^{\mathrm{C}}\varphi(t_0+h,\gamma(t_0+h))\cdot(1,\dot\gamma(t_0+h)).$$

Thus there exist $(p_{0,h},p_{1,h})\in\partial^{\mathrm{C}}\varphi(t_0+h,\gamma(t_0+h))$ such that

$$-1=p_{0,h}+p_{1,h}\cdot\dot\gamma(t_0+h). \tag{4.12}$$

Notice moreover that, since φ is semiconcave, one has $(p_{0,h},p_{1,h})\in D^+\varphi(t_0+h,\gamma(t_0+h))$. Since φ is Lipschitz continuous, the family $((p_{0,h},p_{1,h}))_{h\in(0,\varphi(t_0,x_0))}$ is bounded, and thus there exists a sequence $(h_n)_{n\in\mathbb{N}}$ with $h_n\to 0$ as $n\to\infty$ and a point $(p_0,p_1)\in\mathbb{R}\times\mathbb{R}^d$ such that, as $n\to\infty$, $p_{0,h_n}\to p_0$ and $p_{1,h_n}\to p_1$. Moreover, $(p_0,p_1)\in D^+\varphi(t_0,x_0)$ (see, e.g., [21, Proposition 3.3.4(a)]). One also has that $\dot\gamma(t_0+h)=k(t_0+h,\gamma(t_0+h))u(t_0+h)$ and thus, taking $h=h_n$ in (4.12) and letting $n\to\infty$, one obtains that

$$-1=p_0+p_1\cdot u_0k(t_0,x_0). \tag{4.13}$$

Letting $c>0$ be as in Lemma 4.8, one obtains that $-p_1\cdot u_0k(t_0,x_0)\geq c$, and thus, in particular, $p_1\neq 0$. Since φ admits a normalized gradient at (t_0,x_0), one then concludes that $\frac{p_1}{|p_1|}=\widehat{\nabla\varphi}(t_0,x_0)$.

Now, since $(p_0,p_1)\in D^+\varphi(t_0,x_0)$, there exists a neighborhood V of (t_0,x_0) in $\mathbb{R}_+\times\Omega$ and a function $\phi\in\mathcal{C}^1(V,\mathbb{R})$ such that $\phi(t_0,x_0)=\varphi(t_0,x_0)$, $\phi(t,x)\geq\varphi(t,x)$ for every $(t,x)\in V$, and

$p_0 = \partial_t \phi(t_0,x_0)$, $p_1 = \nabla_x \phi(t_0,x_0)$. Since φ is a viscosity subsolution of (4.7) on $\mathbb{R}_+ \times \Omega$, one concludes that

$$-p_0 + |p_1| k(t_0,x_0) - 1 \leq 0.$$

Combining the above inequality with (4.13), one obtains that

$$|p_1| k(t_0,x_0) + p_1 \cdot u_0 k(t_0,x_0) \leq 0,$$

i.e., $p_1 \cdot u_0 \leq -|p_1|$. Since $u_0 \in \mathbb{S}^{d-1}$, this means that $u_0 = -\frac{p_1}{|p_1|} = -\widehat{\nabla\varphi}(t_0,x_0)$. Since this holds for every $u_0 \in \mathcal{U}_{(t_0,x_0)}$, one concludes that the unique element in $\mathcal{U}_{(t_0,x_0)}$ is $-\widehat{\nabla\varphi}(t_0,x_0)$.

Assume now that $\mathcal{U}_{(t_0,x_0)}$ contains only one element, which we denote by u_0. Since $D^+\varphi(t_0,x_0) = \operatorname{co} D^*\varphi(t_0,x_0)$ is non-empty, $D^*\varphi(t_0,x_0) \neq \emptyset$. Take $p = (p_0,p_1) \in D^*\varphi(t_0,x_0)$. Hence there exists a sequence $((t_{0,n},x_{0,n}))_{n\in\mathbb{N}}$ such that φ is differentiable at $(t_{0,n},x_{0,n})$ for every $n \in \mathbb{N}$ and $t_{0,n} \to t_0$, $x_{0,n} \to x_0$, and $D\varphi(t_{0,n},x_{0,n}) \to p$ as $n \to \infty$. For $n \in \mathbb{N}$, write $(p_{0,n},p_{1,n}) = D\varphi(t_{0,n},x_{0,n})$, and let $\gamma_n \in \operatorname{Opt}(k,t_{0,n},x_{0,n})$ and u_n be an optimal control associated with γ_n. By Proposition 4.5 and Corollary 4.10, there exists $c > 0$ such that, for every $n \in \mathbb{N}$, $|p_{1,n}| \geq c$ and $u_n(t_{0,n}) = -\frac{p_{1,n}}{|p_{1,n}|}$ In particular, $|p_1| \geq c$, and thus $p_1 \neq 0$. Using Corollary 4.17, one obtains that (γ_n,u_n) solves

$$\begin{cases} \dot\gamma_n(t) = k(t,\gamma_n(t))u_n(t), & t \in [t_{0,n}, t_{0,n} + \varphi(t_{0,n},x_{0,n})], \\ \dot u_n(t) = -\operatorname{Pr}^{\perp}_{u_n(t)} \nabla_x k(t,\gamma_n(t)), & t \in [t_{0,n}, t_{0,n} + \varphi(t_{0,n},x_{0,n})], \\ \gamma_n(t_{0,n}) = x_{0,n}, \\ u_n(t_{0,n}) = -\dfrac{p_{1,n}}{|p_{1,n}|}. \end{cases}$$

We modify u_n outside of the interval $[t_{0,n}, t_{0,n} + \varphi(t_{0,n},x_{0,n})]$ by setting $u_n(t) = u_n(t_{0,n})$ for $t \in [0,t_{0,n})$ and $u_n(t) = u_n(t_{0,n} + \varphi(t_{0,n},x_{0,n}))$ for $t > t_{0,n} + \varphi(t_{0,n},x_{0,n})$. We have $\gamma_n \in \operatorname{Lip}_{K_{\max}}(\overline{\Omega})$, $u_n \in \operatorname{Lip}_M(\mathbb{S}^{d-1})$, where M is a Lipschitz constant for k, and thus, by Arzelà–Ascoli Theorem, there exist subsequences, which we still denote by $(\gamma_n)_{n\in\mathbb{N}}$ and $(u_n)_{n\in\mathbb{N}}$ for simplicity, and elements $\gamma \in \operatorname{Lip}_{K_{\max}}(\overline{\Omega})$, $u \in \operatorname{Lip}_M(\mathbb{S}^{d-1})$ such that $\gamma_n \to \gamma$ and $u_n \to u$ uniformly on compact time intervals.

Since φ is Lipschitz continuous, one has $\varphi(t_{0,n},x_{0,n}) \to \varphi(t_0,x_0)$ as $n \to \infty$. Since $\gamma_n(t) = \gamma_n(t_{0,n})$ and $u_n(t) = u_n(t_{0,n})$ for $t \in [0,t_{0,n}]$ and $\gamma_n(t) = \gamma_n(t_{0,n} + \varphi(t_{0,n},x_{0,n}))$ and $u_n(t) = u_n(t_{0,n} + \varphi(t_{0,n},x_{0,n}))$ for $t \geq t_{0,n} + \varphi(t_{0,n},x_{0,n})$, one obtains that $\gamma(t) = \gamma(t_0)$ and $u(t) = u(t_0)$ for $t \in [0,t_0]$ and $\gamma(t) = \gamma(t_0 + \varphi(t_0,x_0))$, $u(t) = u(t_0 + \varphi(t_0,x_0))$ for $t \geq t_0 + \varphi(t_0,x_0)$. Since $\gamma_n(t_{0,n} + \varphi(t_{0,n},x_{0,n})) \in \partial\Omega$, one also obtains that $\gamma(t_0 + \varphi(t_0,x_0)) \in \partial\Omega$, and in particular its exit time satisfies $\tau(t_0,\gamma) \leq \varphi(t_0,x_0)$. From $\dot\gamma_n = k(\cdot,\gamma_n)u_n$, the local uniform convergence of $(\gamma_n)_{n\in\mathbb{N}}$ and $(u_n)_{n\in\mathbb{N}}$ provides, at the limit, $\dot\gamma = k(\cdot,\gamma)u$. Hence $\gamma \in \operatorname{Opt}(k,t_0,x_0)$ and u is an optimal control associated with γ.

Since $u(t_0) = \lim_{n\to\infty} u_n(t_{0,n}) = \lim_{n\to\infty}\left(-\frac{p_{1,n}}{|p_{1,n}|}\right) = -\frac{p_1}{|p_1|}$, one obtains that $-\frac{p_1}{|p_1|} \in \mathcal{U}_{(t_0,x_0)}$, and thus, by assumption, $-\frac{p_1}{|p_1|} = u_0$. We have thus proved that, for every $p = (p_0,p_1) \in D^*\varphi(t_0,x_0)$, one has $p_1 \neq 0$ and $-\frac{p_1}{|p_1|} = u_0$. Since $D^+\varphi(t_0,x_0) = \operatorname{co} D^*\varphi(t_0,x_0)$, one immediately checks that $p_1 \neq 0$ and $-\frac{p_1}{|p_1|} = u_0$ for every $p = (p_0,p_1) \in D^+\varphi(t_0,x_0)$, which proves that φ admits a normalized gradient at (t_0,x_0) and $-\widehat{\nabla\varphi}(t_0,x_0) = u_0$, i.e., $\mathcal{U}_{(t_0,x_0)} = \{-\widehat{\nabla\varphi}(t_0,x_0)\}$. □

As an immediate consequence of Proposition 4.16 and Theorem 4.20, one obtains the following characterization of the optimal control.

Corollary 4.21. *Consider the optimal control problem* $\mathrm{OCP}(X,\Gamma,k)$ *and assume that Hypotheses 3.5(a)–(h) hold. Let* $(t_0,x_0) \in \mathbb{R}_+ \times \Omega$, $\gamma \in \operatorname{Opt}(k,t_0,x_0)$, *and* u *be the optimal control associated with* γ. *Then, for every* $t \in (t_0, t_0 + \varphi(t_0,x_0))$, φ *admits a normalized gradient at* $(t,\gamma(t))$ *and* $u(t) = -\widehat{\nabla\varphi}(t,\gamma(t))$, *i.e.,*

$$\dot\gamma(t) = -k(t,\gamma(t))\widehat{\nabla\varphi}(t,\gamma(t)).$$

Corollary 4.21 can be seen as an improved version of Corollary 4.10(b), where we now characterize the optimal control for every optimal trajectory and every time, replacing the normalization of the gradient of φ by its normalized gradient.

Notice that, combining Corollaries 4.17 and 4.21, we obtain that $t \mapsto \widehat{\nabla\varphi}(t,\gamma(t))$ is $\mathcal{C}^{1,1}$ for every optimal trajectory γ, as long as $\gamma(t) \in \Omega$ and t is larger than the initial time of γ. However, this provides no information on the regularity of $(t,x) \mapsto \widehat{\nabla\varphi}(t,x)$. We are now interested in proving continuity of this map in a set where $\widehat{\nabla\varphi}$ is defined.

Set

$$\Upsilon = \{(t,x) \in \mathbb{R}_+^* \times \Omega \mid \exists t_0 \in [0,t),\ \exists x_0 \in \Omega,\ \exists \gamma \in \mathrm{Opt}(k,t_0,x_0) \text{ s.t. } \gamma(t) = x\}. \tag{4.14}$$

This set (which can be easily seen to be a countable union of compact sets, and in particular a Borel measurable set) contains all points $(t,x) \in \mathbb{R}_+^* \times \Omega$ which are not starting points of optimal trajectories. In particular, it follows from Corollary 4.21 that $\widehat{\nabla\varphi}(t,x)$ exists for every $(t,x) \in \Upsilon$.

Proposition 4.22. *Consider the optimal control problem* $\mathrm{OCP}(X,\Gamma,k)$ *and assume that Hypotheses 3.5(a)–(h) hold. Then the normalized gradient* $\widehat{\nabla\varphi}$ *of* φ *is continuous on* Υ.

Proof. Let $(t_*,x_*) \in \Upsilon$ and $((t_n,x_n))_{n\in\mathbb{N}}$ be a sequence in Υ converging to (t_*,x_*) as $n \to \infty$. For $n \in \mathbb{N}$, let $t_{0,n} \in [0,t_n)$, $x_{0,n} \in \Omega$ and $\gamma_n \in \mathrm{Opt}(k,t_{0,n},x_{0,n})$ be such that $\gamma_n(t_n) = x_n$. Let u_n be an optimal control associated with γ_n, and modify u_n outside of $[t_{0,n}, t_{0,n} + \varphi(t_{0,n},x_{0,n})]$ for u_n to be constant on $[0,t_{0,n}]$ and on $[t_{0,n} + \varphi(t_{0,n},x_{0,n}), +\infty)$. Notice that, since $x_n \in \Omega$, one has $t_n \in (t_{0,n}, t_{0,n} + \varphi(t_{0,n},x_{0,n}))$ and, by Corollary 4.21, $u_n(t_n) = -\widehat{\nabla\varphi}(t_n,x_n)$. We want to prove that $u_n(t_n) \to -\widehat{\nabla\varphi}(t_*,x_*)$ as $n \to \infty$. Since $(u_n(t_n))_{n\in\mathbb{N}}$ is bounded, it is enough to prove that every convergent subsequence of $(u_n(t_n))_{n\in\mathbb{N}}$ converges to $-\widehat{\nabla\varphi}(t_*,x_*)$ as $n \to \infty$.

Assume that $(u_{n_k}(t_{n_k}))_{k\in\mathbb{N}}$ is a converging subsequence of $(u_n(t_n))_{n\in\mathbb{N}}$, and denote by u_* its limit. Up to subsequences (using Arzelà–Ascoli Theorem together with Corollary 4.14), we can assume that γ_{n_k} and u_{n_k} also converge, uniformly on every compact time interval, and that $(t_{0,n_k})_{k\in\mathbb{N}}$ and $(x_{0,n_k})_{k\in\mathbb{N}}$ also converge. Let $\gamma = \lim_{k\to\infty} \gamma_{n_k}$, $u = \lim_{k\to\infty} u_{n_k}$, $t_0 = \lim_{k\to\infty} t_{0,n_k}$, and $x_0 = \lim_{k\to\infty} x_{0,k}$.

Since $\gamma_n(t) = x_{0,n}$ for $t \in [0,t_{0,n}]$, $u_n(t) = u_n(t_{0,n})$ for $t \in [0,t_{0,n}]$, $\gamma_n(t) = \gamma_n(t_{0,n} + \varphi(t_{0,n},x_{0,n}))$ for $t \geq t_{0,n} + \varphi(t_{0,n},x_{0,n})$, and $u_n(t) = u_n(t_{0,n} + \varphi(t_{0,n},x_{0,n}))$ for $t \geq t_{0,n} + \varphi(t_{0,n},x_{0,n})$, one obtains that $\gamma(t) = x_0$ for $t \in [0,t_0]$, $u(t) = u(t_0)$ for $t \in [0,t_0]$, $\gamma(t) = \gamma(t_0 + \varphi(t_0,x_0))$ for $t \geq t_0 + \varphi(t_0,x_0)$, and $u(t) = u(t_0 + \varphi(t_0,x_0))$ for $t \geq t_0 + \varphi(t_0,x_0)$, which proves in particular that $\tau(t_0,\gamma) \leq \varphi(t_0,x_0)$, and one will obtain that $\gamma \in \mathrm{Opt}(k,t_0,x_0)$ as soon as one proves that $\gamma \in \mathrm{Adm}(k)$. This can be done exactly as in the proof of Theorem 4.20, also obtaining that u is an optimal control associated with γ. Notice moreover that $\gamma(t_*) = \lim_{k\to\infty} \gamma_{n_k}(t_{n_k}) = \lim_{k\to\infty} x_{n_k} = x_*$.

Since $(t_*,x_*) \in \Upsilon$, there exist $t_{0,*} \in [0,t_*)$, $x_{0,*} \in \Omega$, and $\gamma_* \in \mathrm{Opt}(k,t_{0,*},x_{0,*})$ such that $\gamma_*(t_*) = x_*$. Let u_* be an optimal control associated with γ_*, assumed to be constant on $[0,t_{0,*}]$ and on $[t_{0,*} + \varphi(t_{0,*},x_{0,*}), +\infty)$. Since $x_* \in \Omega$, one has $t_* \in (t_{0,*}, t_{0,*} + \varphi(t_{0,*},x_{0,*}))$, and one also has, from Corollary 4.21, that $u_*(t_*) = -\widehat{\nabla\varphi}(t_*,x_*)$. Define $\widetilde{\gamma} : \mathbb{R}_+ \to \overline{\Omega}$ and $\widetilde{u} : \mathbb{R}_+ \to \mathbb{S}^{d-1}$ for $t \in \mathbb{R}_+$ by

$$\widetilde{\gamma}(t) = \begin{cases} \gamma_*(t), & \text{if } t < t_*, \\ \gamma(t), & \text{if } t \geq t_*, \end{cases} \qquad \widetilde{u}(t) = \begin{cases} u_*(t), & \text{if } t < t_*, \\ u(t), & \text{if } t \geq t_*. \end{cases}$$

Since $\gamma_* \in \mathrm{Opt}(k,t_{0,*},x_{0,*})$, $\gamma \in \mathrm{Opt}(k,t_0,x_0)$, and $\gamma_*(t_*) = \gamma(t_*) = x_*$, one obtains that $\widetilde{\gamma} \in \mathrm{Opt}(k,t_{0,*},x_{0,*})$, and $\widetilde{u}$ is an optimal control associated with $\widetilde{\gamma}$. Hence, by Corollary 4.17, the restrictions of $\widetilde{\gamma}$ and $\widetilde{u}$ to $[t_{0,*}, t_{0,*} + \varphi(t_{0,*},x_{0,*})]$, still denoted by $\widetilde{\gamma}$ and $\widetilde{u}$ for simplicity, satisfy $\widetilde{\gamma} \in \mathcal{C}^{2,1}([t_{0,*}, t_{0,*} + \varphi(t_{0,*},x_{0,*})], \overline{\Omega})$ and $\widetilde{u} \in \mathcal{C}^{1,1}([t_{0,*}, t_{0,*} + \varphi(t_{0,*},x_{0,*})], \mathbb{S}^{d-1})$. In particular, since $t_* \in (t_{0,*}, t_{0,*} + \varphi(t_{0,*},x_{0,*}))$, $\widetilde{u}$ is continuous at t_*, which finally proves that $-\widehat{\nabla\varphi}(t_*,x_*) = u_*(t_*) = u(t_*) = \lim_{k\to\infty} u_{n_k}(t_{n_k}) = u_*$. $\square$

5. Existence of equilibria for minimal-time MFG

After the preliminary study of the optimal control problem $\mathrm{OCP}(X,\Gamma,k)$ from Section 4, we finally turn to our main problem, the mean field game $\mathrm{MFG}(X,\Gamma,K)$. This section is concerned with the existence of equilibria for $\mathrm{MFG}(X,\Gamma,K)$. We shall need here only some elementary results on the value function from Section 4.1, and our main result is the following.

Theorem 5.1. *Consider the mean field game* $\mathrm{MFG}(X,\Gamma,K)$ *and assume that Hypotheses 3.5(a), (c), and (i) hold. Then, for every* $m_0 \in \mathcal{P}(X)$*, there exists an equilibrium* $Q \in \mathcal{P}(\mathcal{C}_X)$ *for* $\mathrm{MFG}(X,\Gamma,K)$ *with initial condition* m_0*.*

The remainder of the section provides the proof of Theorem 5.1, which is decomposed in several steps. First, for a fixed $Q \in \mathcal{P}(\mathcal{C}_X)$, we consider in Section 5.1 the optimal control problem $\mathrm{OCP}(X,\Gamma,k)$ where k is given by $k(t,x) = K(e_{t\#}Q,x)$, proving first that k is Lipschitz continuous, in order to be able to apply the results from Section 4.1, and then that the value function φ is also Lipschitz continuous with respect to Q. Such properties are then used in Section 5.2 to prove that one may choose an optimal trajectory for each starting point in a measurable way. Finally, Section 5.3 characterizes equilibria of $\mathrm{MFG}(X,\Gamma,K)$ as fixed points of a suitable multi-valued function and concludes the proof of Theorem 5.1 by the application of a suitable fixed point theorem.

5.1. Lipschitz continuity of the value function.

Notice that, under Hypothesis 3.5(i), any equilibrium Q of $\mathrm{MFG}(X,\Gamma,K)$ should be concentrated on $K_{\max}$-Lipschitz continuous trajectories. Define

$$\mathcal{Q} = \{Q \in \mathcal{P}(\mathcal{C}_X) \mid Q(\mathrm{Lip}_{K_{\max}}(X)) = 1\}.$$

Notice that $\mathcal{Q}$ is convex, and one can prove by standard arguments that it is also tight and closed, obtaining, as a consequence of Prohorov's Theorem (see, e.g., [11, Chapter 1, Theorem 5.1]), that $\mathcal{Q}$ is compact. We now prove that the push-forward of a measure $Q \in \mathcal{Q}$ by the evaluation map e_t is Lipschitz continuous in both variables t and Q.

Proposition 5.2. *Assume that Hypothesis 3.5(a) holds.*

(a) *For every* $Q \in \mathcal{Q}$*, the map* $\mathbb{R}_+ \ni t \mapsto e_{t\#}Q \in \mathcal{P}(X)$ *is* $K_{\max}$*-Lipschitz continuous.*
(b) *For every* $t \in \mathbb{R}_+$*, the map* $\mathcal{P}(\mathcal{C}_X) \ni Q \mapsto e_{t\#}Q \in \mathcal{P}(X)$ *is* 2^{t+1}*-Lipschitz continuous.*

Proof. To prove (a), take $t,s \in \mathbb{R}_+$. By (2.1) and using the fact that $Q(\mathrm{Lip}_{K_{\max}}(X)) = 1$, we have

$$\begin{aligned} W_1(e_{t\#}Q, e_{s\#}Q) &= \sup_{\substack{\phi:X\to\mathbb{R}\\ \text{1-Lipschitz}}} \int_X \phi(x)\,\mathrm{d}(e_{t\#}Q - e_{s\#}Q)(x) \\ &= \sup_{\substack{\phi:X\to\mathbb{R}\\ \text{1-Lipschitz}}} \int_{\mathrm{Lip}_{K_{\max}}(X)} (\phi(\gamma(t)) - \phi(\gamma(s)))\,\mathrm{d}Q(\gamma) \le K_{\max}\,|t-s|\,, \end{aligned}$$

and thus $t \mapsto e_{t\#}Q$ is $K_{\max}$-Lipschitz continuous.

In order to prove (b), we just need to observe that, due to the definition of $\mathbf{d}_{\mathcal{C}_X}$, the evaluation map e_t is 2^n-Lipschitz continuous for every $n \in \mathbb{N}$ with $n \ge t$. Hence, this map is 2^{t+1}-Lipschitz continuous and this implies (b). □

As an immediate consequence of Proposition 5.2, we obtain the following.

Corollary 5.3. *Let* $Q \in \mathcal{Q}$ *and assume that Hypothesis 3.5(a) holds and* $K : \mathcal{P}(X) \times X \to \mathbb{R}_+$ *satisfies Hypothesis 3.5(i). Define* $k : \mathbb{R}_+ \times X \to \mathbb{R}_+$ *by* $k(t,x) = K(e_{t\#}Q,x)$*. Then* k *satisfies Hypothesis 3.5(b), i.e.,* k *is Lipschitz continuous on* $\mathbb{R}_+ \times X$ *and* $k(t,x) \in [K_{\min}, K_{\max}]$ *for every* $(t,x) \in \mathbb{R}_+ \times X$*.*

In particular, when k is constructed from K as in Corollary 5.3, the results of Section 4.1 apply to k. Notice that the value function φ from Definition 4.1 depends on Q through k. In the sequel, whenever needed, we make this dependence explicit in the notation of the value function by writing it as φ_Q.

Proposition 5.4. *Consider the mean field game* $\mathrm{MFG}(X,\Gamma,K)$ *and assume that Hypotheses 3.5(a), (c), and (i) hold. Let* $(t,x)\in\mathbb{R}_+\times X$*. Then the map* $\mathcal{Q}\ni Q\mapsto\varphi_Q(t,x)$ *is Lipschitz continuous in* $\mathcal{Q}$*, its Lipschitz constant being independent of* x *and a non-decreasing function of* t*.*

Proof. Consider $\mathrm{OCP}(X,\Gamma,k)$ with k defined from K as in Corollary 5.3. Let $M>0$ be as in Proposition 4.2(a). Let $Q_1,Q_2\in\mathcal{Q}$. Let $\gamma_1\in\mathrm{Opt}(Q_1,t,x)$ and define $\phi:[t,+\infty)\to\mathbb{R}$ as the unique function satisfying

$$\begin{cases}\dot\phi(s)=\dfrac{K(e_{s\#}Q_2,\gamma_1(\phi(s)))}{K(e_{\phi(s)\#}Q_1,\gamma_1(\phi(s)))},\\ \phi(t)=t.\end{cases}\tag{5.1}$$

Similarly to the proof of Proposition 4.3, ϕ is a well-defined $\mathcal{C}^1$ function whose derivative takes values in $\left[\frac{K_{\min}}{K_{\max}},\frac{K_{\max}}{K_{\min}}\right]$. Moreover, ϕ is strictly increasing and maps $[t,+\infty)$ onto itself, its inverse ϕ^{-1} being defined on $[t,+\infty)$. Define $\gamma_2:\mathbb{R}_+\to X$ by

$$\gamma_2(s)=\begin{cases}x, & \text{if } s\le t,\\ \gamma_1(\phi(s)), & \text{if } s>t.\end{cases}$$

Then $\gamma_2\in\mathrm{Lip}(\mathbb{R}_+,X)$, $\gamma_2(s)=x$ for every $s\le t$, $\gamma_2(s)=\gamma_2(\phi^{-1}(t+\varphi_{Q_1}(t,x)))\in\Gamma$ for every $s\ge\phi^{-1}(t+\varphi_{Q_1}(t,x))$, and $|\dot\gamma_2|(s)\le K(e_{s\#}Q_2,\gamma_2(s))$ for almost every $s\ge t$. Thus $\gamma_2\in\mathrm{Adm}(Q_2)$ and $t+\varphi_{Q_2}(t,x)\le\phi^{-1}(t+\varphi_{Q_1}(t,x))$, i.e., $\phi(t+\varphi_{Q_2}(t,x))\le t+\varphi_{Q_1}(t,x)$.

By Hypothesis 3.5(i) and Proposition 5.2, there exists $M_0>0$, independent of (t,x), such that, for every $s\in\mathbb{R}_+$,

$$\left|K(e_{s\#}Q_2,\gamma_1(\phi(s)))-K(e_{\phi(s)\#}Q_1,\gamma_1(\phi(s)))\right|\le M_0\,|\phi(s)-s|+2^{s+1}M_0W_1(Q_1,Q_2).$$

One has, from (5.1),

$$\phi(s)-s=\int_t^s\left[\frac{K(e_{\sigma\#}Q_2,\gamma_1(\phi(\sigma)))}{K(e_{\phi(\sigma)\#}Q_1,\gamma_1(\phi(\sigma)))}-1\right]\mathrm{d}\sigma,$$

and thus, for every $s\in[t,t+M]$,

$$|\phi(s)-s|\le\frac{1}{K_{\min}}\left[2^{t+M+1}MM_0W_1(Q_1,Q_2)+M_0\int_t^s|\phi(\sigma)-\sigma|\,\mathrm{d}\sigma\right],$$

which yields, by Gronwall's inequality, that, for every $s\in[t,t+M]$,

$$|\phi(s)-s|\le C_tW_1(Q_1,Q_2),$$

where C_t is non-decreasing on t. We then conclude that

$$t+\varphi_{Q_2}(t,x)\le\phi(t+\varphi_{Q_2}(t,x))+C_tW_1(Q_1,Q_2)\le t+\varphi_{Q_1}(t,x)+C_tW_1(Q_1,Q_2),$$

and finally

$$\varphi_{Q_2}(t,x)-\varphi_{Q_1}(t,x)\le C_tW_1(Q_1,Q_2).$$

By exchanging Q_1 and Q_2 in the above inequality, one obtains the Lipschitz continuity of $Q\mapsto\varphi_Q(t,x)$ with Lipschitz constant C_t. □

5.2. **Measurable selection of optimal trajectories.** Proposition 4.2(b) and Corollary 5.3 imply that, given $Q\in\mathcal{Q}$, for every $x\in X$, there exists an optimal trajectory γ for $(Q,0,x)$. Such a trajectory depends on x, and one may thus construct a map that, to each point $x\in X$, associates some optimal trajectory for $(Q,0,x)$. The main result of this section, Proposition 5.6, asserts that it is possible to construct such a map in a measurable way.

For $Q\in\mathcal{Q}$, let G_Q be the graph of the set-valued map $x\mapsto\mathrm{Opt}(Q,0,x)$, given by

$$G_Q=\{(x,\gamma)\in X\times\mathcal{C}_X\mid\gamma\in\mathrm{Opt}(Q,0,x)\}.\tag{5.2}$$

Lemma 5.5. *Consider the mean field game* $\mathrm{MFG}(X,\Gamma,K)$ *and assume that Hypotheses 3.5(a), (c), and (i) hold. Let* $Q \in \mathcal{Q}$*. Then the graph* G_Q *is compact.*

Proof. Let us first prove that G_Q is a closed subset of $X \times \mathcal{C}_X$. Let $((x_n,\gamma_n))_{n\in\mathbb{N}}$ be a sequence in G_Q such that, as $n \to \infty$, $x_n \to x$ for some $x \in X$ and $\gamma_n \to \gamma$ for some $\gamma \in \mathcal{C}_X$ (uniformly on compact time intervals). Since $\gamma_n \in \mathrm{Opt}(Q,0,x_n)$, one has that $\gamma_n \in \mathrm{Lip}_{K_{\max}}(X)$, and thus, since $\mathrm{Lip}_{K_{\max}}(X)$ is closed, one has $\gamma \in \mathrm{Lip}_{K_{\max}}(X)$.

Since $\gamma_n(0) = x_n$, it follows that $\gamma(0) = x$. The continuity of the value function φ_Q also implies that $\varphi_Q(0,x_n) \to \varphi_Q(0,x)$ as $n \to \infty$. Since $\gamma_n(\varphi_Q(0,x_n)) \in \Gamma$ and Γ is closed, one concludes that $\gamma(\varphi_Q(0,x)) \in \Gamma$. Moreover, if $t \in \mathbb{R}_+$ is such that $t > \varphi_Q(0,x)$, then $t > \varphi_Q(0,x_n)$ for n large enough, and thus $\gamma_n(t) = \gamma_n(\varphi_Q(0,x_n))$ for n large enough, which shows that $\gamma(t) = \gamma(\varphi_Q(0,x))$ for every $t > \varphi_Q(0,x)$. Notice, in particular, that the exit time of γ satisfies $\tau(0,\gamma) \le \varphi_Q(0,x)$, and thus, to conclude that $\gamma \in \mathrm{Opt}(Q,0,x)$, one is only left to prove that $\gamma \in \mathrm{Adm}(Q)$.

For $n \in \mathbb{N}$, since $\gamma_n \in \mathrm{Adm}(Q)$, one has $|\dot\gamma_n|(t) \le K(e_{t\#}Q,\gamma_n(t))$, a condition which passes to the limit as $n \to \infty$, providing $|\dot\gamma|(t) \le K(e_{t\#}Q,\gamma(t))$ for almost every $t \in \mathbb{R}_+$. Hence $\gamma \in \mathrm{Adm}(Q)$, which concludes the proof that G_Q is closed.

Since $\mathrm{Opt}(Q,0,x) \subset \mathrm{Lip}_{K_{\max}}(X)$ for every $x \in X$, one has $G_Q \subset X \times \mathrm{Lip}_{K_{\max}}(X)$ and, since $X \times \mathrm{Lip}_{K_{\max}}(X)$ is compact and G_Q is closed, one concludes that G_Q is compact. □

As a consequence of Lemma 5.5 and the fact that $\mathrm{Opt}(Q,0,x)$ is non-empty for every $x \in X$, one obtains the following measurable selection of optimal trajectories immediately from von Neumann–Aumann Selection Theorem (see, e.g., [33, Theorem III.22] or [68, Theorem 3]).

Proposition 5.6. *Consider the mean field game* $\mathrm{MFG}(X,\Gamma,K)$ *and assume that Hypotheses 3.5(a), (c), and (i) hold. Let* $Q \in \mathcal{Q}$*,* $m_0 = e_{0\#}Q \in \mathcal{P}(X)$*, and* $\widehat{\mathfrak{B}}$ *be the completion of the Borel* σ*-algebra of* X *with respect to* m_0*. Then there exists a function* $\boldsymbol{\gamma}_Q : X \to \mathcal{C}_X$ *such that* $\boldsymbol{\gamma}_Q(x) \in \mathrm{Opt}(Q,0,x)$ *for every* $x \in X$ *and* $\boldsymbol{\gamma}_Q$ *is measurable with respect to* $\widehat{\mathfrak{B}}$ *in* X *and to the Borel* σ*-algebra in* $\mathcal{C}_X$*.*

5.3. **Fixed point formulation for equilibria.** To prove Theorem 5.1, we reformulate the question of the existence of an equilibrium for $\mathrm{MFG}(X,\Gamma,K)$ into a question of the existence of a fixed point for a set-valued map. Let $m_0 \in \mathcal{P}(X)$ and define

$$\mathcal{Q}_{m_0} = \{Q \in \mathcal{Q} \mid e_{0\#}Q = m_0\}.$$

Lemma 5.7. *Assume that Hypothesis 3.5(a) holds. For every* $m_0 \in \mathcal{P}(X)$*,* $\mathcal{Q}_{m_0}$ *is non-empty, convex, and compact.*

Proof. The convexity of $\mathcal{Q}_{m_0}$ follows immediately from its definition. To prove that it is non-empty, denote by $\psi : X \to \mathcal{C}_X$ the function that associates with each $x \in X$ the curve $\psi(x) \in \mathcal{C}_X$ which remains at x for every time $t \ge 0$. Then $\psi_\# m_0 \in \mathcal{Q}_{m_0}$, and hence $\mathcal{Q}_{m_0}$ is non-empty. One also has that $\mathcal{Q}_{m_0}$ is closed in $\mathcal{Q}$, since, by Proposition 5.2, $Q \mapsto e_{0\#}Q$ is continuous. In particular, since $\mathcal{Q}_{m_0} \subset \mathcal{Q}$ and $\mathcal{Q}$ is compact, one obtains that $\mathcal{Q}_{m_0}$ is compact. □

We now define a set-valued map $F_{m_0} : \mathcal{Q}_{m_0} \rightrightarrows \mathcal{Q}_{m_0}$ by setting, for $Q \in \mathcal{Q}_{m_0}$,

$$F_{m_0}(Q) = \left\{ \widetilde{Q} \in \mathcal{Q}_{m_0} \;\middle|\; \widetilde{Q}\text{-almost every } \gamma \in \mathcal{C}_X \text{ satisfies } \gamma \in \mathrm{Opt}(Q,0,\gamma(0)) \right\}.$$

By the definition of equilibria for $\mathrm{MFG}(X,\Gamma,K)$, $Q \in \mathcal{Q}_{m_0}$ is an equilibrium for $\mathrm{MFG}(X,\Gamma,K)$ with initial condition m_0 if and only if $Q \in F_{m_0}(Q)$. Introduce the set $\mathbf{Opt}(Q)$ of all optimal trajectories for Q starting at time 0, i.e.,

$$\mathbf{Opt}(Q) = \bigcup_{x\in X} \mathrm{Opt}(Q,0,x). \tag{5.3}$$

Notice that $\mathbf{Opt}(Q)$ is compact as the projection into $\mathcal{C}_X$ of the compact set G_Q from (5.2). The set $F_{m_0}(Q)$ can be rewritten in terms of $\mathbf{Opt}(Q)$ as

$$F_{m_0}(Q) = \left\{ \widetilde{Q} \in \mathcal{Q}_{m_0} \;\middle|\; \widetilde{Q}(\mathbf{Opt}(Q)) = 1 \right\}. \tag{5.4}$$

Let us provide some properties of the set-valued map F_{m_0}.

Lemma 5.8. *Consider the mean field game* $\mathrm{MFG}(X,\Gamma,K)$ *and assume that Hypotheses 3.5(a), (c), and (i) hold. Then, for every* $Q \in \mathcal{Q}_{m_0}$, $F_{m_0}(Q)$ *is non-empty and compact.*

Proof. Let $\boldsymbol{\gamma}_Q : X \to \mathcal{C}_X$ be the function from Proposition 5.6. Write $\widehat{m}_0$ for the completion of the measure m_0, which is defined in the σ-algebra $\widehat{\mathfrak{B}}$ from Proposition 5.6. Let $\widetilde{Q} = \boldsymbol{\gamma}_{Q\#}\widehat{m}_0$, which is well-defined thanks to Proposition 5.6.

Notice that $\widetilde{Q} \in \mathcal{Q}_{m_0}$. Indeed, $\widetilde{Q}\left(\mathrm{Lip}_{K_{\max}}(X)\right) = \widehat{m}_0\left(\boldsymbol{\gamma}_Q^{-1}\left(\mathrm{Lip}_{K_{\max}}(X)\right)\right) = \widehat{m}_0(X) = 1$, since $\boldsymbol{\gamma}_Q(x) \in \mathrm{Opt}(Q,0,x) \subset \mathrm{Lip}_{K_{\max}}(X)$ for every $x \in X$. Moreover, for every Borel set B in X, one has $e_{0\#}\widetilde{Q}(B) = \widetilde{Q}\left(e_0^{-1}(B)\right) = \widehat{m}_0\left(\boldsymbol{\gamma}_Q^{-1}\left(e_0^{-1}(B)\right)\right) = \widehat{m}_0(B) = m_0(B)$ since $e_0 \circ \boldsymbol{\gamma}_Q$ is the identity map in X and $\widehat{m}_0$ and m_0 coincide on Borel sets. Hence $\widetilde{Q} \in \mathcal{Q}_{m_0}$.

Since $\boldsymbol{\gamma}_Q(X) \subset \mathbf{Opt}(Q)$, one has $\widetilde{Q}(\mathbf{Opt}(Q)) \geq \widetilde{Q}(\boldsymbol{\gamma}_Q(X)) = \widehat{m}_0(X) = 1$, which proves that $\widetilde{Q} \in F_{m_0}(Q)$, and thus $F_{m_0}(Q)$ is non-empty.

Let us now prove that $F_{m_0}(Q)$ is compact. Since $F_{m_0}(Q) \subset \mathcal{Q}_{m_0}$ and $\mathcal{Q}_{m_0}$ is compact, it suffices to prove that $F_{m_0}(Q)$ is closed in $\mathcal{Q}_{m_0}$. Let $(\widetilde{Q}_n)_{n\in\mathbb{N}}$ be a sequence in $F_{m_0}(Q)$ converging as $n \to \infty$ to some $\widetilde{Q} \in \mathcal{Q}_{m_0}$. Since $\mathbf{Opt}(Q)$ is closed in $\mathcal{C}_X$, one obtains (see, e.g., [11, Chapter 1, Theorem 2.1]) that

$$\widetilde{Q}(\mathbf{Opt}(Q)) \geq \limsup_{n\to\infty} \widetilde{Q}_n(\mathbf{Opt}(Q)) = 1,$$

which proves that $\widetilde{Q}(\mathbf{Opt}(Q)) = 1$, and thus $\widetilde{Q} \in F_{m_0}(Q)$. Hence $F_{m_0}(Q)$ is closed. □

Recall that, for A, B two topological spaces, a set-valued map $H : A \rightrightarrows B$ is said to be *upper semi-continuous* if, for every open set $W \subset B$, the set $\{x \in A \mid H(x) \subset W\}$ is open in A.

Lemma 5.9. *Consider the mean field game* $\mathrm{MFG}(X,\Gamma,K)$ *and assume that Hypotheses 3.5(a), (c), and (i) hold. Let* $m_0 \in \mathcal{P}(X)$ *and* $\mathbf{Opt} : \mathcal{Q}_{m_0} \rightrightarrows \mathrm{Lip}_{K_{\max}}(X)$ *be the set-valued map defined for* $Q \in \mathcal{Q}_{m_0}$ *by* (5.3). *Then* $\mathbf{Opt}$ *is upper semi-continuous.*

Proof. Since $\mathbf{Opt}(Q)$ is a non-empty compact subset of the compact set $\mathrm{Lip}_{K_{\max}}(X)$ for every $Q \in \mathcal{Q}_{m_0}$, it suffices to show that the graph of $\mathbf{Opt}$ is closed (see, e.g., [6, Proposition 1.4.8]).

Let $(Q_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{Q}_{m_0}$ with $Q_n \to Q$ as $n \to \infty$ for some $Q \in \mathcal{Q}_{m_0}$ and $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $\mathrm{Lip}_{K_{\max}}(X)$ with $\gamma_n \in \mathbf{Opt}(Q_n)$ for every $n \in \mathbb{N}$ and $\gamma_n \to \gamma$ as $n \to \infty$ for some $\gamma \in \mathcal{C}_X$. Notice that $\gamma \in \mathrm{Lip}_{K_{\max}}(X)$ since $\mathrm{Lip}_{K_{\max}}(X)$ is closed. Since $\gamma_n \in \mathbf{Opt}(Q_n)$, then $\gamma_n(t) = \gamma_n(\varphi_{Q_n}(0,\gamma_n(0))) \in \Gamma$ for $t \geq \varphi_{Q_n}(0,\gamma_n(0))$ and $|\dot{\gamma}_n|(t) \leq K(e_{t\#}Q_n, \gamma_n(t))$ for almost every $t \geq 0$.

By Propositions 4.4 and 5.4, one has that $\varphi_{Q_n}(0,\gamma_n(0)) \to \varphi_Q(0,\gamma(0))$ as $n \to \infty$, which implies that $\gamma(t) = \gamma(\varphi_Q(0,\gamma(0))) \in \Gamma$ for every $t \geq \varphi_Q(0,\gamma(0))$. In particular, the exit time of γ satisfies $\tau(0,\gamma) \leq \varphi_Q(0,\gamma(0))$. The condition $|\dot{\gamma}_n|(t) \leq K(e_{t\#}Q_n, \gamma_n(t))$ passes to the limit as $n \to \infty$, proving that $\gamma \in \mathrm{Adm}(Q)$. Hence $\gamma \in \mathrm{Opt}(Q,0,\gamma(0)) \subset \mathbf{Opt}(Q)$, and thus the graph of $\mathbf{Opt}$ is closed. □

Lemma 5.10. *Consider the mean field game* $\mathrm{MFG}(X,\Gamma,K)$ *and assume that Hypotheses 3.5(a), (c), and (i) hold. Let* $m_0 \in \mathcal{P}(X)$. *Then the set-valued map* $F_{m_0} : \mathcal{Q}_{m_0} \rightrightarrows \mathcal{Q}_{m_0}$ *defined in* (5.4) *is upper semi-continuous.*

Proof. Using Lemma 5.8 and [6, Proposition 1.4.8], it suffices to show that the graph of F_{m_0} is closed. Let $(Q_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{Q}_{m_0}$ with $Q_n \to Q$ as $n \to \infty$ for some $Q \in \mathcal{Q}_{m_0}$ and $(\widetilde{Q}_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{Q}_{m_0}$ with $\widetilde{Q}_n \in F_{m_0}(Q_n)$ for every $n \in \mathbb{N}$ and $\widetilde{Q}_n \to \widetilde{Q}$ as $n \to \infty$ for some $\widetilde{Q} \in \mathcal{Q}_{m_0}$.

For $k \in \mathbb{N}^*$, let $V_k = \left\{\gamma \in \mathcal{C}_X \,\middle|\, \mathbf{d}(\gamma, \mathbf{Opt}(Q)) \leq \frac{1}{k}\right\}$. Since V_k is a neighborhood of $\mathbf{Opt}(Q)$, it follows from Lemma 5.9 that there exists a neighborhood W_k of Q in $\mathcal{Q}_{m_0}$ such that $\mathbf{Opt}(\widehat{Q}) \subset V_k$ for every $\widehat{Q} \in W_k$. Since $Q_n \to Q$ as $n \to \infty$, there exists $N_k \in \mathbb{N}$ such that, for every $n \geq N_k$, one has $Q_n \in W_k$, and thus $\mathbf{Opt}(Q_n) \subset V_k$. Since $\widetilde{Q}_n(\mathbf{Opt}(Q_n)) = 1$, one obtains that $\widetilde{Q}_n(V_k) = 1$ for every $n \geq N_k$. Since $\widetilde{Q}_n \to \widetilde{Q}$ as $n \to \infty$ and V_k is closed, it follows that $\widetilde{Q}(V_k) \geq \limsup_{n\to\infty} \widetilde{Q}_n(V_k) = 1$,

and thus $\widetilde{Q}(V_k) = 1$. Since this holds for every $k \in \mathbb{N}^*$ and $(V_k)_{k\in\mathbb{N}^*}$ is a non-increasing family of sets with $\bigcap_{k\in\mathbb{N}^*} V_k = \mathbf{Opt}(Q)$, one concludes that $\widetilde{Q}(\mathbf{Opt}(Q)) = \lim_{k\to\infty} \widetilde{Q}(V_k) = 1$. Hence $\widetilde{Q} \in F_{m_0}(Q)$, which proves that the graph of F_{m_0} is closed. □

We can now conclude the proof of Theorem 5.1.

Proof of Theorem 5.1. By Lemmas 5.7, 5.8, and 5.10, $\mathcal{Q}_{m_0}$ is convex, $F_{m_0} : \mathcal{Q}_{m_0} \rightrightarrows \mathcal{Q}_{m_0}$ is upper semi-continuous, and $F_{m_0}(Q)$ is non-empty, compact, and convex for every $Q \in \mathcal{Q}_{m_0}$. This means that F_{m_0} is a Kakutani map (see, e.g., [43, §7, Definition 8.1]), and hence, thanks to Kakutani fixed point theorem (see, e.g., [43, §7, Theorem 8.6]), it admits a fixed point, i.e., an element $Q \in \mathcal{Q}_{m_0}$ such that $Q \in F_{m_0}(Q)$, which means that Q is an equilibrium for $\mathrm{MFG}(X,\Gamma,K)$ with initial condition m_0. □

6. THE MFG SYSTEM

In most of the works on mean field games, such as the classical references [53,58–60], equilibria are characterized as solutions of a system of partial differential equations: the time evolution of the distribution of agents $m : \mathbb{R}_+ \to \mathcal{P}(X)$ satisfies a certain continuity equation, coupled with a Hamilton–Jacobi equation characterizing the value function φ of the optimal control problem solved by each agent. This section provides such a characterization for the mean field game $\mathrm{MFG}(X,\Gamma,K)$:

Theorem 6.1. *Consider the mean field game* $\mathrm{MFG}(X,\Gamma,K)$ *and assume that Hypotheses 3.5(a), (c)–(e), and (h)–(j) hold. Let* $m_0 \in \mathcal{P}(\overline{\Omega})$*,* $Q \in \mathcal{Q}$ *be an equilibrium of* $\mathrm{MFG}(\overline{\Omega},\partial\Omega,K)$ *with initial condition* m_0*,* k *be defined from* K *and* Q *as in Corollary 5.3,* φ *be the value function defined in* (4.1)*, and* $m = \mu^Q$*. Then* (m,φ) *solve the* MFG system

$$\begin{cases} \partial_t m(t,x) - \operatorname{div}\Big[m(t,x)K(m_t,x)\widehat{\nabla\varphi}(t,x)\Big] = 0, & (t,x) \in \mathbb{R}_+^* \times \Omega, \\ -\partial_t\varphi(t,x) + |\nabla_x\varphi(t,x)|K(m_t,x) - 1 = 0, & (t,x) \in \mathbb{R}_+ \times \Omega, \\ m(0,\cdot) = m_0, & \\ \varphi(t,x) = 0, & (t,x) \in \mathbb{R}_+ \times \partial\Omega, \end{cases} \tag{6.1}$$

where the first and second equations are satisfied, respectively, in the sense of distributions and in the viscosity sense.

Recall that $\mu^Q : \mathbb{R}_+ \to \mathcal{P}(X)$ is the measure defined from Q by $\mu_t^Q = e_{t\#}Q$. System (6.1) is composed of a continuity equation on m and a Hamilton–Jacobi equation on the value function φ. The Hamilton–Jacobi equation on φ and its boundary condition follow immediately from Theorem 4.9 and Corollary 5.3, and need only Hypotheses 3.5(a), (c), (d), and (i).

Concerning the continuity equation on m, it is clear that m satisfies an equation of the form $\partial_t m + \operatorname{div}(mv) = 0$ for some velocity field v (Proposition 5.2(a) can be easily adapted to show that $t \mapsto e_{t\#}Q$ is Lipschitz continuous with respect to the W_p Wasserstein distance in $\mathcal{P}(X)$ for every $p \geq 1$, and thus one can apply [5, Theorem 8.3.1]). The point of Theorem 6.1 is to identify this velocity field as $v = -K\widehat{\nabla\varphi}$.

This identification is possible thanks to Corollary 4.21 when k is constructed from K and Q as in Corollary 5.3. However, for a function K satisfying Hypothesis 3.5(i), k does not necessarily satisfies Hypothesis 3.5(g), which is needed for applying Corollary 4.21. The goal of this section is to prove that, if K satisfies the more restrictive Hypothesis 3.5(j) and Q is an equilibrium of $\mathrm{MFG}(X,\Gamma,K)$, then k satisfies Hypothesis 3.5(g), and thus Corollary 5.3 applies. We shall also prove that the measures m_t are concentrated on the set Υ introduced in (4.14), on which $\widehat{\nabla\varphi}$ is continuous.

Proposition 6.2. *Consider the mean field game* $\mathrm{MFG}(X,\Gamma,K)$ *and assume that Hypotheses 3.5(a), (c), (e), and (j) hold. Let* Q *be an equilibrium of* $\mathrm{MFG}(X,\Gamma,K)$ *and define* $k:\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}_+$ *by* $k(t,x)=K(e_{t\#}Q,x)$*. Then* $k\in\mathcal{C}^{1,1}(\mathbb{R}_+\times\mathbb{R}^d,\mathbb{R}_+)$ *and it admits an extension to a* $\mathcal{C}^{1,1}$ *function on* $\mathbb{R}\times\mathbb{R}^d$*.*

Proof. Notice that, as in Corollary 5.3, one has $k\in\mathrm{Lip}(\mathbb{R}_+\times\mathbb{R}^d,\mathbb{R}_+)$, and one may extend k to a Lipschitz continuous function $\widehat{k}$ defined on $\mathbb{R}\times\mathbb{R}^d$ by setting $\widehat{k}(t,x)=k(0,x)$ for $t<0$. This allows one to apply Corollary 4.14 with $\widehat{k}$ to obtain in particular that, for every $\gamma\in\mathbf{Opt}(Q)$, one has $\gamma\in\mathcal{C}^{1,1}([0,\varphi(0,\gamma(0))],\mathbb{R}^d)$. We extend each curve $\gamma\in\mathbf{Opt}(Q)$ to a curve $\widetilde{\gamma}$ in $\mathcal{C}^{1,1}((-\infty,\varphi(0,\gamma(0))],\mathbb{R}^d)$ by setting $\widetilde{\gamma}(t)=\gamma(0)+t\dot{\gamma}(0)$ for $t<0$, and we denote by $\widetilde{Q}$ the measure on $\mathcal{C}(\mathbb{R},\mathbb{R}^d)$ defined as the pushforward of Q by such extension operator. We define $\widetilde{k}(t,x)=K(e_{t\#}\widetilde{Q},x)$, which is defined on $\mathbb{R}\times\mathbb{R}^d$ and coincides with k on $\mathbb{R}_+\times\mathbb{R}^d$.

Let $E:\mathcal{P}(\overline{\Omega})\times\mathbb{R}^d\to\mathbb{R}_+$ be as in Hypothesis 3.5(j). Notice that E can be easily extended to a function defined on $\mathcal{P}(\mathbb{R}^d)\times\mathbb{R}^d$, that we still denote by E. We define $E_0:\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}_+$ by setting $E_0(t,x)=E(e_{t\#}\widetilde{Q},x)$ for every $(t,x)\in\mathbb{R}\times\mathbb{R}^d$. Notice that, for every $(t,x)\in\mathbb{R}\times\mathbb{R}^d$, one has

$$\nabla_x E_0(t,x)=\int_{\mathbb{R}^d}\nabla\chi(x-y)\eta(y)\,\mathrm{d}e_{t\#}\widetilde{Q}(y),$$

and in particular, proceeding as in the proof of Proposition 3.6 and using Proposition 5.2, we obtain that $\nabla_x E_0$ is locally Lipschitz continuous on $\mathbb{R}\times\mathbb{R}^d$. Let us consider now $\partial_t E_0$. One has

$$E_0(t,x)=\int_{\mathbb{R}^d}\chi(x-y)\eta(y)\,\mathrm{d}e_{t\#}\widetilde{Q}(y)=\int_{\mathcal{C}(\mathbb{R},\mathbb{R}^d)}\chi(x-\widetilde{\gamma}(t))\eta(\widetilde{\gamma}(t))\,\mathrm{d}\widetilde{Q}(\widetilde{\gamma}).$$

For given $x\in\mathbb{R}^d$, since $\widetilde{Q}$ is concentrated on curves $\widetilde{\gamma}$ belonging to $\mathcal{C}^{1,1}((-\infty,\varphi(0,\widetilde{\gamma}(0))],\mathbb{R}^d)$ and constant for $t\geq\varphi(0,\widetilde{\gamma}(0))$, then, for $\widetilde{Q}$-almost every $\widetilde{\gamma}$, the function $t\mapsto\chi(x-\widetilde{\gamma}(t))\eta(\widetilde{\gamma}(t))$ is differentiable everywhere, except possibly at $t=\varphi(0,\widetilde{\gamma}(0))$, with

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\chi(x-\widetilde{\gamma}(t))\eta(\widetilde{\gamma}(t))\right]=-\nabla\chi(x-\widetilde{\gamma}(t))\cdot\dot{\widetilde{\gamma}}(t)\eta(\widetilde{\gamma}(t))+\chi(x-\widetilde{\gamma}(t))\nabla\eta(\widetilde{\gamma}(t))\cdot\dot{\widetilde{\gamma}}(t). \tag{6.2}$$

Since $\eta(x)=0$ and $\nabla\eta(x)=0$ for $x\in\partial\Omega$ and $\widetilde{\gamma}(t)\in\partial\Omega$ for $t=\varphi(0,\widetilde{\gamma}(0))$, one can also prove that the above function is differentiable at $t=\varphi(0,\widetilde{\gamma}(0))$. Moreover, thanks to (6.2), its derivative is Lipschitz continuous and upper bounded, and thus $\partial_t E_0(t,x)$ exists, with

$$\partial_t E_0(t,x)=\int_{\mathcal{C}(\mathbb{R},\mathbb{R}^d)}\left[-\nabla\chi(x-\widetilde{\gamma}(t))\cdot\dot{\widetilde{\gamma}}(t)\eta(\widetilde{\gamma}(t))+\chi(x-\widetilde{\gamma}(t))\nabla\eta(\widetilde{\gamma}(t))\cdot\dot{\widetilde{\gamma}}(t)\right]\mathrm{d}\widetilde{Q}(\widetilde{\gamma}),$$

and one immediately verifies using the previous assumptions that $\partial_t E_0$ is Lipschitz continuous in $\mathbb{R}\times\mathbb{R}^d$. Together with the corresponding property for $\nabla_x E_0$, we obtain that $E_0\in\mathcal{C}^{1,1}(\mathbb{R}\times\mathbb{R}^d,\mathbb{R}_+)$. Since $g\in\mathcal{C}^{1,1}(\mathbb{R}_+,\mathbb{R}_+^*)$, we conclude that $\widetilde{k}\in\mathcal{C}^{1,1}(\mathbb{R}\times\mathbb{R}^d,\mathbb{R}_+)$. □

Proposition 6.2 allows one to apply all the results of Section 4, and in particular Corollary 4.21, to the optimal control problem $\mathrm{OCP}(\overline{\Omega},\partial\Omega,k)$ when k is constructed from K as in Proposition 6.2. To conclude the proof of Theorem 6.1, we show that the set where $\widehat{\nabla\varphi}$ is discontinuous has m_t measure zero.

Proposition 6.3. *Consider the mean field game* $\mathrm{MFG}(X,\Gamma,K)$ *and assume that Hypotheses 3.5(a), (c), (e), and (j) hold. Let* $Q\in\mathcal{Q}$ *be an equilibrium of* $\mathrm{MFG}(\overline{\Omega},\partial\Omega,K)$*,* $m=\mu^Q$*,* k *be defined from* K *and* Q *as in Proposition 6.2, and* Υ_t *be defined for* $t\geq 0$ *via* $\Upsilon_t=\{x\in\Omega\mid(t,x)\in\Upsilon\}$*. Then, for every* $t>0$*,* $m_t(\Omega\setminus\Upsilon_t)=0$*.*

Proof. Notice that, since Q is an equilibrium of $\mathrm{MFG}(\overline{\Omega},\partial\Omega,K)$, then $Q(\mathbf{Opt}(Q))=1$. It follows easily from the definition of Υ that, for $t>0$, $\{\gamma\in\mathbf{Opt}(Q)\mid\gamma(t)\in\Omega\setminus\Upsilon_t\}=\emptyset$, and then $m_t(\Omega\setminus\Upsilon_t)=Q(\{\gamma\in\mathbf{Opt}(Q)\mid\gamma(t)\in\Omega\setminus\Upsilon_t\})=Q(\emptyset)=0$. □

Thanks to Propositions 6.2 and 6.3, one can easily obtain from Corollary 4.21 and Proposition 4.22 that m satisfies the continuity equation in (6.1). For the sake of completeness, we detail the proof of this fact.

Proof of Theorem 6.1. The Hamilton–Jacobi equation on φ and its boundary condition follow immediately from Theorem 4.9 and Corollary 5.3, and the initial condition on m follows immediately from its definition. One is then left to show that m satisfies the continuity equation in (6.1).

By Proposition 4.22, $\widehat{\nabla\varphi}$ is continuous on the set Υ defined in (4.14); moreover, it can easily be extended into a Borel-measurable function defined everywhere on $\mathbb{R}_+ \times \overline{\Omega}$, that we will still denote by $\widehat{\nabla\varphi}$.

Let $\phi \in \mathcal{C}^\infty_c(\mathbb{R}^*_+ \times \Omega, \mathbb{R})$ be a test function. Take $\gamma \in \mathbf{Opt}(Q)$ and $x_0 \in \overline{\Omega}$ such that $\gamma \in \mathrm{Opt}(Q, 0, x_0)$ and let $T = \varphi(0, x_0)$. By Corollary 4.21, one has $\dot\gamma(t) = -K(m_t, \gamma(t))\widehat{\nabla\varphi}(t, \gamma(t))$ for every $t \in (0, T)$. Hence, for every $t \in (0, T)$,

$$\frac{\mathrm{d}}{\mathrm{d}t}[\phi(t, \gamma(t))] = \partial_t\phi(t, \gamma(t)) - \nabla_x\phi(t, \gamma(t)) \cdot \widehat{\nabla\varphi}(t, \gamma(t))K(m_t, \gamma(t)). \tag{6.3}$$

Notice that the right-hand side of (6.3) is a continuous function of $(t, \gamma) \in \mathbb{R}^*_+ \times \mathbf{Opt}(Q)$. The only non-trivial term is $\nabla_x\phi(t, \gamma(t)) \cdot \widehat{\nabla\varphi}(t, \gamma(t))$. Take $(\widetilde{t}, \widetilde{\gamma}) \in \mathbb{R}^*_+ \times \mathbf{Opt}(Q)$. If $\widetilde{\gamma}(\widetilde{t}) \in \partial\Omega$ the continuity at $(\widetilde{t}, \widetilde{\gamma})$ is guaranteed by the fact that ϕ is compactly supported on $\mathbb{R}^*_+ \times \Omega$. Otherwise, if $\widetilde{\gamma}(\widetilde{t}) \in \Omega$, let $(t_n, \gamma_n)_{n\in\mathbb{N}}$ be a sequence with $t_n \to \widetilde{t}$ and $\gamma_n \to \widetilde{\gamma}$ in $\mathcal{C}_{\overline{\Omega}}$ as $n \to \infty$. For n large enough, we have $\gamma_n(t_n) \in \Omega$ and $t_n > 0$. Since $\gamma_n \in \mathbf{Opt}(Q)$, there exists $x_{0,n} \in \overline{\Omega}$ such that $\gamma_n \in \mathrm{Opt}(Q, 0, x_{0,n})$. Let $\widetilde{x}_0 = \widetilde{\gamma}(0) = \lim_{n\to\infty}\gamma_n(0) = \lim_{n\to\infty}x_{0,n}$ and notice that $\widetilde{\gamma} \in \mathrm{Opt}(Q, 0, \widetilde{x}_0)$. Since $\widetilde{\gamma}(\widetilde{t}) \in \Omega$ and $\gamma_n(t_n) \in \Omega$, one has $\widetilde{x}_0 \in \Omega$ and $x_{0,n} \in \Omega$. Thus $(\widetilde{t}, \widetilde{\gamma}(\widetilde{t})), (t_n, \gamma_n(t_n)) \in \Upsilon$, and then, using Proposition 4.22, one concludes that $\lim_{n\to\infty}\widehat{\nabla\varphi}(t_n, \gamma_n(t_n)) = \widehat{\nabla\varphi}(\widetilde{t}, \widetilde{\gamma}(\widetilde{t}))$, yielding the desired continuity.

Thanks to the continuity of its right-hand side in $\mathbb{R}_+ \times \mathbf{Opt}(Q)$, one can integrate (6.3) over this set, yielding

$$\begin{aligned}\int_0^{+\infty}\int_{\mathbf{Opt}(Q)}\frac{\mathrm{d}}{\mathrm{d}t}[\phi(t, \gamma(t))]\,\mathrm{d}Q(\gamma)\,\mathrm{d}t &= \int_0^{+\infty}\int_{\mathbf{Opt}(Q)}\partial_t\phi(t, \gamma(t))\,\mathrm{d}Q(\gamma)\,\mathrm{d}t \\ &\quad - \int_0^{+\infty}\int_{\mathbf{Opt}(Q)}\nabla_x\phi(t, \gamma(t)) \cdot \widehat{\nabla\varphi}(t, \gamma(t))K(m_t, \gamma(t))\,\mathrm{d}Q(\gamma)\,\mathrm{d}t.\end{aligned}$$

Since φ is compactly supported, the left-hand side of the above equality is zero. Moreover, using Proposition 6.3 and the facts that $\gamma(t) = e_t(\gamma)$ and $m_t = e_{t\#}Q$, one finally deduces that, for every $\phi \in \mathcal{C}^\infty_c(\mathbb{R}^*_+ \times \Omega, \mathbb{R})$,

$$\int_0^{+\infty}\int_\Omega \partial_t\phi(t, x)\,\mathrm{d}m_t(x)\,\mathrm{d}t = \int_0^{+\infty}\int_\Omega \nabla_x\phi(t, x) \cdot \widehat{\nabla\varphi}(t, x)K(m_t, x)\,\mathrm{d}m_t(x)\,\mathrm{d}t,$$

which is precisely the weak formulation of the continuity equation in (6.1). □

Remark 6.4. Theorem 6.1 shows that, if Q is an equilibrium of $\mathrm{MFG}(\overline{\Omega}, \partial\Omega, K)$, then $m = \mu^Q$ and the corresponding value function φ solve the MFG system (6.1). One may also prove the converse statement, i.e., that any solution (m, φ) to (6.1) comes from an equilibrium Q of $\mathrm{MFG}(\overline{\Omega}, \partial\Omega, K)$. Let us provide a brief idea of such a proof.

Using classical techniques in optimal control (cf., e.g., [7, Chapter IV, Corollary 4.3]), one may prove that any viscosity solution φ of the Hamilton–Jacobi equation in (6.1) satisfying also the corresponding boundary condition is the value function of $\mathrm{OCP}(\overline{\Omega}, \partial\Omega, k)$, where k is obtained from K and m_t as in Corollary 5.3. Moreover, using the superposition principle (cf., e.g., [4, Theorem 3.2]), one may prove that any solution m in the sense of distributions of the continuity equation in (6.1) is a superposition solution, i.e., $m = \mu^Q$ for some $Q \in \mathcal{P}(\mathcal{C}_{\overline{\Omega}})$ concentrated on the solutions of $\dot\gamma(t) = -K(m_t, \gamma(t))\widehat{\nabla\varphi}(t, \gamma(t))$. Finally, one may again use classical techniques in optimal control to show that, since φ is the value function of $\mathrm{OCP}(\overline{\Omega}, \partial\Omega, k)$, any such γ is necessarily an optimal trajectory of $\mathrm{OCP}(\overline{\Omega}, \partial\Omega, k)$, implying that Q is an equilibrium of $\mathrm{MFG}(\overline{\Omega}, \partial\Omega, K)$, as required.

7. EXAMPLES

7.1. **Mean field game on a segment.** Consider the mean field game $\mathrm{MFG}(X,\Gamma,K)$ with $\Omega = (0,1) \subset \mathbb{R}$, $X = \overline{\Omega}$, $\Gamma = \partial\Omega = \{0,1\}$, and K satisfying the assumptions from Hypothesis 3.5(j). This situation can be interpreted as the movement of agents on a long corridor, with one exit at each end, and where we assume that agents have no preference on which exit they take. We first remark that the equilibrium of such mean field game consists on agents on a certain segment $(0,\varkappa)$ going left and agents on $(\varkappa,1)$ going right, the position $\varkappa$ of the splitting point depending on the initial distribution of agents m_0.

Proposition 7.1. *Let $m_0 \in \mathcal{P}([0,1])$, K satisfy Hypothesis 3.5(j), $Q \in \mathcal{Q}$ be an equilibrium of* $\mathrm{MFG}([0,1],\{0,1\},K)$ *with initial condition m_0, and $m = \mu^Q$. Then there exists $\varkappa \in (0,1)$ such that, for every $\gamma \in \mathbf{Opt}(Q)$, γ satisfies, for every $t \in [0,\varphi(0,\gamma(0)))$,*

$$\begin{cases} \dot\gamma(t) = -K(m_t,\gamma(t)), & \text{if } \gamma(0) \in (0,\varkappa), \\ \dot\gamma(t) = K(m_t,\gamma(t)), & \text{if } \gamma(0) \in (\varkappa,1). \end{cases} \tag{7.1}$$

The fact that agents move in a constant direction is a consequence of Corollary 4.14, since the optimal control u is Lipschitz continuous and takes values in $\{-1,1\}$, being thus necessarily constant. Existence of $\varkappa$ can be determined by defining, for a given equilibrium Q, the functions $T_l, T_r : [0,1] \to \mathbb{R}_+$ by setting, for $x \in [0,1]$, $T_l(x)$ and $T_r(x)$ as the times needed for solutions of $\dot\gamma(t) = -K(m_t,\gamma(t))$ and $\dot\gamma(t) = K(m_t,\gamma(t))$, respectively, to reach the boundary when their initial condition is $\gamma(0) = x$. Thanks to a characterization of T_l and T_r as implicit functions in terms of the flows of such differential equations, one can prove that $T_l'(x) > 0$ and $T_r'(x) < 0$ for every $x \in (0,1)$, yielding, thanks to the fact that $T_l(0) = T_r(1) = 0$, the existence of a unique $\varkappa \in (0,1)$ such that $T_l(\varkappa) = T_r(\varkappa)$. One can then verify that such $\varkappa$ satisfies the required properties.

We consider in the sequel the particular case where all agents are concentrated at some point $\ell \in [0,1]$ at the starting time, i.e., $m_0 = \delta_\ell$, in which one can completely describe Lagrangian equilibria. This is one of the interesting points of the non-local congested dynamics K, the fact that it allows to study the case of Dirac masses (local dynamics in MFG are usually not well-defined for non-absolutely continuous measures).

Proposition 7.2. *Let $\ell \in [0,1]$, K satisfy Hypothesis 3.5(j), $Q \in \mathcal{Q}$ be an equilibrium of* $\mathrm{MFG}([0,1],\{0,1\},K)$ *with initial condition δ_ℓ, and $m = \mu^Q$. Then there exists $\alpha \in [0,1]$ such that*

$$Q = \alpha\delta_{\gamma_l} + (1-\alpha)\delta_{\gamma_r}, \tag{7.2}$$

where $\gamma_l, \gamma_r \in \mathbf{Opt}(Q)$ are the unique solutions of

$$\dot\gamma_l(t) = \begin{cases} -K(m_t,\gamma_l(t)), & \text{if } 0 \le t < \varphi(0,\ell), \\ 0, & \text{if } t > \varphi(0,\ell), \end{cases} \qquad \dot\gamma_r(t) = \begin{cases} K(m_t,\gamma_r(t)), & \text{if } 0 \le t < \varphi(0,\ell), \\ 0, & \text{if } t > \varphi(0,\ell), \end{cases} \tag{7.3}$$

with initial condition $\gamma_l(0) = \gamma_r(0) = \ell$. In particular,

$$m_t = \alpha\delta_{\gamma_l(t)} + (1-\alpha)\delta_{\gamma_r(t)} \tag{7.4}$$

for every $t \ge 0$.

Proof. By definition of equilibrium, one has $Q(\mathbf{Opt}(Q)) = 1$. Moreover, since $e_{0\#}Q = \delta_\ell$, one has $Q(\{\gamma \in \mathbf{Opt}(Q) \mid \gamma(0) = \ell\}) = 1$. If $\gamma \in \mathbf{Opt}(Q)$ is such that $\gamma(0) = \ell$, then, by Corollary 4.14, its associated optimal control u satisfies $u \in \mathrm{Lip}([0,\varphi(0,\ell)],\{-1,1\})$, implying that u is constant on $[0,\varphi(0,\ell)]$. Hence, γ solves either $\dot\gamma(t) = -K(m_t,\gamma(t))$ or $\dot\gamma(t) = K(m_t,\gamma(t))$ on $[0,\varphi(0,\ell))$ with initial condition $\gamma(0) = \ell$, and, since $\gamma \in \mathbf{Opt}(Q)$, one has $\dot\gamma(t) = 0$ for $t > \varphi(0,\ell)$. Thus one has either $\gamma = \gamma_l$ or $\gamma = \gamma_r$, which proves that $\{\gamma \in \mathbf{Opt}(Q) \mid \gamma(0) = \ell\} \subset \{\gamma_l,\gamma_r\}$, and then $Q(\{\gamma_l,\gamma_r\}) = 1$, yielding the existence of $\alpha \in [0,1]$ such that (7.2) holds. In particular, (7.4) also holds. □

Remark 7.3. In general, one may have several possible equilibria. Indeed, if $\ell = \frac{1}{2}$ and K is defined by $K(\mu,x) = 1$ for every $(\mu,x) \in \mathcal{P}([0,1]) \times \mathbb{R}$, one can immediately verify that the functions γ_l, γ_r from (7.3) are given by $\gamma_l(t) = \max\left(\frac{1}{2} - t, 0\right)$ and $\gamma_r(t) = \min\left(\frac{1}{2} + t, 1\right)$ for every $t \geq 0$, and $Q = \alpha\delta_{\gamma_l} + (1-\alpha)\delta_{\gamma_r}$ is an equilibrium for every $\alpha \in [0,1]$.

Proposition 7.2 states that every equilibrium of $\mathrm{MFG}([0,1],\{0,1\},K)$ has the form (7.2). Our next result provides sufficient conditions for a measure Q given by (7.2) to be an equilibrium of $\mathrm{MFG}([0,1],\{0,1\},K)$.

Proposition 7.4. *Let K be defined from $g \in \mathcal{C}^{1,1}(\mathbb{R}_+,\mathbb{R}_+^*)$, $\chi \in \mathcal{C}^{1,1}(\mathbb{R},\mathbb{R}_+)$, and $\eta \in \mathcal{C}^{1,1}(\mathbb{R},\mathbb{R}_+)$ as in Hypothesis 3.5(j). Assume further that $\eta(x) = 0$ for every $x \in \mathbb{R} \setminus (0,1)$. Let $\ell, \alpha \in [0,1]$ and $\widetilde{\gamma}_l, \widetilde{\gamma}_r : \mathbb{R}_+ \to \mathbb{R}$ be the unique solutions of*

$$\begin{cases} \dot{\widetilde{\gamma}}_l(t) = -g\left[\alpha\chi(0)\eta(\widetilde{\gamma}_l(t)) + (1-\alpha)\chi(\widetilde{\gamma}_l(t) - \widetilde{\gamma}_r(t))\eta(\widetilde{\gamma}_r(t))\right], & t \geq 0, \\ \dot{\widetilde{\gamma}}_r(t) = \quad g\left[\alpha\chi(\widetilde{\gamma}_r(t) - \widetilde{\gamma}_l(t))\eta(\widetilde{\gamma}_l(t)) + (1-\alpha)\chi(0)\eta(\widetilde{\gamma}_r(t))\right], & t \geq 0, \\ \widetilde{\gamma}_l(0) = \widetilde{\gamma}_r(0) = \ell. \end{cases} \tag{7.5}$$

Define $T_l, T_r \in \mathbb{R}_+$ and $\gamma_l, \gamma_r : \mathbb{R}_+ \to [0,1]$ by

$$\begin{aligned} T_l &= \inf\{t \geq 0 \mid \widetilde{\gamma}_l(t) = 0\}, & \gamma_l(t) &= \max(\widetilde{\gamma}_l(t), 0) \quad \textit{for } t \geq 0, \\ T_r &= \inf\{t \geq 0 \mid \widetilde{\gamma}_r(t) = 1\}, & \gamma_r(t) &= \min(\widetilde{\gamma}_r(t), 1) \quad \textit{for } t \geq 0, \end{aligned} \tag{7.6}$$

and let $Q \in \mathcal{P}(\mathcal{C}_{[0,1]})$ be given by $Q = \alpha\delta_{\gamma_l} + (1-\alpha)\delta_{\gamma_r}$. Then Q is an equilibrium for $\mathrm{MFG}([0,1],\{0,1\},K)$ *with initial condition δ_ℓ if and only if one of the following conditions hold:*

(a) $\alpha \in (0,1)$ *and* $T_l = T_r$*; or*
(b) $\alpha = 0$ *and* $T_l \geq T_r$*; or*
(c) $\alpha = 1$ *and* $T_l \leq T_r$*.*

When Q is an equilibrium, the fact that one of the conditions (a)–(c) is satisfied can be deduced from Proposition 7.2. Conversely, to prove that Q is an equilibrium when one of (a)–(c) is satisfied, it suffices to show that γ_l and γ_r are optimal trajectories, which can be easily done after observing that any other Q-admissible trajectory γ must satisfy $\gamma_l(t) \leq \gamma(t) \leq \gamma_r(t)$ for every $t \geq 0$.

For fixed $\ell \in [0,1]$, one can consider T_l and T_r defined in (7.6) as functions of $\alpha \in [0,1]$. Notice that, denoting by $\widetilde{\gamma}_l(\cdot;\alpha)$ and $\widetilde{\gamma}_r(\cdot;\alpha)$ the solutions of (7.5) for a given $\alpha \in [0,1]$, $T_l(\alpha)$ and $T_r(\alpha)$ can be characterized by the equations $\widetilde{\gamma}_l(T_l(\alpha);\alpha) = 0$ and $\widetilde{\gamma}_r(T_r(\alpha);\alpha) = 1$, and it follows from the implicit function theorem that T_l and T_r belong to $C^1([0,1],\mathbb{R}_+)$ (after a suitable extension to a neighborhood of $[0,1]$).

Given $\ell \in [0,1]$, Proposition 7.4 allows one to numerically approximate equilibria of $\mathrm{MFG}([0,1],\{0,1\},K)$ with initial condition δ_ℓ. Indeed, one can obtain approximations for $T_l(0)$, $T_r(0)$, $T_l(1)$, $T_r(1)$ by numerically integrating (7.5) with $\alpha = 0$ and $\alpha = 1$ and check if conditions (b) or (c) from Proposition 7.4 are satisfied. If this is the case, then one has found an equilibrium of $\mathrm{MFG}([0,1],\{0,1\},K)$. Otherwise, one has $T_l(0) < T_r(0)$ and $T_l(1) > T_r(1)$, and one can thus find $\alpha \in (0,1)$ such that $T_l(\alpha) = T_r(\alpha)$ by using standard numerical methods.

This simulation has been done by approximating the solution of (7.5) using an explicit Euler method with time step $\Delta t = 10^{-4}$. We have chosen g, χ, and η as

$$g(x) = \frac{1}{1 + \left(\frac{2x}{15}\right)^4}, \qquad \chi(x) = \begin{cases} \dfrac{1}{2\varepsilon}\left[1 + \cos\left(\dfrac{\pi x}{\varepsilon}\right)\right], & \text{if } |x| < \varepsilon, \\ 0, & \text{otherwise}, \end{cases} \tag{7.7}$$

$$\eta(x) = \begin{cases} \dfrac{1}{2}\left[1 - \cos\left(\dfrac{\pi \mathbf{d}(x,\{0,1\})}{\varepsilon}\right)\right], & \text{if } \mathbf{d}(x,\{0,1\}) < \varepsilon \text{ and } x \in [0,1], \\ 1, & \text{if } \mathbf{d}(x,\{0,1\}) \geq \varepsilon \text{ and } x \in [0,1], \\ 0, & \text{if } x \in \mathbb{R} \setminus [0,1], \end{cases}$$

where $\varepsilon = \frac{1}{10}$. With this choice, η is a smooth function equal to 1 on $[\varepsilon, 1-\varepsilon]$ and vanishing outside of the interval $[0,1]$, allowing one to not take into account in K agents who already left the domain $[0,1]$ and discounting agents who are close to leaving. The convolution kernel χ is smooth and vanishes outside of $[-\varepsilon, \varepsilon]$, meaning that an agent only takes into account in their congestion other agents at a distance at most ε. Finally, the function g is decreasing, meaning that higher concentrations of agents yield slower velocities. Its precise form allows one to have an important variation on the velocities for the range of agent concentrations for this problem. The functions T_l and T_r are represented in Figure 7.1 for $\ell = 0.2$ and $\ell = 0.4$.

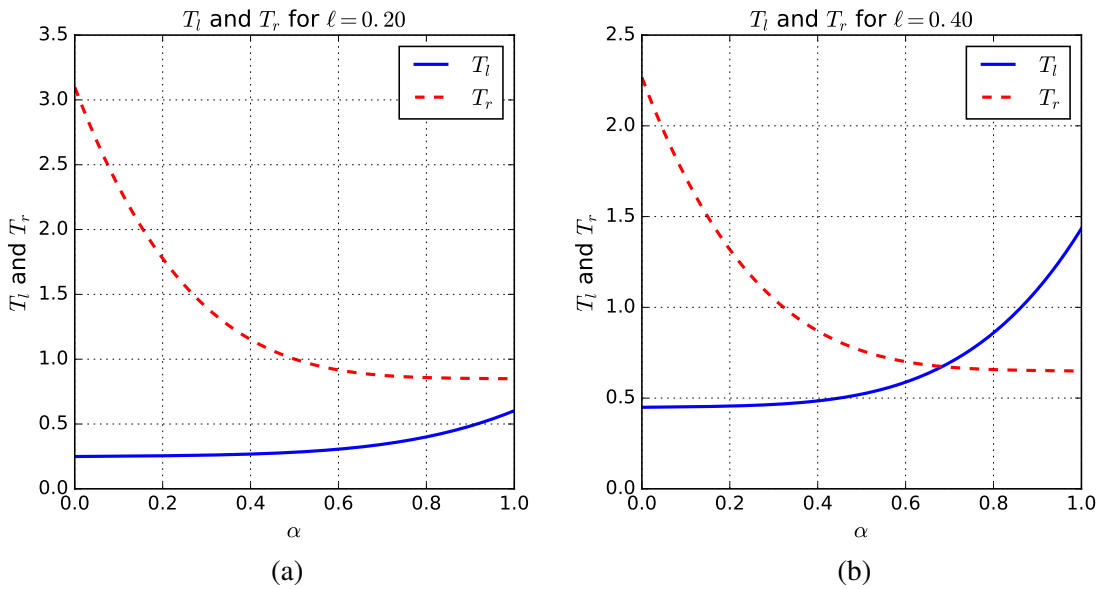


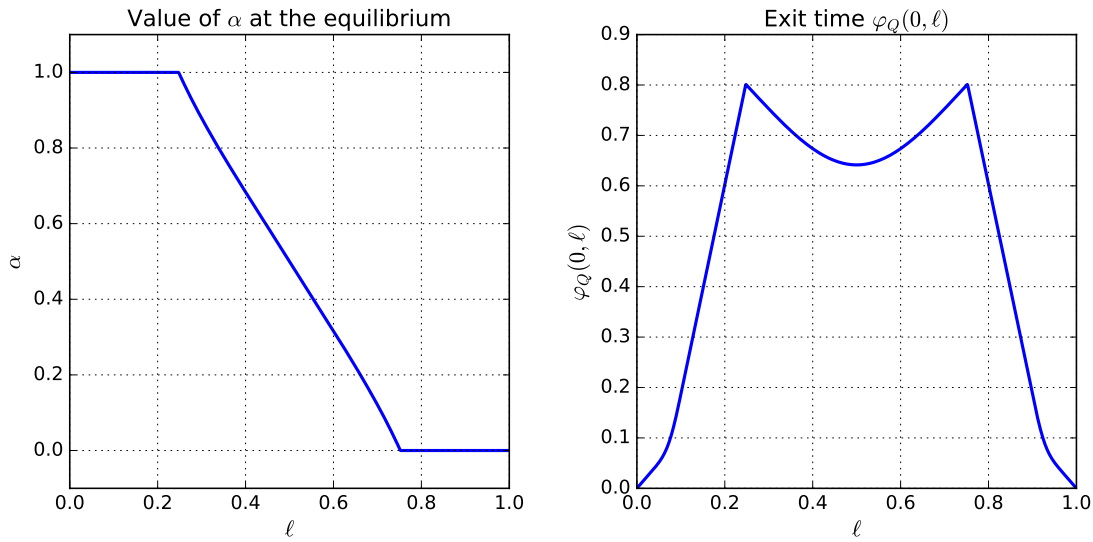
FIGURE 7.1. Functions T_l (solid line) and T_r (dashed line) for (a) $\ell = 0.2$ and (b) $\ell = 0.4$.

For $\ell = 0.2$, Figure 7.1(a) shows that $T_l(\alpha) < T_r(\alpha)$ for every $\alpha \in [0,1]$, and thus the unique equilibrium in this case is $Q = \delta_{\gamma_l}$, which corresponds to all agents moving left. For $\ell = 0.4$, Figure 7.1(b) shows that $T_l(0) < T_r(0)$, $T_l(1) > T_r(1)$, and that there exists a unique $\alpha \approx 0.683$ such that $T_l(\alpha) = T_r(\alpha)$, giving thus a unique equilibrium where approximately 68.3% of the agents move left and 31.7% of the agents move right.

Using this method, one can compute, for each $\ell \in [0,1]$, a value α for which (7.2) is an equilibrium, and the corresponding minimal exit time $\varphi_Q(0,\ell)$. Even though α need not be unique in general, as seen in Remark 7.3, it seems from the simulations that, for every ℓ, T_l is increasing and T_r is decreasing, and thus one has uniqueness of the equilibrium in the framework of our simulations. Figure 7.2 presents the values of α and $\varphi_Q(0,\ell)$ at the equilibrium as functions of ℓ obtained from our simulations.

Figure 7.2(a) corresponds to the behavior one might intuitively expect. The represented curve is symmetric around the point $\left(\frac{1}{2}, \frac{1}{2}\right)$, which one expects since this mean field game model is symmetric with respect to the transformation $x \mapsto 1-x$ of the interval $[0,1]$ and by exchanging α and $1-\alpha$. When ℓ is close to 0, meaning that all agents are initially much closer to the exit at 0 than to the exit at 1, all agents move left to the exit at 0, with a symmetric situation when ℓ is close to 1. For intermediate values of ℓ, agents split in two parts, according to Proposition 7.2, with a higher proportion of agents moving to the closer exit and a smaller proportion of agents moving to the further exit.

From Figure 7.2(b), one remarks that, close to the boundary $\{0,1\}$, the exit time is small, increasing as one gets further away from the boundary, up to the point where agents start splitting



(a) (b)

FIGURE 7.2. Values of (a) α and (b) $\varphi_Q(0,\ell)$ at the equilibrium in function of ℓ.

instead of going all in the same direction. At these points where splitting starts to occur, a seemingly counter-intuitive situation happens: the further agents start from the boundary, the faster they arrive at the exit, and the points where splitting starts to occur are maxima of the exit time.

As a final remark for these simulations, notice that Figure 7.2(b) is *not* the graph of the value function at time 0, since the equilibrium Q used to compute $\varphi_Q(0,\ell)$ depends on ℓ.

7.2. **A Braess-type paradox.** Consider the mean field game $\mathrm{MFG}(X,\Gamma,K)$ where X is a network whose set of vertices and edges are, respectively, $V=\{A,B,C,D\}$ and $E=\{\{A,C\},\{A,D\},\{C,D\},\{B,C\},\{B,D\}\}$, the lengths L_e of the edges $e\in E$ being $L_{\{A,C\}}=L_{\{B,D\}}=1$, $L_{\{A,D\}}=L_{\{B,C\}}=L>1$, and $L_{\{C,D\}}=\ell\in(0,1)$, and $\Gamma=\{B\}$. We identify each edge $e\in E$ with a segment of length L_e and consider X as the union of such segments with the suitable endpoints identified, endowed with its natural distance $\mathbf{d}$ along the edges. This network is represented in Figure 7.3. We are interested in numerically simulating equilibria of $\mathrm{MFG}(X,\Gamma,K)$ with initial condition δ_A.

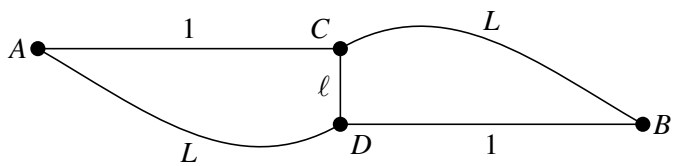


FIGURE 7.3. Network considered in the example of Section 7.2.

The motivation for considering this model comes from Braess' paradox in traffic flow. The original model studied by Braess in [13] is a static traffic flow on an oriented network similar to X where agents move from A to B and, instead of prescribing lengths, one prescribes the travel time of an edge as an increasing affine function of its vehicular flow. The paradox consists on the fact that the total travel time from A to B at equilibrium may be reduced (according to how the travel times on individual edges are prescribed) when the edge $\{C,D\}$ is suppressed from the network. Similar phenomena have later been observed on other traffic flow models [70, 71] as well as on mechanical, electrical, or computer networks [36, 67].

We assume in this section that K is given by

$$K(\mu,x) = g\left[\int_X \chi(\mathbf{d}(x,y))\,\mathrm{d}\mu(y)\right], \tag{7.8}$$

with $\chi : \mathbb{R}_+ \to \mathbb{R}_+$ a smooth decreasing convolution kernel sufficiently concentrated around 0 and g a smooth decreasing function. Notice that, by Proposition 3.6, this K satisfies Hypothesis 3.5(i). Contrarily to Section 7.1, we provide here only an informal description of the behavior of equilibria.

Let $Q \in \mathcal{P}(X)$ be an equilibrium of $\mathrm{MFG}(X,\Gamma,K)$ with initial condition δ_A and $m = \mu^Q$. For $t > 0$, the Dirac mass at A splits in two, one mass $\alpha_1 \in [0,1]$ moving along a trajectory γ_{AC} on the edge $\{A,C\}$, and the remaining mass $1-\alpha_1$ moving along γ_{AD} on the edge $\{A,D\}$. If γ_{AD} reaches D before γ_{AC} reaches C, then, since g is decreasing, one has $1-\alpha_1 < \alpha_1$. Since $\mathbf{d}(D,B) < \mathbf{d}(C,B)$ and $1-\alpha_1 < \alpha_1$, one concludes that the mass $1-\alpha_1$ arriving at D will reach the boundary before the mass α_1 traveling to C, which contradicts the fact that Q is an equilibrium. Hence, γ_{AC} reaches C before γ_{AD} reaches D. Let $T_1 > 0$ be the time at which γ_{AC} reaches C.

The mass α_1 arriving at C splits into a mass $\alpha_1\alpha_2$ moving along a trajectory γ_{CD} on the edge $\{C,D\}$ and a mass $\alpha_1(1-\alpha_2)$ moving along γ_{CB} on the edge $\{B,C\}$, where $\alpha_2 \in [0,1]$. Clearly, γ_{CB} cannot arrive at B before γ_{CD} arrives at D, for otherwise Q would not be an equilibrium. Let $T_2 > T_1$ be the time at which γ_{CD} reaches D.

If γ_{AD} reaches D before γ_{CD} does, then any part of mass from γ_{AD} moving on $\{B,D\}$ would be in advance with respect to a part of the mass from γ_{CD} moving on the same edge, and would thus arrive at B before, since optimal trajectories cannot merge (see Proposition 4.5). Similarly, γ_{CD} cannot reach D before γ_{AD}, and thus one concludes that both must arrive at D at the same time T_2.

At time T_2, one has a proportion $\alpha_1(1-\alpha_2)$ of the mass moving on γ_{CB} at some point in the edge $\{B,C\}$, and a proportion $1-\alpha_1+\alpha_1\alpha_2$ at the point D. For the latter mass, it is not optimal for any part of it to take the edges $\{A,D\}$ or $\{C,D\}$, since it would certainly arrive at B after the mass moving on γ_{CB}. Hence, for $t > T_2$, the mass $1-\alpha_1+\alpha_1\alpha_2$ moves along a trajectory γ_{DB} on the edge $\{B,D\}$. Since Q is an equilibrium, both masses arrive at D at the same time $T > T_2$.

The previous arguments show that $m = \mu^Q$ is given by

$$m_t = \begin{cases} \alpha_1\delta_{\gamma_{AC}(t)} + (1-\alpha_1)\delta_{\gamma_{AD}(t)}, & \text{if } 0 \le t < T_1, \\ \alpha_1\alpha_2\delta_{\gamma_{CD}(t)} + \alpha_1(1-\alpha_2)\delta_{\gamma_{CB}(t)} + (1-\alpha_1)\delta_{\gamma_{AD}(t)}, & \text{if } T_1 \le t < T_2, \\ \alpha_1(1-\alpha_2)\delta_{\gamma_{CB}(t)} + (1-\alpha_1+\alpha_1\alpha_2)\delta_{\gamma_{DB}(t)}, & \text{if } T_2 \le t < T, \\ \delta_B, & \text{if } t \ge T. \end{cases} \tag{7.9}$$

Notice moreover that, since agents move at maximal speed,

$$\dot{\gamma}_i(t) = K(m_t, \gamma_i(t))u_i(t), \tag{7.10}$$

where $i \in \{AC, AD, CD, CB, DB\}$ and $u_i(t) \in \{-1,1\}$, according to the direction of the movement and the orientation chosen for the edges.

The expression (7.9) allows one to numerically simulate the equilibrium by searching for $\alpha_1, \alpha_2 \in [0,1]$ such that, when solving (7.9)–(7.10), one obtains $\gamma_{CD}(T_2) = \gamma_{AD}(T_2) = D$ and $\gamma_{DB}(T) = \gamma_{CB}(T) = B$. We describe a first method of implementing the numerical simulation, which is decomposed in two steps. As a first step, we implement a function that, for each $\alpha_2 \in [0,1]$, finds $\alpha_1 \in [0,1]$ such that, similarly to Proposition 7.4, one is in one of the following situations:

- $\alpha_1 \in (0,1)$ and $\gamma_{CD}(T_2) = \gamma_{AD}(T_2) = D$; or
- $\alpha_1 = 0$ and γ_{AD} reaches D before γ_{CD}; or
- $\alpha_1 = 1$ and γ_{CD} reaches D before γ_{AD}.

The second step consists on finding $\alpha_2 \in [0,1]$ such that, when α_1 is computed from α_2 using the first step, one is in one of the following situations:

- $\alpha_2 \in (0,1)$ and $\gamma_{DB}(T) = \gamma_{CB}(T) = B$; or
- $\alpha_2 = 0$ and γ_{CB} reaches B before γ_{DB}; or

- $\alpha_2 = 1$ and γ_{DB} reaches B before γ_{CB}.

The searches for α_1 and α_2 in both steps are implemented using a bisection method. Notice that, in the case $\alpha_2 = 0$, the only possible equilibrium is when $\alpha_1 = \frac{1}{2}$.

A second method of implementing the numerical simulation, which is much faster, can be obtained if one is in a situation where the equilibrium is unique. Indeed, notice that, by transforming t into $T-t$, one obtains an equilibrium of a mean field game on the same network with initial condition δ_B and exit at $\Gamma = \{A\}$. Up to relabeling the vertices, this is an equilibrium of $\mathrm{MFG}(X,\Gamma,K)$, and, by uniqueness, one concludes that $\alpha_1 = 1-\alpha_1+\alpha_1\alpha_2$ and $1-\alpha_1 = \alpha_1(1-\alpha_2)$, yielding the relation

$$\alpha_1 = \frac{1}{2-\alpha_2}. \tag{7.11}$$

Hence, the second method for the simulation consists on finding $\alpha_2 \in [0,1]$ such that, if α_1 is computed from α_2 by (7.11), then, similarly to Proposition 7.4, one is in one of the following situations:

- $\alpha_2 \in (0,1)$ and $\gamma_{CD}(T_2) = \gamma_{AD}(T_2) = D$; or
- $\alpha_2 = 0$ and γ_{AD} reaches D before γ_{CD}; or
- $\alpha_2 = 1$ and γ_{CD} reaches D before γ_{AD}.

Notice that it suffices here to guarantee that γ_{CD} and γ_{AD} arrive at D at the same time (or one of the corresponding masses is zero), since (7.11) guarantees the symmetry of the equilibrium.

We have performed this simulation for $L = 1.25$ and different values of ℓ on the interval $(0,0.35]$, approximating the solutions of (7.10) by an explicit Euler method with time step $\Delta t = 10^{-4}$. The functions g and χ we chose were those from (7.7) with $\varepsilon = \frac{1}{10}$. We have first selected a few values of ℓ and computed an equilibrium using the first method, verifying that the values of α_1 and α_2 at equilibrium satisfy (7.11). We have then used the second method to simulate the equilibrium for 350 equally spaced values of ℓ on the interval $(0,0.35]$. Figure 7.4 presents the results of this simulation.

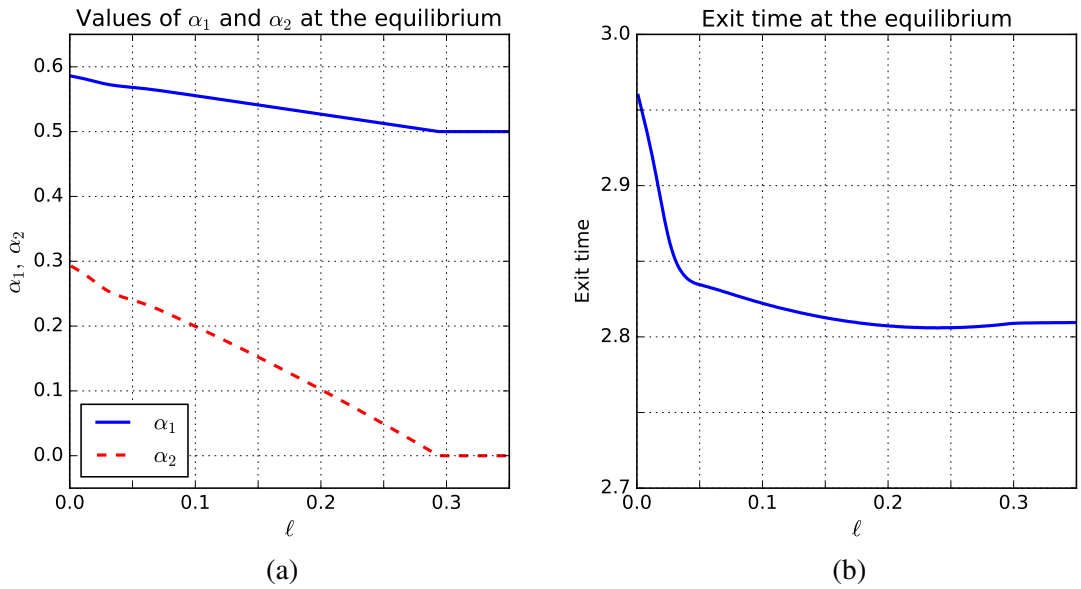


FIGURE 7.4. Values of (a) α_1 and α_2 and (b) the exit time at the equilibrium in function of the length ℓ of the edge CD.

We observe in Figure 7.4(a) that, as expected, α_2 decreases with ℓ, eventually reaching zero when $\ell \approx 0.294$: as the length of CD increases, fewer people will go through it, until CD is so long

that no agent wants to move on this edge. The value of α_1 also decreases with ℓ, since it increases with α_2: when the length of CD is small, people anticipate the fact that they will go through the edge CD and so a higher proportion of people start by taking the edge AC in order to move through CD later. When $\ell \geq 0.294$, no one travels on the edge CD, and the equilibrium is the same as the one we would get if the edge CD did not exist and we had only two edges connecting A to B, both of length $1+L$.

As regards the behavior of the exit time at the equilibrium, shown in Figure 7.4(b), for $\ell \geq 0.294$, the exit time is constant, which is expected since no agents move on CD and hence the presence and the length of CD do not influence the equilibrium. As ℓ decreases, one observes a small decrease of the exit time, until it reaches its global minimum at $\ell_* \approx 0.239$. As ℓ decreases below ℓ_*, the exit time increases (very sharply when $\ell < 0.05$), which may at first look counter-intuitive since decreasing ℓ means decreasing the length of the shortest path from A to B, which is the sequence of edges (AC, CD, DB) when $\ell < 0.25$. This exactly corresponds to a Braess-type paradox.

As in the original Braess paradox, this situation can be explained by the congestion terms. Indeed, as ℓ decreases below ℓ_*, more agents will go through the path CD, which means that more agents will also move through AC and DB, increasing congestion on those edges.

7.3. **Mean field game on a disk.** As a last example, consider the mean field game $\mathrm{MFG}(X,\Gamma,K)$ with $\Omega = B_d(0,1)$, $X = \overline{\Omega}$, $\Gamma = \partial\Omega = \mathbb{S}^{d-1}$, and K satisfying Hypothesis 3.5(j). We assume further that the functions χ and η from Hypothesis 3.5(j) are radial, i.e., there exist $\chi_0, \eta_0 : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\chi(x) = \chi_0(|x|)$ and $\eta(x) = \eta_0(|x|)$ for every $x \in \mathbb{R}^d$. When the initial condition is a Dirac mass at the center of the ball, one can obtain an explicit expression for an equilibrium, given in the following proposition.

Proposition 7.5. *Let K, g, χ, and η satisfy Hypothesis 3.5(j) and assume that there exist $\chi_0, \eta_0 : \mathbb{R}_+ \to \mathbb{R}_+$ such that χ and η are given by $\chi(x) = \chi_0(|x|)$ and $\eta(x) = \eta_0(|x|)$ for every $x \in \mathbb{R}^d$. Let $e \in \mathbb{S}^{d-1}$, ν be the normalized uniform measure on $\mathbb{S}^{d-1}$, and $\widetilde{r} : \mathbb{R}_+ \to \mathbb{R}_+$ be the unique solution of*

$$\begin{cases} \dot{\widetilde{r}}(t) = g\left[\eta_0(\widetilde{r}(t)) \int_{\mathbb{S}^{d-1}} \chi_0(\widetilde{r}(t)\,|e-\omega|)\,\mathrm{d}\nu(\omega)\right], \\ \widetilde{r}(0) = 0. \end{cases} \tag{7.12}$$

Then $\widetilde{r}$ does not depend on $e \in \mathbb{S}^{d-1}$. Define $r : \mathbb{R}_+ \to [0,1]$ by $r(t) = \min\{\widetilde{r}(t), 1\}$ and let $R : \mathbb{S}^{d-1} \to \mathcal{C}_{\overline{\Omega}}$ be the function defined by $R(\omega)(t) = r(t)\omega$ for every $\omega \in \mathbb{S}^{d-1}$ and $t \in \mathbb{R}_+$. Then $Q = R_\# \nu$ is an equilibrium of $\mathrm{MFG}(X,\Gamma,K)$ with initial condition δ_0, and, letting $m = \mu^Q$, m_t is the normalized uniform measure on the sphere of radius $r(t)$ and centered at 0 for every $t > 0$.

As in the proof of Proposition 7.4, the main idea for the proof of Proposition 7.5 is to show that $R(\omega)$ is an optimal trajectory for every $\omega \in \mathbb{S}^{d-1}$, which is done by observing that $|R(\omega)(t)| = r(t)$ and $|\gamma(t)| \leq r(t)$ for every admissible trajectory γ and every $t \geq 0$.

Notice that the equilibrium from Proposition 7.5 is not necessarily unique. Indeed, if g is constant and equal to 1, then the measure Q_ω concentrated on the curve $t \mapsto \min\{t,1\}\omega$ is an equilibrium for every $\omega \in \mathbb{S}^{d-1}$.

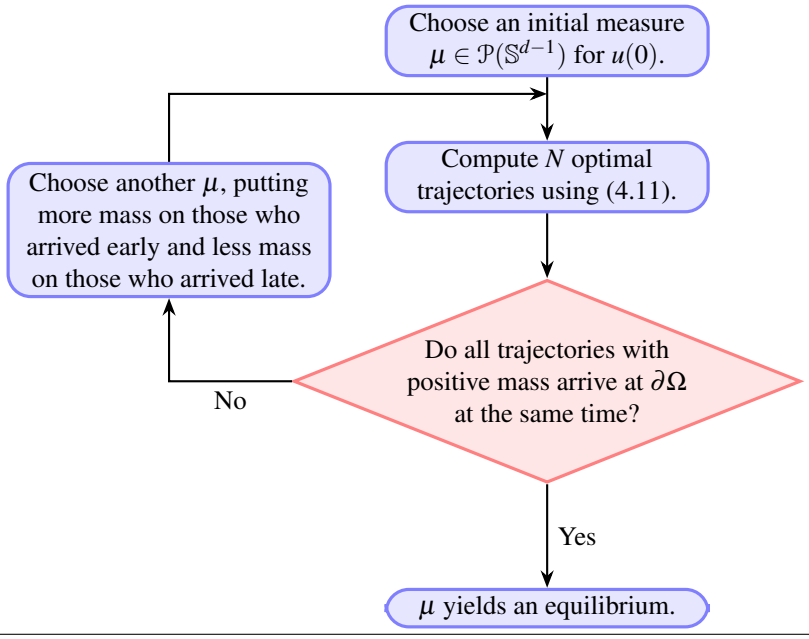
We are now interested in the case where the initial condition is δ_p for some point $p \in B_d(0,1)$. Notice that, when $g = 1$ and $p \neq 0$, there exists a unique equilibrium Q of $\mathrm{MFG}(X,\Gamma,K)$ with initial condition δ_p, which is concentrated on the segment joining p to its closest point on the boundary with unit speed. However, when g is decreasing and χ is a sufficiently concentrated radially decreasing convolution kernel, one may expect that agents will avoid congestion, concentrating on a surface evolving in time, and arriving at the boundary at a same time. Moreover, if p is close to the origin, one may expect to find equilibria with closed surfaces similar to the circles from the equilibrium of Proposition 7.5.

Since an explicit characterization of an equilibrium similar to Proposition 7.5 is too hard to obtain with initial condition δ_p, $p \neq 0$, we validate the previous intuition of the behavior of equilibria by a numerical simulation. The idea for our simulation goes as follows.

Given an equilibrium Q, let μ be the measure on $\mathbb{S}^{d-1}$ obtained as the pushforward of Q by the map that, to each optimal trajectory γ, associates the value of the corresponding optimal control at time 0, $u(0)$. If μ is known, then one can obtain an approximation of m_t using (4.11). Indeed, for $N \in \mathbb{N}$, we can approximate μ as a sum of finitely many Dirac masses, $\mu^N = \sum_{k=1}^N \mu_k \delta_{z_k}$, where $z_k \in \mathbb{S}^{d-1}$ and $\mu_k \in [0,1]$, $\sum_{k=1}^N \mu_k = 1$. One can thus compute N trajectories $\gamma_1, \dots, \gamma_N$ and their associated optimal controls $u_1, \dots, u_N$ by solving (4.11) with initial conditions $\gamma_k(0) = p$, $u_k(0) = z_k$ for $k \in \{1, \dots, N\}$, where we approximate m_t by $m_t^N = \sum_{k=1}^N \mu_k \delta_{\gamma_k(t)}$. Since all optimal trajectories arrive at $\partial\Omega$ at the same time, one expects that the approximated trajectories $\gamma_1, \dots, \gamma_N$ corresponding to a weight $\mu_k > 0$ will arrive at $\partial\Omega$ at approximately the same time.

One may thus search for an equilibrium by searching for $\mu \in \mathcal{P}(\mathbb{S}^{d-1})$ such that the above approximated trajectories with positive weigth arrive at $\partial\Omega$ at the same time. This motivates the introduction of Algorithm 7.1, based on a fixed-point strategy, to compute an equilibrium for $\mathrm{MFG}(X, \Gamma, K)$ with initial condition δ_p.

Algorithm 7.1 Algorithm used to find an equilibrium for $\mathrm{MFG}(\overline{\Omega}, \partial\Omega, K)$ with initial condition δ_p.

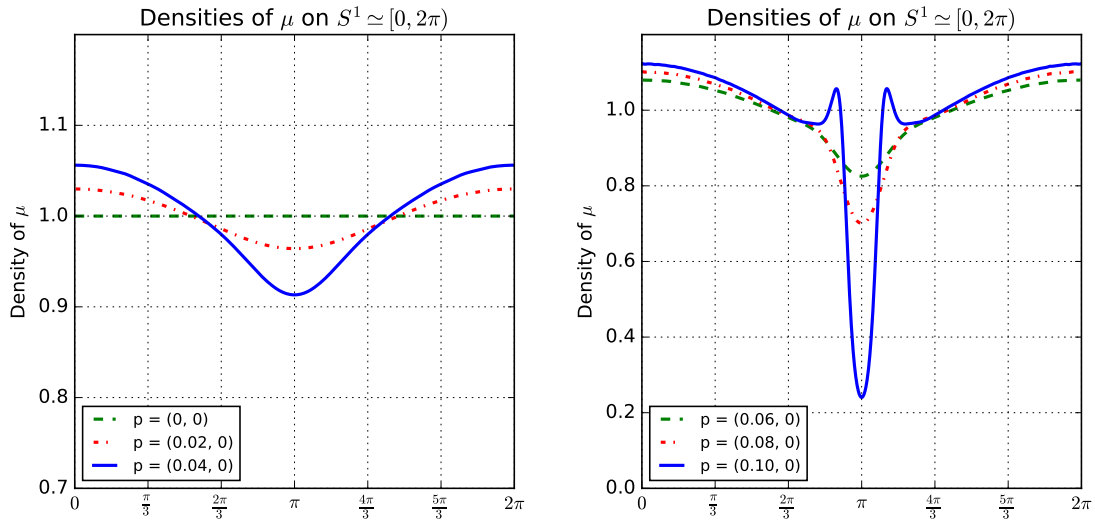


We have implemented Algorithm 7.1 in dimension $d = 2$, identifying the plane $\mathbb{R}^2$ with the complex plane $\mathbb{C}$ for simplicity, and choosing g, χ_0, and η_0 as

$$g(x) = \frac{1}{1 + \left(\frac{x}{2}\right)^2}, \qquad \chi_0(x) = \begin{cases} \dfrac{1}{2\varepsilon}\left[1 + \cos\left(\dfrac{\pi x}{\varepsilon}\right)\right], & \text{if } 0 \le x < \varepsilon, \\ 0, & \text{if } x \ge \varepsilon, \end{cases}$$

$$\eta_0(x) = \begin{cases} 1, & \text{if } 0 \le x < 1 - \varepsilon, \\ \dfrac{1}{2}\left[1 - \cos\left(\dfrac{\pi(1-x)}{\varepsilon}\right)\right], & \text{if } 1 - \varepsilon \le x < 1, \\ 0, & \text{if } x \ge 1, \end{cases}$$

with $\varepsilon = \frac{1}{10}$. Figure 7.5 presents the results of the numerical simulations, showing the densities, with respect to the normalized uniform measure on $\mathbb{S}^1$, of the measures μ yielding an equilibrium for different initial conditions of the form δ_p. Figure 7.6 shows the support of the measures m_t and some optimal trajectories at different times for the initial condition $\delta_{(0.1,0)}$. Before commenting on the results shown in Figures 7.5 and 7.6, let us describe some details of the implementation of Algorithm 7.1.



(a) (b)

FIGURE 7.5. Densities of the measures μ yielding an equilibrium with initial condition δ_p for (a) $p = (0,0)$, $p = (0.02,0)$, and $p = (0.04,0)$, and (b) $p = (0.06,0)$, $p = (0.08,0)$, and $p = (0.1,0)$. Notice that the scales of the vertical axes are not the same.

The choice of the initial measure $\mu \in \mathcal{P}(\mathbb{S}^{d-1})$ is important: from the numerical simulations, the algorithm seems to converge only if the initial measure is sufficiently close to a measure yielding an equilibrium. When p is close to 0, a good choice, based on Proposition 7.5, is to take μ as the normalized uniform measure on $\mathbb{S}^{d-1}$. For p further from 0, a possible technique is to choose points $p_0, p_1, \dots, p_n$ with $p_0 = 0$, $p_n = p$, and $|p_i - p_{i-1}|$ small for every $i \in \{1, \dots, n\}$, and apply the algorithm to compute successively an equilibrium with initial condition δ_{p_i} using as initial measure the one that yields an equilibrium with initial condition $\delta_{p_{i-1}}$ (in the spirit of the so-called *continuation method*).

The initial measure μ was numerically approximated by a finite sum of N equally spaced Dirac masses on $\mathbb{S}^1$, $\mu \approx \sum_{k=1}^N \mu_k \delta_{z_k}$, where $z_k = e^{i\frac{2\pi}{N}\left(k+\frac{1}{2}\right)}$ and $\mu_k = \mu(\{e^{i\theta} \mid \frac{2\pi}{N}k \le \theta < \frac{2\pi}{N}(k+1)\})$ for $k \in \{1, \dots, N\}$. For our simulations, we chose $N = 1000$. In order to compute an equilibrium for the initial condition δ_p with $p = (0.1,0)$, as described before, we have computed equilibria with initial conditions δ_{p_k} with $p_k = \left(\frac{k}{100}, 0\right)$ for $k \in \{1, \dots, 10\}$, using at each step the measure μ yielding an equilibrium for p_k as an initial measure for p_{k+1}.

Concerning the computations of the optimal trajectories from (4.11), the time discretization was done by an explicit Euler method with time step $\Delta t = 10^{-2}$. The simulation is stopped at the time T when the first trajectory γ_k with positive mass reaches the boundary, and we say that all trajectories have reached the boundary at the same time if $\max_{k \in \{1,\dots,N\} | \mu_k > 0} \big||\gamma_k(T)| - 1\big| < \zeta$, where the tolerance ζ was chosen as $\zeta = 3 \cdot 10^{-3}$.

The intuition for the construction of a new initial measure μ' from μ and the final positions $(\gamma_k(T))_{k=1}^N$ is that μ' should contain less mass at the points that ended further from the boundary. To do so, our choice was to first construct a measure $\mu'' = \sum_{k=1}^N \mu''_k \delta_{z_k}$ given by

$$\mu''_k = \frac{\mu_k \left|\gamma_k(T)\right|}{\sum_{k=1}^N \mu_j \left|\gamma_j(T)\right|}, \tag{7.13}$$

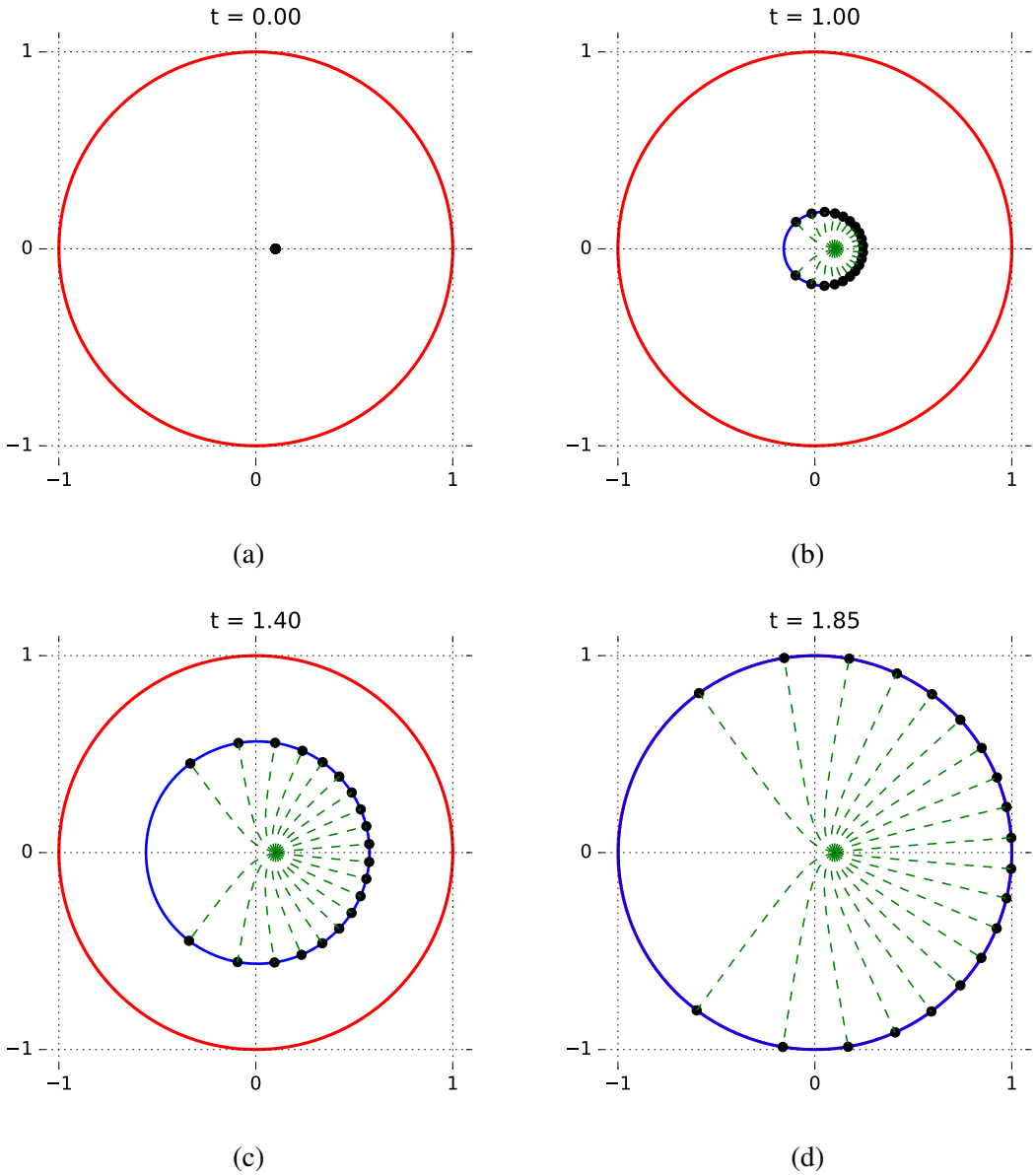


FIGURE 7.6. Support of m_t and optimal trajectories for $p = (0.1, 0)$ at times (a) $t = 0$, (b) $t = 1$, (c) $t = 1.4$, and (d) the final time $t \approx 1.85$.

and then define $\mu' = \sum_{k=1}^{N} \mu'_k \delta_{z_k}$ by

$$\mu'_k = \frac{\mu_k + \mu''_k}{2}. \tag{7.14}$$

The construction of μ'' in (7.13) uses the condition $|\gamma_k(T)| < 1$ for trajectories that did not arrive at the boundary as a mean to reduce the mass given to such trajectories. The renormalization will then increase the mass of trajectories that arrived first at the boundary. The average computed in (7.14) was used in order to improve stability of the method, even though it reduces convergence speed.

Figure 7.5 shows that, as expected, the measure μ yielding an equilibrium is uniform when $p = (0, 0)$ and its density around the angle π decreases as the x-component of p increases, with a very sharp decrease when $p = (0.1, 0)$. Figure 7.6 shows the behavior of the support of m_t^N and 20 optimal trajectories at different times from 0 until the final time $t \approx 1.85$. The 20 trajectories

represented were chosen in such a way that the total mass of agents between any two neighbor trajectories is the same, and thus their positions at time t provide an illustration of the measure m_t^N. One can observe that, even though p is close to the origin in this case, agents are much more concentrated on the right half-ball, with a few proportion of agents moving into the left half-ball.

REFERENCES

[1] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games: numerical methods. *SIAM J. Numer. Anal.*, 48(3):1136–1162, 2010.

[2] Y. Achdou and A. Porretta. Convergence of a finite difference scheme to weak solutions of the system of partial differential equations arising in mean field games. *SIAM J. Numer. Anal.*, 54(1):161–186, 2016.

[3] Y. Achdou and A. Porretta. Mean field games with congestion. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 35(2):443–480, 2018.

[4] L. Ambrosio. Transport equation and Cauchy problem for non-smooth vector fields. In *Calculus of variations and nonlinear partial differential equations*, volume 1927 of *Lecture Notes in Math.*, pages 1–41. Springer, Berlin, 2008.

[5] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[6] J.-P. Aubin and H. Frankowska. *Set-valued analysis*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2009. Reprint of the 1990 edition.

[7] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.

[8] J.-D. Benamou, G. Carlier, and F. Santambrogio. Variational mean field games. In *Active particles. Vol. 1. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 141–171. Birkhäuser/Springer, Cham, 2017.

[9] A. Bensoussan, J. Frehse, and S. C. P. Yam. On the interpretation of the Master Equation. *Stochastic Process. Appl.*, 127(7):2093–2137, 2017.

[10] M. Bernot, V. Caselles, and J.-M. Morel. *Optimal transportation networks*, volume 1955 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009. Models and theory.

[11] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[12] N. Bourbaki. *Topologie Générale. Chapitres 5 à 10*. Éléments de Mathématique. Springer, 2007.

[13] D. Braess. Über ein Paradoxon aus der Verkehrsplanung. *Unternehmensforschung*, 12:258–268, 1968.

[14] Y. Brenier. The least action principle and the related concept of generalized flows for incompressible perfect fluids. *J. Amer. Math. Soc.*, 2(2):225–255, 1989.

[15] M. Burger, M. D. Francesco, P. A. Markowich, and M.-T. Wolfram. On a mean field game optimal control approach modeling fast exit scenarios in human crowds. In *52nd IEEE Conference on Decision and Control*. IEEE, dec 2013.

[16] S. Cacace, F. Camilli, and C. Marchi. A numerical method for mean field games on networks. *ESAIM Math. Model. Numer. Anal.*, 51(1):63–88, 2017.

[17] F. Camilli, E. Carlini, and C. Marchi. A model problem for mean field games on networks. *Discrete Contin. Dyn. Syst.*, 35(9):4173–4192, 2015.

[18] P. Cannarsa and R. Capuani. Existence and uniqueness for mean field games with state constraints. Preprint arXiv:1711.01063.

[19] P. Cannarsa and M. Castelpietra. Lipschitz continuity and local semiconcavity for exit time problems with state constraints. *J. Differential Equations*, 245(3):616–636, 2008.

[20] P. Cannarsa, M. Castelpietra, and P. Cardaliaguet. Regularity properties of attainable sets under state constraints. In *Geometric control and nonsmooth analysis*, volume 76 of *Ser. Adv. Math. Appl. Sci.*, pages 120–135. World Sci. Publ., Hackensack, NJ, 2008.

[21] P. Cannarsa and C. Sinestrari. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*. Progress in Nonlinear Differential Equations and their Applications, 58. Birkhäuser Boston, Inc., Boston, MA, 2004.

[22] P. Cardaliaguet. Long time average of first order mean field games and weak KAM theory. *Dyn. Games Appl.*, 3(4):473–488, 2013.

[23] P. Cardaliaguet. Notes on mean field games (from P.-L. Lions' lectures at Collège de France). Available at `https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf`, 2013.

[24] P. Cardaliaguet. Weak solutions for first order mean field games with local coupling. In *Analysis and geometry in control theory and its applications*, volume 11 of *Springer INdAM Ser.*, pages 111–158. Springer, Cham, 2015.

[25] P. Cardaliaguet. The convergence problem in mean field games with local coupling. *Applied Mathematics & Optimization*, 76(1):177–215, jun 2017.

[26] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions. The master equation and the convergence problem in mean field games. Preprint arXiv:1509.02505.

[27] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, and A. Porretta. Long time average of mean field games with a nonlocal coupling. *SIAM J. Control Optim.*, 51(5):3558–3591, 2013.

[28] P. Cardaliaguet, A. R. Mészáros, and F. Santambrogio. First order mean field games with density constraints: pressure equals price. *SIAM J. Control Optim.*, 54(5):2672–2709, 2016.

[29] G. Carlier, C. Jimenez, and F. Santambrogio. Optimal transportation with traffic congestion and Wardrop equilibria. *SIAM J. Control Optim.*, 47(3):1330–1350, 2008.

[30] G. Carlier and F. Santambrogio. A continuous theory of traffic congestion and Wardrop equilibria. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 390(Teoriya Predstavleniĭ, Dinamicheskie Sistemy, Kombinatornye Metody. XX):69–91, 307–308, 2011.

[31] E. Carlini and F. J. Silva. A fully discrete semi-Lagrangian scheme for a first order mean field game problem. *SIAM J. Numer. Anal.*, 52(1):45–67, 2014.

[32] R. Carmona and F. Delarue. The master equation for large population equilibriums. In *Stochastic analysis and applications 2014*, volume 100 of *Springer Proc. Math. Stat.*, pages 77–128. Springer, Cham, 2014.

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

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

[35] F. H. Clarke. *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.

[36] J. E. Cohen and P. Horowitz. Paradoxical behaviour of mechanical and electrical networks. *Nature*, 352 (6337):699–701, aug 1991.

[37] M. G. Crandall and P.-L. Lions. Viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 277(1):1–42, 1983.

[38] E. Cristiani, B. Piccoli, and A. Tosin. Multiscale modeling of granular flows with application to crowd dynamics. *Multiscale Model. Simul.*, 9(1):155–182, 2011.

[39] S. Faure and B. Maury. Crowd motion from the granular standpoint. *Math. Models Methods Appl. Sci.*, 25(3):463–493, 2015.

[40] S. Fischer, H. Räcke, and B. Vöcking. Fast convergence to Wardrop equilibria by adaptive sampling methods. *SIAM J. Comput.*, 39(8):3700–3735, 2010.

[41] D. A. Gomes, J. Mohr, and R. R. Souza. Continuous time finite state mean field games. *Appl. Math. Optim.*, 68(1):99–143, 2013.

[42] D. A. Gomes and J. Saúde. Mean field games models—a brief survey. *Dyn. Games Appl.*, 4(2):110–154, 2014.

[43] A. Granas and J. Dugundji. *Fixed point theory*. Springer Monographs in Mathematics. Springer-Verlag, New York, 2003.

[44] O. Guéant. New numerical methods for mean field games with quadratic costs. *Netw. Heterog. Media*, 7(2):315–336, 2012.

[45] O. Guéant. Existence and uniqueness result for mean field games with congestion effect on graphs. *Appl. Math. Optim.*, 72(2):291–303, 2015.

[46] O. Guéant, J.-M. Lasry, and P.-L. Lions. Mean field games and applications. In *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Math.*, pages 205–266. Springer, Berlin, 2011.

[47] A. Haurie and P. Marcotte. On the relationship between Nash-Cournot and Wardrop equilibria. *Networks*, 15(3):295–308, 1985.

[48] D. Helbing, I. Farkas, and T. Vicsek. Simulating dynamical features of escape panic. *Nature*, 407(6803):487, 2000.

[49] L. F. Henderson. The statistics of crowd fluids. *Nature*, 229(5284):381–383, feb 1971.

[50] J.-B. Hiriart-Urruty. Extension of Lipschitz functions. *J. Math. Anal. Appl.*, 77(2):539–554, 1980.

[51] M. Huang, P. E. Caines, and R. P. Malhamé. Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions. In *42nd IEEE Conference on Decision and Control, 2003. Proceedings*, volume 1, pages 98–103. IEEE, 2003.

[52] M. Huang, P. E. Caines, and R. P. Malhamé. Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ε-Nash equilibria. *IEEE Trans. Automat. Control*, 52(9):1560–1571, 2007.

[53] M. Huang, R. P. Malhamé, and P. E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.*, 6(3):221–251, 2006.

[54] R. L. Hughes. A continuum theory for the flow of pedestrians. *Transportation Research Part B: Methodological*, 36(6):507–535, jul 2002.

[55] R. L. Hughes. The flow of human crowds. In *Annual review of fluid mechanics, Vol. 35*, volume 35 of *Annu. Rev. Fluid Mech.*, pages 169–182. Annual Reviews, Palo Alto, CA, 2003.

[56] V. N. Kolokoltsov, M. Troeva, and W. Yang. On the rate of convergence for the mean-field approximation of controlled diffusions with large number of players. *Dyn. Games Appl.*, 4(2):208–230, 2014.

[57] A. Lachapelle and M.-T. Wolfram. On a mean field game approach modeling congestion and aversion in pedestrian crowds. *Transportation Research Part B: Methodological*, 45(10):1572–1589, dec 2011.
[58] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006.
[59] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006.
[60] J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.
[61] B. Maury, A. Roudneff-Chupin, and F. Santambrogio. A macroscopic crowd motion model of gradient flow type. *Math. Models Methods Appl. Sci.*, 20(10):1787–1821, 2010.
[62] B. Maury, A. Roudneff-Chupin, F. Santambrogio, and J. Venel. Handling congestion in crowd motion modeling. *Netw. Heterog. Media*, 6(3):485–519, 2011.
[63] A. R. Mészáros and F. J. Silva. A variational approach to second order mean field games with density constraints: the stationary case. *J. Math. Pures Appl. (9)*, 104(6):1135–1159, 2015.
[64] B. Piccoli and A. Tosin. Time-evolving measures and macroscopic modeling of pedestrian flow. *Arch. Ration. Mech. Anal.*, 199(3):707–738, 2011.
[65] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The mathematical theory of optimal processes*. Translated from the Russian by K. N. Trirogoff; edited by L. W. Neustadt. Interscience Publishers John Wiley & Sons, Inc. New York-London, 1962.
[66] A. Prosinski and F. Santambrogio. Global-in-time regularity via duality for congestion-penalized mean field games. *Stochastics*, 89(6-7):923–942, 2017.
[67] T. Roughgarden and E. Tardos. How bad is selfish routing? *J. ACM*, 49(2):236–259, 2002.
[68] M.-F. Sainte-Beuve. On the extension of von Neumann–Aumann's theorem. *J. Functional Analysis*, 17:112–129, 1974.
[69] F. Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.
[70] M. Smith. In a road network, increasing delay locally can reduce delay globally. *Transportation Research*, 12(6):419–422, dec 1978.
[71] A. Taguchi. Braess' paradox in a two-terminal transportation network. *J. Oper. Res. Soc. Japan*, 25(4):376–389, 1982.
[72] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.
[73] J. G. Wardrop. Some theoretical aspects of road traffic research. *Proc. Inst. Civ. Eng.*, 1(3):325–362, 1952.

INDEX OF NOTATIONS

The following table provides the main notations used in this paper, together with their brief descriptions and the pages of their definitions.

Notation	Description	Page
$\mathcal{C}_X$	$\mathcal{C}(\mathbb{R}_+, X)$ with the topology of uniform convergence on compact sets.	3
$\mathrm{Lip}_c(X)$	Set of c-Lipschitz continuous functions in $\mathcal{C}_X$.	4
e_t	Evaluation map, associates with a curve γ its value at time t.	4
$\mathcal{P}(X)$	Set of Borel probability measures on X.	4
W_1	Wasserstein distance on $\mathcal{P}(X)$.	4
$\lvert\dot{\gamma}\rvert$	Metric derivative of γ.	4
$\partial^{\mathrm{C}}\phi$	Generalized gradient of ϕ.	4
$\mathrm{MFG}(X,\Gamma,K)$	Minimal-time mean field game in the space X with target set Γ and maximal speed given by the function K.	5
$\mathrm{OCP}(X,\Gamma,k)$	Minimal-time optimal control problem in the space X with target set Γ and maximal speed given by the function k.	5

Notation	Description	Page
$\mathrm{Adm}(k)$ $\mathrm{Adm}(m)$ $\mathrm{Adm}(Q)$	Sets of k-, m-, and μ^Q-admissible curves, respectively.	5, 6, 6
$\tau(t_0,\gamma)$	First exit time after t_0 of the curve γ.	5
$\mathrm{Opt}(k,t_0,x_0)$ $\mathrm{Opt}(m,t_0,x_0)$ $\mathrm{Opt}(Q,t_0,x_0)$	Set of optimal curves for k, m, and μ^Q, respectively, starting at the position x_0 at time t_0.	5, 6, 6
$\varphi(t,x)$	Value function of $\mathrm{OCP}(X,\Gamma,k)$.	8
$D^+w(x)$	Superdifferential of the function x at the point x.	11
$\mathcal{U}_{(t_0,x_0)}$	Set of optimal directions at (t_0,x_0).	15
$\widehat{\nabla w}(t,x)$	Normalized gradient of the function w with respect to x at (t,x).	16
Υ	Set of points (t,x) which are not starting points of optimal trajectories.	18
$\mathcal{Q}$	Set of measures in $\mathcal{P}(\mathcal{C}_X)$ concentrated on $K_{\max}$-Lipschitz continuous curves.	19
$\mathbf{Opt}(Q)$	Set of optimal trajectories for Q starting at time 0.	21

LABORATOIRE DE MATHÉMATIQUES D'ORSAY, UNIV. PARIS-SUD, CNRS, UNIV. PARIS-SACLAY, 91405 ORSAY, FRANCE.

Email address: guilherme.mazanti@math.u-psud.fr

INSTITUT CAMILLE JORDAN, UNIVERSITÉ CLAUDE BERNARD LYON 1, 43 BOULEVARD DU 11 NOVEMBRE 1918, 69622 VILLEURBANNE CEDEX, FRANCE.

Email address: santambrogio@math.univ-lyon1.fr