PhD Programme in Mathematics

Department of Mathematics

Time-optimal Control Problems in the Space of Measures

Giulia Cavagnari

UNIVERSITÀ DEGLI STUDI DI TRENTO

2016

Doctoral thesis in **Mathematics, XXIX cycle**

Department of Mathematics, **University of Trento**

Academic year 2015/2016

Supervisor: **Antonio Marigonda, University of Verona**

University of Trento

Trento, Italy

2016

Contents

Introduction

Classical minimum time problem in finite-dimension deals with the minimization of the time needed to steer a point $x_0 \in \mathbb{R}^d$ to a given closed subset S of $\mathbb{R}^d$, called the *target set*, along the trajectories of a controlled dynamics that can be presented by mean of a differential inclusion as follows:

$$\begin{cases} \dot{x}(t) \in F(x(t)), & t > 0, \\ \\ x(0) = x_0, \end{cases} \tag{0.1}$$

where F is a given set-valued map from $\mathbb{R}^d$ to $\mathbb{R}^d$, satisfying some structural assumptions, and whose value at each point denotes the set of admissible velocities at that point.

In this way it is possible to define the *minimum time function* T: given $x \in \mathbb{R}^d$, we define $T(x)$ to be the minimum time needed to steer such point to the given target S along trajectories of (0.1). The study of regularity properties of T is a central topic in optimal control theory and it has been extensively treated in literature. In particular, we refer to [20, 23] and to references therein, for recent results on the regularity of T in the framework of differential inclusions.

The present work aims to generalize the finite-dimensional time-optimal control problem to the infinite-dimensional setting of Borel measures. The main motivation for such a formulation is to model situations in which the knowledge of the initial state x_0 is only probabilistic, for example it can be obtained only by an averaging of many measurement processes, e.g. when measurements are affected by noises, or also in cases in which we are interested in modeling *multi-agent systems*, where the number of agents is so huge to make viable only a statistical (*macroscopic*) description of the system. In the first case, the time-evolving measure represents our probabilistic knowledge about the state of the particle, while, in the second case, it represents the statistical distribution of the agents. It is worth noticing that this situations can happen even if we assume a pure deterministic evolution of the system as it is in our case of study.

In the framework of crowd dynamics, several studies have been made to provide mathematical models and numerical simulations to take into account different kinds of behaviour of pedestrians, related also to mutual interactions. For instance, a possible application comes from the *evacuation problem* in the pedestrian dynamics, where the objective is to drive a crowd of people outside of a room in the minimum amount of time.

A very recent survey on this topic is the monograph [39], providing a new and unified multiscale description based on measure theory for the modeling of the crowd dynamics, which usually follows two main points of view, a *microscopic* and a *macroscopic* one, in order to analyze the relations between *individual* and *collective* behaviours, respectively.

To model real-world situations, it is also needed to consider situations where the evolving total mass is not conserved in time, as it happens for instance in the evacuation problems where the pedestrians are removed from the system once they get outside of the room. In this case, the evolving mass solves a continuity equation with sink. To treat cases of transport equation with source/sinks, and more precisely to compare measures with different total mass, the classical Wasserstein distance between probability measures cannot be used, thus in [64, 65] a generalized Wasserstein distance between positive finite Borel measures is introduced.

A measure theoretic approach for transportation problems can be found also in [66] where the modeling approach relies on the concept of discrete-time evolving measures and in [19] in which authors focus mainly on concentration and congestion effects.

For other possible references regarding the study of *multi-agents systems*, we address the reader to [24] in which the target is not a physical object, indeed the aim is to find the *sparsest* control strategy (i.e. action concentrated on the fewest number of agents) to achieve a state in which the evolving group will reach an alignment consensus by self-organization. The notes [27] presents instead a summary on the mean-filed limit for a huge number of interacting particles with applications to swarming models, while in [44, 45] the authors introduce and develope the concept of *mean-field optimal control* in which the individuals are not freely interacting but influenced by an external *policy maker* so that the moving population is divided into leaders and followers.

Due to this reasons, other authors have investigated different problems studying systems for which the initial conditions are given by a probability distribution, instead of a deterministic point, e.g. in [17] in which a stocastic approach is presented, or in [49] in which the authors adopt a random variable approach.

Motivated by the previous considerations and considering a deterministic dynamics, in Chapter 2 we will give a general description of a control problem in the space of positive Borel measures studying basic properties on very general cost functionals stating the problem both in a mass-preserving setting and in a non-isolated case with instantaneous annihilation of the evolving mass.

More specifically, in a mass-preserving setting, a time-optimal control problem in the space of probability measures endowed with the topology induced by the Wasserstein metric will be introduced in Chapter 3 (see [28, 30–32]), where the dynamics is given by a *controlled continuity equation* in the space of probability measures, which naturally arises as an infinite-dimensional counterpart of a finite-dimensional differential inclusion.

Indeed, a natural choice to model our knowledge about the particle's starting position is to consider it as a Borel probability measure $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$, looking to a new *macroscopic* control system made by a suitable *superposition* of a continuum of weighted solutions of the classical differential inclusion (0.1) starting from each point of the support of μ_0 (*microscopic* point of view). The case in which

μ_0 is a Dirac delta concentrated at a point x_0 corresponds of course to the classical case in which perfect knowledge of the starting position is assumed.

The *deterministic* time evolution of the macroscopic system in the space of probability measures, under suitable assumptions, can be thought as ruled by the *(controlled) continuity equation* to be understood in the distributional sense

$$\begin{cases} \partial_t \mu(t,x) + \operatorname{div}(v_t(x)\mu(t,x)) = 0, & \text{for } 0 < t < T,\ x \in \mathbb{R}^d, \\ \\ \mu(0,\cdot) = \mu_0, \end{cases} \tag{0.2}$$

which represents the conservation of the total mass $\mu_0(\mathbb{R}^d)$ during the evolution. The resulting admissible mass-preserving trajectories $\boldsymbol{\mu} := \{\mu_t\}_{t\in[0,T]}$, $\mu_{|t=0} = \mu_0$, are time-depending Borel probability measures on $\mathbb{R}^d$. Here $v_t(x)$ is a suitable time-depending Eulerian vector field, representing the velocity of the mass crossing position x at time t.

In order to reflect the original control system (0.1) at a microscopic level, a natural requirement on the vector field $v_t(\cdot)$ is to be a $L^1_{\mu_t}$-Borel selection of the set-valued map $F(\cdot)$: this means that the microscopic particles/agents still obey the nonholonomic constraints coming from (0.1). On the other hand, since the conservation of the mass gives us the property $\mu(t,\mathbb{R}^d) = \mu_0(\mathbb{R}^d)$ for all t, we are entitled – according to our motivation – to say that the measure $\mu(t,\cdot)$ actually represents the probability distribution in the space $\mathbb{R}^d$ of the evolving particles at time t.

The analysis of (0.2) by mean of the superposition of ODEs of the form $\dot{x}(t) = v(x(t))$, or $\dot{x}(t) = v(t, x(t))$, has been extensively studied in the past years by many authors mainly inspired by a result appearing in the appendix of [75]: for a general introduction, an overview of known results and open problems, and a comprehensive bibliography, we refer to the recent survey [1]. The main issue in these problems is to study existence, uniqueness and regularity of the solution of (0.2), for μ_0 in a suitable class of measures, when the vector field v has low regularity and, hence, it does not ensure that the corresponding ODEs have a (possibly not unique) solution among absolutely continuous functions, for every initial data x_0. In this case, the solution of (0.2) provides existence and uniqueness not in a pointwise sense, but rather *generically*.

Moreover, also the links between continuity equation (0.2) and optimal transport theory have been investigated recently by many authors. One can prove that suitable subsets of $\mathscr{P}(\mathbb{R}^d)$ can be endowed with a metric structure – the *Wasserstein metric* – whose absolutely continuous curves turn out to be precisely the solutions of (0.2). This has been applied to solve many variational problems, among which we recall optimal transport problems, asymptotic limit for gradient flows of integral functionals, and calculus of variations in infinite dimensional spaces. We refer to [9, 15, 41, 74] for an introduction to the subject, and for generalizations from $\mathbb{R}^d$ to infinite dimensional metric spaces. However, we will not address this problem in this work.

It is well known that, in the case in which $v_t(\cdot)$ is locally Lipschitz in x uniformly w.r.t. t, the solution of the continuity equation (0.2) can be represented as the *push forward* of the initial state μ_0 through the *unique* solution T_t of the *characteristic system*

$$\begin{cases}\dot\gamma(t)=v_t(\gamma(t)), & \text{for } \mathscr{L}^1\text{-a.e. } t\in(0,T),\\ \gamma(0)=x,\end{cases}\tag{0.3}$$

i.e. $\mu_t = T_t\sharp\mu_0$ for all $t\in[0,T)$, where the push-forward of μ_0 through T_t (called *transport map*) is defined by $T_t\sharp\mu_0(B) := \mu_0(T_t^{-1}(B))$, for all Borel sets $B\subseteq\mathbb{R}^d$. Regularity properties of v_t are crucial to have such a representation formula.

However (0.2) has been proven to be well-posed even in situations in which the regularity of the vector field v_t is not sufficient to guarantee uniqueness of the solutions of (0.3). Heuristically, this is due to the fact that the evolution of the measure is not affected by singularities in a μ_t-negligible set. Following [9], we recall that the integrability assumption $\|v_t\|_{L^p_\mu(\mathbb{R}^d)}\in L^1([0,T])$ yields the existence of a solution of (0.2) in the sense of a continuous curve $t\mapsto\mu_t$ in the space of probability measures endowed with the weak* topology induced by the duality with continuous and bounded functions $\varphi\in C^0_b(\mathbb{R}^d)$ (i.e., a *narrowly continuous curve* in the space of probability measures).

In Theorem 8.2.1 in [9] and Theorem 5.8 in [15], the so called *Superposition Principle* states that, if we require much milder assumptions on v_t, the solution μ_t of the continuity equation can be characterized by the push-forward $e_t\sharp\boldsymbol{\eta}$, where $e_t:\mathbb{R}^d\times\Gamma_T\to\mathbb{R}^d, (x,\gamma)\mapsto\gamma(t)$, $\Gamma_T := C^0([0,T];\mathbb{R}^d)$ and $\boldsymbol{\eta}$ is a probability measure in the infinite-dimensional space $\mathbb{R}^d\times\Gamma_T$ concentrated on those pairs $(x,\gamma)\in\mathbb{R}^d\times\Gamma_T$ such that γ is an integral solution of the underlying characteristic system (0.3). We refer the reader to the surveys [1, 9] and the references therein for a deep analysis of this approach that is at the basis of the present work.

Pursuing the goal of facing control systems involving measures, we define a generalization of the target set S by duality. We consider an observer that is interested in measuring some quantities $\phi(\cdot)\in\Phi$ (*observables*); the results of this measurements are the average of these quantities w.r.t. the state of the system. The elements of the generalized target set $\tilde S^\Phi$ are the states for which the results of all these measurements are below a fixed thershold.

Another possible interpretation of our framework in this case can be given in terms of *pedestrian dynamics*: suppose to have initially a crowd of people represented by a (normalized) probability measure μ_0 and to be able to identify a *safety zone* $S\subseteq\mathbb{R}^d$, while $F(\cdot)$ represents some (possible) nonholonomic constraints to the motion. Then if our aim in case of danger is to steer all the crowd to the safety zone in the minimum amount of time, we can choose $\Phi=\{d_S(\cdot)\}$. In a more realistic situation, it may not be possible to steer *all* the crowd to S. If we fix $\alpha\in[0,1]$ and choose $\Phi=\{d_S(\cdot)-\alpha\}$, we are still satisfied for example if the ratio between the number of people in the safe zone and all the people is above $1-\alpha$, or if we can take the people sufficiently near to the safe zone.

Having defined the set of admissible trajectories and the target set in the space of probability measures, the definition of generalized minimum time function at a probability measure μ_0 is the straigthforwardly generalization of the classical one, i.e., the infimum of all the times T for which there exists an admissible trajectory defined on $[0,T]$ and satisfying $\mu_T\in\tilde S^\Phi$.

Our main results for Chapter 3 can be summarized as follows:

- a theorem of existence of time–optimal curves in the space of probability measures;
- a Dynamic Programming Principle;
- a comparison result between classical and generalized minimum time functions in some cases;
- some attainability results and sufficient conditions yielding Lipschitz continuity of the generalized minimum time function (see [28]);
- the proof that the generalized minimum time function is a viscosity solution in a suitable sense of an Hamilton-Jacobi-Bellman equation analoguos to the classical one;
- the definition of a correspondent quantity for the Lie bracket in a measure-theoretic setting for nonsmooth vector fields (see [29]) in order to open the door to the study of higher order controllability conditions in this framework.

Since classical minimum time function can be characterized as unique viscosity solution of a Hamilton-Jacobi-Bellman equation, the problem to study a similar formulation for the generalized setting would be quite interesting. Several authors have treated a similar problem in the space of probability measures or in a general metric space, giving different definitions of sub-/super differentials and viscosity solutions (see e.g. [7, 9, 26, 46, 47], or [48] for a new notion of viscosity solution for Eikonal equations in a general metric space). For example, the theory presented in [47] is quite complete: indeed there are proved also results on time-dependent problems, comparison principles granting uniqueness of the viscosity solutions under very reasonable assumptions.

However, when we consider as metric space the space $\mathscr{P}_2(\mathbb{R}^d)$, i.e. the space of probability measures with 2-moment finite, we notice that the class of equations that can be solved is quite small: the general structure of metric space of [47] allows only to rely on the metric gradient, while $\mathscr{P}_2(\mathbb{R}^d)$ enjoys a much more richer structure in the tangent space (which, at many points, can be identified with a subset of L^2).

Dealing with the definition of sub-/superdifferential given in [26], the major bond is that the "perturbed" measure is assumed to be of the form $(\mathrm{Id}_{\mathbb{R}^d} + \phi)\sharp\mu$ in which a (rescaled) transport plan is used. It is well known that, by Brenier's Theorem, if $\mu \ll \mathscr{L}^d$ in this way we can describe *all* the measures near to μ. However in general this is not true. Thus if the set of admissible trajectories contains curves whose points are not all a.c. w.r.t. Lebesgue measure (as in our case), the definition in [26] cannot be used.

In order to fully exploit the richer structure of the tangent space of $\mathscr{P}_2(\mathbb{R}^d)$, recalling that AC curves in $\mathscr{P}_2(\mathbb{R}^d)$ are characterized to be weak solutions of the continuity equation (Theorem 8.3.1 in [9]), we considered a different definition than the one presented in [26] using the *Superposition Principle*.

In this work, we just proved that the generalized minimum time function solves in a suitable viscosity sense a natural Hamilton-Jacobi-Bellman equation, which presents strong analogies with the finite-dimensional case. However, a

Comparison Principle for the generalized HJB equation is still the principal open problem in this framework, as well as to give a Pontryagin's maximum principle comparable with the classical one.

Related to such a problem, a further application could be the theory of mean field games [54, 55]. According to this theory, in games with a continuum of agents, having the same dynamics and the same performance criteria, the value function for an average player can be retrieved by solving an infinite dimensional Hamilton–Jacobi equation, coupled with the continuity equation describing how the mass of players evolves in time.

Another application might be in the context of discontinuous feedback controls for general nonlinear control systems $\dot{x} = f(x, u)$. Here, the construction of stabilizing or nearly optimal controls $x \mapsto u(x)$ cannot be performed, even for smooth dynamics, among continuous controls [72]. However, it is possible to construct discontinuous feedback controls which are stabilizing or nearly optimal, and whose discontinuities are sufficiently tame to ensure the existence of Carathéodory solutions for the closed loop system $\dot{x} = f(x, u(x))$, the so–called *patchy feedback controls* [10, 11, 16], but uniqueness only holds for a set of full measure of initial data.

Finally, in Chapter 4 (see [33]) we move from the framework presented in Chapter 3, but with a different formulation of the time-optimal problem and allowing the loss of mass during the evolution, which turns out to be closer to applications in pedestrian dynamics or general multi-agent systems.

More precisely, in this chapter we consider an admissible mass-preserving trajectory $\boldsymbol{\mu} \subseteq \mathscr{P}(\mathbb{R}^d)$ in the space of probability measures coupled with a density decreasing in time.

The problem we have in mind can be seen as a problem of *optimal equipment*. Indeed, we consider a target set $S \subseteq \mathbb{R}^d$, strongly invariant for the underlying differential inclusion driven by F, which represents for example a region of the space where we want to steer our initial state $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$ describing the given initial distribution of agents (ex. cars). To every admissible mass-preserving trajectory $\boldsymbol{\mu}$ starting by μ_0, it is assigned an admissible function $f_0 : \mathbb{R}^d \to [0, +\infty]$ called *clock-function*, which expresses the amount of goods (ex. fuel) that has to be assigned to each agent/car in the support of μ_0 in order to reach the target following the trajectory $\boldsymbol{\mu}$. We treated the case in which we have a time-linear consumption of goods for our problem.

From a macroscopic point of view, this defines a new concept of admissible trajectory in the space of positive Borel measures that we call *clock-trajectory*, which is no more mass-preserving but it looses its mass linearly in time.

Our aim is to minimize the average of f_0 w.r.t. the initial distribution of agents, μ_0, among all the Borel functions f_0 keeping nonnegative the density associated to $\boldsymbol{\mu}$ along all the evolution.

Equivalently, in a time-optimal context, the problem can be interpreted as follows thinking about the *evacuation problem*. The target S stands for the doors through which we want to drive a mass of people whose initial distribution is described by μ_0. The strong invariance of S means that, from a microscopic point of view, once a single agent has reached the target we remove it from the system. Here, f_0 represents the time assigned to the agents to reach the target, so the density associated to $\boldsymbol{\mu}$ works as a countdown. The cost to minimize is then $\int_{\mathbb{R}^d} f_0(x)\, d\mu_0(x)$.

We will show also that the best clock-function can be interpreted as the minimum amount of time that has to be assigned at the beginning to each agent in order to reach the target. In this sense the optimal vector field for the problem in the space of measures can be seen as a *measurable feedback strategy* for the underlying finite-dimensional control problem.

The main results of Chapter 4 are as follows:

· an approximation and representation result in the mass-preserving setting;

· a theorem of existence of an optimal clock-trajectory for the system, which proves also that the optimal clock-function turns out to be the classical minimum time function;

· a Dynamic Programming Principle and some regularity results on the value function;

· an Hamilton-Jacobi-Bellman equation, solved in a suitable viscosity sense by the value function, in analogy with the problem discussed in Chapter 3.

To conclude, in the last Chapter 5 we list the main open problems.

Notation

$\mathscr{P}(X)$	Space of probability measures on a separable metric space X
$\mathscr{P}_p(X)$	Space of probability measures with finite p-moment (see Definition 1.1.5)
$\mathscr{M}(X)$	Space of finite Radon measures on a separable metric space X
$\mathscr{M}^+(X)$	Subspace of $\mathscr{M}(X)$ made of positive measures
$\mathscr{M}(X;\mathbb{R}^d)$	Space of Radon $\mathbb{R}^d$-valued measures on a separable metric space X
$\lvert\nu\rvert$	Total variation of $\nu \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$
$\mathscr{L}^d$	d-dimensional Lebesgue's measure
$\operatorname{supp}\mu$	Support of a measure μ
$r\sharp\mu$	Push-forward of μ through r (see Definition 1.1.3)
$\Pi(\mu_1,\mu_2)$	Set of admissible transport plans with marginals μ_1,μ_2
$\Pi_o(\mu_1,\mu_2)$	Set of optimal transport plans with marginals μ_1,μ_2
$W_p(\mu_1,\mu_2)$	p-th Wasserstein distance between μ_1 and μ_2)
$\mathrm{m}_p(\mu)$	p-th moment of a measure μ
$L^p_\mu(X)$	L^p space of μ-measurable real maps defined on X
$L^p_\mu(\mathbb{R}^d;\mathbb{R}^d)$	L^p space of μ-measurable maps from $\mathbb{R}^d$ to $\mathbb{R}^d$
cl_{W_p}	Closure in p-Wasserstein topology
$\mathrm{cl}_{d_{\mathscr{P}}}$	Weak*-closure
$\mathrm{dom}(g)$	Domain of the function g
$\mathrm{Lip}(g,D)$	Lipschitz constant of the function $g:\mathbb{R}^d\to\mathbb{R}$ on the set $D\subseteq\mathbb{R}^d$
$C^0_b(X;Y)$	Space of continuous and bounded functions from X to Y
$C^0_b(X)$	Space of continuous and bounded real functions defined on X
$C^0_C(X;Y)$	Space of continuous functions from X to Y with compact support in X
$C^0_C(X)$	Space of continuous real functions with compact support in X
Γ_I	Space of continuous functions from $I=[a,b]\subseteq\mathbb{R}$ to $\mathbb{R}^d$
Γ_T	Space of continuous functions from $[0,T]\subseteq\mathbb{R}$ to $\mathbb{R}^d$
Γ_T^x	Space of maps in Γ_T starting at $x\in\mathbb{R}^d$
$\mathrm{AC}^p([a,b];\mathbb{R}^d)$	Space of absolutely continuous maps $\gamma:[a,b]\to\mathbb{R}^d$ with $\dot\gamma\in L^p([a,b])$
$\mathrm{Bor}(X)$	Set of Borel maps from a separable metric space X to $\mathbb{R}$

$\mathrm{Bor}_b(X)$	Subset of $\mathrm{Bor}(X)$ made of bounded maps
$\mathrm{Bor}(\mathbb{R}^d;\mathbb{R}^d)$	Set of Borel maps from $\mathbb{R}^d$ to $\mathbb{R}^d$
$\mathrm{SC}(A;\mathbb{R})$	Space of semiconcave functions from an open set $A \subseteq \mathbb{R}^d$ to $\mathbb{R}$
$Id_{\mathbb{R}^d}$	Identity map on $\mathbb{R}^d$
$I_A(\cdot)$	Indicator function of $A \subseteq X$ (see Definition 1.0.4)
$\chi_A(\cdot)$	Characteristic function of $A \subseteq X$ (see Definition 1.0.4)
$\sigma_A(\cdot)$	Support function to $A \subseteq X$ (see Definition 1.0.5)
$d_A(\cdot)$	Distance function from a closed, nonempty set $A \subseteq \mathbb{R}^d$
pr^i	Projection operator on the i-th component defined on a product space X^N, $N \geq 1$
$\partial^+ f(x)$	Fréchet superdifferential of a function $f : A \to \mathbb{R}$ at $x \in A$
$B(x,r)$	Open ball of radius r centered at $x \in \mathbb{R}^d$
A^c	Complementary set of a subset $A \subseteq \mathbb{R}^d$, i.e. $\mathbb{R}^d \setminus A$
$\mathrm{co}\, A$	Convex hull of $A \subseteq \mathbb{R}^d$

Chapter 1

Preliminaries

In this chapter we review some concepts from measure theory, optimal transport, and control theory.

Let us begin by listing some preliminary definitions and notations.

Throughout this work, if X is a separable metric space, we will denote with $\mathrm{Bor}(X)$ the set of Borel maps from X to $\mathbb{R}$, with $\mathrm{Bor}_b(X)$ the set of bounded Borel maps from X to $\mathbb{R}$, and with $\mathrm{Bor}(\mathbb{R}^d;\mathbb{R}^d)$ the set of Borel maps from $\mathbb{R}^d$ to $\mathbb{R}^d$.

We will denote with $\mathscr{L}^d$ the d-dimensional Lebesgue's measure.

Definition 1.0.1.

(i) A *modulus of continuity* is a function $\omega : [0,+\infty] \to [0,+\infty]$ such that $\lim_{t\to 0^+} \omega(t) = \omega(0) = 0$.

(ii) Given $x \in \mathbb{R}^d$, we say that a function $f : \mathbb{R}^d \to \mathbb{R}^d$ admits $\omega_x(\cdot)$ as modulus of continuity at the point x if and only if for all $y \in \mathbb{R}^d$

$$|f(y) - f(x)| \leq \omega_x(|y - x|).$$

(iii) Given a function $g : \mathbb{R}^d \to \mathbb{R}$ and $D \subseteq \mathbb{R}^d$, we define the *Lipschitz constant of g on D* to be

$$\mathrm{Lip}(g, D) := \sup_{\substack{x,y\in D\\ x\neq y}} \frac{|g(x) - g(y)|}{|x - y|}.$$

When $D = \mathbb{R}^d$ we will omit it, thus $\mathrm{Lip}(g) := \mathrm{Lip}(g, \mathbb{R}^d)$.

Let us now recall the following definitions of semiconcave function and superdifferential given in [22].

Definition 1.0.2 (Superdifferential)**.** Let $A \subseteq \mathbb{R}^d$ be open, $x \in A$. We define the *(Fréchet) superdifferential of a function $f : A \to \mathbb{R}$ at x* by

$$\partial^+ f(x) := \left\{\xi(x) \in \mathbb{R}^d : \limsup_{y\to x} \frac{f(y) - f(x) - \langle \xi(x), y - x\rangle}{|y - x|} \leq 0\right\}.$$

Definition 1.0.3 (Semiconcave function)**.** Let $K > 0$, $A \subseteq \mathbb{R}^d$ be open, $x \in A$. A function $f : A \to \mathbb{R}$ is said to be *semiconcave at x with constant $K > 0$* if for all $\xi(x) \in \partial^+ f(x)$ we have

$$f(y) - f(x) \leq \langle \xi(x), y - x \rangle + K\,|y - x|^2$$

for any point $y \in A$ such that $[y, x] \subset A$.

If f is semiconcave for all $x \in A$ we write $f \in SC(A; \mathbb{R})$.

Definition 1.0.4. Let X be a set, $A \subseteq X$.

1. The *indicator function of* A is the function $I_A : X \to \{0, +\infty\}$ defined as $I_A(x) = 0$ for all $x \in A$ and $I_A(x) = +\infty$ for all $x \notin A$.
2. The *characteristic function of* A is the function $\chi_A : X \to \{0, 1\}$ defined as $\chi_A(x) = 1$ for all $x \in A$ and $\chi_A(x) = 0$ for all $x \notin A$.

Definition 1.0.5 (Support function)**.** Let X be a Banach space, X' be its topological dual, $A \subseteq X$ be nonempty. We define the *support function* to A at $x^* \in X'$ by setting

$$\sigma_A(x^*) := \sup_{x \in A} \langle x^*, x \rangle_{X',X} . \tag{1.1}$$

It turns out that $\sigma_A(x^*) = \sigma_{\overline{\mathrm{co}}(A)}(x^*)$ for every $x^* \in X'$ and that $\sigma_A : X' \to \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous.

Definition 1.0.6. Given $T \in [0, +\infty[$, the *evaluation map* $e_t : \mathbb{R}^d \times \Gamma_T \to \mathbb{R}^d$ is defined as $e_t(x, \gamma) = \gamma(t)$ for all $0 \leq t \leq T$. Moreover, we set

$$\Gamma_T := C^0([0, T]; \mathbb{R}^d), \qquad \Gamma_T^x := \{\gamma \in \Gamma_T : \gamma(0) = x\},$$

where $x \in \mathbb{R}^d$. We endow all the above spaces with the usual sup-norm, recalling that Γ_T is a separable Banach space for every $0 < T < +\infty$.

1.1 Measure theory

In this section we recall some essential definitions and results on measure theory. Our main references for this part are [9, 74].

Definition 1.1.1 (Probability measures)**.** Let X be a complete separable metric space, $\mathscr{P}(X)$ be the set of Borel probability measures on X. Since $\mathscr{P}(X)$ can be identified with a convex subset of the unitary ball of $(C_b^0(X))'$ (the dual space of the space of bounded continuous functions on X), we can equip $\mathscr{P}(X)$ with the weak* topology induced by $(C_b^0(X))'$. In particular, we say that a sequence of probability measures $\{\mu_n\}_{n \in \mathbb{N}}$ is *w^*-convergent* (or *narrowly converges*) to a probability measure $\mu \in \mathscr{P}(X)$, and write $\mu_n \rightharpoonup^* \mu$, if and only if for every $f \in C_b^0(X)$ it holds

$$\lim_{n \to \infty} \int_X f(x)\, d\mu_n(x) = \int_X f(x)\, d\mu(x).$$

We will consider on $\mathscr{P}(X)$ the σ-algebra of Borel sets generated by the w^*-open subsets of $\mathscr{P}(X)$.

We have that the space $\mathscr{P}(X)$, endowed with the w^*-topology, is metrizable (for instance by the Prokhorov's metric). We will denote by $\mathrm{d}_{\mathscr{P}}$ any metric on $\mathscr{P}(X)$ inducing the w^*-topology on $\mathscr{P}(X)$.

Definition 1.1.2 (Tightness)**.** Let X be a metric space and $\mathscr{K} \subseteq \mathscr{P}(X)$. We say that $\mathscr{K}$ is *tight* if for every $\varepsilon > 0$ there exists a compact subset K_ε of X such that $\mu(X \setminus K_\varepsilon) \leq \varepsilon$ for every $\mu \in \mathscr{K}$. Every tight subset of $\mathscr{P}(X)$ is relatively compact in $\mathscr{P}(X)$. The converse is true if there exists an equivalent complete metric on X.

This last result is known as Prokhorov's theorem (see for instance [9, 73, 74] or the recent books [8, 71]).

Given a separable metric space X, we denote by $\mathscr{M}(X)$ the set of finite Radon measures on X, with $\mathscr{M}^+(X) \subset \mathscr{M}(X)$ the measures that are also positive and with $\mathscr{M}(X;\mathbb{R}^d)$ the set of Radon $\mathbb{R}^d$-valued measures on X.

Definition 1.1.3 (Push forward)**.** If X, Y are separable metric spaces, $\mu \in \mathscr{M}(X)$, and $r : X \to Y$ is a Borel (or, more generally, μ-measurable) map, we denote by $r\sharp\mu \in \mathscr{M}(Y)$ the push-forward of μ through r, defined by

$$r\sharp\mu(B) := \mu(r^{-1}(B)), \text{ for all Borel sets } B \subseteq Y.$$

Equivalently, we have

$$\int_X f(r(x))\, d\mu(x) = \int_Y f(y)\, dr\sharp\mu(y),$$

for every bounded (or $r\sharp\mu$-integrable) Borel function $f : Y \to \mathbb{R}$.

Observe that, by definition, the push-forward operator is mass-preserving.

Proposition 1.1.4 (Properties of push forward)**.** *Let X, Y, Z be separable metric spaces, $\mu \in \mathscr{P}(X)$, and let $r : X \to Y$ be a Borel map.*

1. *If $\nu \in \mathscr{P}(X)$ satisfies $\nu \ll \mu$, then $r\sharp\nu \ll r\sharp\mu$.*
2. *Given a Borel map $s : Y \to Z$, the following composition rule holds*
$$(s \circ r)\sharp\mu = s\sharp(r\sharp\mu).$$
3. *If $r \in C^0(X;Y)$ then $r\sharp : \mathscr{P}(X) \to \mathscr{P}(Y)$ is continuous with respect to the narrow convergence and*
$$r(\operatorname{supp}\mu) \subseteq \operatorname{supp}(r\sharp\mu) = \overline{r(\operatorname{supp}\mu)}.$$
4. *Let $\{r_n : X \to Y\}_{n\in\mathbb{N}}$ be a sequence of Borel maps uniformly convergent to r on compact subsets of X, and let $\{\mu_n\}_{n\in\mathbb{N}} \subseteq \mathscr{P}(X)$ be a tight sequence narrowly convergent to μ. Then if r is continuous, we have that $r_n\sharp\mu_n \rightharpoonup^* r\sharp\mu$.*

Proof. See [9], Chapter 5, Section 2. □

Definition 1.1.5 (p-moment)**.** Let X be a separable Banach space, $\mu \in \mathscr{P}(X)$, $p \geq 1$. We say that μ has finite p-moment if

$$\mathrm{m}_p(\mu) := \int_X |x|^p\, d\mu(x) < +\infty.$$

Equivalently, we have that μ has p-moment finite if and only if for every $x_0 \in X$ we have

$$\int_X |x - x_0|^p \, d\mu(x) < +\infty.$$

We denote by $\mathscr{P}_p(X)$ the subset of $\mathscr{P}(X)$ consisting of probability measures with finite p-moment.

Definition 1.1.6 (Uniform integrability)**.** Let X be a separable Banach space, $\mathscr{K} \subseteq \mathscr{P}(X)$, $g : X \to [0, +\infty]$ be a Borel function. We say that

1. g is *uniformly integrable with respect to* $\mathscr{K}$ if

$$\lim_{k\to\infty} \sup_{\mu\in\mathscr{K}} \int_{\{x\in X:\, g(x)>k\}} g(x) \, d\mu(x) = 0.$$

2. the set $\mathscr{K}$ *has uniformly integrable p-moments*, $p \geq 1$, if $|x|^p$ is uniformly integrable with respect to $\mathscr{K}$.

Lemma 1.1.7 (Uniform integrability criterion)**.** *Let X be a separable Banach space, $\mathscr{K} = \{\mu_n\}_{n\in\mathbb{N}} \subseteq \mathscr{P}(X)$, $p \geq 1$, $\mu_n \rightharpoonup^* \mu \in \mathscr{P}(X)$. Then the set $\mathscr{K}$ has uniformly integrable p-moments if and only if*

$$\lim_{n\to\infty} \int_X f(x) \, d\mu_n(x) = \int_X f(x) \, d\mu(x),$$

for every continuous function $f : X \to \mathbb{R}$ such that there exist $a, b \geq 0$ and $x_0 \in X$ with $|f(x)| \leq a + b|x - x_0|^p$ for every $x \in X$.

Proof. See Lemma 5.1.7 of [9]. □

1.2 Optimal transport and Wasserstein distances

This section is devoted to recall the very basic definitions and results in transport theory. We mention that a first research attempt in this field was proposed by Monge in 1781 in [62] and then reformulated by Kantorovich in 1942 in [53]. We refer the reader to [73, 74] or to the recent books [8, 71] for an introduction and a deep study in this field.

For the following, let X be a separable Banach space.

Definition 1.2.1 (Wasserstein distance)**.** Given $\mu_1, \mu_2 \in \mathscr{P}(X)$, $p \geq 1$, we define the *p-Wasserstein distance* between μ_1 and μ_2 by setting

$$W_p(\mu_1, \mu_2) := \left(\inf \left\{ \iint_{X\times X} |x_1 - x_2|^p \, d\pi(x_1, x_2) \,:\, \pi \in \Pi(\mu_1, \mu_2) \right\} \right)^{1/p}, \tag{1.2}$$

where the set of *admissible transport plans* $\Pi(\mu_1, \mu_2)$ is defined by

$$\Pi(\mu_1, \mu_2) := \left\{ \pi \in \mathscr{P}(X \times X) : \begin{array}{l} \pi(A_1 \times X) = \mu_1(A_1), \\ \pi(X \times A_2) = \mu_2(A_2), \end{array} \right.$$
$$\left. \text{for all } \mu_i\text{-measurable sets } A_i,\ i = 1, 2 \right\}.$$

We also denote with $\Pi_o^p(\mu_1, \mu_2)$ the subset of $\Pi(\mu_1, \mu_2)$ consisting of optimal transport plans, i.e. the set of all plans π for which the infimum in (1.2) is attained. We will also use the notation $\Pi_o(\mu_1, \mu_2)$ when the context makes clear which distance W_p is being considered.

In the following, we summarize some properties of the Wasserstein metric. For a detailed discussion on Wasserstein distance we refer to chapter 7 in [74], chapter 6 in [73], or section 7.1 in [9].

Proposition 1.2.2. *$\mathscr{P}_p(X)$ endowed with the p-Wasserstein metric $W_p(\cdot, \cdot)$ is a complete separable metric space. Moreover, given a sequence $\{\mu_n\}_{n\in\mathbb{N}} \subseteq \mathscr{P}_p(X)$ and $\mu \in \mathscr{P}_p(X)$, we have that the following are equivalent*

1. $\lim_{n\to\infty} W_p(\mu_n, \mu) = 0$,
2. $\mu_n \rightharpoonup^* \mu$ *and* $\{\mu_n\}_{n\in\mathbb{N}}$ *has uniformly integrable* p*-moments.*

Proof. See Proposition 7.1.5 in [9]. □

Proposition 1.2.3. *The Wasserstein distance defined above satisfies the following properties:*

- Metric character. *W_p is a pseudo-distance on $\mathscr{P}(X)$, i.e. it satisfies the axioms of the distance, but it can assume the value $+\infty$. Namely, for all $\mu_0, \mu_1, \mu_2 \in \mathscr{P}(X)$ we have*

  (i) $W_p(\mu_0, \mu_1) \geq 0$, *and* $W_p(\mu_0, \mu_1) = 0$ *if and only if* $\mu_0 = \mu_1$ (positive definiteness)*;*

  (ii) $W_p(\mu_0, \mu_1) = W_p(\mu_1, \mu_0)$ (symmetry)*;*

  (iii) $W_p(\mu_0, \mu_2) \leq W_p(\mu_0, \mu_1) + W_p(\mu_1, \mu_2)$ (triangle inequality).

  When restricted to $\mathscr{P}_p(X)$, W_p is actually finite, so it is a metric.

- Topological properties. *The topology induced by W_p on $\mathscr{P}_p(X)$ is finer (equivalently stronger) than or equal to the narrow one.*

- Lower semicontinuity. *If $\mu_n^0 \rightharpoonup^* \mu^0$, $\mu_n^1 \rightharpoonup^* \mu^1$ in $\mathscr{P}(X)$ when $n \to +\infty$, then*
$$W_p(\mu^0, \mu^1) \leq \liminf_{n\to+\infty} W_p(\mu_n^0, \mu_n^1).$$

- Gronwall-like property. *Let X, Y be separable Banach spaces. If $f : X \to Y$ is a Lipschitz continuous map, then $W_p(f\sharp\mu_1, f\sharp\mu_2) \leq \mathrm{Lip}(f)\, W_p(\mu_1, \mu_2)$, for all $\mu_1, \mu_2 \in \mathscr{P}(X)$.*

Proposition 1.2.4 (Monge–Kantorovich duality)**.** *Given $\mu_1, \mu_2 \in \mathscr{P}(X)$, $p \geq 1$, the following dual representation holds*

$$W_p^p(\mu_1, \mu_2) = \tag{1.3}$$
$$= \sup \left\{ \int_X \varphi(x_1)\, d\mu_1(x_1) + \int_X \psi(x_2)\, d\mu_2(x_2) \ : \ \begin{array}{l} \varphi, \psi \in C_b^0(X) \\ \varphi(x_1) + \psi(x_2) \leq |x_1 - x_2|^p \\ \textit{for } \mu_i\textit{–a.e. } x_i \in X \end{array} \right\}.$$

Proof. See Theorem 6.1.1 in [9]. □

1.3 Continuity equation

For this part the main reference is [9].

Definition 1.3.1 (Continuity equation)**.** Given $\tau > 0$, a Borel family of probability measures $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\tau]} \subseteq \mathscr{P}(\mathbb{R}^d)$ and a Borel map $v : [0,\tau] \times \mathbb{R}^d \to \mathbb{R}^d$ (we will write also $v_t(x) = v(t,x)$), we say that $\boldsymbol{\mu}$ solves the *continuity equation*

$$\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0, \tag{1.4}$$

if for every $\varphi \in C_C^\infty(\mathbb{R}^d)$ there holds

$$\frac{d}{dt}\int_{\mathbb{R}^d} \varphi(x) d\mu_t(x) = \int_{\mathbb{R}^d} \langle v_t(x), \nabla\varphi(x)\rangle \, d\mu_t(x),$$

in the sense of distributions on $]0,\tau[$.

According to Lemma 8.1.2 in [9], if the above v satisfies

$$\int_0^\tau \int_{\mathbb{R}^d} |v_t(x)| \, d\mu_t(x) \, dt < +\infty, \tag{1.5}$$

then there exists a curve $\tilde{\mu} : [0,\tau] \to \mathscr{P}(\mathbb{R}^d)$ which is continuous with respect to narrow convergence and such that $\tilde{\mu}(t) = \mu_t$ for $\mathscr{L}^1$-a.e. $t \in (0,\tau)$, i.e. each solution of the continuity equation admits a unique narrowly continuous representative.

The following gluing lemma will be also used.

Lemma 1.3.2. *Let $T_1, T_2 > 0$ be given. For $i = 1,2$, assume that $\boldsymbol{\mu}^i = \{\mu_t^i\}_{t\in[0,T_i]}$ are narrowly continuous families of probability measures on $\mathbb{R}^d$, and $v^i : [0,T_i] \times \mathbb{R}^d \to \mathbb{R}^d$ are Borel maps such that $\mu^1_{|t=T_1} = \mu^2_{|t=0}$ and*

$$\begin{cases} \partial_t \mu_t^i + \operatorname{div}(v_t^i \mu_t^i) = 0, \\ \\ \displaystyle\int_0^{T_i}\int_{\mathbb{R}^d} |v_t^i(x)| \, d\mu_t^i(x) \, dt < +\infty, \end{cases} \qquad i = 1,2\,.$$

Then if we set

$$(\mu_t, v_t) = \begin{cases} (\mu_t^1, v_t^1), & \text{for } 0 \le t \le T_1, \\ (\mu^2_{t-T_1}, v^2_{t-T_1}), & \text{for } T_1 \le t \le T_1 + T_2, \end{cases}$$

we have that $\boldsymbol{\mu} := \{\mu_t\}_{t\in[0,T_1+T_2]}$ solves the continuity equation $\partial_t\mu_t + \operatorname{div}(v_t\mu_t) = 0$.

Proof. See Lemma 4.4 in [41]. □

Under very mild assumptions on the vector field v_t, the following important result gives us the possibility to characterize a solution of the continuity equation by mean of a measure concentrated on the pairs (x,γ), where γ is an integral solution of the underlying ODE, $\dot{\gamma}(t) = v_t(\gamma(t))$ for a.e. $0 < t \le T$, with $\gamma(0) = x$.

Theorem 1.3.3 (Superposition Principle)**.** *Let $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ be a solution of the continuity equation $\partial_t\mu_t + \mathrm{div}(v_t\mu_t) = 0$ for a suitable Borel vector field $v : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ satisfying*

$$\int_0^T \int_{\mathbb{R}^d} \frac{|v_t(x)|}{1+|x|}\, d\mu_t(x)\, dt < +\infty\,.$$

Then there exists a probability measure $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ such that

(i) $\boldsymbol{\eta}$ is concentrated on the pairs $(x,\gamma) \in \mathbb{R}^d\times\Gamma_T$ such that γ is an absolutely continuous solution of

$$\begin{cases} \dot\gamma(t) = v_t(\gamma(t)), & \text{for } \mathscr{L}^1\text{-a.e } t \in (0,T) \\ \gamma(0) = x, \end{cases}$$

(ii) for all $t \in [0,T]$ and all $\varphi \in C^0_b(\mathbb{R}^d)$ we have

$$\int_{\mathbb{R}^d} \varphi(x)d\mu_t(x) = \iint_{\mathbb{R}^d\times\Gamma_T} \varphi(\gamma(t))\, d\boldsymbol{\eta}(x,\gamma).$$

Conversely, given any $\boldsymbol{\eta}$ satisfying (i) above and defined $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ as in (ii) above, we have that $\partial_t\mu_t + \mathrm{div}(v_t\mu_t) = 0$ and $\mu_{|t=0} = \gamma(0)\sharp\boldsymbol{\eta}$.

Proof. See Theorem 5.8 in [15], Theorem 8.2.1 in [9] and Theorem 3.2 in [2]. □

1.4 Differential inclusions and classical minimum time

We recall now some concepts about the classical optimal control problem with dynamics represented as a differential inclusion in $\mathbb{R}^d$. For this part, our main references are [12, 13].

Definition 1.4.1 (Standing Assumptions)**.** We will say that a set-valued function $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ satisfies the assumption (F_j), $j = 0,1,2,3,4$ if the following hold true

(F_0) $F(x) \neq \emptyset$ is compact and convex for every $x \in \mathbb{R}^d$, moreover $F(\cdot)$ is continuous with respect to the Hausdorff metric, i.e. given $x \in X$, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|y-x| \le \delta$ implies $F(y) \subseteq F(x)+B(0,\varepsilon)$ and $F(x) \subseteq F(y) + B(0,\varepsilon)$.

(F_1) $F(\cdot)$ has linear growth, i.e. there exists a constant $C > 0$ such that $F(x) \subseteq \overline{B(0, C(|x|+1))}$ for every $x \in \mathbb{R}^d$.

(F_2) $F(\cdot)$ is uniformly continuous, i.e. for every $\varepsilon > 0$ there exists $\delta > 0$ such that $F(y) \subseteq F(x) + B(0,\varepsilon)$ for all $x, y \in \mathbb{R}^d$ such that $|x-y| \le \delta$.

(F_3) $F(\cdot)$ is Lipschitz continuous with respect to the Hausdorff metric, i.e., there exists $L > 0$, $L \in \mathbb{R}$, such that for all $x, y \in \mathbb{R}^d$ it holds

$$F(x) \subseteq F(y) + L|y-x|\overline{B(0,1)}.$$

(F_4) $F(\cdot)$ is bounded, i.e. there exist $M > 0$ such that $|y| \leq M$ for all $x \in \mathbb{R}^d$, $y \in F(x)$.

Theorem 1.4.2. *Under assumptions* (F_0) *and* (F_1)*, the differential inclusion*

$$\dot{x}(t) \in F(x(t)), \tag{1.6}$$

has at least one Carathéodory solution defined in $[0, +\infty[$ *for every initial data* $x(0)$ *in* $\mathbb{R}^d$*, i.e., an absolutely continuous function* $x(\cdot)$ *satisfying* (1.6) *for a.e.* $t \geq 0$.
Moreover, the set of trajectories of the differential inclusions (1.6) *is closed in the topology of uniform convergence.*

Proof. See e.g. Theorem 2 p. 97 in [12] and Theorem 1.11 p.186 in Chapter 4 of [36]. □

The following simple classical lemma will be used.

Lemma 1.4.3 (A priori estimate on differential inclusions)**.** *Assume* (F_0) *and* (F_1)*. Let* $K \subset \mathbb{R}^d$ *be compact and* $T > 0$ *and set* $|K| := \max_{y \in K} |y|$*. Then, for all Carathéodory solutions* $\gamma : [0, T] \to \mathbb{R}^d$ *of* (1.6) *we have*

- *(i) forward estimate: if* $\gamma(0) \in K$ *then* $|\gamma(t)| \leq (|K| + CT)\, e^{CT}$ *for all* $t \in [0, T]$*;*
- *(ii) backward estimate: if* $\gamma(T) \in K$ *then* $|\gamma(t)| \leq (|K| + CT)\, e^{CT}$ *for all* $t \in [0, T]$*,*

where C *is the constant in* (F_1).

Proof. Recalling that $\dot{\gamma}(s) \in F(\gamma(s))$ for a.e. $s \in [0, T]$ and that $F(\gamma(s)) \subseteq B(0, C(|x| + 1))$, we have

$$|\gamma(t)| \leq |\gamma(0)| + \int_0^t |\dot{\gamma}(s)|\, ds \leq |K| + CT + C \int_0^t |\gamma(s)|\, ds\,.$$

According to Gronwall's inequality, we then have $|\gamma(t)| \leq (|K| + CT)e^{Ct}$, whence ($i$) follows.
Next, we define $w(t) = \gamma(T - t)$ and observe that w is a solution of $\dot{w}(t) \in -F(w(t))$. Since $-F(\cdot)$ still satisfies (F_0) and (F_1) and $w(0) \in K$, the previous analysis implies

$$|\gamma(t)| = |w(T - t)| \leq (|K| + CT)\, e^{C(T-t)},$$

whence (ii) follows. □

Definition 1.4.4 (Weak invariance)**.** Given a set-valued map $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$, we say that $S \subseteq \mathbb{R}^d$ is *weakly invariant* for $F(\cdot)$ if for every $x \in S$ there exists a Carathéodory solution $x(\cdot)$ of (1.6), defined in $[0, +\infty[$, such that $x(0) = x$ and $x(t) \in S$ for every $t \geq 0$.

For conditions on S and F ensuring weak invariance, we refer to Theorem 2.10 in Chapter 4 of [36].

Definition 1.4.5 (Strong invariance)**.** Given a set-valued map $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$, we say that $S \subseteq \mathbb{R}^d$ is *strongly invariant for* $F(\cdot)$ if for any Carathéodory solution $x(\cdot)$ of (1.6), defined in $[0, +\infty[$, such that there exists $t > 0$ with $x(t) \in S$, we have also that $x(s) \in S$ for all $s \geq t$.

Definition 1.4.6 (Classical admissible trajectory)**.** We say that an absolutely continuous curve γ is an *admissible trajectory for* F *starting from* $x \in \mathbb{R}^d$ *defined on* $[0, T]$ if $\gamma \in AC([0, T]; \mathbb{R}^d)$ and

$$\begin{cases} \dot{\gamma}(t) \in F(\gamma(t)), & \text{for a.e } 0 < t \leq T \\ \gamma(0) = x. \end{cases} \tag{1.7}$$

Definition 1.4.7 (Minimum time function)**.** Let $F(\cdot)$ be a set-valued function satisfying (F_0), S be a nonempty closed subset of $\mathbb{R}^d$. We define the *minimum time function* $T : \mathbb{R}^d \to [0, +\infty]$ by setting

$$T(x) = \inf \{\bar{t} > 0 \,:\, \exists \gamma(\cdot) \text{ adm. traj. for } F \text{ starting from } x \text{ s.t. } \gamma(\bar{t}) \in S\},$$

where by convention $\inf \emptyset = +\infty$. $T(\cdot)$ is the minimum amount of time needed to steer x to the target set S following an admissible trajectory for F. An admissible trajectory $\bar{\gamma}$ is called *optimal for* x if $\bar{\gamma}(0) = x$ and it realizes the infimum in the above functional.

Theorem 1.4.8 (Classical Dynamic Programming Principle)**.** *Let* $s \geq 0$, $x \in \mathbb{R}^d$. *Let* γ *be any admissible trajectory for* F *starting from* x. *Then*

$$T(\gamma(0)) \leq s + T(\gamma(s)). \tag{1.8}$$

Moreover, γ *is an optimal trajectory starting from* x *if and only if the above equality holds for all* $s \in [0, T]$ *s.t.* $\gamma(\tau) \notin S$ *for* $\tau \in [0, s]$.

We refer the reader to Chapter I, Section 2 of [14] for this fundamental result.

Chapter 2

A general overview on control problems in the space of positive finite Borel measures

In this chapter we discuss some aspects related to control problems in the space of positive and finite Borel measures on $\mathbb{R}^d$. The interest in this argument comes from applications to pedestrian dynamics or, more generally, from *multi-agent systems*, i.e. systems with a number of agents so large that only a *macroscopic* (i.e. statistical) description can be provided. In many cases, the *interaction* between the agents prevents a simple reduction of the macroscopic behaviour of the agents to the superposition of the optimal behaviour for each agent, leading possibly to complex behaviours (e.g. self-organization, flocking...).

The main ingredients of this study will be as follows:

1. a *microscopic dynamics*, providing an *Eulerian* description of the available velocities for the agents;

2. a *superposition principle*, providing a connection between the microscopic dynamics of each agent and a macroscopic dynamics describing the evolution of the system;

3. a *micro/macroscopic cost functional*, embedding the main characteristics of the model in which we are interested.

We will choose the microscopic dynamics to be a controlled dynamics in form of an autonomous differential inclusion $\dot{x}(t) \in F(x(t))$, where the set-valued map $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ satisfies some standard assumptions (namely, nonempty compact convex values, continuity w.r.t. Pompeiu-Hausdorff metric, linear growth).

Two main cases have to be considered to connect the microscopic dynamics to the macroscopic evolution of the system:

1. there is neither agent loss nor creation, i.e., the total population considered is preserved throughout the whole evolution;

2. there may be a loss or creation of agents during the evolution.

The first case amounts to make the assumption that the system is *isolated*, without any interaction with the rest of the universe. In this case it is always convenient to normalize the size of the population (i.e., the *total mass*) to be 1 throughout the whole evolution.

The second case can occurr, for example, in problems with a *boundary* in the underlying finite-dimensional state space, where, as soon as an agent crosses the boundary, it is immediately removed from the system and does not affect the system any more (e.g. studying the behaviour of the pedestrians entering and exiting from a room).

In any case, the evolution of the macroscopic system can be expressed by a possibly non-homogenous continuity equation

$$\begin{cases} \partial_t \mu_t + \operatorname{div}(v_t \mu_t) = \omega_t, \\ \mu_{|t=0} = \mu_0 \end{cases} \tag{2.1}$$

where

1. μ_t is a time-dependent measure giving the macroscopic description of the system, in the following sense: given a domain $\Omega \subseteq \mathbb{R}^d$ the quantity

$$\mu_t(\Omega) = \int_\Omega d\mu_t$$

gives the size of the population encompassed in the domain Ω at time t,

2. μ_0 represents the initial distribution of the agents,

3. the vector-valued measure $\nu_t = v_t \mu_t$ describes the macroscopic flux of the mass during the evolution,

4. the term ω_t is the rate of creation ($\omega_t > 0$)/destruction ($\omega_t < 0$) of the agents during the evolution. Under the assumption of isolated system, we have $\omega_t \equiv 0$.

The main results we proved in this general framework are:

· a Dynamic Programming Principle for a generic value function in the mass-preserving case (Proposition 2.2.6 and Corollary 2.2.7);

· analysis of lower semicontinuity of some kinds of cost functionals interesting from an applicative point of view (Lemma 2.2.12, Lemma 2.2.15, Corollary 2.2.16 and Lemma 2.2.21);

· definition and probabilistic representation of an admissible trajectory with mass annihilation in a given space region (Lemma 2.3.2 and Lemma 2.3.3) and derivation of an associated continuity equation with sink described by an absorption measure in $[0, T] \times \mathbb{R}^d$ (Proposition 2.3.7);

· correspondence between the cost functionals defined in the annihilation case with the ones in the mass-preserving case (discussed at the end of this chapter).

2.1 Semicontinuity of functionals depending on measures

Here we recall some notations and a result that will be used in this chapter to prove lower semicontinuity of cost functionals depending on measures.

Given a l.s.c. function $f : \mathbb{R}^n \times \mathbb{R}^m \to [0, +\infty]$ such that $f(x, \cdot)$ is convex for all x, and denoted by

$$f_\infty(x, v) := \lim_{\alpha \to +\infty} \frac{f(x, w + \alpha v) - f(x, w)}{\alpha}$$

the *recession function* of $f(x, \cdot)$,we are concerned with functionals

$$G : \mathscr{M}^+(\mathbb{R}^n) \times \mathscr{M}(\mathbb{R}^n; \mathbb{R}^m) \to [0, +\infty]$$

of the form

$$G(\xi, \zeta) = \int_{\mathbb{R}^n} f\left(x, \frac{\zeta}{\xi}(x)\right) d\xi(x) + \int_{\mathbb{R}^n} f_\infty\left(x, \frac{\zeta^s}{|\zeta^s|}(x)\right) d|\zeta^s|(x), \tag{2.2}$$

where ζ^s is the singular part of ζ w.r.t. ξ.

Notice that, if $f(x, \cdot)$ has bounded domain (or, more generally, superlinear growth) we have $f_\infty(x, v) = 0$ if $v = 0$, and $f_\infty(x, v) = +\infty$ if $v \neq 0$. This means that, in those situations, the functional G becomes

$$G(\xi, \zeta) = \begin{cases} \displaystyle\int_{\mathbb{R}^n} f\left(x, \frac{\zeta}{\xi}(x)\right) d\xi(x), & \text{if } |\zeta| \ll \xi, \\ +\infty, & \text{otherwise.} \end{cases}$$

In the present chapter, we will often use the following result (see Lemma 2.2.3, p. 39, Theorem 3.4.1, p.115, and Corollary 3.4.2 in [18]).

Lemma 2.1.1. *Consider the functional G defined as in* (2.2)*. Assume that at least one of the two conditions below holds true:*

(i) there exists a continuous function $z_0 : \mathbb{R}^m \to \mathbb{R}^n$ such that the function $s \mapsto f(s, z_0(s))$ is continuous and finite;

(ii) there exists a function $\theta : \mathbb{R} \to \mathbb{R}$ such that $\lim_{t \to +\infty} \frac{\theta(t)}{t} = +\infty$ and $f(s, z) > \theta(|z|)$ for every $s \in \mathbb{R}^n$, $z \in \mathbb{R}^m$.

Then if $\{\zeta_h\}_{h \in \mathbb{N}} \subseteq \mathscr{M}(\mathbb{R}^n; \mathbb{R}^m)$ and $\{\xi_h\}_{h \in \mathbb{N}} \subseteq \mathscr{M}^+(\mathbb{R}^n)$ are sequences of measures such that $\zeta_h \rightharpoonup^* \zeta \in \mathscr{M}(\mathbb{R}^n; \mathbb{R}^m)$ and $\xi_h \rightharpoonup^* \xi \in \mathscr{M}^+(\mathbb{R}^n)$, we have

$$G(\xi, \zeta) \leq \liminf_{h \to +\infty} G(\xi_h, \zeta_h).$$

2.2 The isolated (mass-preserving) case ($\omega_t = 0$)

2.2.1 Description of the macroscopic dynamics

Consider now the isolated case $\omega_t = 0$ and let us normalize the mass to 1 for simplicity by taking $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$. According to the Superposition Principle (see Theorem 8.2.1 in [9]), under mild integrability assumptions on v_t, we may express the solution $t \mapsto \mu_t \in \mathscr{P}(\mathbb{R}^d)$ of (2.1) in $[0,T]$ by

$$\mu_t = e_t \sharp \boldsymbol{\eta},$$

where we recall that

- $e_t : \mathbb{R}^d \times \Gamma_T \to \mathbb{R}^d$ is the evaluation operator defined by $e_t(x,\gamma) = \gamma(t)$, with $\Gamma_T = C^0([0,T];\mathbb{R}^d)$,
- $\boldsymbol{\eta}$ is *any* probability measure satisfying the following property: $(x,\gamma) \in \operatorname{supp} \boldsymbol{\eta}$ iff $\gamma \in \Gamma_T$ is an absolutely continuous curve which is an integral solution of the characteristic system

$$\begin{cases} \dot\gamma(t) = v_t \circ \gamma(t), & \text{for a.e. } t \in]0,T] \\ \gamma(0) = x. \end{cases} \tag{2.3}$$

- the initial condition is satisfied, i.e., $e_0 \sharp \boldsymbol{\eta} = \mu_0$.

By using the disintegration theorem (Theorem 5.3.1 in [9]), we have

$$\boldsymbol{\eta} = \mu_0 \otimes \eta_x,$$

where $\{\eta_x\}_{x\in\mathbb{R}^d}$ is a family of probability measures on $\Gamma_T^x := \{\gamma \in \Gamma_T : \gamma(0) = x\}$, which is μ_0-a.e. uniquely determined. In other words, η_x assigns a weigth on each (possible non unique) characteristic curve starting from x.

To establish the link between the microscopic and the macroscopic dynamics it is thus enough to assume that $v_t(x) \in F(x)$ for a.e. $t \in [0,T]$ and μ_t-a.e. $x \in \mathbb{R}^d$. In this way, the agents move along admissible curves of the underlying finite-dimensional control system. When v_t is locally Lipschitz continuous, we have that $\eta_x = \delta_{\gamma_x}$, where $\gamma_x(\cdot)$ is the *unique solution* of the characteristic system (2.3), thus the formula simplifies, becoming $\mu_t = T_t \sharp \mu_0$ where T_t is the flow of v_t at time t, i.e., $\dot T_t(x) = v_t \circ T_t(x)$, $T_0(x) = x$.

Definition 2.2.1. Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be a Borel set-valued map, $p \geq 1$.

1. Given $\mu \in \mathscr{P}(\mathbb{R}^d)$, we define
$$\mathscr{V}_F^p(\mu) = \left\{ \nu \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d) : |\nu| \ll \mu, \frac{\nu}{\mu}(x) \in F(x) \text{ for } \mu\text{-a.e. } x \in \mathbb{R}^d, \left\|\frac{\nu}{\mu}\right\|_{L^p_\mu} < +\infty \right\}.$$
2. Given $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$, we define $\mathscr{C}_F^p(\boldsymbol{\eta})$ to be the set of Borel maps $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ such that
    (a) $v(t,x) \in F(x)$ for all $x \in \mathbb{R}^d$ and a.e. $t \in [0,T]$;
    (b) $\boldsymbol{\eta}$ is concentrated on the pairs $(x,\gamma) \in \mathbb{R}^d \times AC^p([0,T];\mathbb{R}^d)$ with $\gamma(0) = x$ and $\dot\gamma(t) = v_t(\gamma(t))$ for a.e. $t \in [0,T]$;

(c) $\displaystyle\int_0^T \iint_{\mathbb{R}^d\times\Gamma_T} |v_t\circ\gamma(t)|^p\,d\boldsymbol{\eta}(x,\gamma)\,dt < +\infty.$

We will often write $v_t(x)$ instead of $v(t,x)$.

Definition 2.2.2. Given $p\geq 1$, a Borel set-valued map $F:\mathbb{R}^d \rightrightarrows \mathbb{R}^d$, and a family of measures $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ we say that $\boldsymbol{\mu}$ is a *p-admissible curve* if there exists a family of measures $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]} \subseteq \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ such that

1. $t\mapsto\mu_t$ is Borel, i.e., $t\mapsto\mu_t(B)$ is a Borel map for every Borel set $B\subseteq\mathbb{R}^d$,
2. $\partial_t\mu_t + \mathrm{div}(\nu_t) = 0$ in the sense of distributions in $[0,T]\times\mathbb{R}^d$,
3. $\nu_t\in\mathscr{V}^p_F(\mu_t)$ for a.e. $t\in[0,T]$,
4. $\displaystyle\int_0^T \left\|\frac{\nu_t}{\mu_t}\right\|^p_{L^p_{\mu_t}} dt < +\infty.$

In this case, we will say also that $\boldsymbol{\mu}$ is driven by $\boldsymbol{\nu}$.

Remark 2.2.3. We precise that in Chapter 3 and 4 we will give a slightly different definition of *(mass-preserving) admissible curve in the space* $\mathscr{P}(\mathbb{R}^d)$. Indeed, there we will ask condition 3 in Definition 2.2.2 with $p=1$, and we will not require condition 4, substituting this last requirement, when necessary, with the condition (F_1) on the growth of F and the boundedness of the p-moments for the evolving measure.

In Proposition 3.2.17 in Chapter 3 we will see another alternative proof of the following result regarding the closedness of the set of admissible trajectories.

Lemma 2.2.4 (Closedness of the set of admissible trajectories)**.** *Assume hypothesis (F_0). Let $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$ be a sequence of p-admissible curves on $[0,T]$ such that $\boldsymbol{\mu}^n$ is driven by $\boldsymbol{\nu}^n$. Assume that there exists $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ and $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$ such that for a.e. $t\in[0,T]$ we have $\mu^n_t \rightharpoonup^* \mu_t$ and $\nu^n_t\rightharpoonup^*\nu_t$, and*

$$\liminf_{n\to+\infty}\int_0^T \left\|\frac{\nu^n_t}{\mu^n_t}\right\|_{L^p_{\mu^n_t}} dt < +\infty.$$

Then $\boldsymbol{\mu}$ is a p-admissible trajectory driven by $\boldsymbol{\nu}$.

Proof. Define

$$\mathcal{C}^{(1)}_T(\boldsymbol{\mu},\boldsymbol{\nu}) := \sup_{\varphi\in C^\infty_C([0,T]\times\mathbb{R}^d)} \left\{\iint_{[0,T]\times\mathbb{R}^d} \partial_t\varphi(t,x)\,d\mu_t(x)\,dt + \iint_{[0,T]\times\mathbb{R}^d}\nabla\varphi(t,x)\,d\nu_t(x)\,dt\right\},$$

$$\mathcal{C}^{(2)}(\mu,\nu) := \begin{cases} \displaystyle\int_{\mathbb{R}^d}\left(\left|\frac{\nu}{\mu}(x)\right|^p + I_{F(x)}\left(\frac{\nu}{\mu}(x)\right)\right)d\mu, & \text{if } |\nu|\ll\mu,\\ +\infty, & \text{otherwise.}\end{cases}$$

We notice that $\mathcal{C}^{(1)}_T : [0,T]^{\mathscr{P}(\mathbb{R}^d)}\times[0,T]^{\mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)} \to [0,+\infty]$ is convex l.s.c. since it can be written as supremum of linear and continuous maps (we endow the domain with pointwise convergence a.e. w.r.t. weak* convergence)

$$(\boldsymbol{\mu},\boldsymbol{\nu})\mapsto \iint_{[0,T]\times\mathbb{R}^d}\partial_t\varphi(t,x)\,d\mu_t(x)\,dt + \iint_{[0,T]\times\mathbb{R}^d}\nabla\varphi(t,x)\,d\nu_t(x)\,dt.$$

Moreover, $\mathcal{C}_T^{(1)}$ takes only the values 0 or $+\infty$.

Set $f : \mathbb{R}^d \times \mathbb{R}^d \to [0, +\infty]$, $f(x, v) = |v|^p + I_{F(x)}(v)$. Since F is u.s.c. with convex values, we have that $f(\cdot, \cdot)$ is l.s.c., and $f(x, \cdot)$ is convex. By compactness of $F(x)$, we have that the domain of $f(x, \cdot)$ is bounded, thus following the notation in Section 2.1, $f_\infty(x, v) = 0$ if $v = 0$ and $f_\infty(x, v) = +\infty$ if $v \neq 0$. Thus for all $t \in [0, T]$ we have that $(\mu, \nu) \mapsto \mathcal{C}^{(2)}(\mu, \nu)$ can be written in the form of (2.2) for this choice of f. By l.s.c. of F, there exists a continuous selection $z_0 : \mathbb{R}^d \to \mathbb{R}^d$ of F, i.e., there exists $z_0 \in C^0(\mathbb{R}^d; \mathbb{R}^d)$ satisfying $z_0(x) \in F(x)$ for all $x \in \mathbb{R}^d$. Thus $x \mapsto f(x, z_0(x))$ is continuous and finite. The functional $\mathcal{C}^{(2)}(\cdot, \cdot)$ satisfies now the assumptions of Lemma 2.1.1, and so it is l.s.c.

Define now

$$\mathscr{C}(T, \boldsymbol{\mu}, \boldsymbol{\nu}) := \mathcal{C}_T^{(1)}(\boldsymbol{\mu}, \boldsymbol{\nu}) + \int_0^T \mathcal{C}^{(2)}(\mu_t, \nu_t)\, dt,$$

and notice that $\mathscr{C}(T, \boldsymbol{\mu}, \boldsymbol{\nu}) < +\infty$ if and only if we have that $\boldsymbol{\mu}$ is a p-admissible trajectory driven by $\boldsymbol{\nu}$.

If we have sequences $\boldsymbol{\mu}^n = \{\mu_t^n\}_{t \in [0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$, $\boldsymbol{\nu}^n = \{\nu_t^n\}_{t \in [0,T]} \subseteq \mathscr{M}(\mathbb{R}^d; \mathbb{R}^d)$ such that $\mu_t^n \rightharpoonup^* \mu_t$ and $\nu_t^n \rightharpoonup^* \nu_t$ for a.e. $t \in [0, T]$, recalling the l.s.c. of $\mathcal{C}_T^{(1)}$ and $\mathcal{C}^{(2)}$, we have

$$\mathscr{C}(T, \boldsymbol{\mu}, \boldsymbol{\nu}) \leq \liminf_{n \to +\infty} \mathscr{C}(T, \boldsymbol{\mu}^n, \boldsymbol{\nu}^n).$$

Since the right hand side is finite by assumption, we have $\mathscr{C}(T, \boldsymbol{\mu}, \boldsymbol{\nu}) < +\infty$, i.e., $\boldsymbol{\mu}$ is a p-admissible trajectory driven by $\boldsymbol{\nu}$. □

2.2.2 The cost functional

Now we turn our attention to the cost functional to be minimized during the evolution described by (2.1).

We can distinguish two kinds of contribution to the final cost, i.e.,

- a part related to the *superposition of the (microscopic) costs of each agent*, which depends basically only on the microscopic dynamics;
- a part related to the *macroscopic effects* of the evolution.

We can write

$$J(T, \boldsymbol{\mu}, \boldsymbol{\nu}, \boldsymbol{\eta}) = J^{mic}(T, \boldsymbol{\eta}) + J^{mac}(T, \boldsymbol{\mu}, \boldsymbol{\nu}), \tag{2.4}$$

in order to distinguish between these two contributions (the link between $\boldsymbol{\eta}$ and $(\boldsymbol{\mu}, \boldsymbol{\nu})$ is given by the Superposition Principle).

Some cost terms admit a description both in terms of superposition of the costs of each agent and of macroscopic effects, but in general this is not true.

Roughly speaking, we have that the contribution $J^{mac}(T, \boldsymbol{\mu}, \boldsymbol{\nu})$ can be written in the form

$$J^{mac}(T, \boldsymbol{\mu}, \boldsymbol{\nu}) = \int_0^T \mathcal{L}_M(t, \mu_t, \nu_t)\, dt + g(T, \mu_T), \tag{2.5}$$

i.e., we are considering the macroscopic description of the system as a curve in the space of measures, assigning a running cost and computing the final cost as an integral over the time interval plus an exit cost, in analogy with the finite-dimensional case. Notice that only macroscopic quantities and Eulerian description of the velocities are involved.

The contribution given by the superposition of the microscopic costs of each agent is obtained by considering

$$J^{mic}(T, \boldsymbol{\eta}) = \iint_{\mathbb{R}^d \times \Gamma_T} \mathcal{L}_m(T, x, \gamma)\, d\boldsymbol{\eta}(x, \gamma), \tag{2.6}$$

where $\mathcal{L}_m(T, x, \gamma)$ is the total contribution given by a single agent starting at x and moving along the curve γ. Notice that in this case we are interested in the microscopic description only. Moreover, it is important to notice that the cost $\mathcal{L}_m(T, x, \gamma)$ depends *on the whole of the trajectory* γ.

We can give a natural definition of value function of the minimization problem.

Definition 2.2.5 (Value function)**.** Let $p > 1$. Given a cost functional

$$J(T, \boldsymbol{\mu}, \boldsymbol{\nu}, \boldsymbol{\eta}) = \int_0^T \mathcal{L}_M(t, \mu_t, \nu_t)\, dt + \iint_{\mathbb{R}^d \times \Gamma_T} \int_0^T \mathcal{L}_m(t, \gamma(t), \dot{\gamma}(t))\, dt\, d\boldsymbol{\eta} + g(T, \mu_T)$$

we define the *value function* $V : [0, T] \times \mathscr{M}^+(\mathbb{R}^d) \to [0, +\infty]$ by

$$V(s, \mu) = \inf \left\{ \int_s^T \mathcal{L}_M(t, \mu_t, \nu_t)\, dt + \iint_{\mathbb{R}^d \times \Gamma_T} \int_s^T \mathcal{L}_m(t, \gamma(t), \dot{\gamma}(t))\, dt\, d\boldsymbol{\eta} + g(T, \mu_T) \right\},$$

where the infimum is taken on the families $\boldsymbol{\mu} = \{\mu_t\}_{t \in [s,T]}$, $\boldsymbol{\nu} = \{\nu_t\}_{t \in [s,T]}$, $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ such that $\mu_t = e_t \sharp \boldsymbol{\eta}$ for a.e. $t \in [s, T]$, $\boldsymbol{\mu}$ is a p-admissible trajectory driven by $\boldsymbol{\nu}$, and $\mu_s = \mu$.

A p-admissible trajectory $\boldsymbol{\mu}$ is called *optimal for* μ if it realizes the previous infimum.

In the following proposition and subsequent corollary we will prove that a dynamic programming principle holds also in our infinite-dimensional setting for the generic value function just defined.

Proposition 2.2.6 (Dynamic Programming Principle)**.** *Let $p > 1$. For all $0 \le s \le \tau \le T$ we have*

$$V(s, \mu) = \inf \left\{ \int_s^\tau \mathcal{L}_M(t, \mu_t, \nu_t)\, dt + \iint_{\mathbb{R}^d \times \Gamma_T} \int_s^\tau \mathcal{L}_m(t, \gamma(t), \dot{\gamma}(t))\, dt\, d\boldsymbol{\eta} + V(\tau, \mu_\tau) \right\}, \tag{2.7}$$

where the infimum is taken on the families $\boldsymbol{\mu} = \{\mu_t\}_{t \in [s,T]}$, $\boldsymbol{\nu} = \{\nu_t\}_{t \in [s,T]}$, $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ such that $\mu_t = e_t \sharp \boldsymbol{\eta}$ for a.e. $t \in [s, T]$, $\boldsymbol{\mu}$ is a p-admissible trajectory driven by $\boldsymbol{\nu}$, and $\mu_s = \mu$.

Proof. For all $0 \le s \le \tau \le T$, $\varepsilon > 0$ there exist $\boldsymbol{\mu}^\varepsilon = \{\mu_t^\varepsilon\}_{t \in [s,T]}$, $\boldsymbol{\nu}^\varepsilon = \{\nu_t^\varepsilon\}_{t \in [s,T]}$, $\boldsymbol{\eta}^\varepsilon \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ such that $\mu_t^\varepsilon = e_t \sharp \boldsymbol{\eta}^\varepsilon$ for a.e. $t \in [s, T]$ and $\boldsymbol{\mu}^\varepsilon$ is a p-admissible trajectory driven by $\boldsymbol{\nu}^\varepsilon$, with

$$\begin{aligned}V(s,\mu)+\varepsilon &\geq \int_s^\tau \mathcal{L}_M(t,\mu_t^\varepsilon,\nu_t^\varepsilon)\,dt+\iint_{\mathbb{R}^d\times\Gamma_T}\int_s^\tau \mathcal{L}_m(t,\gamma(t),\dot\gamma(t))\,dt\,d\boldsymbol{\eta}^\varepsilon+\\&\quad+\int_\tau^T \mathcal{L}_M(t,\mu_t^\varepsilon,\nu_t^\varepsilon)\,dt+\iint_{\mathbb{R}^d\times\Gamma_T}\int_\tau^T \mathcal{L}_m(t,\gamma(t),\dot\gamma(t))\,dt\,d\boldsymbol{\eta}^\varepsilon+g(T,\mu_T)\\&\geq \int_s^\tau \mathcal{L}_M(t,\mu_t^\varepsilon,\nu_t^\varepsilon)\,dt+\iint_{\mathbb{R}^d\times\Gamma_T}\int_s^\tau \mathcal{L}_m(t,\gamma(t),\dot\gamma(t))\,dt\,d\boldsymbol{\eta}^\varepsilon+V(\tau,\mu_\tau^\varepsilon),\end{aligned}$$

where we used the fact that, since we have that $\{\mu_t^\varepsilon\}_{t\in[\tau,T]}$ is a p-admissible curve driven by $\{\nu_t^\varepsilon\}_{t\in[\tau,T]}$ with $\mu_t^\varepsilon=e_t\sharp\boldsymbol{\eta}^\varepsilon$ for all $t\in[\tau,T]$ and $\mu^\varepsilon_{|t=\tau}=\mu^\varepsilon_\tau$, we have

$$V(\tau,\mu_\tau^\varepsilon)\leq \int_\tau^T \mathcal{L}_M(t,\mu_t^\varepsilon,\nu_t^\varepsilon)\,dt+\iint_{\mathbb{R}^d\times\Gamma_T}\int_\tau^T \mathcal{L}_m(t,\gamma(t),\dot\gamma(t))\,dt\,d\boldsymbol{\eta}^\varepsilon+g(T,\mu_T).$$

Thus we have

$$V(s,\mu)+\varepsilon\geq \inf\left\{\int_s^\tau \mathcal{L}_M(t,\mu_t,\nu_t)\,dt+\iint_{\mathbb{R}^d\times\Gamma_T}\int_s^\tau \mathcal{L}_m(t,\gamma(t),\dot\gamma(t))\,dt\,d\boldsymbol{\eta}+V(\tau,\mu_\tau)\right\},$$

where the infimum is taken on the families $\boldsymbol{\mu}=\{\mu_t\}_{t\in[s,T]}$, $\boldsymbol{\nu}=\{\nu_t\}_{t\in[s,T]}$, $\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$ such that $\mu_t=e_t\sharp\boldsymbol{\eta}$ for a.e. $t\in[s,T]$, $\boldsymbol{\mu}$ is a p-admissible trajectory driven by $\boldsymbol{\nu}$, and $\mu_s=\mu$.

By letting $\varepsilon\to0^+$ we obtain that

$$V(s,\mu)\geq \inf\left\{\int_s^\tau \mathcal{L}_M(t,\mu_t,\nu_t)\,dt+\iint_{\mathbb{R}^d\times\Gamma_T}\int_s^\tau \mathcal{L}_m(t,\gamma(t),\dot\gamma(t))\,dt\,d\boldsymbol{\eta}+V(\tau,\mu_\tau)\right\},$$

Conversely, let $\boldsymbol{\mu}^{(1)}=\{\mu_t^{(1)}\}_{t\in[s,T]}$, $\boldsymbol{\nu}^{(1)}=\{\nu_t^{(1)}\}_{t\in[s,T]}$, $\boldsymbol{\eta}^{(1)}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$ $\boldsymbol{\mu}^{(\varepsilon)}=\{\mu_t^{(\varepsilon)}\}_{t\in[\tau,T]}$, $\boldsymbol{\nu}^{(\varepsilon)}=\{\nu_t^{(\varepsilon)}\}_{t\in[\tau,T]}$, $\boldsymbol{\eta}^{(\varepsilon)}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$ be such that

1. $\mu_t^{(1)}=e_t\sharp\boldsymbol{\eta}^{(1)}$ for a.e. $t\in[s,T]$, $\boldsymbol{\mu}^{(1)}$ is a p-admissible trajectory driven by $\boldsymbol{\nu}^{(1)}$, and $\mu_s=\mu$;
2. $\mu_t^{(\varepsilon)}=e_t\sharp\boldsymbol{\eta}^{(\varepsilon)}$ for a.e. $t\in[\tau,T]$, $\boldsymbol{\mu}^{(\varepsilon)}$ is a p-admissible trajectory driven by $\boldsymbol{\nu}^{(\varepsilon)}$, and $\mu_\tau^{(\varepsilon)}=\mu_\tau^{(1)}$
3. for all $\varepsilon>0$ we have

$$\begin{aligned}V(\tau,\mu_\tau^{(1)})+\varepsilon&\geq\int_\tau^T \mathcal{L}_M(t,\mu_t^{(\varepsilon)},\nu_t^{(\varepsilon)})\,dt+\\&\qquad+\iint_{\mathbb{R}^d\times\Gamma_T}\int_\tau^T \mathcal{L}_m(t,\gamma(t),\dot\gamma(t))\,dt\,d\boldsymbol{\eta}^{(\varepsilon)}+g(T,\mu_T^{(\varepsilon)}).\end{aligned}$$

Then we define $\boldsymbol{\mu}=\{\mu_t\}_{t\in[s,T]}$, $\boldsymbol{\nu}=\{\nu_t\}_{t\in[s,T]}$, by setting $\mu_t=\mu_t^{(1)}$ and $\nu_t=\nu_t^{(1)}$ for $t\in[s,\tau]$, and $\mu_t=\mu_t^{(\varepsilon)}$ and $\nu_t=\nu_t^{(\varepsilon)}$ for $t\in[\tau,T]$. Thus we have that $\boldsymbol{\mu}$ is a p-admissible trajectory driven by $\boldsymbol{\nu}$ and there exists $\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$ such that $\mu_t=e_t\sharp\boldsymbol{\eta}$ for $t\in[s,T]$. We then have

$$V(s,\mu)\leq\int_s^\tau \mathcal{L}_M(t,\mu_t,\nu_t)\,dt+\iint_{\mathbb{R}^d\times\Gamma_T}\int_s^\tau \mathcal{L}_m(t,\gamma(t),\dot\gamma(t))\,dt\,d\boldsymbol{\eta}+V(\tau,\mu_\tau)+\varepsilon,$$

By letting $\varepsilon \to 0^+$ and taking the infimum on the families $\boldsymbol{\mu} = \{\mu_t\}_{t\in[s,T]}$, $\boldsymbol{\nu} = \{\nu_t\}_{t\in[s,T]}$, $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ such that $\mu_t = e_t\sharp\boldsymbol{\eta}$ for a.e. $t \in [s,T]$, $\boldsymbol{\mu}$ is a p-admissible trajectory driven by $\boldsymbol{\nu}$, with $\mu_s = \mu$, we obtain

$$V(s,\mu) \le \inf\left\{\int_s^\tau \mathcal{L}_M(t,\mu_t,\nu_t)\,dt + \iint_{\mathbb{R}^d\times\Gamma_T}\int_s^\tau \mathcal{L}_m(t,\gamma(t),\dot\gamma(t))\,dt\,d\boldsymbol{\eta} + V(\tau,\mu_\tau)\right\},$$

and so equality holds. □

Corollary 2.2.7. *Let $p > 1$. Given the families $\boldsymbol{\mu} = \{\mu_t\}_{t\in[s,T]}$, $\boldsymbol{\nu} = \{\nu_t\}_{t\in[s,T]}$, $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ such that $\mu_t = e_t\sharp\boldsymbol{\eta}$ for a.e. $t \in [s,T]$, $\boldsymbol{\mu}$ is a p-admissible trajectory driven by $\boldsymbol{\nu}$, we have that the map*

$$h(\tau,\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\eta}) := \int_s^\tau \mathcal{L}_M(t,\mu_t,\nu_t)\,dt + \iint_{\mathbb{R}^d\times\Gamma_T}\int_s^\tau \mathcal{L}_m(t,\gamma(t),\dot\gamma(t))\,dt\,d\boldsymbol{\eta} + V(\tau,\mu_\tau)$$

is nondecreasing for $\tau \in [s,T]$. Moreover, it is constant if and only if $\boldsymbol{\mu}$ is optimal.

Proof. The first assertion comes directly from (2.7), indeed for all $s \le \tau_1 \le \tau_2 \le T$ we have

$$V(\tau_1,\mu_{\tau_1}) \le h(\tau_2,\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\eta}) - \left[\int_s^{\tau_1} \mathcal{L}_M(t,\mu_t,\nu_t)\,dt + \iint_{\mathbb{R}^d\times\Gamma_T}\int_s^{\tau_1} \mathcal{L}_m(t,\gamma(t),\dot\gamma(t))\,dt\,d\boldsymbol{\eta}\right],$$

and so $h(\tau_1,\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\eta}) \le h(\tau_2,\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\eta})$.

Suppose now that $\boldsymbol{\mu}$ is an optimal trajectory, i.e.,

$$V(s,\mu_s) = \int_s^T \mathcal{L}_M(t,\mu_t,\nu_t)\,dt + \iint_{\mathbb{R}^d\times\Gamma_T}\int_s^T \mathcal{L}_m(t,\gamma(t),\dot\gamma(t))\,dt\,d\boldsymbol{\eta} + g(T,\mu_T).$$

Recalling that $V(T,\mu_T) = g(T,\mu_T)$, we have

$$V(s,\mu_s) = h(s,\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\eta}) \le h(T,\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\eta}) = V(s,\mu_s),$$

and so $h(\cdot,\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\eta})$ is constant on $[s,T]$.

Conversely, assume that $h(\cdot,\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\eta})$ is constant on $[s,T]$. Then in particular we have

$$h(s,\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\eta}) = h(T,\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\eta}),$$

which amounts to say

$$V(s,\mu_s) = \int_s^T \mathcal{L}_M(t,\mu_t,\nu_t)\,dt + \iint_{\mathbb{R}^d\times\Gamma_T}\int_s^T \mathcal{L}_m(t,\gamma(t),\dot\gamma(t))\,dt\,d\boldsymbol{\eta} + g(T,\mu_T),$$

i.e., $\boldsymbol{\mu}$ is optimal. □

In this way we proved the Dynamic Programming Principle for a generic functional in the framework of curves in the space of probability measures.

We will pass to analyze now some kinds of cost terms which can be useful from an applicative point of view. Basically we will consider cost terms of the following type:

I): cost terms expressing the *superposition of the costs of each agent* travelling along the admissible trajectories of the underlying finite-dimensional control system;

II): cost terms due to the evolution of the *macroscopic distribution* of the mass and of the velocities of the agents;

III): cost terms taking into account the *interactions* between the agents.

As a general rule, we will put assumptions in order to ensure that each cost term is *nonnegative and l.s.c.* Together with some compactness assumptions, this will ensure *existence* of minimizers.

Let us begin by analyzing the first case of cost functionals.

Definition 2.2.8 (Instantanous microscopic cost functional)**.** Let $L_c^a : \mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^d \to [0,+\infty]$ be a Borel function. We define the functional

$$J_{\mathrm{sys}}(T,\boldsymbol{\eta}) = \iint_{\mathbb{R}^d\times\Gamma_T} \left(\int_0^T L_c^a\left(t,\gamma(t),\dot\gamma(t)\right)\,dt\right)\,d\boldsymbol{\eta}(x,\gamma), \tag{2.8}$$

which, recalling (2.6), is the superposition of a microscopicagent cost in the form

$$\mathcal{L}_m(T,x,\gamma) := \int_0^T L_c^a\left(t,\gamma(t),\dot\gamma(t)\right)\,dt.$$

This amounts to say that there exists a current cost given by $L_c^a(\cdot)$ that all the agents pay instantaneously along their trajectories.

We can give an analogous definition from a macroscopic point of view.

Definition 2.2.9 (Instantanous macroscopic cost functional)**.** Let $L_c^a : \mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^d \to [0,+\infty]$ be a Borel function. We define the functional

$$\hat J_{\mathrm{sys}}(T,\boldsymbol{\mu},\boldsymbol{\nu}) := \begin{cases} \displaystyle\int_0^T\int_{\mathbb{R}^d} L_c^a\left(t,x,\frac{\nu_t}{\mu_t}(x)\right)\,d\mu_t(x)\,dt, & \text{if } \nu_t\in\mathscr{V}_F^p(\mu_t) \\ & \text{for a.e. } t\in[0,T], \\ \\ +\infty, & \text{otherwise,} \end{cases} \tag{2.9}$$

which represents a current cost for the curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ in the space of probability measures.

The following Lemma proves that, under the assumptions of the Superposition Principle, these two costs agrees.

Lemma 2.2.10 (Equivalence)**.** *Let $p>1$, $\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$, and $v\in\mathscr{C}_F^p(\boldsymbol{\eta})$. Define $\mu_t = e_t\sharp\boldsymbol{\eta}$, $\nu_t = v_t\mu_t$, $\boldsymbol{\mu}=\{\mu_t\}_{t\in[0,T]}$, $\boldsymbol{\nu}=\{\nu_t\}_{t\in[0,T]}$. Then*

$$\partial_t\mu_t + \mathrm{div}\nu_t = 0, \qquad\qquad J_{\mathrm{sys}}(T,\boldsymbol{\eta}) = \hat J_{\mathrm{sys}}(T,\boldsymbol{\mu},\boldsymbol{\nu}).$$

Conversely, let $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$, $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$ *be satisfying* $\nu_t \in \mathscr{V}_F^p(\mu_t)$ *for a.e.* $t \in [0,T]$, $\partial_t\mu_t + \mathrm{div}\nu_t = 0$, *and such that*

$$\hat{J}_{\mathrm{sys}}(T, \boldsymbol{\mu}, \boldsymbol{\nu}) + \int_0^T \int_{\mathbb{R}^d} \left|\frac{\nu_t}{\mu_t}(x)\right|^p d\mu_t\, dt < +\infty.$$

Then there exists a measure $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$, *and* $v \in \mathscr{C}_F^p(\boldsymbol{\eta})$ *such that* $\mu_t = e_t \sharp \boldsymbol{\eta}$ *for all* $t \in [0,T]$, $v_t(x) = \dfrac{\nu_t}{\mu_t}(x)$ *for a.e.* $t \in [0,T]$ *and* μ_t*-a.e.* $x \in \mathbb{R}^d$, *and*

$$J_{\mathrm{sys}}(T, \boldsymbol{\eta}) = \hat{J}_{\mathrm{sys}}(T, \boldsymbol{\mu}, \boldsymbol{\nu}).$$

Proof. In both cases the assumptions of the Superposition Principle (Theorem 8.2.1 in [9]) holds, allowing to pass from one description to the other. □

Remark 2.2.11. Choosing $L_c^a(t, x, v) \equiv 1$ amounts to say that the cost for a single agent is the total time traveled, i.e., T. This is the choice, e.g., if we are interested in the problem of minimizing the time needed to steer an initial measure μ_0 to a target set $\tilde{S} \subseteq \mathscr{P}(\mathbb{R}^d)$ along the admissible trajectories of the system (see Chapter 3). A sligthly more general situation is to take $L_a^c(t, x, v) = \chi_{\mathbb{R}^d\setminus V}(x)$, where $V \subseteq \mathbb{R}^d$ is a given closed set. In this case for each agent we count only the time travelled outside V.

We pass now to analyze the regularity of the cost terms.

Lemma 2.2.12 (L.s.c. of the instantaneous cost)**.** *Let* $L_c^a : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \to [0, +\infty]$ *be a Borel map. Assume hypothesis* (F_0), *that* $L_c^a(t, \cdot, \cdot)$ *is continuous, and* $L_c^a(t, x, \cdot)$ *is convex. Then*

1. *the functional* $\mathcal{L}_{\mathrm{sys}}(t, \cdot, \cdot)$ *defined as*

$$\mathcal{L}_{\mathrm{sys}}(t, \mu, \nu) := \begin{cases} \displaystyle\int_{\mathbb{R}^d} L_c^a\left(t, x, \frac{\nu}{\mu}(x)\right) d\mu(x), & \text{if } |\nu| \ll \mu \text{ and } \dfrac{\nu}{\mu}(x) \in F(x) \\ & \text{for } \mu - \text{a.e. } x \in \mathbb{R}^d, \\ \\ +\infty, & \text{otherwise,} \end{cases} \tag{2.10}$$

   is l.s.c. w.r.t. narrow convergence.

2. *given* $\{T_n\}_{n\in\mathbb{N}} \subseteq [0, +\infty[$, *a sequence of measurable curves* $\boldsymbol{\mu}^n = \{\mu_t^n\}_{t\in[0,T_n]}$ *in* $\mathscr{P}(\mathbb{R}^d)$, *and a sequence of Borel vector-valued measures* $\boldsymbol{\nu}^n = \{\nu_t^n\}_{t\in[0,T_n]}$ *in* $\mathscr{M}(\mathbb{R}^d; \mathbb{R}^d)$, $T_n \to T^+$, $\mu_t^n \rightharpoonup^* \mu_t$, $\nu_t^n \rightharpoonup^* \nu_t$ *for a.e.* $t \in [0,T]$, $\nu_t^n \in \mathscr{V}_F^p(\mu_t^n)$ *for all* $n \in \mathbb{N}$ *and a.e.* $t \in [0, T_n]$, *we have for a.e.* $t \in [0,T]$

$$\left\|\frac{\nu_t}{\mu_t}\right\|_{L^p_{\mu_t}} \leq \liminf_{n\to+\infty} \left\|\frac{\nu_t^n}{\mu_t^n}\right\|_{L^p_{\mu_t^n}}.$$

   Moreover, if the left hand side of the above inequality is finite, we have $\nu_t \in \mathscr{V}_F^p(\mu_t)$ *for a.e.* $t \in [0,T]$, *and*

$$\hat{J}_{\mathrm{sys}}(T, \boldsymbol{\mu}, \boldsymbol{\nu}) \leq \liminf_{n\to+\infty} \hat{J}_{\mathrm{sys}}(T_n, \boldsymbol{\mu}^n, \boldsymbol{\nu}^n).$$

Proof.

1. Fix t and define $f(x,v) = L_c^a(t,x,v) + I_{F(x)}(v)$. Since F is u.s.c. with convex values, and recalling the assumptions on L_c^a, we have that $f(\cdot,\cdot)$ is l.s.c. and $f(x,\cdot)$ is convex. By compactness of $F(x)$, we have that the domain of $f(x,\cdot)$ is bounded, thus following the notation in Section 2.1 we have $f_\infty(x,v) = 0$ if $v = 0$ and $f_\infty(x,v) = +\infty$ if $v \neq 0$. Thus (2.10) can be written in the form of (2.2) for this choice of f. By l.s.c. of F, there exists a continuous selection $z_0 : \mathbb{R}^d \to \mathbb{R}^d$ of F, i.e., there exists $z_0 \in C^0(\mathbb{R}^d;\mathbb{R}^d)$ satisfying $z_0(x) \in F(x)$ for all $x \in \mathbb{R}^d$. Thus $x \mapsto f(x, z_0(x))$ is continuous and finite. The functional (2.10) satisfies now the assumptions of Lemma 2.1.1, and so it is l.s.c.

2. Consider the functional defined as

$$(\mu,\nu) \mapsto \begin{cases} \displaystyle\int_{\mathbb{R}^d} \left|\frac{\nu}{\mu}(x)\right|^p d\mu(x), & \text{if } \nu \ll \mu, \\ +\infty, & \text{otherwise.} \end{cases}$$

   This functional clearly satisfies the assumptions of Lemma 2.1.1, and so it is l.s.c. Fix $t \in [0,T]$ such that we have that $\nu_t^n \in \mathscr{V}_F^p(\mu_t^n)$ for all $n \in \mathbb{N}$ (this is a full measure set on $[0,T]$). By the l.s.c. of the above functional, we then have for a.e. $t \in [0,T]$ that

$$\left\|\frac{\nu_t}{\mu_t}\right\|_{L^p_{\mu_t}} \leq \liminf_{n\to+\infty} \left\|\frac{\nu_t^n}{\mu_t^n}\right\|_{L^p_{\mu_t^n}}.$$

   Assume now that the left hand side of the above inequality is finite and that

$$\liminf_{n\to+\infty} \hat{J}_{\text{sys}}(T_n, \boldsymbol{\mu}^n, \boldsymbol{\nu}^n) < +\infty,$$

   otherwise there is nothing to prove. According to Fatou's Lemma,

$$\begin{aligned}\int_0^T \liminf_{n\to+\infty} \mathcal{L}_{\text{sys}}(t,\mu_t^n,\nu_t^n)\,dt &\leq \liminf_{n\to+\infty} \hat{J}_{\text{sys}}(T,\boldsymbol{\mu}^n,\boldsymbol{\nu}^n) \\ &\leq \liminf_{n\to+\infty} \hat{J}_{\text{sys}}(T_n,\boldsymbol{\mu}^n,\boldsymbol{\nu}^n) < +\infty.\end{aligned}$$

   According to item (1), we have that for a.e. $t \in [0,T]$,

$$\mathcal{L}_{\text{sys}}(t,\mu_t,\nu_t) \leq \liminf_{n\to+\infty} \mathcal{L}_{\text{sys}}(t,\mu_t^n,\nu_t^n) < +\infty,$$

   which implies that $\dfrac{\nu_t}{\mu_t}(x) \in F(x)$ for μ_t-a.e. $x \in \mathbb{R}^d$, thus $\nu_t \in \mathscr{V}_F^p(\mu_t)$ for a.e. $t \in [0,T]$. This implies that we have

$$\hat{J}_{\text{sys}}(T,\boldsymbol{\mu},\boldsymbol{\nu}) = \int_0^T \mathcal{L}_{\text{sys}}(t,\mu_t,\nu_t)\,dt \leq \int_0^T \liminf_{n\to+\infty} \mathcal{L}_{\text{sys}}(t,\mu_t^n,\nu_t^n)\,dt,$$

   which completes the proof.

□

We pass now to examine the second case of cost functionals, i.e. a cost term involving the macroscopic behaviour of the agents during the evolution, taking into account the density of their positions and distribution of velocities w.r.t. a fixed reference measure. Notice that this is a term dealing with some global properties of the system which cannot derived simply by the superposition of the behaviours of single agents.

Definition 2.2.13 (Density cost)**.** Let $L_{\text{dens}} : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to [0, +\infty]$ be a Borel map. Given $\sigma \in \mathscr{M}^+(\mathbb{R}^d)$, we define the functional

$$\hat{J}^\sigma_{\text{dens}}(T, \boldsymbol{\mu}, \boldsymbol{\nu}) := \begin{cases} \displaystyle\int_0^T \int_{\mathbb{R}^d} L_{\text{dens}} \left(t, x, \frac{\mu_t}{\sigma}(x), \frac{\nu_t}{\sigma}(x)\right) d\sigma\, dt, & \text{if } \mu_t \ll \sigma \text{ and } |\nu_t| \ll \sigma \\ & \text{for a.e. } t \in [0, T], \\ & \\ +\infty, & \text{otherwise,} \end{cases} \tag{2.11}$$

with $T \geq 0$, $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ and $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]} \subseteq \mathscr{M}(\mathbb{R}^d \times \mathbb{R}^d)$.

Remark 2.2.14. We can use this term to model some pointwise constraints of the maximum allowed density w.r.t. a reference measure σ of the distribution of agents and of their velocities. For instance, imagine that the agents are moving on a frozen lake: clearly, the thickness of the ice is related to the maximum affordable load, and in any case there is a constraint on the maximum density tolerable by the agents (*overcrowding threshold*). To model this situation, we can simply take $\sigma = \mathscr{L}^d$ and $L_{\text{dens}}(t, x, d_x, d_v) = I_{[0,d_{\max}(x)\wedge d_{\text{thre}}]}(d_x)$, where $d_{\max}(x)$ is the maximum affordable load at point x and d_{thre} is the overcrowding threshold. Similarly, we can penalize distribution in the velocities: assume for example that in a certain point the agents are allowed to travel in a certain direction, but the road heading to that direction is very narrow. Clearly, if we have many agents all trying to move in a direction where the road is narrow, the travelling cost will be higher with respect to the same situation in which we have only few agents.

The following results provide some sufficient conditions ensuring the l.s.c. of $\hat{J}_{\text{dens}}$.

Lemma 2.2.15. *Let $L_{\text{dens}} : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to [0, +\infty]$ be a Borel map such that*

(i) $L_{\text{dens}}(t, \cdot, \cdot, \cdot)$ *is l.s.c.,*

(ii) $L_{\text{dens}}(t, x, \cdot, \cdot)$ *is convex satisfying*

$$\lim_{|(d_x,d_v)|\to+\infty} \frac{L_{\text{dens}}(t, x, d_x, d_v)}{|(d_x, d_v)|} = +\infty,$$

(iii) for any fixed $t \in \mathbb{R}$, one of these two properties holds

a) there exists a continuous map $x \mapsto (d_x(x), d_v(x))$ such that the map $x \mapsto L_{\text{dens}}(t, x, d_x(x), d_v(x))$ is finite and continuous,

b) there exists a function $\theta : \mathbb{R} \to \mathbb{R}$ such that $\displaystyle\lim_{s\to+\infty} \frac{\theta(s)}{s} = +\infty$ and $L_{\text{dens}}(t, x, d_x, d_v) > \theta(|(d_x, d_v)|)$ for every $x, d_x, d_v \in \mathbb{R}^d$.

Then the functional $\hat{J}^\sigma_{\text{dens}}$ *defined in* (2.11) *is lower semicontinuous.*

Proof. The functional

$$\mathcal{L}^\sigma_{\text{dens}}(t,\mu,\nu) := \begin{cases} \displaystyle\int_{\mathbb{R}^d} L_{\text{dens}}\left(t, x, \frac{\mu}{\sigma}(x), \frac{\nu}{\sigma}(x)\right) d\sigma, & \text{if } \mu \ll \sigma \text{ and } |\nu| \ll \sigma, \\ \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.12}$$

is l.s.c. according to Lemma 2.1.1 recalling that $L^\infty_{\text{dens}}(t,x,d_x,d_v) = +\infty$ if $(d_x,d_v) \neq (0,0)$ and $L^\infty_{\text{dens}}(t,x,0,0) = 0$. By using Fatou's Lemma, we can conclude arguing as in the last part of the proof of Lemma 2.2.12 (2). □

However, as already noticed, from a modelling point of view, it is more realistic to fix a uniform upper bound on the density of agents (*overcrowding threshold*) and also on the density of their velocity distribution. Moreover, we introduce explicitly a constraint on the agents' density and agents' velocity distribution density depending on the point. In this way, the above results simplifies as follows.

Corollary 2.2.16. *Let* $d_{\max} > 0$ *be a constant,* $D_v \subseteq \mathbb{R}^d$ *be compact and convex. We set*

$$L_{\text{dens}}(t,x,d_x,d_v) = L^a_{\text{dens}}(t,x,d_x,d_v) + I_{[0,d_{\max}]\times D_v}(d_x,d_v),$$

where $L_{\text{dens}} : \mathbb{R}\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d \to [0,+\infty]$ *is a Borel map such that* $L^a_{\text{dens}}(t,\cdot,\cdot,\cdot)$ *is l.s.c.,* $L_{\text{dens}}(t,x,\cdot,\cdot)$ *is convex. Then the functional* $\hat{J}^\sigma_{\text{dens}}(t,\cdot,\cdot)$ *defined in* (2.11) *is lower semicontinuous.*

Proof. All the assuptions of Lemma 2.2.15 are satisfied, by taking $\theta(s) = s^2\chi_{[R,+\infty[}(s) - 1$ where $R > 0$ is constant such that $[0,d_{\max}] \times D_v \subseteq B(0,R)$. Indeed, we have that $\displaystyle\lim_{s\to+\infty} \frac{\theta(s)}{s} = +\infty$, and

$$L_{\text{dens}}(t,x,d_x,d_v) \geq 0 > \theta(|(d_x,d_v)|), \text{ for every } (t,x,d_x,d_v) \text{ with } |(d_x,d_v)| < R,$$

while

$$L_{\text{dens}}(t,x,d_x,d_v) = +\infty > \theta(|(d_x,d_v)|), \text{ for every } (t,x,d_x,d_v) \text{ with } |(d_x,d_v)| \geq R.$$

□

Remark 2.2.17. Notice that in Corollary 2.2.16 we are not asking L^a_{dens} to be neither continuous, nor finite at every point, thus we are completely free to add for instance further constraints on the density of agents depending on the point: it is enough to add a term $I_{[0,\tilde{d}_x(x)]\times\tilde{D}_v(x)}(d_x,d_v)$ to $L^a_{\text{dens}}(t,x,d_x,d_v)$, where $\tilde{d}_x : \mathbb{R}^d \to [0,+\infty]$ is a measurable function, and $\tilde{D}_v : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is a measurable set-valued map with closed convex values (not necessarily bounded). We can model also obstacles by imposing that $d_x(x) = 0$ in a region $\Omega \subseteq \mathbb{R}^d$. In this way, when the functional is finite, we have that no mass can flow through Ω.

We pass now to consider the third case of cost functionals, i.e. a cost term dealing with the interaction between the agents.

Example 2.2.18. The simplest interaction model between agents is provided by assuming that the interaction between two agents depends only on their mutual distance. In this case, we have

$$\hat{J}_{\text{inter}}(T, \boldsymbol{\mu}, \boldsymbol{\nu}) = \int_0^T \iint_{\mathbb{R}^d \times \mathbb{R}^d} W(y - x)\, d\mu_t(x)\, d\mu_t(y)\, dt,$$

with $T \geq 0$, $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$, $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]} \subseteq \mathscr{M}(\mathbb{R}^d \times \mathbb{R}^d)$, and where $W : \mathbb{R}^d \to [0, +\infty]$ is a radial function, i.e. $W(z) = \tilde{W}(|z|)$ for all $z \in \mathbb{R}^d$. The above integral can be expressed also in convolution form as

$$\hat{J}_{\text{inter}}(T, \boldsymbol{\mu}, \boldsymbol{\nu}) = \int_0^T \int_{\mathbb{R}^d} W * \mu_t(x)\, d\mu_t(x).$$

More complex interactions may involve also the velocities of the agents. This occurs, for instance, in modeling the *consensus phenomenon* in flocking, i.e. the alignment to a global common speed of all the agents (see for instance [45]).

For example in [63] the authors studied some models for self-organized dynamics focusing on concentration around an emerging consensus: each agent adjusts its state according to the state of its neighbors. The paper focus its attention on the role of mid-range alignment which stands between the short-range attraction and the long-range repulsion, i.e. on the tendency to adjust to agents' "environmental averages", studying conditions for flocking and the formation of *clusters*. In [50] a kinetic description of such models is given.

We give the following definition.

Definition 2.2.19 (Interaction cost)**.** Let $L_{\text{inter}} : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \to [0, +\infty]$ be a given Borel map. Define $X_{\text{inter}} := \mathbb{R}^d \times \Gamma_T \times \mathbb{R}^d \times \Gamma_T$, and the *interaction cost functionals*

$$J_{\text{inter}}(T, \boldsymbol{\eta}) = \int_{X_{\text{inter}}} \int_0^T L_{\text{inter}}\left(t, \gamma_x(t), \gamma_y(t), \dot{\gamma}_x(t), \dot{\gamma}_y(t)\right)\, dt\, d\boldsymbol{\eta}(x, \gamma_x)\, d\boldsymbol{\eta}(y, \gamma_y), \tag{2.13}$$

$$\hat{J}_{\text{inter}}(T, \boldsymbol{\mu}, \boldsymbol{\nu}) = \begin{cases} \displaystyle\int_0^T \iint_{\mathbb{R}^d \times \mathbb{R}^d} L_{\text{inter}}\left(t, x, y, \frac{\nu_t}{\mu_t}(x), \frac{\nu_t}{\mu_t}(y)\right) d\mu_t(x)\, d\mu_t(y)\, dt, & \text{if } \nu_t \in \mathscr{V}^p_F(\mu_t) \\ & \text{for a.e. } t \in [0, T], \\ +\infty, & \text{otherwise,} \end{cases} \tag{2.14}$$

in microscopic and macroscopic point of view, respectively.

Remark 2.2.20. Note that, under the assumptions of the Superposition Principle, these two costs agrees. The proof follows the same argument used for the functionals J_{sys} and $\hat{J}_{\text{sys}}$ in Lemma 2.2.10.

Lemma 2.2.21 (L.s.c. of the interaction cost)**.** *Let $L_{\text{inter}} : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \to [0, +\infty]$ be a given Borel map. Assume hypothesis (F_0), that $L_{\text{inter}}(t, \cdot, \cdot, \cdot, \cdot)$ is continuous, and $L_{\text{inter}}(t, x, y, \cdot, \cdot)$ is convex. Then given $\{T_n\}_{n\in\mathbb{N}} \subseteq [0, +\infty[$, a sequence of measurable curves $\boldsymbol{\mu}^n = \{\mu^n_t\}_{t\in[0,T_n]}$ in $\mathscr{P}(\mathbb{R}^d)$, and a sequence of Borel vector-valued measures $\boldsymbol{\nu}^n = \{\nu^n_t\}_{t\in[0,T_n]}$ in $\mathscr{M}(\mathbb{R}^d; \mathbb{R}^d)$, $T_n \to T^+$, $\mu^n_t \rightharpoonup^* \mu_t$, $\nu^n_t \rightharpoonup^* \nu_t$ for a.e. $t \in [0, T]$, $\nu^n_t \in \mathscr{V}^p_F(\mu^n_t)$ for all $n \in \mathbb{N}$ and a.e. $t \in [0, T_n]$, we have for a.e. $t \in [0, T]$*

$$\left\| \frac{\nu_t}{\mu_t} \right\|_{L^p_{\mu_t}} \leq \liminf_{n\to+\infty} \left\| \frac{\nu^n_t}{\mu^n_t} \right\|_{L^p_{\mu^n_t}}.$$

Moreover, if the left hand side of the above inequality is finite, we have $\nu_t \in \mathscr{V}_F^p(\mu_t)$ *for a.e.* $t \in [0,T]$*, and*

$$\hat{J}_{\text{inter}}(T, \boldsymbol{\mu}, \boldsymbol{\nu}) \leq \liminf_{n\to+\infty} \hat{J}_{\text{inter}}(T_n, \boldsymbol{\mu}^n, \boldsymbol{\nu}^n).$$

Proof. The proof follows the very same argument of Lemma 2.2.12, by replacing μ_t and ν_t by $\mu_t \otimes \mu_t$ and $\nu_t \otimes \nu_t$, respectively. □

Example 2.2.22. In the pedestrian dynamics case, we may assume that the agent at $x \in \mathbb{R}^2$, heading to the direction $v \in \mathbb{S}^1$, can interact only with the agents in its *vision cone*

$$\text{Vis}(x, v) := \{y \in \mathbb{R}^2 : \langle y - x, v\rangle > |y - x| \cos\alpha\},$$

where α is an angle that, for human beings, can be taken as $\alpha \simeq \pi/2$. In the simplest case we can take the *interaction domain* $C(x, v) = \text{Vis}(x, v)$, however the presence of *obstacles* in the environment can *reduce* the interaction domain: for example, a pedestrian cannot be aware of the presence of another pedestrian behind a solid wall. If $\mathcal{O} \subseteq \mathbb{R}^d$ is a closed set representing an obstacle, to compute the interaction domain we must remove from the vision cone all the points *hidden* by the obstacle itself, i.e.

$$C(x, v) = \text{Vis}(x, v) \setminus \{x + \lambda(p - x) : \lambda > 1, p \in \mathcal{O}\}.$$

Moreover, other factors (such that for example fog, or darkness) can affect the vision field and so the *intensity* of the interaction itself. In such a situation, the simplest choice is to take $L_{\text{inter}}(t, x, y, v_x, v_y) = \chi_{C(x,v_x)}(y)$ where the set $C(x, v_x)$ is the interaction domain of the agent at point x with speed v_x, however with such a choice the functional fails in general to be convex in the last two variables. This can cause some problems, for example if we allow the speed to switch instantaneously its direction, it is possible to have agents who follow a trajectory which doesn't allow them to see an obstacle located in front of the target even if they are very closed to it and so, by passing to the limit we can face the inconvenience that the optimal trajectory goes through the obstacle. To tackle this situation, a second-order approach in the dynamics is needed in order to ask more regularity on the velocities.

Definition 2.2.23 (Terminal cost)**.** We assume that the map

$$g : \mathbb{R} \times \mathscr{P}(\mathbb{R}^d) \to [0, +\infty],$$

appearing in (2.5), is a l.s.c. function. It will be called *terminal (or exit) cost.*

Remark 2.2.24. The terminal cost can be used to model terminal constraints. For instance, $g(T, \mu_T) = I_{\tilde{S}}(\mu_T)$ for a given closed set $\tilde{S} \subseteq \mathscr{P}(\mathbb{R}^d)$ models the terminal costraint $\mu_T \in \tilde{S}$ in the generalized minimum time problem. More generally, it can be used also to model a less sharp penalization or simply an *exit cost*: for example we relax the constraint $\mu_T \in \tilde{S}$ penalizing the generalized distance of the final measure from the target $\tilde{S}$ by taking $g(T, \mu_T) = k\tilde{d}^2_{\tilde{S}}(\mu_T)$, where $k > 0$ is a suitable constant and $\tilde{d}_{\tilde{S}}(\mu_T) := \inf_{\sigma\in\tilde{S}} W_2(\mu_T, \sigma)$.

In the mass-preserving case, gluing of admissible trajectories holds. Moreover, the functional is (sub)additive w.r.t. gluing of the trajectories. As already seen in Proposition 2.2.6, these two ingredients yields a Dynamic Programming Principle, whose form depends on the structure of $\mathcal{L}_M$, $\mathcal{L}_m$ and $g(T, \mu_T)$.

Concerning this general treatment, we leave some open problems (see Chapter 5) that we will face only for two specific cases of cost functions: the minimum time function for a mass-preserving case which will be defined in Chapter 3 and for a non-isolated case studied in Chapter 4.

2.3 Non-isolated case

In this case we have a nontrivial creation/destruction of mass during the evolution, so we have $\omega_t \neq 0$, hence the total mass is not preserved during the evolution. For instance, we consider a room with some doors, and the agents may enter or exit these doors at some rate. The problem can be complicated assuming that the rate depends on the concentration of the agents near to the doors, and maybe also on the direction.

The main difficulty in this case is to provide a Superposition Principle comparable to the one holding in the mass-preserving case. For a work on this subject we refer to [57]. Another possibility to circumvent this difficulty, that is the one adopted here, is *to drop completely* the continuity equation, and thus working only with the functional, *building* a solution as a superposition of characteristics, as we will do in Section 2.3.1.

This amounts to consider for instance functionals of the following kind

$$J(T, \hat{\boldsymbol{\eta}}) = \iiint_{[0,T]\times\mathbb{R}^d\times\Gamma_T} L(t, x, \gamma)\, d\hat{\boldsymbol{\eta}}(t, x, \gamma),$$

where $T \geq 0$, and $\hat{\boldsymbol{\eta}} \in \mathscr{M}^+([0,T] \times \mathbb{R}^d \times \Gamma_T)$ is concentrated on $(t, x, \gamma) \in [0,T] \times \mathbb{R}^d \times \Gamma_T$ satisfying $\gamma(0) = x$.

Notice that $L = L(t, x, \gamma)$ in this case depends not only on the current value of t and x, but also on *the whole history* of the trajectory. The first issue is how to embed the underlying finite-dimensional system in this setting.

A natural choice, that we will deepen in Section 2.3.1, is sketched below. We disintegrate $\hat{\boldsymbol{\eta}}$ w.r.t. the first component, i.e., we define a measure $\tau \in \mathscr{M}^+([0,T])$ such that

$$\int_{[0,T]} \varphi(t)\, d\tau := \iiint_{[0,T]\times\mathbb{R}^d\times\Gamma_T} \varphi(t)\, d\hat{\boldsymbol{\eta}}(t, x, \gamma),$$

for all $\varphi \in C^0_C([0,T])$. Then, we can write $\hat{\boldsymbol{\eta}} = \tau \otimes \boldsymbol{\eta}_t$, where $\boldsymbol{\eta}_t \in \mathscr{M}^+(\mathbb{R}^d \times \Gamma_T)$, and we define the measure $\mu_t = e_t \sharp \boldsymbol{\eta}_t \in \mathscr{M}^+(\mathbb{R}^d)$ as before. The measure τ may be used to take into account a variation of the total mass during the time. Usually, we will restrict our attention to $\tau \ll \mathscr{L}^1$ with $\left\| \dfrac{\tau}{\mathscr{L}^1} \right\|_{L^\infty} \leq 1$, i.e. we consider only mass loss due to the position in space. Consequently, $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ is a trajectory which can loose its mass depending on the crossed region of the space.

However, it does not appear immediately evident how to define ν_t, i.e. the corresponding Borel vecor-valued measures. A possibility is to write $\boldsymbol{\eta}_t = \mu_t \otimes \sigma_{t,x}$ with $\sigma_{t,x} \in \mathscr{M}^+(\Gamma_T)$ and define

$$\nu_t := v_t\mu_t := \left(\int_{\{\gamma\in\Gamma_T:\,\gamma(t)=x\}} \dot\gamma(t)\,d\sigma_{t,x}(\gamma)\right)\mu_t, \tag{2.15}$$

provided that the set where $\dot\gamma(t)$ does not exists is τ-negligible. This is actually what we will get from Proposition 2.3.7, where $\boldsymbol{\eta}_t^V$, μ_t^V and $\eta_{t,x}^V$ are respectively what here we just have called $\boldsymbol{\eta}_t$, μ_t and $\sigma_{t,x}$.

If the above construction in (2.15) holds, we can use the same terms of the mass-preserving case to embed the constraints $\dfrac{\nu_t}{\mu_t}(x) \in F(x)$ for μ_t-a.e. $x \in \mathbb{R}^d$ and a.e. $t \in [0,T]$.

In order to have well-posedness of the definition of ν_t, it is sufficient to observe that the set

$$\mathscr{N} := \Big\{(t,x,\gamma) \in [0,T]\times\mathbb{R}^d\times\Gamma_T \,:\, \gamma(0)\neq x \text{ or } \dot\gamma(t) \text{ does not exists}\Big\}$$

is $\mathscr{L}^1\otimes\boldsymbol{\eta}_t$-negligible. Indeed, let us call with $\tilde{\boldsymbol{\eta}}_t^n$ a convex combination of n Dirac deltas concentrated in points belonging to $\operatorname{supp}\boldsymbol{\eta}_t$. We have that $\tilde{\boldsymbol{\eta}}_t^n \rightharpoonup^* \boldsymbol{\eta}_t$, $n \to +\infty$, hence $\mathscr{L}^1\otimes\tilde{\boldsymbol{\eta}}_t^n \rightharpoonup^* \mathscr{L}^1\otimes\boldsymbol{\eta}_t$. By construction, $\mathscr{L}^1\otimes\tilde{\boldsymbol{\eta}}_t^n(\mathscr{N}) = 0$, indeed $\mathscr{N}\cap\operatorname{supp}(\mathscr{L}^1\otimes\tilde{\boldsymbol{\eta}}_t^n)$ is a finite union of sets with zero measure w.r.t. $\mathscr{L}^1\otimes\tilde{\boldsymbol{\eta}}_t^n$. Thus, $\mathscr{L}^1\otimes\boldsymbol{\eta}_t(\mathscr{N}) = 0$ and by projection on the first component, we have that $\dot\gamma(t)$ exists for $\mathscr{L}^1$-a.e. $t\in[0,T]$ and $\sigma_{t,x}$-a.e. $\gamma\in\Gamma_T$.

The major difference with the previous case is that the functional is taking into account *the whole history* of the trajectory γ. In particular, this allows us to consider for instance, to make the mass disappear for all the time after the first time in which the characteristic has entered into a region of the space (this fact implies that at every time, knowing the whole history of the curve, we know if the characteristic curve has already entered the region or not).

Gluing of two trajectories $\boldsymbol{\eta}_i$, $i = 1,2$, seems not to be straightforward in this case, since not only we must have $e_{T_1}\sharp\boldsymbol{\eta}_1 = e_0\sharp\boldsymbol{\eta}_2$, but maybe we should ask something also on the weights of the characteristics. The Dynamic Programming Principle is then far from being trivial.

2.3.1 A more rigorous construction for the annihilation case

For what concerns the treatment about the non-isolated case, here we will consider the particular situation in which the mass is annihilated as soon as it enters into a region V of the space and this represents the only cause of destruction of the mass during the evolution.

Definition 2.3.1 (Absorption time)**.** Let $V \subseteq \mathbb{R}^d$ be closed. Define the map $\tau : \Gamma_T \to [0,T]$ to be the first time in which the curve γ enters in V, i.e.,

$$\tau(\gamma) := \inf\{0 \le t \le T :\, \gamma(t)\in V\} = \min\{0\le t\le T :\, \gamma(t)\in V\},$$

where the infimum is attained by the closedness of V and the continuity of γ. We will call $\tau(\gamma)$ the *absorption time* of $\gamma\in\Gamma_T$. We define now a map from Γ_T

to $\mathscr{M}^+([0,T])$. Given $\gamma \in \Gamma_T$, we set

$$\tau_\gamma(B) = \mathscr{L}^1(B \cap [0,\tau(\gamma)[) = \int_B \chi_{[0,\tau(\gamma)[}(t)\,dt, \text{ for any measurable } B \subseteq [0,T].$$

Lemma 2.3.2. *The following properties hold:*

1. *the map $\tau(\cdot)$ is l.s.c.;*
2. *the map $\gamma \mapsto \tau_\gamma(B)$ is l.s.c. for any fixed measurable $B \subseteq [0,T]$;*
3. *for any Borel measure $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ the product measure*
$$\hat{\boldsymbol{\eta}} := \boldsymbol{\eta} \otimes \tau_\gamma \in \mathscr{M}^+([0,T] \times \mathbb{R}^d \times \Gamma_T)$$
is well-defined and for any bounded Borel map $f : [0,T] \times \mathbb{R}^d \times \Gamma_T \to \mathbb{R}$ we have
$$\iiint_{[0,T]\times\mathbb{R}^d\times\Gamma_T} f(t,x,\gamma)\,d\hat{\boldsymbol{\eta}}(t,x,\gamma) = \iint_{\mathbb{R}^d\times\Gamma_T} \left(\int_0^T f(t,x,\gamma)\,d\tau_\gamma(t)\right) d\boldsymbol{\eta}(x,\gamma).$$

Proof.

1. Let $\{\gamma_n\}_{n\in\mathbb{N}} \subseteq \Gamma_T$ be a sequence uniformly convergent to $\gamma \in \Gamma_T$. Since τ is nonnegative, the result is trivial if $\tau(\gamma) = 0$. Otherwise fix $0 < \varepsilon \le \tau(\gamma)$. By definition, for any $0 \le s \le \tau(\gamma) - \varepsilon$ we have that $\gamma(s) \in \mathbb{R}^d \setminus V$, which is open. Let
$$\delta = \inf\{d_V(\gamma(s)) : s \in [0, \tau(\gamma) - \varepsilon]\},$$
and notice that the infimum is a minimum by Weierstrass Theorem, and moreover that $\delta > 0$. Since for n sufficiently large we have
$$|\gamma_n(s) - \gamma(s)| \le \|\gamma_n - \gamma\|_\infty < \delta/2,$$
we conclude that for n sufficiently large we have
$$\delta \le d_V(\gamma(s)) \le |\gamma_n(s) - \gamma(s)| + d_V(\gamma_n(s)) < \frac{\delta}{2} + d_V(\gamma_n(s)),$$
and so $\gamma_n(s) \notin V$ for all $0 \le s \le \tau(\gamma) - \varepsilon$. Thus for n sufficiently large we obtain $\tau(\gamma_n) \ge \tau(\gamma) - \varepsilon$. By taking the $\liminf$ as $n \to +\infty$ and then letting $\varepsilon \to 0^+$ we have that $\tau(\cdot)$ is l.s.c.

2. Since for n sufficiently large we have $\tau(\gamma_n) \ge \tau(\gamma) - \varepsilon$, we have also that
$$\chi_{[0,\tau(\gamma_n)[}(s) \ge \chi_{[0,\tau(\gamma)-\varepsilon[}(s) = \chi_{[0,\tau(\gamma)[}(s+\varepsilon),$$
and taking again the $\liminf$ as $n \to +\infty$, and $\varepsilon \to 0^+$, recalling the l.s.c. of $\chi_{[0,\tau(\gamma)[}(\cdot)$ on $[0,+\infty[$, we have
$$\liminf_{n\to+\infty} \chi_{[0,\tau(\gamma_n)[}(s) \ge \chi_{[0,\tau(\gamma)[}(s).$$
Given a measurable $B \subseteq [0,T]$ we have by Fatou's Lemma
$$\begin{aligned}\tau_\gamma(B) &= \int_B \chi_{[0,\tau(\gamma)[}(t)\,dt \le \int_B \liminf_{n\to+\infty} \chi_{[0,\tau(\gamma_n)[}(t)\,dt \\ &\le \liminf_{n\to+\infty} \int_B \chi_{[0,\tau(\gamma_n)[}(t)\,dt = \liminf_{n\to+\infty} \tau_{\gamma_n}(B),\end{aligned}$$
which yields the l.s.c. of $\gamma \mapsto \tau_\gamma(B)$ for any measurable $B \subseteq [0,T]$.

3. Since we have that $(x,\gamma)\mapsto\tau_\gamma(B)$ is a Borel map for any measurable $B\subseteq[0,T]$, according to Section 5.3 in [9], if $T<+\infty$ we can define the product measure $\hat{\boldsymbol{\eta}}:=\boldsymbol{\eta}\otimes\tau_\gamma$ on $[0,T]\times\mathbb{R}^d\times\Gamma_T$ for any Borel measure $\boldsymbol{\eta}$ on $\mathbb{R}^d\times\Gamma_T$ (which is separable since $T<+\infty$).

□

In the following result we state probabilistic representations for an admissible trajectory with mass annihilation in the space region V under examination, through the measure $\hat{\boldsymbol{\eta}}$ just defined in the previous lemma.

Lemma 2.3.3 (Representations)**.** *We have the following representations*

$$\hat{\boldsymbol{\eta}}=\mathscr{L}^1\otimes\boldsymbol{\eta}_t^V=\mathscr{L}^1\otimes\mu_t^V\otimes\eta_{t,x}^V,$$

where $\eta_t^V\in\mathscr{M}^+(\mathbb{R}^d\times\Gamma_T)$ *and* $\mu_t^V:=e_t\sharp\boldsymbol{\eta}_t^V\in\mathscr{M}^+(\mathbb{R}^d)$ *are defined for* $\mathscr{L}^1$*-a.e.* $t\in[0,T]$ *by*

$$\iint_{\mathbb{R}^d\times\Gamma_T}\psi(x,\gamma)\,d\boldsymbol{\eta}_t^V(x,\gamma)=\iint_{\mathbb{R}^d\times\Gamma_T}\psi(x,\gamma)\chi_{[0,\tau(\gamma)[}(t)\,d\boldsymbol{\eta}(x,\gamma),$$
$$\int_{\mathbb{R}^d}\varphi(x)\,d\mu_t^V(x)=\int_{\mathbb{R}^d\times\Gamma_T}\varphi(\gamma(t))\chi_{[0,\tau(\gamma)[}(t)\,d\boldsymbol{\eta}(x,\gamma),$$

for all $\psi\in C^0_C(\mathbb{R}^d\times\Gamma_T)$ *and* $\varphi\in C^0_C(\mathbb{R}^d)$*, and for a* $\mathscr{L}^1\otimes\mu_t^V$*-a.e. uniquely defined family of Borel measures* $\{\eta^V_{t,x}\}_{(t,x)\in[0,T]\times\mathbb{R}^d}\subseteq\mathscr{M}^+(\Gamma_T)$.

Finally, for a.e. $t\in[0,T]$ *we have* $\mu_t^V\ll e_t\sharp\boldsymbol{\eta}$ *and for a.e.* $t\in[0,T]$ *and* $(e_t\sharp\boldsymbol{\eta})$*-a.e.* $x\in\mathbb{R}^d$ *it holds*

$$\frac{\mu_t^V}{e_t\sharp\boldsymbol{\eta}}(x)=\int_{(e_t)^{-1}(x)}\chi_{[0,\tau(\gamma)[}(t)\,d\eta_{t,x}(\gamma),$$

where $\{\eta_{t,x}\}_{(t,x)\in[0,T]\times\mathbb{R}^d}\subseteq\mathscr{P}(\Gamma_T)$ *is a family of probability measure uniquely defined for a.e.* $t\in[0,T]$ *and* $(e_t\sharp\boldsymbol{\eta})$*-a.e.* $x\in\mathbb{R}^d$ *and such that* $\boldsymbol{\eta}=(e_t\sharp\boldsymbol{\eta})\otimes\eta_{t,x}$.

Proof. For any fixed t, the map

$$G_t(\psi):=\iint_{\mathbb{R}^d\times\Gamma_T}\psi(x,\gamma)\chi_{[0,\tau(\gamma)[}(t)\,d\boldsymbol{\eta}(x,\gamma),$$

is trivially linear and continuous from $C^0_C(\mathbb{R}^d\times\Gamma_T)\to\mathbb{R}$, since we have

$$|G_t(\psi_1)-G_t(\psi_2)|\le\|\psi_1-\psi_2\|_\infty,$$

thus we have that

$$G_t(\psi)=\iint_{\mathbb{R}^d\times\Gamma_T}\psi(x,\gamma)\,d\boldsymbol{\eta}_t^V$$

for a uniquely defined $\boldsymbol{\eta}_t^V\in\mathscr{M}^+(\mathbb{R}^d\times\Gamma_T)$. Thus we have that $\boldsymbol{\eta}_t^V$ is well defined. By the disintegration theorem (Theorem 5.3.1 in [9]), we define $\mu_t^V=e_t\sharp\boldsymbol{\eta}_t^V$ and the family $\{\eta^V_{t,x}\}_{(t,x)\in[0,T]\times\mathbb{R}^d}$, which satisfy the properties of the statement. To prove that the product $\mathscr{L}^1 \otimes \boldsymbol{\eta}_t^V$ is well-defined, we fix a Borel set $B \subseteq \mathbb{R}^d \times \Gamma_T$, and notice that the map

$$t \mapsto \boldsymbol{\eta}_t^V(B) := \int_B d\boldsymbol{\eta}_t^V(x,\gamma) := \int_B \chi_{[0,\tau(\gamma)[}(t)\, d\boldsymbol{\eta}(x,\gamma)$$

is l.s.c., hence Borel measurable. Thus the product $\mathscr{L}^1 \otimes \boldsymbol{\eta}_t^V$ is well defined, and moreover it coincides with $\hat{\boldsymbol{\eta}}$ since for all bounded Borel functions f we have

$$\int_0^T \iint_{\mathbb{R}^d \times \Gamma_T} f(t,x,\gamma)\, d\boldsymbol{\eta}_t^V(x,\gamma)\, dt = \iiint_{[0,T]\times\mathbb{R}^d \times \Gamma_T} f(t,x,\gamma)\, d\hat{\boldsymbol{\eta}}(t,x,\gamma).$$

The last assertion comes from the fact that for every $\varphi \in C_b^0(\mathbb{R}^d)$ with $\varphi \geq 0$, we have

$$\int_{\mathbb{R}^d} \varphi(x)\, d\mu_t^V(x) \leq \iint_{\mathbb{R}^d \times \Gamma_T} \varphi(\gamma(t))\, d\boldsymbol{\eta}(x,\gamma) = \int_{\mathbb{R}^d} \varphi(x)\, d(e_t \sharp \boldsymbol{\eta})(x),$$

and so for every Borel set B we have $\mu_t^V(B) \leq e_t \sharp \boldsymbol{\eta}(B)$, moreover by taking the disintegration of $\boldsymbol{\eta}$ w.r.t. the Borel map e_t we obtain

$$\begin{aligned}
\int_{\mathbb{R}^d} \varphi(x)\, d\mu_t^V(x) &= \iint_{\mathbb{R}^d \times \Gamma_T} \varphi(\gamma(t))\chi_{[0,\tau(\gamma)[}(t)\, d\boldsymbol{\eta} \\
&= \int_{\mathbb{R}^d} \int_{(e_t)^{-1}(x)} \varphi(\gamma(t))\chi_{[0,\tau(\gamma)[}(t)\, d\eta_{t,x}(\gamma)\, d(e_t \sharp \boldsymbol{\eta})(x) \\
&= \int_{\mathbb{R}^d} \varphi(x) \left(\int_{(e_t)^{-1}(x)} \chi_{[0,\tau(\gamma)[}(t)\, d\eta_{t,x}(\gamma) \right) d(e_t \sharp \boldsymbol{\eta})(x),
\end{aligned}$$

which yields the statement on the density. □

Remark 2.3.4. Since $(e_t)^{-1}(x) = \{\gamma \in \Gamma_T : \gamma(t) = x\}$, we can interpret $\dfrac{\mu_t^V}{(e_t \sharp \boldsymbol{\eta})}(x)$ as the fraction of the characteristic curves passing through x at time t that never passed before through the sink V.

Remark 2.3.5. In general, due to instantaneous mass loss, we cannot expect absolute continuity of the trajectory $[0,T] \to \mathscr{M}^+(\mathbb{R}^d)$, $t \mapsto \mu_t^V$, w.r.t. narrow convergence and consequently w.r.t. W_p^{gen} convergence, where W_p^{gen} denotes the generalized p-Wasserstein distance defined in [64, 65] for finite Borel measures with possibly different masses.

Our aim now is to describe the instantaneous annihilation of the mass when it reaches the sink V and use this defined object to study a continuity equation satisfied by the measure μ_t^V which looses its mass as soon as the underlying characteristics touch the sink V.

Lemma 2.3.6 (Absorption measure)**.** *Let $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$. There exists a unique Radon measure $\mathscr{A}^{\boldsymbol{\eta}} \in \mathscr{M}([0,T] \times \mathbb{R}^d)$ such that*

$$\iint_{[0,T]\times\mathbb{R}^d} \varphi(t,x) d\mathscr{A}^{\boldsymbol{\eta}}(t,x) = \iint_{\mathbb{R}^d \times \Gamma_T} \varphi(\tau(\gamma), \gamma(\tau(\gamma)))\, d\boldsymbol{\eta}(x,\gamma),$$

for all $\varphi \in C_b^0([0,T] \times \mathbb{R}^d)$. We will call $\mathscr{A}^{\boldsymbol{\eta}}$ the absorption measure *associated to $\boldsymbol{\eta}$.*

Proof. Indeed, since for any $\varphi_1, \varphi_2 \in C^0_b([0,T]\times\mathbb{R}^d)$ we have

$$\iint_{\mathbb{R}^d\times\Gamma_T} |\varphi_1(\tau(\gamma),\gamma(\tau(\gamma))) - \varphi_2(\tau(\gamma),\gamma(\tau(\gamma)))|\, d\boldsymbol{\eta}(x,\gamma) \le \|\varphi_1-\varphi_2\|_\infty,$$

we have that the map from $C^0_b([0,T]\times\mathbb{R}^d)$ to $\mathbb{R}$ defined as

$$\varphi \mapsto \iint_{\mathbb{R}^d\times\Gamma_T} \varphi(\tau(\gamma),\gamma(\tau(\gamma)))\, d\boldsymbol{\eta}(x,\gamma)$$

is linear and 1-Lipschitz continuous, thus $\mathscr{A}^{\boldsymbol{\eta}} \in [C^0_b([0,T]\times\mathbb{R}^d)]'$. □

We want now to apply the previous consideration to find a PDE satisfied by μ_t^V when $\boldsymbol{\eta}$ is chosen in order to have that $t\mapsto \mu_t := e_t\sharp\boldsymbol{\eta}$ satisfies the homogeneous continuity equation $\partial_t\mu_t + \mathrm{div}(v_t\mu_t) = 0$, in the spirit of the Superposition Principle in Theorem 8.2.1 in [9].

Proposition 2.3.7. *Let $p>1$. Assume that $v : [0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ is a Borel time-depending vector field and $\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$ is a measure such that $\boldsymbol{\eta}$ is concentrated on the pairs (x,γ) where γ is an AC^p solution of $\dot\gamma(t) = v_t\circ\gamma(t)$, $\gamma(0) = x\notin V$ and*

$$\int_0^T \iint_{\mathbb{R}^d\times\Gamma_T} |v_t\circ\gamma(t)|^p\, d\boldsymbol{\eta}\, dt < +\infty.$$

Define $\hat{\boldsymbol{\eta}} = \boldsymbol{\eta}\otimes\tau_\gamma$ as in Lemma 2.3.2, and $\{\mu_t^V\}_{t\in[0,T]}$ as in Lemma 2.3.3. Then in the sense of distributions we have

$$\partial_t\mu_t^V + \mathrm{div}(v_t\mu_t^V) = -\mathscr{A}^{\boldsymbol{\eta}}. \tag{2.16}$$

Moreover, if $\left\|\dfrac{1+\mathrm{Id}_{\mathbb{R}^d}}{d_V}\right\| \in L^\infty_{e_0\sharp\boldsymbol{\eta}}$ and there exists $C>0$ such that $\left|\dfrac{v_t(x)}{1+|x|}\right| \le C$ for all $(t,x)\in[0,T]\times\mathbb{R}^d$, then there exists $\varepsilon>0$ and a family $\{\tilde\mu_t^V\}_{t\in[0,\varepsilon]}\subseteq\mathscr{P}(\mathbb{R}^d)$ such that $\tilde\mu_t^V = \mu_t^V$ for a.e. $t\in[0,\varepsilon]$, $t\mapsto\tilde\mu_t^V$ is narrowly continuous, and $\tilde\mu^V_{|t=0} = e_0\sharp\boldsymbol{\eta}$.

Proof. Consider a test function $\varphi\in C^\infty_C([0,T]\times\mathbb{R}^d)$, we have

$$\begin{aligned}
\int_0^T\int_{\mathbb{R}^d} \partial_t\varphi(t,x)\, d\mu_t^V\, dt &= \int_0^T \iint_{\mathbb{R}^d\times\Gamma_T} \partial_t\varphi(t,\gamma(t))\, d\boldsymbol{\eta}_t^V(x,\gamma)\, dt \\
&= \iiint_{[0,T]\times\mathbb{R}^d\times\Gamma_T} \partial_t\varphi(t,\gamma(t))\, d\hat{\boldsymbol{\eta}}(t,x,\gamma) \\
&= \iiint_{[0,T]\times\mathbb{R}^d\times\Gamma_T} \left(\frac{d}{dt}[\varphi(t,\gamma(t))] - \langle\nabla\varphi(t,\gamma(t)),\dot\gamma(t)\rangle\right) d\hat{\boldsymbol{\eta}}(t,x,\gamma)
\end{aligned}$$

Recalling that $\boldsymbol{\eta}$ is supported on $(\gamma(0),\gamma)$ where $\dot\gamma(t) = v_t\circ\gamma(t)$ for a.e. $t\in[0,T]$

and so for τ_γ-a.e. $t \in [0,T]$, we have

$$\begin{aligned}\iiint_{[0,T]\times\mathbb{R}^d\times\Gamma_T} \nabla\varphi(t,\gamma(t))\dot\gamma(t)\,d\hat{\boldsymbol{\eta}}(t,x,\gamma) &=\\ &= \iint_{\mathbb{R}^d\times\Gamma_T}\int_0^T \langle\nabla\varphi(t,\gamma(t)), v_t\circ\gamma(t)\rangle\,d\tau_\gamma\,d\boldsymbol{\eta}(x,\gamma)\\ &= \iiint_{[0,T]\times\mathbb{R}^d\times\Gamma_T}\langle\nabla\varphi(t,\gamma(t)), v_t\circ\gamma(t)\rangle\,d\hat{\boldsymbol{\eta}}(t,x,\gamma)\\ &= \int_0^T\iint_{\mathbb{R}^d\times\Gamma_T}\langle\nabla\varphi(t,\gamma(t)), v_t\circ\gamma(t)\rangle\,d\boldsymbol{\eta}_t^V(x,\gamma)\,dt\\ &= \int_0^T\int_{\mathbb{R}^d}\langle\nabla\varphi(t,x), v_t(x)\rangle\,d\mu_t^V(x)\,dt\end{aligned}$$

Since $\varphi \in C_C^\infty([0,T]\times\mathbb{R}^d)$ we have $\varphi(0,y)\equiv 0$ for all $y\in\mathbb{R}^d$, and so

$$\begin{aligned}\iiint_{[0,T]\times\mathbb{R}^d\times\Gamma_T}\frac{d}{dt}\left[\varphi(t,\gamma(t))\right]\,d\hat{\boldsymbol{\eta}}(t,x,\gamma) &= \iint_{\mathbb{R}^d\times\Gamma_T}\int_0^T\frac{d}{dt}\left[\varphi(t,\gamma(t))\right]\,d\tau_\gamma\,d\boldsymbol{\eta}(x,\gamma)\\ &= \iint_{\mathbb{R}^d\times\Gamma_T}\int_0^{\tau(\gamma)}\frac{d}{dt}\left[\varphi(t,\gamma(t))\right]\,dt\,d\boldsymbol{\eta}(x,\gamma)\\ &= \iint_{\mathbb{R}^d\times\Gamma_T}\varphi(\tau(\gamma),\gamma(\tau(\gamma)))\,d\boldsymbol{\eta}(x,\gamma)\\ &= \iint_{[0,T]\times\mathbb{R}^d}\varphi(t,x)\,d\mathscr{A}^{\boldsymbol{\eta}}(t,x).\end{aligned}$$

We have that (2.16) follows.

To prove the last assertion, we recall that since $|v_t(x)|\le C(|x|+1)$, then for all $(x,\gamma)\in\operatorname{supp}\boldsymbol{\eta}$ we have

$$\begin{aligned}|\gamma(t)-\gamma(0)| &\le \int_0^t|\dot\gamma(s)|\,ds = \int_0^t|v_t\circ\gamma(s)|\,ds \le C\int_0^t|\gamma(s)|\,ds + Ct\\ &\le C\int_0^t|\gamma(s)-\gamma(0)|\,ds + Ct(1+|\gamma(0)|),\end{aligned}$$

and so by Gronwall's inequality

$$|\gamma(t)-\gamma(0)|\le Ct(1+|\gamma(0)|)e^{Ct}\le Ct(1+|\gamma(0)|)e^{CT} = Ct(1+|x|)e^{CT}.$$

By assumption, for $e_0\sharp\boldsymbol{\eta}$-a.e. $x\in\mathbb{R}^d$ we have $1+|x|<C'd_V(x)$ for a constant $C'>0$, thus

$$|\gamma(t)-\gamma(0)|\le C\cdot C'td_V(x)e^{CT},\ \text{ for } \boldsymbol{\eta}\text{-a.e } (x,\gamma)\in\mathbb{R}^d\times\Gamma_T.$$

This implies that if $t<\dfrac{e^{-CT}}{CC'}$ we have $\gamma(t)\notin V$, and so $\tau(\gamma)\ge\dfrac{e^{-CT}}{CC'}$. Set $\varepsilon = e^{-CT}/(2CC')$. Then for every $\varphi\in C_C^\infty([0,T]\times\mathbb{R}^d)$ with $\operatorname{supp}\varphi\subseteq[0,\varepsilon]\times\mathbb{R}^d$ we have

$$\iint_{[0,T]\times\mathbb{R}^d}\varphi(t,x)\,d\mathscr{A}^{\boldsymbol{\eta}}(t,x) = 0,$$

and so we have that the restriction of $t \mapsto \mu_t^V$ to $[0,\varepsilon]$ solves the homogeneous continuity equation with initial data $e_0\sharp\boldsymbol{\eta}$. Thus the existence of a continuous representative follows from Lemma 8.1.2 in [9]. □

The previous result provides us with a continuity equation in the non-isolated case with annihilation. We precise that here the sink is described by an absorption measure that is defined in $[0,T]\times\mathbb{R}^d$, hence μ_t^V satisfies the continuity equation with sink in the sense of distributions with integration also in time. This allows us to hide the impulsive term in the absorption measure $\mathscr{A}^{\boldsymbol{\eta}}$.

We pass now to consider some cost functionals defined on curves $\{\mu_t^V\}_{t\in[0,T]}$ which are constructed as in Lemma 2.3.3 by mean of Lemma 2.3.2. We will see that it is possible to write such functionals defined for the non-isolated case with annihilation, in the form of the functionals of the mass-preserving case. In this way we inherit the results of the isolated case.

Let $p>1$. Assume that $v:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ is a Borel time-depending vector field and $\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$ is a measure such that $\boldsymbol{\eta}$ is concentrated on the pairs (x,γ) where γ is an AC^p solution of $\dot\gamma(t)=v_t\circ\gamma(t)$, $\gamma(0)=x$ and

$$\int_0^T\iint_{\mathbb{R}^d\times\Gamma_T}|v_t\circ\gamma(t)|^p\,d\boldsymbol{\eta}\,dt<+\infty.$$

Define $\hat{\boldsymbol{\eta}}=\boldsymbol{\eta}\otimes\tau_\gamma$ as in Lemma 2.3.2, $\boldsymbol{\mu}^V=\{\mu_t^V\}_{t\in[0,T]}$ as in Lemma 2.3.3, and set $\boldsymbol{\nu}^V=\{\nu_t^V:=v_t\mu_t^V\}_{t\in[0,T]}$, $\boldsymbol{\mu}=\{\mu_t\}_{t\in[0,T]}$, with $\mu_t:=e_t\sharp\boldsymbol{\eta}$, and $\boldsymbol{\nu}=\{\nu_t:=v_t\mu_t\}_{t\in[0,T]}$. Then we get the following relations for the three different types of cost terms already analyzed for the mass-preserving case.

- Let $L:\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^d\to[0,+\infty]$ be a Borel function, and consider the functional

$$J_{\mathrm{sys}}(T,\boldsymbol{\mu}^V,\boldsymbol{\nu}^V):=\int_0^T\int_{\mathbb{R}^d}L\left(t,x,\frac{\nu_t^V}{\mu_t^V}(x)\right)\,d\mu_t^V(x)\,dt.$$

  Then
$$J_{\mathrm{sys}}(T,\boldsymbol{\mu}^V,\boldsymbol{\nu}^V)=\tilde J_{\mathrm{sys}}(T,\boldsymbol{\eta}),$$
  where
$$\tilde J_{\mathrm{sys}}(T,\boldsymbol{\eta}):=\iint_{\mathbb{R}^d\times\Gamma_T}\int_0^T\chi_{[0,\tau(\gamma)[}(t)\,L\left(t,\gamma(t),\dot\gamma(t)\right)\,dt\,d\boldsymbol{\eta}(x,\gamma).$$

- Let $L^V:\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d\to[0,+\infty]$ be a Borel map, $\sigma\in\mathscr{M}^+(\mathbb{R}^d)$, and consider the functional

$$J_{\mathrm{dens}}(T,\boldsymbol{\mu}^V,\boldsymbol{\nu}^V):=\begin{cases}\displaystyle\int_0^T\int_{\mathbb{R}^d}L^V\left(t,x,\frac{\mu_t^V}{\sigma}(x),\frac{\nu_t^V}{\sigma}(x)\right)d\sigma\,dt, & \text{if } \mu_t^V\ll\sigma \text{ and } |\nu_t|^V\ll\sigma \text{ for a.e.} t\in[0,T],\\ +\infty, & \text{otherwise.}\end{cases}\tag{2.17}$$

Recalling Lemma 2.3.3, we have that $\mu_t^V \ll \mu_t$, more precisely, we have

$$\mu_t^V = \left(\int_{(e_t)^{-1}(x)} \chi_{[0,\tau(\gamma)[}(t)\, d\eta_{t,x}(\gamma) \right) \mu_t.$$

We can set

$$L_{\mathrm{dens}}(t,x,d_x,d_v) := L^V\left(t, x, \frac{\mu_t^V}{\mu_t}(x)\cdot d_x, \frac{\nu_t^V}{\nu_t}(x)\cdot d_v\right)$$

$$\tilde{J}_{\mathrm{dens}}(T,\boldsymbol{\mu},\boldsymbol{\nu}) := \begin{cases} \displaystyle\int_0^T \int_{\mathbb{R}^d} L_{\mathrm{dens}}\left(t,x,\frac{\mu_t}{\sigma}(x),\frac{\nu_t}{\sigma}(x)\right) d\sigma\, dt, & \text{if for a.e.} t\in[0,T], \\ & \text{either } \left\| \dfrac{\mu_t^V}{\mu_t} \right\|_{L^1_{\mu_t}} = 0 \\ & \text{or } \mu_t \ll \sigma,\ |\nu_t| \ll \sigma. \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.18}$$

We thus obtain

$$J_{\mathrm{dens}}(T,\boldsymbol{\mu}^V,\boldsymbol{\nu}^V) = \tilde{J}_{\mathrm{dens}}(T,\boldsymbol{\mu},\boldsymbol{\nu}).$$

- Let $L_{\mathrm{inter}} : \mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d \to [0,+\infty]$ be a Borel function, and consider the functional

$$J_{\mathrm{inter}}(T,\boldsymbol{\mu}^V,\boldsymbol{\nu}^V) := \int_0^T \iint_{\mathbb{R}^d\times\mathbb{R}^d} L_{\mathrm{inter}}\left(t,x,y,\frac{\nu_t^V}{\mu_t^V}(x),\frac{\nu_t^V}{\mu_t^V}(y)\right) d\mu_t^V(x)\, d\mu_t^V(y)\, dt.$$

Then

$$J_{\mathrm{inter}}(T,\boldsymbol{\mu}^V,\boldsymbol{\nu}^V) = \tilde{J}_{\mathrm{inter}}(T,\boldsymbol{\eta}),$$

where

$$\tilde{J}_{\mathrm{inter}}(T,\boldsymbol{\eta}) := \int_{X_{\mathrm{inter}}} \int_0^T \chi_{[0,\min\{\tau(\gamma_y),\tau(\gamma_x)\}[}(t)\cdot \\ \cdot L_{\mathrm{inter}}\left(t,\gamma_x(t),\gamma_y(t),\dot\gamma_x(t),\dot\gamma_y(t)\right)\, dt\, d\boldsymbol{\eta}(x,\gamma_x)\, d\boldsymbol{\eta}(y,\gamma_y).$$

Chapter 3

Time-optimal control problem in the mass-preserving case

In this chapter we investigate a time-optimal control problem in the space of positive and finite Borel measures dealing with a mass-preserving situation. The dynamics is thus described by an homogeneous continuity equation. Without loss of generality we choose to normalize the total mass to 1, dealing with Borel probability measures.

This study can be found also in [28, 30–32].

The main results obtained in this Chapter can be summarized as follows:

1. a theorem of existence of time-optimal curves in the space of probability measures (Theorem 3.2.20);

2. a Dynamic Programming Principle (Theorem 3.2.25);

3. comparison results between classical and generalized minimum time function (Proposition 3.2.12, Corollary 3.2.22 and Corollary 3.2.23);

4. sufficient conditions providing upper bounds of the generalized minimum time function (attainability results) (Theorem 3.2.26, Theorem 3.2.32 and Theorem 3.2.35),

5. sufficient conditions yielding Lipschitz continuity of the generalized minimum time function (Theorem 3.2.42);

6. the introduction of a natural Hamilton-Jacobi-Bellman equation for the generalized minimum time function, which turns out to be a solution in a suitable infinite-dimensional viscosity sense (Theorem 3.3.9).

7. some tools which would lead to the study of higher order attainability conditions (Section 3.4).

3.1 Generalized targets

In this section we propose some suitable generalizations of the classical target set in $\mathbb{R}^d$ that can be used in our framework in the space of probability measures and we analyse some properties (convexity, closedness, compactness) and relations with the classical target, when possible. Also regularity properties of the correspondent generalized distance from the target are studied.

Definition 3.1.1 (Generalized targets)**.** Let $p \geq 1$, Φ be a given set of lower semicontinuous maps from $\mathbb{R}^d$ to $\mathbb{R}$, such that the following property holds

(T_E) there exists $x_0 \in \mathbb{R}^d$ with $\phi(x_0) \leq 0$ for all $\phi \in \Phi$, and all $\phi \in \Phi$ are bounded from below.

We define the *generalized targets* $\tilde{S}^\Phi$ and $\tilde{S}^\Phi_p$ as follows

$$\begin{aligned}
\tilde{S}^\Phi :=& \left\{ \mu \in \mathscr{P}(\mathbb{R}^d) : \phi \in L^1_\mu \text{ and } \int_{\mathbb{R}^d} \phi(x)\, d\mu(x) \leq 0 \text{ for all } \phi \in \Phi \right\}, \\
\tilde{S}^\Phi_p :=& \tilde{S}^\Phi \cap \mathscr{P}_p(\mathbb{R}^d).
\end{aligned}$$

We define also the *generalized distance* from $\tilde{S}^\Phi_p$ as

$$\tilde{d}_{\tilde{S}^\Phi_p}(\cdot) := \inf_{\mu \in \tilde{S}^\Phi_p} W_p(\cdot, \mu).$$

Notice that $\tilde{S}^\Phi_p \neq \emptyset$ because $\delta_{x_0} \in \tilde{S}^\Phi_p$, hence $\tilde{S}^\Phi \neq \emptyset$. The 1-Lipschitz continuity of $\tilde{d}_{\tilde{S}^\Phi_p}(\cdot)$ follows from the structure of metric space: indeed let $\mu, \nu \in \mathscr{P}_p(\mathbb{R}^d)$, and fix $\varepsilon > 0$. Choose $\sigma_\nu \in \tilde{S}^\Phi_p$ such that $\tilde{d}_{\tilde{S}^\Phi_p}(\nu) \geq W_p(\nu, \sigma_\nu) - \varepsilon$. Then we have by triangular inequality

$$\tilde{d}_{\tilde{S}^\Phi_p}(\mu) - \tilde{d}_{\tilde{S}^\Phi_p}(\nu) \leq W_p(\mu, \sigma_\nu) - W_p(\nu, \sigma_\nu) + \varepsilon \leq W_p(\mu, \nu) + \varepsilon.$$

By switching the role of μ, ν and letting $\varepsilon \to 0^+$, we obtain the desired Lipschitz continuity property.

For further use, we will say that Φ satisfies property (T_p) with $p \geq 1$ if the following holds true

(T_p) for all $\phi \in \Phi$ there exist $A_\phi, C_\phi > 0$ such that $\phi(x) \geq A_\phi |x|^p - C_\phi$.

Remark 3.1.2. Roughly speaking, a physical interpretation of the generalized target can be given as follows: to describe the state of the system, an observer chooses to measure some quantities ϕ. The results of the measurements are the average of the quantities ϕ with respect to the measure μ_t representing the state of the system at time t. Our aim is to steer the system to states where the result of such measurements is below a fixed threshold (that without loss of generality we assume to be 0).

Remark 3.1.3. Given a nonempty and closed set $S \subseteq \mathbb{R}^d$ and $\alpha \in]0, 1]$, a natural choice for Φ can be for example $\Phi = \{d_S(\cdot) - \alpha\}$. In this case, a measure belonging to $\tilde{S}^\Phi$ corresponds to the state of a particle which is on S with probability $1 - \alpha$. If $\alpha = 0$, i.e. $\Phi = \{d_S(\cdot)\}$, then $\tilde{S}^\Phi$ reduces to the set of all probability measures supported on S.

The following proposition establishes some straightforward properties of the generalized targets.

Proposition 3.1.4 (Properties of the generalized targets)**.** *Let $p \geq 1$ and Φ be a given set of lower semicontinuous maps from $\mathbb{R}^d$ to $\mathbb{R}$ such that (T_E) holds. Then:*

(1) $\tilde{S}^\Phi$ and $\tilde{S}^\Phi_p$ are convex;

(2) $\tilde{S}^\Phi$ is w^*-closed in $\mathscr{P}(\mathbb{R}^d)$;

(3) $\tilde{S}^\Phi_p$ is closed in $\mathscr{P}_p(\mathbb{R}^d)$ endowed with the p-Wasserstein metric $W_p(\cdot,\cdot)$;

(4) for every $\mu \in \mathscr{P}_p(\mathbb{R}^d)$ we have $\tilde{d}_{\tilde{S}^\Phi_p}(\mu) = 0$ if and only if $\mu \in \tilde{S}^\Phi_p$;

(5) if there exists $\bar{\phi} \in \Phi$, $A, C > 0$ and $p \geq 1$ such that $\bar{\phi}(x) \geq A|x|^p - C$, then $\tilde{S}^\Phi = \tilde{S}^\Phi_p$ is compact in the w^*-topology and in the W_p-topology. In particular, this holds if Φ satisfies property (T_p).

Proof.

1. The convexity property is trivial from the definition.

2. Let $\{\mu_n\}_{n\in\mathbb{N}}$ be a sequence in $\tilde{S}^\Phi$, and $\mu \in \mathscr{P}(\mathbb{R}^d)$ be such that $\mu_n \rightharpoonup^* \mu$. Since for any fixed $\phi \in \Phi$, ϕ is a l.s.c. function bounded from below, we have $\phi(x) = \sup_{k\in\mathbb{N}} \phi_k(x)$, $x \in \mathbb{R}^d$, where
$$\phi_k(x) := \min\left\{\inf_{y\in\mathbb{R}^d}\{\phi(y) + k|x-y|\}, k\right\},$$
$k \in \mathbb{N}$, and ϕ_k is a bounded Lipschitz continuous function for every $k \in \mathbb{N}$. Then by Monotone Convergence Theorem we have for all $n \in \mathbb{N}$,
$$\begin{aligned} 0 \geq \int_{\mathbb{R}^d} \phi(x)\, d\mu_n(x) &= \int_{\mathbb{R}^d} \left[\sup_{k\in\mathbb{N}} \phi_k(x)\right] d\mu_n(x) \\ &= \sup_{k\in\mathbb{N}} \int_{\mathbb{R}^d} \phi_k(x)\, d\mu_n(x) \geq \int_{\mathbb{R}^d} \phi_k(x)\, d\mu_n(x), \end{aligned}$$
for all $k \in \mathbb{N}$. By letting $n \to +\infty$, recalling the weak* convergence of μ_n to μ, we obtain that $0 \geq \int_{\mathbb{R}^d} \phi_k(x)\, d\mu(x)$, for all $k \in \mathbb{N}$. Hence, by passing to the supremum on $k \in \mathbb{N}$ we get $0 \geq \int_{\mathbb{R}^d} \phi(x)\, d\mu(x)$, and so $\mu \in \tilde{S}^\Phi$.

3. It follows from the fact that convergence in $W_p(\cdot,\cdot)$ implies w^*-convergence, and that $\mathscr{P}_p(\mathbb{R}^d)$ endowed with the p-Wasserstein metric $W_p(\cdot,\cdot)$ is a complete separable metric space according to Proposition 1.2.2.

4. It is obvious that if $\mu \in \tilde{S}^\Phi_p$ then $\tilde{d}_{\tilde{S}^\Phi_p}(\mu) = 0$. Conversely, if $\tilde{d}_{\tilde{S}^\Phi_p}(\mu) = 0$ there exists a sequence $\{\mu_n\}_{n\in\mathbb{N}} \subseteq \tilde{S}^\Phi_p$ such that $\lim\limits_{n\to\infty} W_p(\mu_n, \mu) = 0$, and, by the closedness of $\tilde{S}^\Phi_p$, we conclude that $\mu \in \tilde{S}^\Phi_p$.

5. Given $p \geq 1$, trivially we have that $\tilde{S}^\Phi_p \subseteq \tilde{S}^\Phi$. Conversely, given $\mu \in \tilde{S}^\Phi$, we have
$$\int_{\mathbb{R}^d} (A|x|^p - C)\, d\mu \leq \int_{\mathbb{R}^d} \bar{\phi}(x)\, d\mu \leq 0,$$

where $\bar{\phi}$, A, C, p, are as in the assumptions. Thus for all $\mu \in \tilde{S}^\Phi$ we have

$$\int_{\mathbb{R}^d} |x|^p \, d\mu \leq \frac{C}{A} < +\infty,$$

hence $\mu \in \tilde{S}^\Phi_p$. So all the measures in $\tilde{S}^\Phi_p = \tilde{S}^\Phi$ have uniformly bounded p-moments. Hence, if $\{\mu_n\}_{n\in\mathbb{N}} \subseteq \tilde{S}^\Phi$ and $\mu_n \rightharpoonup^* \mu$, by the w^*-closure of $\tilde{S}^\Phi$ we have that $\mu \in \tilde{S}^\Phi = \tilde{S}^\Phi_p$ and it has finite p-moment. Thus, the family $\{\mu_n\}_{n\in\mathbb{N}}$ has equiuniformly integrable p-moments, and $W_p(\mu_n, \mu) \to 0$ by Proposition 1.2.2. This means that the w^*-topology and W_p-topology coincide on $\tilde{S}^\Phi = \tilde{S}^\Phi_p$, which turns out to be tight, according to Remark 5.1.5 in [9], and w^*-closed, hence w^*-compact and W_p-compact.

□

Given a nonempty closed set $S \subseteq \mathbb{R}^d$, and set $\Phi = \{d_S(\cdot)\}$, a natural problem is to express the generalized distance $\tilde{d}_{\tilde{S}^\Phi_p}(\cdot)$ in terms of $d_S(\cdot)$. More generally, we give the following definition.

Definition 3.1.5 (Classical counterpart of generalized target)**.** Let $p \geq 1$ and $\Phi \subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1.1. Given a set $S \subseteq \mathbb{R}^d$, we say that

1. S is a *classical counterpart of the generalized target* $\tilde{S}^\Phi$ if the following equality holds
$$\tilde{S}^\Phi = \{\mu \in \mathscr{P}(\mathbb{R}^d) : \operatorname{supp}\mu \subseteq S\}.$$
2. S is a *classical counterpart of the generalized target* $\tilde{S}^\Phi_p$ if the following equality holds
$$\tilde{S}^\Phi_p = \{\mu \in \mathscr{P}_p(\mathbb{R}^d) : \operatorname{supp}\mu \subseteq S\}.$$

Proposition 3.1.6 (Existence, uniqueness and properties of the classical counterpart)**.** *Let $p \geq 1$ and $\Phi \subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1.1. Then*

1. *if $\tilde{S}^\Phi$ admits a classical counterpart S, then $\tilde{S}^\Phi_p$ admits S as a classical counterpart for all $p \geq 1$.*
2. *if S, S', are classical counterparts of the generalized targets $\tilde{S}^\Phi$, $\tilde{S}^\Phi_p$, respectively, then $S = S'$;*
3. *if S is a classical counterpart of $\tilde{S}^\Phi$ or of $\tilde{S}^\Phi_p$, then S is closed;*
4. *if S is the classical counterpart of $\tilde{S}^\Phi$ then $\phi(x) \leq 0$ for all $\phi \in \Phi$, $x \in S$;*
5. *if $\phi(x) \geq 0$ for all $\phi \in \Phi$ and $x \in \mathbb{R}^d$ then the set*
$$S := \{x \in \mathbb{R}^d : \phi(x) = 0 \text{ for all } \phi \in \Phi\}$$
is the classical counterpart of $\tilde{S}^\Phi$ and of $\tilde{S}^\Phi_p$ (uniqueness follows from item (2) above);
6. *if S is the classical counterpart of $\tilde{S}^\Phi$, then there exists a representation of $\tilde{S}^\Phi$ as $\tilde{S}^{\Phi'}$, where $\phi'(x) \geq 0$ $\forall x \in \mathbb{R}^d$, $\phi' \in \Phi'$. In particular we can take $\Phi' = \{d_S\}$ and we have $\tilde{S}^\Phi = \tilde{S}^{\{d_S\}}$ and $\tilde{S}^\Phi_p = \tilde{S}^{\{d_S\}}_p$, i.e., we can replace Φ with the set $\{d_S\}$;*

7. *if for every $\phi \in \Phi$ we have either $\phi(x) \geq 0$ or $\phi(x) \leq 0$ for all $x \in \mathbb{R}^d$, then $\tilde{S}^\Phi$ and $\tilde{S}^\Phi_p$ admit as classical counterpart the set*

$$S = \bigcap_{\phi\in\Phi} \{x \in \mathbb{R}^d : \phi(x) \leq 0\} = \bigcap_{\phi\in\Phi^+} \{x \in \mathbb{R}^d : \phi(x) = 0\},$$

where $\Phi^+ = \{\phi \in \Phi : \phi(x) \geq 0$ for all $x \in \mathbb{R}^d\}$, and if $\Phi^+ = \emptyset$ we set $S = \mathbb{R}^d$.

Proof.

1. By definition, for all $p \geq 1$ we have

$$\begin{aligned}\tilde{S}^\Phi_p := \tilde{S}^\Phi \cap \mathscr{P}_p(\mathbb{R}^d) &= \{\mu \in \mathscr{P}(\mathbb{R}^d) : \operatorname{supp}\mu \subseteq S\} \cap \mathscr{P}_p(\mathbb{R}^d) \\ &= \{\mu \in \mathscr{P}_p(\mathbb{R}^d) : \operatorname{supp}\mu \subseteq S\}.\end{aligned}$$

2. Let S and S' be two classical counterparts of $\tilde{S}^\Phi$ and of $\tilde{S}^\Phi_p$, respectively. For every $x \in S$ we have that $\delta_x \in \tilde{S}^\Phi_p \subseteq \tilde{S}^\Phi$ for all $p \geq 1$, hence we must have also $x \in S'$ since S' is a classical counterpart of the generalized target $\tilde{S}^\Phi_p$. So $S \subseteq S'$. By reversing the roles of S and S' we obtain $S = S'$.

3. Let S be the classical counterpart of $\tilde{S}^\Phi$ (the proof is analoguos for $\tilde{S}^\Phi_p$). Let $\{x_n\}_{n\in\mathbb{N}} \subseteq S$ be s.t. $x_n \to \bar{x}$ for some $\bar{x} \in \partial S$. By contradiction, let us suppose $\bar{x} \notin S$, thus $\delta_{\bar{x}} \notin \tilde{S}^\Phi$. Then there exists $\bar{\phi} \in \Phi$ s.t. $\bar{\phi}(\bar{x}) > 0$, and thus for n sufficiently large we have $\bar{\phi}(x_n) > 0$ by continuity of $\bar{\phi}$. It follows that $\delta_{x_n} \notin \tilde{S}^\Phi$ for n sufficiently large, thus $x_n \notin S$ by definition of classical counterpart and we get a contradiction.

4. Immediate by definition of generalized target and of classical counterpart, in fact we have $\delta_{\bar{x}} \in \tilde{S}^\Phi$ for all $\bar{x} \in S$.

5. Obviously we have

$$\{\mu \in \mathscr{P}(\mathbb{R}^d) : \operatorname{supp}\mu \subseteq S\} \subseteq \tilde{S}^\Phi.$$

Let us prove the other inclusion. Note that by hypothesis $\phi \geq 0$ for every $\phi \in \Phi$, hence

$$\tilde{S}^\Phi = \left\{\mu \in \mathscr{P}(\mathbb{R}^d) : \phi \in L^1_\mu \text{ and } \int_{\mathbb{R}^d} \phi(x)\,d\mu(x) = 0 \text{ for all } \phi \in \Phi\right\}.$$

Let $\mu \in \tilde{S}^\Phi$, then

$$\int_{\mathbb{R}^d} \phi(x)\,d\mu(x) = 0 \quad \forall \phi \in \Phi,$$

i.e. $\phi(x) = 0$ for μ-a.e. $x \in \mathbb{R}^d$, $\forall\phi \in \Phi$, i.e. $\phi(x) = 0$ for all $x \in \operatorname{supp}\mu$, $\forall\phi \in \Phi$. Thus $\operatorname{supp}\mu \subseteq S$. By item (1), S is the classical counterpart also of $\tilde{S}^\Phi_p$.

6. Let us prove that $\tilde{S}^{\{d_S\}} = \tilde{S}^\Phi$. First $\tilde{S}^{\{d_S\}} \subseteq \tilde{S}^\Phi$, in fact if $\mu \in \tilde{S}^{\{d_S\}}$ then $\mu(\mathbb{R}^d \setminus S) = 0$, and so $\mu \in \tilde{S}^\Phi$ by definition of classical counterpart. Moreover, $\tilde{S}^{\{d_S\}} \supseteq \tilde{S}^\Phi$, in fact if $\mu \in \tilde{S}^\Phi$, then $\operatorname{supp}\mu \subseteq S$ and it follows that $\int_{\mathbb{R}^d} d_S(x)\,d\mu(x) = 0$, thus $\mu \in \tilde{S}^{\{d_S\}}$.

7. By item (1), it is sufficient to prove that S is the classical counterpart of $\tilde{S}^\Phi$. Assume that $\Phi^+ = \emptyset$. This means that $\phi(x) \leq 0$ for all $x \in \mathbb{R}^d$ and for all $\phi \in \Phi$. In this case we have that $\tilde{S}^\Phi = \mathscr{P}(\mathbb{R}^d)$ since for every $\mu \in \mathscr{P}(\mathbb{R}^d)$ we have
$$\int_{\mathbb{R}^d} \phi(x)\, d\mu(x) \leq 0.$$
Thus we have trivially $S = \mathbb{R}^d$.

   Suppose now $\Phi^+ \neq \emptyset$. Clearly, every measure supported in S belongs to $\tilde{S}^\Phi$, since all the elements of Φ are nonpositive on S, i.e. $\tilde{S}^{\{d_S\}} \subseteq \tilde{S}^\Phi$. Conversely, let $\mu \in \tilde{S}^\Phi$ and by contradiction assume that there exists $\bar{x} \in \operatorname{supp} \mu \setminus S$. This implies that there exists an open neighborhood A of $\bar{x}$ such that $\mu(A) > 0$, and an element $\phi \in \Phi^+$ such that $\phi(\bar{x}) \neq 0$. By continuity of ϕ, we can assume that $\phi > 0$ on the whole of A, thus, recalling that $\phi(x) \geq 0$ for all $x \in \mathbb{R}^d$, we obtain
$$\int_{\mathbb{R}^d} \phi(x)\, d\mu(x) \geq \int_A \phi(x)\, d\mu(x) > 0,$$
contradicting the fact that $\mu \in \tilde{S}^\Phi$.

□

Example 3.1.7.

1. In general $\tilde{S}^\Phi$ may fail to possess a classical counterpart: in $\mathbb{R}$, take $\Phi = \{\phi\}$ where $\phi : \mathbb{R} \to \mathbb{R}$, $\phi(x) := |x+1| - 1$ (notice that ϕ is bounded from below). Then if $\tilde{S}^\Phi$ or $\tilde{S}^\Phi_p$ admitted a classical counterpart S, we should have $S \subseteq [-2, 0]$ by item (4) of the Proposition above. Define $\mu_0 := \frac{1}{2}(\delta_{-1} + \delta_1)$. Thus we have $\mu_0 \in \tilde{S}^\Phi_p$, in fact $\int_{\mathbb{R}} \phi(x)\, d\mu_0(x) = 0$, but $\operatorname{supp} \mu_0 = \{-1, 1\} \nsubseteq S$ for any possible S. So neither $\tilde{S}^\Phi$ nor $\tilde{S}^\Phi_p$ admit a classical counterpart.

2. The converse of item (7) of Proposition 3.1.6 is not true: in $\mathbb{R}$, take $\Phi = \{\phi_1, \phi_2, \phi_3\}$ where $\phi_i : \mathbb{R} \to \mathbb{R}$, $i = 1, 2, 3$ are defined to be $\phi_1(x) = \max\{x, 0\}$, $\phi_2(x) = \min\{\max\{-x, -1\}, 0\}$, $\phi_3(x) = \max\{x, -1\}$. Then both $\tilde{S}^\Phi_p$ and $\tilde{S}^\Phi$ admits S as their classical counterpart, with $S =]-\infty, 0]$, but ϕ_3 can change its sign.

We are now ready to state some comparison results between the generalized distance and the classical one.

Proposition 3.1.8 (Comparison with classical distance). *Let $p \geq 1$, $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$, $\Phi \subseteq C^0(\mathbb{R}^d; \mathbb{R})$ satisfying (T_E) in Definition 3.1.1, and set*
$$C := \{x \in \mathbb{R}^d : \phi(x) \leq 0 \text{ for all } \phi \in \Phi\}.$$
Then

1. $\tilde{d}_{\tilde{S}^\Phi_p}(\mu_0) \leq \|d_C\|_{L^p_{\mu_0}}$,

2. *if there exists $\tilde{\phi}(\cdot) \in \Phi$ such that $\tilde{\phi}(x) \geq 0$ for all $x \in \mathbb{R}^d$, then $\tilde{d}_{\tilde{S}^\Phi_p}(\mu_0) \geq \|d_D\|_{L^p_{\mu_0}}$, where*
$$D := \{x \in \mathbb{R}^d : \tilde{\phi}(x) = 0\}.$$

3. *if* $\tilde{S}^{\Phi}_p$ *admits a classical counterpart* S*, then* $C = S$ *and* $\tilde{d}_{\tilde{S}^{\Phi}_p}(\mu_0) = \|d_S\|_{L^p_{\mu_0}}$*, moreover* $\tilde{d}^p_{\tilde{S}^{\Phi}_p} : \mathscr{P}_p(\mathbb{R}^d) \to [0,+\infty[$ *is convex.*

Proof. Clearly, according to assumption (T_E) in Definition 3.1.1 we have $C \neq \emptyset$.

1. If $\|d_C\|_{L^p(\mu_0)} = +\infty$ then there is nothing to prove. So let us assume that $\|d_C\|_{L^p(\mu_0)} < +\infty$.

   Define the multifunction

$$G(x) := \{y \in \mathbb{R}^d : |x-y| = d_C(x)\} \cap C = \partial B(x, d_C(x)) \cap C.$$

   Since the map $f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ defined by setting $f(x,y) := |x-y| - d_C(x)$ is continuous, we have that $G(\cdot)$ has closed graph in $\mathbb{R}^d \times \mathbb{R}^d$, and in particular $G(\cdot)$ is measurable. According to Theorem 8.1.3 in [13], there exists a Borel map $g : \mathbb{R}^d \to C$ such that $|x - g(x)| = d_C(x)$ for all $x \in \mathbb{R}^d$ (that is $g(x) \in G(x)$ for all $x \in \mathbb{R}^d$).

   We define $\nu_0 := g\sharp\mu_0$ and prove now that $\nu_0 \in \tilde{S}^{\Phi}_p$. Indeed, since $g(x) \in C$ for all $x \in \mathbb{R}^d$, we have

$$\int_{\mathbb{R}^d} \phi(x)\, dg\sharp\mu_0(x) = \int_{\mathbb{R}^d} \phi(g(x))\, d\mu_0(x) \leq 0, \text{ for all } \phi(\cdot) \in \Phi,$$

   whence $\nu_0 \in \tilde{S}^{\Phi}$.

   It remains to prove that the p-moment of ν_0 is finite. Owing to

$$\begin{aligned}\left(\int_{\mathbb{R}^d} |x|^p\, d\nu_0\right)^{1/p} &= \left(\int_{\mathbb{R}^d} |g(x)|^p\, d\mu_0\right)^{1/p} \\ &= \|g\|_{L^p(\mu_0)} \leq \|g - \mathrm{Id}\|_{L^p(\mu_0)} + \|\mathrm{Id}\|_{L^p(\mu_0)},\end{aligned}$$

   we have to prove that the sum in the right hand side is finite. But $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$ implies $\|\mathrm{Id}\|_{L^p(\mu_0)} < +\infty$ and $|g(x) - x| = d_C(x)$ holds by construction, so that $\|g - \mathrm{Id}\|_{L^p(\mu_0)} = \|d_C\|_{L^p(\mu_0)} < +\infty$. Therefore, we conclude $\nu_0 \in \tilde{S}^{\Phi}_p$ and we have

$$\begin{aligned}\tilde{d}_{\tilde{S}^{\Phi}_p}(\mu_0) &\leq W_p(\mu_0, \nu_0) \leq \left(\iint_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^p\, d(\mathrm{Id} \times g)\sharp\mu_0\right)^{1/p} \\ &= \left(\int_{\mathbb{R}^d} |x - g(x)|^p\, d\mu_0\right)^{1/p} = \left(\int_{\mathbb{R}^d} d^p_C(x)\, d\mu_0\right)^{1/p},\end{aligned}$$

   as desired.

2. Let us now assume that there exists $\tilde{\phi}(\cdot) \in \Phi$ such that $\tilde{\phi}(x) \geq 0$ for all $x \in \mathbb{R}^d$ and prove that $\tilde{d}_{\tilde{S}^{\Phi}_p}(\mu_0) \geq \|d_D\|_{L^p_{\mu_0}}$. Let $\{\varphi_n\}_{n\in\mathbb{N}} \subset C^0_C(\mathbb{R}^d; [0,1])$ be such that

$$\varphi_n(x) = \begin{cases} 1, & \text{if } x \in \overline{B(0,n)}, \\ 0, & \text{if } x \notin B(0, n+1). \end{cases}$$

Set $\psi_2^n(y) = \varphi_n(y)\tilde{\phi}(y)$ and $\psi_1^n(x) = \varphi_n(x)d_D^p(x)$, hence we have $\psi_1^n, \psi_2^n \in C_b^0(\mathbb{R}^d)$. Given $\theta \in \tilde{S}_p^\Phi$, we notice that for θ-a.e. $y \in \mathbb{R}^d$ we must have $\tilde{\phi}(y) = 0$, and so $y \in D$ thus for θ-a.e. $y \in \mathbb{R}^d$ and μ_0-a.e. $x \in \mathbb{R}^d$ it holds

$$\psi_1^n(x) + \psi_2^n(y) = \varphi_n(x)d_D^p(x) \leq d_D^p(x) \leq |x-y|^p.$$

So, according to Kantorovich duality (1.3), we have

$$\begin{aligned} W_p^p(\mu_0,\theta) &= \sup_{\substack{\psi_1,\psi_2 \in C_b^0(\mathbb{R}^d) \\ \psi_1(x)+\psi_2(y)\leq|x-y|^p}} \left\{ \int_{\mathbb{R}^d} \psi_1(x)\,d\mu_0(x) + \int_{\mathbb{R}^d} \psi_2(y)\,d\theta(y) \right\} \\ &\geq \int_{\mathbb{R}^d} \varphi_n(x)d_D^p(x)\,d\mu_0(x), \end{aligned}$$

Since $\{\psi_1^n(\cdot)\}_{n\in\mathbb{N}} \subseteq C_b^0(\mathbb{R}^d)$ is an increasing sequence of nonnegative functions pointwise convergent to $d_D^p(\cdot)$, by letting $n \to +\infty$ and applying the Monotone Convergence Theorem we obtain

$$W_p^p(\mu_0,\theta) \geq \int_{\mathbb{R}^d} d_D^p(x)\,d\mu_0(x),$$

for all $\theta \in \tilde{S}_p^\Phi$.

3. The equality $C = S$ is trivial: from item (4) in Proposition 3.1.6 we have $S \subseteq C$, moreover if μ is a measure supported in C we have that $\mu \in \tilde{S}_p^\Phi$, since all the functions of Φ are nonpositive on C, thus $C \subseteq S$, and so equality holds. By item (1) above we have already $\tilde{d}_{\tilde{S}_p^\Phi}(\mu_0) \leq \|d_S\|_{L^p_{\mu_0}}$. By item (6) in Proposition 3.1.6, we have $\tilde{S}_p^\Phi = \tilde{S}_p^{\{d_C\}}$, hence by applying item (2) above with $D = C = S$ and $\tilde{\phi} = d_S$ we obtain $\tilde{d}_{\tilde{S}_p^\Phi}(\mu_0) \geq \|d_S\|_{L^p_{\mu_0}}$, thus equality holds. Finally, the last statement is trivial, and it follows from the fact that
$$\tilde{d}^p_{\tilde{S}_p^\Phi}(\mu) = \int_{\mathbb{R}^d} d_C^p(x)\,d\mu,$$
is linear in μ.

□

Without the assumption of existence of a classical counterpart for $\tilde{S}_p^\Phi$, the inequality $\tilde{d}_{\tilde{S}_p^\Phi}(\mu_0) \leq \|d_C\|_{L^p_{\mu_0}}$ may be strict.

Example 3.1.9. In $\mathbb{R}$, take $\Phi = \{\phi_1, \phi_2, \phi_3\}$ where $\phi_i : \mathbb{R} \to \mathbb{R}$ are defined by

$$\phi_1(x) = |x-1| - 1, \qquad \phi_2(x) = |x+1| - 1, \qquad \phi_3(x) = |x(x^2-1)|.$$

Define also $\mu_0 = \dfrac{1}{2}(\delta_{-1} + \delta_1)$. For any $x \in \mathbb{R}$, we have $\phi_i(x) \geq -1$ for $i = 1, 2$ and $\phi_3(x) \geq 0$ (thus ϕ is uniformly bounded from below for $i = 1, 2, 3$), moreover

$$\begin{aligned} C :=& \{x \in \mathbb{R} : \phi_i \leq 0, \text{ for } i = 1,2,3\} \\ =& \{x \in \mathbb{R} : \phi_i = 0, \text{ for } i = 1,2,3\} = \{0\}, \\ & \int_{\mathbb{R}} \phi_i(x)\,d\mu_0(x) = 0, \qquad\qquad i = 1,2,3, \end{aligned}$$

hence, $\mu_0 \in \tilde{S}^{\Phi}_p$ for all $p \geq 1$, thus $\tilde{d}_{\tilde{S}^{\Phi}_p}(\mu_0) = 0$. However, since $d^p_C(x) = |x|^p$, we have

$$\int_{\mathbb{R}} d^p_C(x) d\mu_0(x) = 1 > 0.$$

We notice that $\tilde{S}^{\Phi}_p$ does not admit a classical counterpart: indeed if a classical counterpart would exist, it would be reduced to $C = \{0\}$, however $\mu_0 \in \tilde{S}^{\Phi}_p \subseteq \tilde{S}^{\Phi}$ and $\operatorname{supp} \mu_0 \not\subseteq C$, thus no classical counterpart may exist.

Without the p-th power, the generalized distance in the case of the Proposition 3.1.8 above may fail to be convex.

Example 3.1.10. Let $p > 1$. In $\mathbb{R}^2$, consider $P = (0,0)$, $Q_1 = (1,0)$, $Q_2 = \big(0, 2^{1/p}\big)$. Set $S = \{P\}$, $\Phi = \{d_S(\cdot)\}$, hence $\tilde{S}^{\Phi}_p := \{\delta_P\}$, and define $\nu_\lambda = \lambda\delta_{Q_1} + (1-\lambda)\delta_{Q_2}$, $\lambda \in [0,1]$. By Proposition 3.1.8, we have

$$\tilde{d}^p_{\tilde{S}^{\Phi}_p}(\nu_\lambda) = W^p_p(\delta_P, \nu_\lambda) = \lambda + 2(1-\lambda) = 2 - \lambda,$$

whence $\tilde{d}_{\tilde{S}^{\Phi}_p}(\nu_\lambda) = \sqrt[p]{2-\lambda}$, which is not convex.

In the metric space $\mathscr{P}_p(\mathbb{R}^d)$ endowed with W_p-distance, another concept of convexity can be given, related more to the metric structure rather than to the linear one.

Given any product space X^N ($N \geq 1$), in the following we denote with $\mathrm{pr}^i \colon X^N \to X$ the projection on the i–th component, i.e., $\mathrm{pr}^i(x_1, \dots, x_N) = x_i$.

Definition 3.1.11 (Geodesics)**.** Given a curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$, we say that it is a *(constant speed) geodesic* if for all $0 \leq s \leq t \leq 1$ we have

$$W_p(\mu_s, \mu_t) = (t-s) W_p(\mu_0, \mu_1).$$

In this case, we will also say that the curve $\boldsymbol{\mu}$ is a *geodesic connecting* μ_0 *and* μ_1.

Theorem 3.1.12 (Characterization of geodesics)**.** *Let* $\mu_0, \mu_1 \in \mathscr{P}_p(\mathbb{R}^d)$ *and let* $\pi \in \Pi^p_o(\mu_0, \mu_1)$ *be an optimal transport plan between* μ_0 *and* μ_1*, i.e.*

$$W^p_p(\mu_0, \mu_1) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x_1 - x_2|^p \, d\pi(x_1, x_2)\,.$$

Then the curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]}$ *defined by*

$$\mu_t := \big((1-t)\,\mathrm{pr}^1 + t\,\mathrm{pr}^2\big)\sharp\pi \;\in\; \mathscr{P}_p(\mathbb{R}^d) \tag{3.1}$$

is a (constant speed) geodesic connecting μ_0 *and* μ_1*.*

Conversely, any (constant speed) geodesic $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]}$ *connecting* μ_0 *and* μ_1 *admits the representation* (3.1) *for a suitable plan* $\pi \in \Pi^p_o(\mu_0, \mu_1)$*.*

Proof. See Theorem 7.2.2 in [9]. □

Definition 3.1.13 (Geodesically and strongly geodesically convex sets)**.** A subset $A \subseteq \mathscr{P}_p(\mathbb{R}^d)$ is said to be

1. *geodesically convex* if for every pair of measures μ_0, μ_1 in A, there exists a geodesic connecting μ_0 and μ_1 which is contained in A.

2. *strongly geodesically convex* if for every pair of measures μ_0, μ_1 in A and for every admissible transport plan $\pi \in \Pi(\mu_0, \mu_1)$, the curve $t \mapsto \mu_t$ defined by (3.1) is contained in A.

The interest in this alternative concept of convexity comes from the fact that, in many problems, functionals defined on probability measures are convex along geodesics (a notion related to geodesically convex sets) and not convex with respect to the linear structure in the usual sense. We refer to Section 9.1 in [9] for further details.

Remark 3.1.14. Notice that, even if the notations does not highlight this fact, the notions of *geodesic* and *geodesical convexity* depend on the exponent p which has been fixed.

Proposition 3.1.15 (Strong geodesic convexity of $\tilde{S}^{\Phi}_p$)**.** *Let $p \geq 1$, Φ satisfying (T_E) in Definition 3.1.1. Assume that all the elements of Φ are continuous and convex. Then the generalized target $\tilde{S}^{\Phi}_p$ is strongly geodesically convex.*

Proof. Let $\mu_0, \mu_1 \in \tilde{S}^{\Phi}_p$ and let $\pi \in \Pi(\mu_0, \mu_1)$ be an admissible transport plan between μ_0 and μ_1. Consider the corresponding curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]}$ defined by (3.1), and fix $t \in [0,1]$. We have for every $\phi(\cdot) \in \Phi$

$$\begin{aligned}
&\int_{\mathbb{R}^d} \phi(x)\, d\mu_t(x) \leq \\
&\quad \leq (1-t) \iint_{\mathbb{R}^d\times\mathbb{R}^d} \phi\left(\mathrm{pr}^1(\xi,\eta)\right)\, d\pi(\xi,\eta) + t \iint_{\mathbb{R}^d\times\mathbb{R}^d} \phi\left(\mathrm{pr}^2(\xi,\eta)\right)\, d\pi(\xi,\eta) \\
&\quad = (1-t) \int_{\mathbb{R}^d} \phi(x)\, d\mu_0(x) + t \int_{\mathbb{R}^d} \phi(y)\, d\mu_1(y) \leq 0\,,
\end{aligned}$$

since $\mathrm{pr}^i \sharp \pi$ are the marginal measures of π, which belong to $\tilde{S}^{\Phi}_p$. The conclusion follows from the arbitrariness of $\phi(\cdot) \in \Phi$. □

Remark 3.1.16. In particular, the above result holds for $\Phi := \{d_S(\cdot) - \alpha\}$ when S is nonempty, closed and convex, and $\alpha \in [0,1]$. In this case, since in the above proof we use only the convexity property of $d_S(\cdot)$, the statement holds also if we equip $\mathbb{R}^d$ with a different norm than the Euclidean one.

We conclude this section by investigating the semiconcavity properties of the generalized distance along geodesics. The case $p = 2$ is particularly easy thanks to the geometric structure of the metric space $\mathscr{P}_2(\mathbb{R}^d)$.

Proposition 3.1.17 (Semiconcavity of $\tilde{d}^2_{\tilde{S}^{\Phi}_2}$)**.** *Let $\tilde{S}^{\Phi}_2$ be the generalized target in $\mathscr{P}_2(\mathbb{R}^d)$ corresponding to $\Phi \subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1.1. Then the square of the generalized distance satisfies the following* global semiconcavity inequality *for every $\mu_0, \mu_1 \in \mathscr{P}_2(\mathbb{R}^d)$ and every $t \in [0,1]$*

$$\tilde{d}^2_{\tilde{S}^{\Phi}_2}(\mu_t) \geq (1-t)\, \tilde{d}^2_{\tilde{S}^{\Phi}_2}(\mu_0) + t\, \tilde{d}^2_{\tilde{S}^{\Phi}_2}(\mu_1) - t(1-t)\, W_2^2(\mu_0,\mu_1),$$

where $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]}$ is any constant speed geodesic for W_2 joining μ_0 and μ_1.

Proof. Owing to Theorem 7.3.2 in [9], we have that for any measure $\sigma \in \mathscr{P}_2(\mathbb{R}^d)$ the function $\mu \mapsto W_2^2(\mu,\sigma)$ is semiconcave along geodesics, with semiconcavity constant independent by σ, i.e. it satisfies for every $t \in [0,1]$

$$W_2^2(\mu_t,\sigma) + t(1-t)\, W_2^2(\mu_0,\mu_1) \geq (1-t)\, W_2^2(\mu_0,\sigma) + t\, W_2^2(\mu_1,\sigma).$$

By passing to the infimum on $\sigma \in \tilde{S}_2^\Phi$, we have

$$\tilde{d}^2_{\tilde{S}_2^\Phi}(\mu_t) + t(1-t)\, W_2^2(\mu_0,\mu_1) \geq (1-t)\, \tilde{d}^2_{\tilde{S}_2^\Phi}(\mu_0) + t\, \tilde{d}^2_{\tilde{S}_2^\Phi}(\mu_1)\,,$$

whence the conclusion follows. □

In the case $p \neq 2$ we need additional requirements on Φ. We start with a technical lemma.

Lemma 3.1.18. *Given $p \geq 1$, define the map $h_p : \mathbb{R} \to \mathbb{R}$ by setting $h_p(\xi) := \operatorname{sign}(\xi)\,|\xi|^p$. Then*

1. *$h_p \in C^1(\mathbb{R})$ is increasing, and $h_p'(\xi) = p|h_{p-1}(\xi)| \geq 0$,*
2. *for every $\xi_0, \xi_1 \in \mathbb{R}$ we have*

$$|h_p(\xi_1) - h_p(\xi_0)| \leq p\, \max\{|\xi_0|, |\xi_1|\}^{p-1}\, |\xi_1 - \xi_0|$$

3. *for every $\xi_0, \xi_1 \in \mathbb{R}$, $t \in [0,1]$ we have that the quantity*

$$(1-t)\, h_p(\xi_0) + t\, h_p(\xi_1) - h_p\big((1-t)\xi_0 + t\xi_1\big)$$

is bounded above by

$$t\,(1-t)\, p\,(p-1)\, \max\{|\xi_0|, |\xi_1|\}^{\max\{p,2\}-2}\, |\xi_0 - \xi_1|^{\min\{p,2\}}.$$

Proof. The proof of (1) is trivial. Property (2) follows from the equality

$$|h_p(\xi_1) - h_p(\xi_0)| = |h_p'(\xi)(\xi_1 - \xi_0)| = p|\xi|^{p-1}|\xi_1 - \xi_0|,$$

for some ξ in the interval joining ξ_1 and ξ_0, and from the monotonicity of $s \mapsto s^{p-1}$ on $\mathbb{R}^+$.
To prove (3), we adapt the argument of Proposition 2.1.2 in [22]. By regularity of h_p we have for all $t \in [0,1]$

$$\begin{aligned}
&(1-t)\, h_p(\xi_0) + t\, h_p(\xi_1) - h_p((1-t)\xi_0 + t\xi_1) = \\
&\quad = (1-t)\, [h_p(\xi_0) - h_p(\xi_0 + t(\xi_1 - \xi_0))] + t\, [h_p(\xi_1) - h_p(\xi_1 + (1-t)(\xi_0 - \xi_1))] \\
&\quad = t\,(1-t)\, \big(h_p'(\eta_0) - h_p'(\eta_1)\big)(\xi_1 - \xi_0) \leq t\,(1-t)\, \big|h_p'(\eta_0) - h_p'(\eta_1)\big|\, |\xi_0 - \xi_1| \\
&\quad \leq pt\,(1-t)\, \big||\eta_0|^{p-1} - |\eta_1|^{p-1}\big|\, |\xi_0 - \xi_1|\,,
\end{aligned}$$

where η_0, η_1 are suitable points in the interval joining ξ_0 and ξ_1. In particular, they satisfy also $|\eta_0 - \eta_1| \leq |\xi_0 - \xi_1|$ and $\max\{|\eta_0|, |\eta_1|\} \leq \max\{|\xi_0|, |\xi_1|\}$. Now we distinguish two cases.

a. For $p \geq 2$ we have that $s \mapsto s^{p-1}$ is convex on $\mathbb{R}^+$ (thus its derivative is monotone increasing), hence by combining (1) and (2) we have

$$\begin{aligned}
p\, \big||\eta_0|^{p-1} - |\eta_1|^{p-1}\big| &\leq p\,(p-1)\, \max\{|\eta_0|, |\eta_1|\}^{p-2}\, |\eta_0 - \eta_1| \\
&\leq p\,(p-1)\, \max\{|\xi_0|, |\xi_1|\}^{p-2}\, |\xi_0 - \xi_1|\,.
\end{aligned}$$

b. For $1 \leq p < 2$, we have that

$$p\,\left||\eta_0|^{p-1}-|\eta_1|^{p-1}\right| \leq p\,\frac{\left||\eta_0|^{p-1}-|\eta_1|^{p-1}\right|}{\left||\eta_0|-|\eta_1|\right|^{p-1}}\cdot|\eta_0-\eta_1|^{p-1}$$
$$= p\,\frac{1-\left(\dfrac{\min\{|\eta_0|,|\eta_1|\}}{\max\{|\eta_0|,|\eta_1|\}}\right)^{p-1}}{\left(1-\dfrac{\min\{|\eta_0|,|\eta_1|\}}{\max\{|\eta_0|,|\eta_1|\}}\right)^{p-1}}\cdot|\eta_0-\eta_1|^{p-1}.$$

Since for $t \in [0,1]$, the map $t \mapsto \dfrac{1-t^{p-1}}{(1-t)^{p-1}}$ has derivative that is less or equal than $\dfrac{p-1}{(1-t)^2}\left(1-\dfrac{t^p}{t^2}\right) \leq 0$, then it attains its maximum (over $[0,1]$) at $t=0$ and such maximum is equal to 1, so that

$$|h_p'(\eta_0)-h_p'(\eta_1)| \leq p\,|\eta_0-\eta_1|^{p-1} \leq p\,|\xi_0-\xi_1|^{p-1}.$$

Combining a. and b., the proof is concluded. □

Proposition 3.1.19 (Semiconcavity of $\tilde{d}^p_{\tilde{S}^\Phi_p}$)**.** *Let $p \geq 1$, and assume that $\Phi \subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfies (T_E) in Definition 3.1.1 and that $\tilde{S}^\Phi_p$ admits a classical counterpart $S \subseteq \mathbb{R}^d$. Let $K \subseteq \mathbb{R}^d \setminus S$ be compact and convex. Then the p-th power of the generalized distance $\tilde{d}_{\tilde{S}^\Phi_p}(\cdot)$ from the generalized target $\tilde{S}^\Phi_p$ corresponding to Φ, satisfies the following* local semiconcavity inequality*: there exists a constant $C = C(p,K) > 0$ such that for every $\mu_0, \mu_1 \in \mathscr{P}_p(K)$ we have*

$$\tilde{d}^p_{\tilde{S}^\Phi_p}(\mu_t) \geq (1-t)\,\tilde{d}^p_{\tilde{S}^\Phi_p}(\mu_0) + t\,\tilde{d}^p_{\tilde{S}^\Phi_p}(\mu_1) - Ct(1-t)\,W_p^{\min\{p,2\}}(\mu_0,\mu_1), \tag{3.2}$$

where $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]}$ is any constant speed geodesic for W_p joining μ_0 and μ_1.

Proof. In this proof to make clearer the notation we will omit the superscript Φ, since Φ is fixed. Under the above assumptions, and recalling Proposition 3.1.8, we have $\tilde{d}_{\tilde{S}_p}(\mu_0) = \|d_S\|_{L^p_{\mu_0}}$.

Given $x_0, x_1 \in \mathbb{R}^d$ and $t \in [0,1]$ we set

$$x_t := (1-t)x_0 + tx_1, \qquad d_t := (1-t)d_S(x_0) + td_S(x_1).$$

Let $D > 0$ such that $D^{-1} < d_S(y) \leq D$ for any $y \in K$ and denote with $M = \operatorname{diam}(K) := \max\limits_{z_1,z_2\in K} |z_1 - z_2|$.

According to Proposition 2.2.2 in [22], there exists $c = c(K) > 0$ such that d_S satisfies the following inequality for all $x_0, x_1 \in K$:

$$d_S(x_t) \geq d_t - ct(1-t)|x_0 - x_1|^2,$$

i.e., $d_S(\cdot)$ is semiconcave (with linear modulus) of constant c according to Definition 2.1.1 in [22]. Without loss of generality, we can assume $c > 1$ and $D > 1$.

Define $h_p(\cdot)$ as in Lemma 3.1.18. Given $x_0, x_1 \in K$ and $t \in [0,1]$, we have

$$\begin{aligned} d_S^p(x_t) &= h_p(d_S(x_t)) \geq h_p\left(d_t - ct(1-t)|x_0 - x_1|^2\right) \\ &\geq h_p(d_t) - p\max\left\{d_t, \left|d_t - ct(1-t)|x_0-x_1|^2\right|\right\}^{p-1} ct(1-t)|x_0-x_1|^2 \\ &\geq h_p(d_t) - c_1 t(1-t)|x_0 - x_1|^{\min\{p,2\}}, \end{aligned}$$

where $c_1 = c_1(p,K) := cp\left(D + cM^2\right)^{p-1} M^{\max\{0,2-p\}}$ and we have used Lemma 3.1.18–(2). Relying on Lemma 3.1.18–(3), we also obtain

$$\begin{aligned} h_p(d_t) &\geq (1-t)\, h_p(d_S(x_0)) + t\, h_p(d_S(x_1)) \\ &\qquad - t(1-t)\, p(p-1)\, D^{\max\{p,2\}-2} |d_S(x_0) - d_S(x_1)|^{\min\{p,2\}} \\ &\geq (1-t)\, d_S^p(x_0) + t\, d_S^p(x_1) - c_2\, t(1-t)\, |x_0 - x_1|^{\min\{p,2\}}, \end{aligned}$$

where $c_2 = c_2(p,K) := p(p-1)\, D^{\max\{p,2\}-2}$ and we used the 1–Lipschitz continuity of d_S. Combining the estimates above, we finally conclude that

$$d_S^p(x_t) \geq (1-t)\, d_S^p(x_0) + t\, d_S^p(x_1) - C' t(1-t)|x_0 - x_1|^{\min\{p,2\}}, \tag{3.3}$$

with $C' = C'(p,K) := c_1 + c_2$.

For any Borel sets $A, B \subseteq \mathbb{R}^d$ and $\pi \in \Pi(\mu_0, \mu_1)$, we now have

$$A \times B \subseteq [(A \times B) \cap (K \times K)] \cup [(A \setminus K) \times \mathbb{R}^d] \cup [\mathbb{R}^d \times (B \setminus K)],$$

so that

$$\begin{aligned} \pi(A \times B) &\leq \pi((A \times B) \cap (K \times K)) + \mu_0(A \setminus K) + \mu_1(B \setminus K) \\ &= \pi((A \times B) \cap (K \times K)), \end{aligned}$$

because μ_0 and μ_1 are concentrated on K. In particular, $\mathrm{supp}(\pi) \subseteq K \times K$. Therefore, we choose a transport plan $\pi \in \Pi_o^p(\mu_0, \mu_1)$ realizing the p-Wasserstein distance between μ_0 and μ_1, so that the representation in formula (3.1) holds, and we integrate the estimate (3.3) to find that

$$\begin{aligned} \int_{\mathbb{R}^d} d_S^p(x)\, d\mu_t = \iint_{\mathbb{R}^d \times \mathbb{R}^d} d_S^p(x_t)\, d\pi &\geq (1-t) \int_{\mathbb{R}^d} d_S^p(x)\, d\mu_0 + t \int_{\mathbb{R}^d} d_S^p(x)\, d\mu_1 \\ &\qquad - C'\, t\,(1-t) \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x_0 - x_1|^{\min\{p,2\}}\, d\pi, \end{aligned}$$

where $\boldsymbol{\mu} = \{\mu_t\}_{t \in [0,1]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ is the constant speed geodesic corresponding to π. But according to Proposition 3.1.8, there holds

$$\tilde{d}^p_{\tilde{S}_p}(\mu_t) = \int_{\mathbb{R}^d} d_S^p(x)\, d\mu_t(x), \qquad \text{and} \qquad \tilde{d}^p_{\tilde{S}_p}(\mu_i) = \int_{\mathbb{R}^d} d_S^p(x)\, d\mu_i(x), \quad i = 0, 1,$$

and applying Jensen's inequality to the concave map $\xi \mapsto \xi^{\gamma/p}$ on $\mathbb{R}^+$, with $\gamma = \min\{p, 2\}$, we obtain that

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} |x_0 - x_1|^{\min\{p,2\}}\, d\pi \leq \begin{cases} \displaystyle\iint_{\mathbb{R}^d \times \mathbb{R}^d} |x_0 - x_1|^p\, d\pi, & \text{for } 1 \leq p < 2, \\[2ex] \displaystyle\left(\iint_{\mathbb{R}^d \times \mathbb{R}^d} |x_0 - x_1|^p\, d\pi\right)^{2/p}, & \text{for } p \geq 2. \end{cases}$$

We thus conclude that

$$\tilde{d}^p_{\tilde{S}_p}(\mu_t) \geq (1-t)\,\tilde{d}^p_{\tilde{S}_p}(\mu_0) + t\,\tilde{d}^p_{\tilde{S}_p}(\mu_1) - C'\,t\,(1-t)\,W_p^{\min\{p,2\}}(\mu_0,\mu_1),$$

and the proof is completed. □

Remark 3.1.20. Notice that inequality (3.2) implies that, for $p \geq 2$ and under the assumption of Proposition 3.1.19, the functional $-\tilde{d}^p_{\tilde{S}_p}(\cdot)\colon \mathscr{P}_p(K) \to]-\infty, 0]$ is λ-geodesically convex, in the sense of Definition 9.1.1 in [9], with $\lambda = -2C'$.

3.2 Generalized minimum time problem

In this section we define a suitable notion of minimum time function, modeled on the finite-dimensional case.

Definition 3.2.1 (Admissible curves)**.** Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be a set-valued function, $I = [a,b]$ a compact interval of $\mathbb{R}$, $\alpha, \beta \in \mathscr{P}(\mathbb{R}^d)$. We say that a Borel family of probability measures $\boldsymbol{\mu} = \{\mu_t\}_{t\in I}$ is an *admissible trajectory (curve) defined in I for the system Σ_F joining α and β*, if there exists a family of Borel vector-valued measures $\boldsymbol{\nu} = \{\nu_t\}_{t\in I} \subseteq \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ such that

1. $\boldsymbol{\mu}$ is a narrowly continuous solution in the distributional sense of
$$\partial_t \mu_t + \mathrm{div}\nu_t = 0,$$
with $\mu_{|t=a} = \alpha$ and $\mu_{|t=b} = \beta$.

2. $J_F(\boldsymbol{\mu},\boldsymbol{\nu}) < +\infty$, where $J_F(\cdot,\cdot)$ is defined as
$$J_F(\boldsymbol{\mu},\boldsymbol{\nu}) := \begin{cases} \displaystyle\int_I \int_{\mathbb{R}^d} \left(1 + I_{F(x)}\left(\frac{\nu_t}{\mu_t}(x)\right)\right) d\mu_t(x)\,dt, & \text{if } |\nu_t| \ll \mu_t \text{ for a.e. } t \in I,\\ +\infty, & \text{otherwise.}\end{cases} \tag{3.4}$$
where $I_{F(x)}$ is the indicator function of the set $F(x)$, i.e., $I_{F(x)}(\xi) = 0$ for all $\xi \in F(x)$ and $I_{F(x)}(\xi) = +\infty$ for all $\xi \notin F(x)$.

In this case, we will also shortly say that $\boldsymbol{\mu}$ is *driven* by $\boldsymbol{\nu}$.

Remark 3.2.2. The finiteness of $J(\boldsymbol{\mu},\boldsymbol{\nu})$ forces the elements of $\boldsymbol{\nu}$ to have the form $\nu_t = v_t\mu_t$ for a vector field $v_t \in L^1_{\mu_t}$ for a.e. $t \in I$, and moreover we have $v_t(x) \in F(x)$ for μ_t–a.e. $x \in \mathbb{R}^d$ and a.e. $t \in I$. When $J_F(\cdot,\cdot)$ is finite, this value expresses the time needed by the system Σ_F to steer α to β along the trajectory $\boldsymbol{\mu}$ with family of velocity vector fields $v = \{v_t\}_{t\in I}$.

In view of the superposition principle stated at Theorem 1.3.3, we can give the following alternative equivalent definition.

Definition 3.2.3 (Admissible curves (alternative definition))**.** Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be a set-valued function, $I = [a,b]$ a compact interval of $\mathbb{R}$, $\alpha, \beta \in \mathscr{P}(\mathbb{R}^d)$. We say that a Borel family of probability measures $\boldsymbol{\mu} = \{\mu_t\}_{t\in I}$ is an *admissible trajectory (curve) defined in I for the system Σ_F joining α and β*, if there exist a probability measure $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_I)$ and a Borel vector field $v : I \times \mathbb{R}^d \to \mathbb{R}^d$ such that:

1. $\boldsymbol{\eta}$ is concentrated on the pairs (x, γ) such that γ is an absolutely continuous solution of $\dot{x}(t) = v_t(x(t))$ with initial condition $\gamma(a) = x$;
2. for every $\varphi \in C_b^0(\mathbb{R}^d)$, $t \in I$ we have
$$\int_{\mathbb{R}^d} \varphi(x)\, d\mu_t(x) = \iint_{\mathbb{R}^d \times \Gamma_I} \varphi(\gamma(t))\, d\boldsymbol{\eta}(x, \gamma),$$
3. $\gamma(a)\sharp\boldsymbol{\eta} = \alpha$, $\gamma(b)\sharp\boldsymbol{\eta} = \beta$,
4. $v_t(x) \in F(x)$ for μ_t-a.e. $x \in \mathbb{R}^d$ and a.e. $t \in I$ and $v_t \in L^1_{\mu_t}$ for a.e. $t \in I$.

In this case, we can define $\nu_t = v_t\mu_t$ thus we have simply $J_F(\boldsymbol{\mu}, \boldsymbol{\nu}) = b - a$.

In the following, we will mainly focus our attention on admissible curves defined in $[0, T]$, for some suitable $T > 0$. We recall Definition 1.0.6 and introduce the following notation.

Definition 3.2.4. Given $T \in [0, +\infty[$, we set

$$\begin{aligned} \mathscr{T}_F(\mu_0) := \{\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T) : T > 0, \boldsymbol{\eta} \text{ concentrated on trajectories of} \\ \dot{\gamma}(t) \in F(\gamma(t)) \text{ and satisfies } \gamma(0)\sharp\boldsymbol{\eta} = \mu_0\}, \end{aligned}$$

where $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$.

Remark 3.2.5. By the Superposition Principle (Theorem 1.3.3), given $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ satisfying (F_1), a Borel family of probability measures $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ is an *admissible trajectory* if and only if there exists $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$ such that $\mu_t = e_t\sharp\boldsymbol{\eta}$ for all $t \in [0, T]$, i.e., $\boldsymbol{\eta} = \mu_0 \otimes \eta_x$ where for μ_0-a.e. $x \in \mathbb{R}^d$ we have that $\eta_x \in \mathscr{P}(\Gamma_T^x)$ is concentrated on the solutions of $\dot{x}(t) \in F(x(t))$, $x(0) = x$.

In this case, we will shortly say that the admissible trajectory $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ *is represented by* $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$.

The following lemma states that under some regularity hypothesis for the multifunction F, it is possible to construct a regularization of an admissible (mass-preserving) curve with the property to be driven by a smooth velocity field which is closed to be admissible.

Lemma 3.2.6 (Approximation with almost-admissible smooth curves)**.** *Assume hypothesis (F_0) and (F_2). Let $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ be an admissible (mass-preserving) trajectory driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$. Consider a family of mollifiers $\{\rho_\varepsilon\}_{\varepsilon\geq 0} \subseteq C_C^\infty(\mathbb{R}^d)$ in the x-variable with* $\operatorname{supp}\rho_\varepsilon \subseteq \overline{B(0,\varepsilon)}$*, and set*

$$\mu_t^\varepsilon = \mu_t * \rho_\varepsilon, \qquad \nu_t^\varepsilon = \nu_t * \rho_\varepsilon, \qquad \textit{for } t \in [0, T].$$

Then for all $\delta > 0$ there exists $\bar{\varepsilon} = \bar{\varepsilon}_\delta > 0$ such that for all $0 < \varepsilon < \bar{\varepsilon}$ we have that $\boldsymbol{\mu}^\varepsilon = \{\mu_t^\varepsilon\}_{t\in[0,T]}$ is a mass-preserving trajectory driven by $\boldsymbol{\nu}^\varepsilon = \{\nu_t^\varepsilon\}_{t\in[0,T]}$ satisfying $\dfrac{\nu_t^\varepsilon}{\mu_t^\varepsilon}(x) \in F(x) + \delta\overline{B(0,1)}$ for a.e. $t > 0$ and μ_t^ε-a.e. $x \in \mathbb{R}^d$.

Proof. Fix $\delta > 0$. Clearly the equation $\partial_t\mu_t^\varepsilon + \operatorname{div}\nu_t^\varepsilon = 0$ is satisfied in the sense of distributions for all $\varepsilon > 0$, and so we have only to check that there exists

$\bar\varepsilon = \bar\varepsilon_\delta > 0$ such that for all $0 < \varepsilon < \bar\varepsilon$ we have $\dfrac{\nu_t^\varepsilon}{\mu_t^\varepsilon}(x) \in F(x) + \delta\overline{B(0,1)}$ for a.e. $t > 0$ and μ_t^ε-a.e. $x \in \mathbb{R}^d$.

To this aim, in the spirit of Lemma 8.1.10 in [9], we prove the following claim: let $\rho \in C^\infty(\mathbb{R}^d)$ be any convolution kernel, and let $\mu \in \mathscr{P}(\mathbb{R}^d)$, $\nu \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ with $\nu \ll \mu$, then

$$\int_{\{x\in\mathbb{R}^d:\,\mu*\rho(x)\neq 0\}} I_{F(x)+\delta\overline{B(0,1)}}\left(\frac{\nu*\rho}{\mu*\rho}(x)\right)\mu*\rho(x)\,dx \leq$$
$$\leq \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} I_{F(x)+\delta\overline{B(0,1)}}\left(\frac{\nu}{\mu}(y)\right)\rho(x-y)\,d\mu(y)\,dx.$$

Indeed, define the map $\Phi : \mathbb{R}^{d+1} \to [0,+\infty]$

$$\Phi(z,t) = \begin{cases} I_{F(x)+\delta\overline{B(0,1)}}\left(\dfrac{z}{t}\right)t, & \text{if } t > 0,\\[2ex] 0, & \text{if } (z,t) = (0,0),\\[2ex] +\infty, & \text{if either } t < 0 \text{ or } t = 0 \text{ and } z \neq 0.\end{cases}.$$

We notice that $\Phi(\cdot)$ is convex, l.s.c., nonnegative, and 1-positively homogeneous, indeed we have

$$\Phi(z,t) = \sup_{\xi\in\mathbb{R}^d}\left\{\langle z,\xi\rangle - t\sigma_{F(x)+\delta\overline{B(0,1)}}(\xi)\right\} + I_{[0,+\infty[}(t).$$

By Jensen's inequality, for any Borel map $\psi : \mathbb{R}^d \to \mathbb{R}^{d+1}$ and any finite positive measure θ on $\mathbb{R}^d$, we have

$$\Phi\left(\int_{\mathbb{R}^d}\psi(y)\,d\theta(y)\right) \leq \int_{\mathbb{R}^d}\Phi(\psi(y))\,d\theta(y).$$

We fix $x \in \mathbb{R}^d$ such that $\mu * \rho(x) \neq 0$ and apply the above inequality by setting $\psi = \left(\dfrac{\nu}{\mu}, 1\right)$ and $\theta = \rho(x-\cdot)\mu$. We obtain

$$\begin{aligned}\Phi\left(\int_{\mathbb{R}^d}\psi(y)\,d\theta(y)\right) &= \Phi\left(\int_{\mathbb{R}^d}\frac{\nu}{\mu}(y)\rho(x-y)\,d\mu(y), \int_{\mathbb{R}^d}\rho(x-y)\mu(y)\right)\\ &= \Phi\left(\int_{\mathbb{R}^d}\rho(x-y)\,d\nu(y), \int_{\mathbb{R}^d}\rho(x-y)\mu(y)\right)\\ &= I_{F(x)+\delta\overline{B(0,1)}}\left(\frac{\nu*\rho}{\mu*\rho}(x)\right)\mu*\rho(x)\\ &\leq \int_{\mathbb{R}^d}\Phi\left(\frac{\nu}{\mu}(y),1\right)d\theta(y)\\ &= \int_{\mathbb{R}^d} I_{F(x)+\delta\overline{B(0,1)}}\left(\frac{\nu}{\mu}(y)\right)\rho(x-y)\,d\mu(y).\end{aligned}$$

Integrating w.r.t. x we have

$$\int_{\mathbb{R}^d} I_{F(x)+\delta\overline{B(0,1)}}\left(\frac{\nu*\rho}{\mu*\rho}(x)\right)\mu*\rho(x)\,dx \leq \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} I_{F(x)+\delta\overline{B(0,1)}}\left(\frac{\nu}{\mu}(y)\right)\rho(x-y)\,d\mu(y)\,dx,$$

as desired. Note that there exists $\bar{\varepsilon} = \bar{\varepsilon}_\delta > 0$ such that for all $0 < \varepsilon < \bar{\varepsilon}$ we have $\dfrac{\nu_t}{\mu_t}(y) \in F(y) \subseteq F(x) + \delta B(0,1)$ for all $y \in \overline{B(x,\varepsilon)}$ by uniform continuity of F, and furthermore $\operatorname{supp} \rho_\varepsilon(x - \cdot) \subseteq \overline{B(x,\varepsilon)}$. Thus, to conclude the proof, we just apply the claim to μ_t and ν_t with $\rho = \rho_\varepsilon$. $\square$

For later use we state the following technical lemma.

Lemma 3.2.7 (Basic estimates)**.** *Assume (F_0) and (F_1), and let C be the constant as in (F_1). Let $T > 0$, $p \geq 1$, $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$ and $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ be an admissible trajectory driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$ and represented by $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$. Then we have:*

(i) $|e_t(x,\gamma)| \leq (|e_0(x,\gamma)| + CT)\, e^{CT}$ *for all* $t \in [0,T]$ *and* $\boldsymbol{\eta}$*-a.e.* $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$*;*

(ii) $e_t \in L^p_{\boldsymbol{\eta}}(\mathbb{R}^d \times \Gamma_T; \mathbb{R}^d)$ *for all* $t \in [0,T]$*;*

(iii) there exists $D > 0$ depending only on C, T, p such that for all $t \in [0,T]$ we have
$$\left\| \frac{e_t - e_0}{t} \right\|^p_{L^p_{\boldsymbol{\eta}}} \leq D\,(\mathrm{m}_p(\mu_0) + 1)\,;$$

(iv) there exist $D', D'' > 0$ depending only on C, T, p such that for all $t \in [0,T]$ we have
$$\begin{aligned} \mathrm{m}_p(\mu_t) &\leq D'(\mathrm{m}_p(\mu_0) + 1), \\ \mathrm{m}_p(|\nu_t|) &\leq D''(\mathrm{m}_{p+1}(\mu_0) + 1). \end{aligned}$$

In particular, we have $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$.

Proof. We set $\varphi_t(x,\gamma) = \dfrac{e_t(x,\gamma) - e_0(x,\gamma)}{t}$, notice that for all $t \geq 0$ the map $(x,\gamma) \mapsto \varphi_t(x,\gamma)$ does not depend on x-variable.

Item (i) follows from Lemma 1.4.3. To prove (ii) it is enough to show $e_0 \in L^p_{\boldsymbol{\eta}}(\mathbb{R}^d \times \Gamma_T)$ and then apply item (i). Indeed, recalling that $(a+b)^p \leq 2^{p-1}(a^p + b^p)$ for any $a, b \geq 0$, we have
$$\begin{aligned} &\iint_{\mathbb{R}^d\times\Gamma_T} |e_0(x,\gamma)|^p \, d\boldsymbol{\eta}(x,\gamma) = \int_{\mathbb{R}^d} |z|^p \, d(\gamma(0)\sharp\boldsymbol{\eta})(z) = \mathrm{m}_p(\mu_0) < +\infty, \\ &\iint_{\mathbb{R}^d\times\Gamma_T} |e_t(x,\gamma)|^p \, d\boldsymbol{\eta}(x,\gamma) \leq \\ &\qquad \leq 2^{p-1} e^{CTp} \left(\iint_{\mathbb{R}^d\times\Gamma_T} |e_0(x,\gamma)|^p \, d\boldsymbol{\eta} + C^p T^p \right) \\ &\qquad \leq K\,(\mathrm{m}_p(\mu_0) + 1)\,, \end{aligned}$$

for a suitable constant $K > 0$ depending only on C, T, p.

We prove now (iii). For all $t \in]0,T[$ we have
$$\begin{aligned} |\varphi_t(x,\gamma)| &= \frac{1}{t}|\gamma(t) - \gamma(0)| = \frac{1}{t}\int_0^t |\dot{\gamma}(s)|\, ds \leq \frac{C}{t}\int_0^t |\gamma(s)|\, ds + C \\ &\leq C\,(|e_0(x,\gamma)| + CT)\; e^{CT} + C \leq \tilde{K}(|\gamma(0)| + 1), \end{aligned}$$

for a suitable $\tilde{K} > 0$ depending only on C, T.

Squaring and integrating w.r.t. $\boldsymbol{\eta}$ we get

$$\begin{aligned}\left\|\frac{e_t - e_0}{t}\right\|^p_{L^p_{\boldsymbol{\eta}}} &= \iint_{\mathbb{R}^d\times\Gamma_T} \left|\frac{e_t(x,\gamma) - e_0(x,\gamma)}{t}\right|^p d\boldsymbol{\eta}(x,\gamma)\\ &\le \iint_{\mathbb{R}^d\times\Gamma_T} \tilde{K}^p(|\gamma(0)|+1)^p\, d\boldsymbol{\eta}(x,\gamma)\\ &\le 2^{p-1}\tilde{K}^p\left[\iint_{\mathbb{R}^d\times\Gamma_T}|\gamma(0)|^p\, d\boldsymbol{\eta}(x,\gamma)+1\right]\\ &\le D(\mathrm{m}_p(\mu_0)+1).\end{aligned}$$

Since

$$\mathrm{m}_p(\mu_t) = \int_{\mathbb{R}^d}|x|^p\, d\mu_t = \iint_{\mathbb{R}^d\times\Gamma_T}|e_t(x,\gamma)|^p\, d\boldsymbol{\eta} = \|e_t\|^p_{L^p_{\boldsymbol{\eta}}},$$

from the above estimate we have also

$$\begin{aligned}\mathrm{m}_p(\mu_t) &\le \left[\mathrm{m}_p^{1/p}(\mu_0) + t\left(D(\mathrm{m}_p(\mu_0)+1)\right)^{1/p}\right]^p \le 2^{p-1}\left(\mathrm{m}_p(\mu_0) + t^p D(\mathrm{m}_p(\mu_0)+1)\right)\\ &\le D'(\mathrm{m}_p(\mu_0)+1).\end{aligned}$$

The estimate for $\mathrm{m}_p(|\nu_t|)$ follows recalling that

$$\begin{aligned}\mathrm{m}_p(|\nu_t|) &= \int_{\mathbb{R}^d}|x|^p\left|\frac{\nu_t}{\mu_t}(x)\right| d\mu_t(x) \le C\int_{\mathbb{R}^d}(|x|+1)^{p+1}\, d\mu_t(x)\\ &\le 2^p C\left(\mathrm{m}_{p+1}(\mu_t)+1\right)\\ &\le 2^p C\left[\tilde{D}(\mathrm{m}_{p+1}(\mu_0)+1)+1\right]\\ &\le D''(\mathrm{m}_{p+1}(\mu_0)+1).\end{aligned}$$

□

Corollary 3.2.8 (Uniform p-integrability)**.** *Assume hypothesis* (F_0), (F_1). *Let* $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ *be an admissible trajectory driven by* $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$, $p \ge 1$, *and set* $v_t(x) = \dfrac{\nu_t}{\mu_t}(x)$. *Assume that* $\mathrm{m}_p(\mu_0) < +\infty$, *then*

$$\int_0^T\int_{\mathbb{R}^d}|v_t(x)|^p\, d\mu_t\, dt < +\infty.$$

Proof. We have

$$\begin{aligned}\int_0^T\int_{\mathbb{R}^d}|v_t(x)|^p\, d\mu_t\, dt &\le TC^p\int_{\mathbb{R}^d}(|x|+1)^p\, d\mu_t \le 2^{p-1}TC^p(\mathrm{m}_p(\mu_t)+1),\\ &\le K(\mathrm{m}_p(\mu_0)+1),\end{aligned}$$

for a suitable constant $K > 0$ depending only on C, T, p and where the last inequality comes from Lemma 3.2.7(iv). □

The following definitions are the natural counterpart of the classical case.

Definition 3.2.9 (Reachable set)**.** Let $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$, and $T > 0$. Define the set of *admissible curves defined on* $[0,T]$ *and starting from* μ_0 by setting

$$\mathscr{A}_T(\mu_0) := \{\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d) : \boldsymbol{\mu} \text{ is an admissible trajectory with } \mu_{|t=0} = \mu_0\}.$$

The *reachable set* from μ_0 in time T is

$$\mathscr{R}_T(\mu_0) := \{\mu \in \mathscr{P}(\mathbb{R}^d) : \text{ there exists } \boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \mathscr{A}_T(\mu_0) \text{ with } \mu = \mu_T\}.$$

Definition 3.2.10 (Generalized minimum time)**.** Let $p \geq 1$, $\Phi \subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1.1, and $\tilde{S}^\Phi$, $\tilde{S}^\Phi_p$ be the corresponding generalized targets defined in Definition 3.1.1. In analogy with the classical case, we define the *generalized minimum time function* $\tilde{T}^\Phi : \mathscr{P}(\mathbb{R}^d) \to [0,+\infty]$ by setting

$$\tilde{T}^\Phi(\mu_0) := \inf\left\{J_F(\boldsymbol{\mu},\boldsymbol{\nu}) : \boldsymbol{\mu} \in \mathscr{A}_T(\mu_0),\, \boldsymbol{\mu} \text{ is driven by } \boldsymbol{\nu},\, \mu_{|t=T} \in \tilde{S}^\Phi\right\}, \quad (3.5)$$

where, by convention, $\inf\emptyset = +\infty$.

Given $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$ with $T^\Phi(\mu_0) < +\infty$, an admissible curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\tilde{T}^\Phi(\mu_0)]} \subseteq \mathscr{P}(\mathbb{R}^d)$, driven by a family of Borel vector-valued measures $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,\tilde{T}^\Phi(\mu_0)]}$ and satisfying $\mu_{|t=0} = \mu_0$ and $\mu_{|t=\tilde{T}^\Phi(\mu_0)} \in \tilde{S}^\Phi$ is *optimal for* μ_0 if

$$\tilde{T}^\Phi(\mu_0) = J_F(\boldsymbol{\mu},\boldsymbol{\nu}).$$

Given $p \geq 1$, we define also a generalized minimum time function $\tilde{T}^\Phi_p : \mathscr{P}_p(\mathbb{R}^d) \to [0,+\infty]$ by replacing in the above definitions $\tilde{S}^\Phi$ by $\tilde{S}^\Phi_p$ and $\mathscr{P}(\mathbb{R}^d)$ by $\mathscr{P}_p(\mathbb{R}^d)$. Since $\tilde{S}^\Phi_p \subseteq \tilde{S}^\Phi$, it is clear that $\tilde{T}^\Phi(\mu_0) \leq \tilde{T}^\Phi_p(\mu_0)$.

Remark 3.2.11. In view of the characterization in Theorem 8.3.1 in [9], and of Remark 3.2.2, one can think to $\tilde{T}^\Phi$ as the minimum time needed by the system to steer μ_0 to a measure in $\tilde{S}^\Phi$, along absolutely continuous curves in $\mathscr{P}_p(\mathbb{R}^d)$.

When the generalized target $\tilde{S}^\Phi$ admits a classical counterpart S, it is natural to ask for a comparison between the generalized minimum time function and the classical minimum time needed to reach S.

Proposition 3.2.12 (First comparison between $\tilde{T}^\Phi$ and T)**.** *Consider the generalized minimum time problem for Σ_F as in Definition 3.2.10 assuming (F_0), (F_1), and suppose that the corresponding generalized target $\tilde{S}^\Phi$ admits S as classical counterpart. Then for all $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$ we have*

$$\tilde{T}^\Phi(\mu_0) \geq \|T\|_{L^\infty_{\mu_0}},$$

where $T : \mathbb{R}^d \to [0,+\infty]$ is the classical minimum time function for the system $\dot{x}(t) \in F(x(t))$ with target S.

Proof. For sake of clarity, in this proof we will simply write $\tilde{T}$ and $\tilde{S}$, thus omitting Φ, since we can always replace the set Φ by $\{d_S\}$ by the assumption of existence of the classical counterpart S for $\tilde{S}^\Phi$.

If $\tilde{T}(\mu_0) = +\infty$ there is nothing to prove, so assume $\tilde{T}(\mu_0) < +\infty$. Fix $\varepsilon > 0$ and let $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ be an admissible curve starting from μ_0, driven by a family of Borel vector-valued measures $\boldsymbol{\nu} = \{\nu_t\}_{t\in I}$ such that $T = J_F(\boldsymbol{\mu},\boldsymbol{\nu}) < \tilde{T}(\mu_0) + \varepsilon$ and $\mu_{|t=T} \in \tilde{S}$. In particular, we have that $v_t(x) :=$

$\dfrac{\nu_t}{\mu_t}(x) \in F(x)$ for μ_t-a.e. $x \in \mathbb{R}^d$ and a.e. $t \in [0,T]$, hence $|v_t(x)| \leq C(1+|x|)$ for μ_t-a.e $x \in \mathbb{R}^d$. Accordingly,

$$\int_0^T \int_{\mathbb{R}^d} \frac{|v_t(x)|}{1+|x|}\, d\mu_t\, dt \leq CT < +\infty.$$

By the Superposition Principle (Theorem 1.3.3), recalling Definition 1.0.6 and 3.2.4, we have that there exists a probability measure $\boldsymbol{\eta} = \mu_0 \otimes \eta_x \in \mathscr{T}_F(\mu_0)$ such that for μ_0-a.e $x \in \mathbb{R}^d$, the measure $\eta_x \in \mathscr{P}(\Gamma_T^x)$ is concentrated on absolutely continuous curves γ satisfying $\dot\gamma(t) = v_t(\gamma(t))$ for a.e. t, and $\mu_t = e_t\sharp\mu_0$. In particular, if $x \notin \operatorname{supp}\mu_0$ or $\gamma(0) \neq x$, then $(x,\gamma) \notin \operatorname{supp}\boldsymbol{\eta}$.

Let $\{\psi_n\}_{n\in\mathbb{N}} \in C_C^\infty(\mathbb{R}^d;[0,1])$ with $\psi_n(x) = 0$ if $x \notin B(0,n+1)$ and $\psi_n(x) = 1$ if $x \in \overline{B(0,n)}$. By Monotone Convergent Theorem, since $\{\psi_n(\cdot)d_S(\cdot)\}_{n\in\mathbb{N}} \subseteq C_b^0(\mathbb{R}^d)$ is an increasing sequence of nonnegative functions pointwise convergent to $d_S(\cdot)$, we have for every $t \in [0,T]$

$$\begin{aligned}\iint_{\mathbb{R}^d\times\Gamma_T} d_S(\gamma(t))\, d\boldsymbol{\eta}(x,\gamma) &= \lim_{n\to\infty}\iint_{\mathbb{R}^d\times\Gamma_T} \psi_n(\gamma(t))d_S(\gamma(t))\, d\boldsymbol{\eta}(x,\gamma)\\ &= \lim_{n\to\infty}\int_{\mathbb{R}^d} \psi_n(x)d_S(x)\, d\mu_t(x)\end{aligned}$$

By taking $t = T$, we have that the last term vanishes because $\mu_{|t=T} \in \tilde{S}$ and so $\operatorname{supp}\mu_{|t=T} \subseteq S$, therefore

$$\iint_{\mathbb{R}^d\times\Gamma_T} d_S(\gamma(T))\, d\boldsymbol{\eta}(x,\gamma) = 0.$$

In particular, we necessarily have that $\gamma(T) \in S$ and $\gamma(0) = x$ for $\boldsymbol{\eta}$-a.e. $(x,\gamma) \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$, whence $T \geq T(x)$ for μ_0-a.e. $x \in \mathbb{R}^d$, since $T(x)$ is the infimum of the times needed to steer x to S along trajectories of the system. Thus, $\tilde{T}(\mu_0) + \varepsilon \geq T(x)$ for μ_0-a.e. $x \in \mathbb{R}^d$ and, by letting $\varepsilon \to 0$, we conclude that $\tilde{T}(\mu_0) \geq \|T\|_{L^\infty_{\mu_0}}$. □

We notice that the inequality appearing in Proposition 3.2.12 may be strict without further assumptions.

Example 3.2.13. In $\mathbb{R}$, let $F(x) = \{1\}$ for all $x \in \mathbb{R}$ and set $\Phi = \{|\cdot|\}$, thus $S = \{0\}$ is the classical counterpart of $\tilde{S}^\Phi = \{\delta_0\}$. Moreover, we have $T(x) = |x|$ for $x \leq 0$ and $T(x) = +\infty$ for $x > 0$. Define $\mu_0 = \dfrac{1}{2}(\delta_{-2} + \delta_{-1})$. We have $\|T\|_{L^\infty_{\mu_0}} = \max\{T(-1), T(-2)\} = 2$. However there are no solutions of $\dot{x}(t) = 1$ steering any two different points to the origin in the *same* time, thus the set of admissible trajectories joining μ_0 and δ_0 is empty, hence $\tilde{T}^\Phi(\mu_0) = +\infty$.

Remark 3.2.14. This implies that in general the problem of the generalized minimum time *cannot be reduced* to the underlying finite dimensional control problem, even in the cases where the underlying control problem is particulary simple. A consequence of this fact is that even if the underlying system enjoys some properties as closure and relative compactness of the set of admissible trajectories (provided for instance by good assumptions on the set-valued map F), which lead to the existence of optimal trajectories for the problem, in our generalized framework all these results must be proved.

Definition 3.2.15 (Convergence of curves in $\mathscr{P}(\mathbb{R}^d)$)**.** We say that a family of curves $\boldsymbol{\mu}^n = \{\mu_t^n\}_{t\in[0,T]}$ in $\mathscr{P}(\mathbb{R}^d)$

1. *pointwise converges* to a curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ in $\mathscr{P}(\mathbb{R}^d)$ if and only if $\mu_t^n \rightharpoonup^* \mu_t$ for all $t \in [0,T]$. In this case we will write $\boldsymbol{\mu}^n \rightharpoonup^* \boldsymbol{\mu}$.

2. *pointwise converges* to a curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ in $\mathscr{P}_p(\mathbb{R}^d)$ if and only if $\boldsymbol{\mu}^n = \{\mu_t^n\}_{t\in[0,T]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ and $\lim_{n\to+\infty} W_p(\mu_t^n, \mu_t) = 0$ for all $t \in [0,T]$. In this case we will write $\boldsymbol{\mu}^n \to^p \boldsymbol{\mu}$.

3. *uniformly converges* to a curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ in $\mathscr{P}_p(\mathbb{R}^d)$ if and only if $\boldsymbol{\mu}^n = \{\mu_t^n\}_{t\in[0,T]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ and
$$\lim_{n\to+\infty} \sup_{t\in[0,T]} W_p(\mu_t^n, \mu_t) = 0.$$
In this case we will write $\boldsymbol{\mu}^n \rightrightarrows^p \boldsymbol{\mu}$.

The following results will be used to prove l.s.c. of the generalized minimum time function in Theorem 3.2.19 and existence of optimal trajectories in Theorem 3.2.20.

Lemma 3.2.16. *Assume that $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ satisfies (F_0). Then the functional $\mathscr{F} : \mathscr{P}(\mathbb{R}^d) \times \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d) \to \{0, +\infty\}$ defined by*
$$\mathscr{F}(\mu, E) := \begin{cases} \displaystyle\int_{\mathbb{R}^d} I_{F(x)}\left(\frac{E}{\mu}(x)\right) d\mu(x), & \text{if } E \ll \mu, \\ \\ +\infty, & \text{otherwise} \end{cases} \tag{3.6}$$
is l.s.c. w.r.t. narrow convergence.

Proof. Define $f(x,v) = I_{F(x)}(v)$. Since F is u.s.c. with convex values, we have that $f(\cdot,\cdot)$ is l.s.c. and $f(x,\cdot)$ is convex. By compactness of $F(x)$, we have that the domain of $f(x,\cdot)$ is bounded, thus following the notation in Section 2.1 we have $f_\infty(x,v) = 0$ if $v = 0$ and $f_\infty(x,v) = +\infty$ if $v \neq 0$. Thus (3.6) can be written in the form of (2.2) for this choice of f. By l.s.c. of F, there exists a continuous selection $z_0 : \mathbb{R}^d \to \mathbb{R}^d$ of F, i.e., there exists $z_0 \in C^0(\mathbb{R}^d;\mathbb{R}^d)$ satisfying $z_0(x) \in F(x)$ for all $x \in \mathbb{R}^d$. Thus $x \mapsto f(x, z_0(x))$ is continuous and finite. The functional (3.6) satisfies now the assumptions of Lemma 2.1.1, and so it is l.s.c. $\square$

Proposition 3.2.17 (Convergence of admissible trajectories)**.** *Assume (F_0). Let $\boldsymbol{\mu}^n = \{\mu_t^n\}_{t\in[0,T]}$ be a sequence of admissible curves defined on $[0,T]$ such that $\boldsymbol{\mu}^n$ is driven by $\boldsymbol{\nu}^n = \{\nu_t^n\}_{t\in[0,T]}$ and suppose that there exist $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ and $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]} \subseteq \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ such that for a.e. $t \in [0,T]$ it holds $(\mu_t^n, \nu_t^n) \rightharpoonup^* (\mu_t, \nu_t)$. Then $\boldsymbol{\mu}$ is an admissible trajectory driven by $\boldsymbol{\nu}$.*

Proof. Fix $t \in [0,T]$ such that $(\mu_t^n, \nu_t^n) \rightharpoonup^* (\mu_t, \nu_t)$ and $\mathscr{F}(\mu_t^n, \nu_t^n) = 0$ for all $n \in \mathbb{N}$. By l.s.c. of $\mathscr{F}$ and recalling that $\mathscr{F} \geq 0$, we have
$$0 \leq \mathscr{F}(\mu_t, \nu_t) \leq \liminf_{n\to+\infty} \mathscr{F}(\mu_t^n, \nu_t^n) = 0,$$

and so for a.e. $t \in [0,T]$ we have $\dfrac{\nu_t}{\mu_t}(x) \in F(x)$ for μ_t-a.e. $x \in \mathbb{R}^d$.

Since for every $\varphi \in C^1_C(\mathbb{R}^d)$ we have in the sense of distributions on $[0,T]$,

$$\frac{d}{dt}\int_{\mathbb{R}^d} \varphi(x) d\mu_t^n(x) = \int_{\mathbb{R}^d} \nabla\varphi(x)\, d\nu_t^n(x),$$

and for the last term we have

$$\lim_{n\to\infty}\int_{\mathbb{R}^d} \nabla\varphi(x)\, d\nu_t^n(x) = \int_{\mathbb{R}^d} \nabla\varphi(x)\, d\nu_t(x),$$

due to the w^*-convergence of ν_t^n to ν_t, thanks to Lemma 8.1.2 in [9], we deduce that, up to changing μ_t and ν_t for all t belonging to a $\mathscr{L}^1$-negligible set of $[0,T]$, we have that $\boldsymbol{\mu}$ is an admissible curve driven by $\boldsymbol{\nu}$. □

The previous Proposition is the key ingredient to prove the following theorem which, in analogy with the classical case, establish a sufficient condition to have relative compactness of a set of admissible trajectories.

Theorem 3.2.18. *Assume (F_0), (F_1). Let $\mathscr{A}$ be a set of admissible trajectories defined on $[0,T]$ and $C_1 > 0$, $p > 1$ be constants such that for all $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \mathscr{A}$ it holds $\mathrm{m}_p(\mu_t) \leq C_1$ for a.e. $t \in [0,T]$. Then the pointwise w^*-closure of $\mathscr{A}$ is a set of admissible trajectories.*

In particular, this holds if $\{\mathrm{m}_p(\mu_0) :$ there exists $\boldsymbol{\mu} \in \mathscr{A}$ with $\mu_{|t=0} = \mu_0\}$ is bounded, and, in particular, it holds for $\mathscr{A}_T(\mu_0)$ when $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$.

Proof. Let $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$ be a sequence in $\mathscr{A}$. Since $\boldsymbol{\mu}^n$ is an admissible trajectory, it is driven by $\boldsymbol{\nu}^n = \{v_t^n\mu_t^n\}_{t\in[0,T]}$ with $v_t^n \in L^1_{\mu_t^n}$ and $v_t^n(x) \in F(x)$ for a.e. $t \in [0,T]$ and μ_t^n-a.e. $x \in \mathbb{R}^d$. Since for a.e. $t \in [0,T]$

$$\int_{\mathbb{R}^d} |x|^p\, d\mu_t^n(x) \leq C_1,$$

according to Remark 5.1.5 in [9], we have that for a.e. $t \in [0,T]$ there exists $\mu_t \in \mathscr{P}(\mathbb{R}^d)$ such that $\mu_t^n \rightharpoonup^* \mu_t$. Similarly,

$$\int_{\mathbb{R}^d} |x|^{p-1}|d\nu_t^n(x)| = \int_{\mathbb{R}^d} |x|^{p-1}|v_t^n(x)|\, d\mu_t^n(x) \leq LC_1 + 1,$$

for a constant $L > 0$. Thus there exists $\nu_t \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ such that $\nu_t^n \rightharpoonup^* \nu_t$. By Proposition 3.2.17, we have that $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ is an admissible trajectory defined on $[0,T]$ driven by $\boldsymbol{\nu}$. The last assertion comes from Lemma 3.2.7, which allows to estimate the moments of μ_t and ν_t in terms of the moments of μ_0. □

Theorem 3.2.19 (L.s.c. of the generalized minimum time)**.** *Assume (F_0) and (F_1). Then $\tilde{T}_p^\Phi : \mathscr{P}_p(\mathbb{R}^d) \to [0,+\infty]$ is l.s.c. for all $p > 1$.*

Proof. Let $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$, we have to prove that $\tilde{T}_p^\Phi(\mu_0) \leq \liminf\limits_{W_p(\mu,\mu_0)\to 0} \tilde{T}_p^\Phi(\mu)$. Taken a sequence $\{\mu_0^n\}_{n\in\mathbb{N}} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ s.t. $W_p(\mu_0^n,\mu_0) \to 0$ for $n \to +\infty$, and $\liminf\limits_{W_p(\mu,\mu_0)\to 0} \tilde{T}_p^\Phi(\mu) = \lim\limits_{n\to+\infty} \tilde{T}_p^\Phi(\mu_0^n) =: T$, we want to prove that $\tilde{T}_p^\Phi(\mu_0) \leq T$.

If $T = +\infty$ there is nothing to prove, so let us assume $T < +\infty$. Then there exists a sequence $\{T_n\}_{n\in\mathbb{N}}$ such that $T_n \to T$, and a sequence of admissible trajectories $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$, with $\boldsymbol{\mu}^n = \{\mu^n_t\}_{t\in[0,T_n]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$, such that $\mu^n_{|t=T_n} \in \tilde{S}^\Phi_p$ for all $n \in \mathbb{N}$.

Without loss of generality, we can assume that all $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$ are defined in an interval containing $[0,T]$, since if $T_n < T$ we can use Lemma 1.3.2 and extend $\boldsymbol{\mu}^n$ to a trajectory defined in $[0,T]$ simply by taking any Borel selection $\bar{v}$ of $F(\cdot)$ (which exists by (F_0) and by Theorem 8.1.3 in [13]), and considering the solution of the continuity equation $\partial_t\mu_t + \operatorname{div}\bar{v}\mu_t = 0$ in $]T_n,T]$ with $\mu_{|t=T_n} = \mu^n_{T_n}$. Now, since μ^n_0 converges in W_p to μ_0, we have that there exists $\bar{n} > 0$ such that the set $\{\mathrm{m}_p(\mu^n_0) : n > \bar{n}\}$ is uniformly bounded by $\mathrm{m}_p(\mu_0)+1$. Then, by Lemma 3.2.7 and by Theorem 3.2.18 there exists an admissible trajectory $\boldsymbol{\mu} := \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ such that $\boldsymbol{\mu}^n \to^p \boldsymbol{\mu}$, $n \to +\infty$, up to subsequences and $\mu_{|t=0} = \mu_0$. Recalling Theorem 8.3.1 in [9], for all $n \in \mathbb{N}$ we have

$$\begin{aligned}\tilde{d}_{\tilde{S}^\Phi_p}(\mu_T) &\le W_p(\mu_T,\mu^n_{T_n}) \le W_p(\mu_T,\mu^n_T) + W_p(\mu^n_T,\mu^n_{T_n})\\ &\le W_p(\mu_T,\mu^n_T) + \left|\int_{T_n}^T \left\|\frac{\nu^n_t}{\mu^n_t}\right\|_{L^p_{\mu^n_t}} dt\right|.\end{aligned}$$

If we show a uniform bound on $\left\|\frac{\nu^n_t}{\mu^n_t}\right\|_{L^p_{\mu^n_t}}$, then by letting $n \to +\infty$ we have that $\mu_T \in \tilde{S}^\Phi_p$, thus $\tilde{T}^\Phi_p(\mu_0) \le T$ and the proof is concluded.

For a.e. $t \in [0,T]$ and μ^n_t-a.e. x we have $\dfrac{\nu^n_t}{\mu^n_t}(x) \in F(x)$. By (F_1) there exists $C > 0$ such that

$$\left\|\frac{\nu^n_t}{\mu^n_t}\right\|_{L^p_{\mu^n_t}} \le C\left(\mathrm{m}_p^{1/p}(\mu^n_t)+1\right).$$

We conclude by using the Lemma 3.2.7 to estimate $\mathrm{m}_p(\mu^n_t)$ in terms of $\mathrm{m}_p(\mu^n_0)$ and recalling that since μ^n_0 converges to μ_0 in W_p, for n sufficiently large we have $\mathrm{m}_p(\mu^n_0) \le \mathrm{m}_p(\mu_0)+1$. □

Thanks to the preliminary result of Theorem 3.2.18 about relative compactness of a set of admissible trajectories in the space of Borel probability measures, together with the lower semicontinuity of the time functional J_F coming from Lemma 3.2.16, we can prove the following result.

Theorem 3.2.20 (Existence of minimizers)**.** *Assume (F_0), (F_1), and let $p > 1$. Let $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$, $\Phi \subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1.1, and let $\tilde{S}^\Phi$ be the corresponding generalized target. Let $\tilde{T}^\Phi(\mu_0) < \infty$. Then there exists an admissible curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$ which is optimal for μ_0, that is $\tilde{T}^\Phi(\mu_0) = J_F(\boldsymbol{\mu},\boldsymbol{\nu})$. Moreover, we have also $\tilde{T}^\Phi(\mu_0) = \tilde{T}^\Phi_p(\mu_0)$.*

Proof. By the hypothesis of finiteness of $\tilde{T}^\Phi(\mu_0)$ and by definition of infimum we have that there exist $\{t_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ and a sequence of admissible trajectories $\boldsymbol{\mu}^n = \{\mu^n_t\}_{t\in[0,t_n]}$, such that $\mu^n|_{t=0} = \mu_0$, $\mu^n|_{t=t_n} =: \sigma^n \in \tilde{S}^\Phi$, $t_n \to \tilde{T}^\Phi(\mu_0)^+$. Moreover, by Lemma 3.2.7, we have that $\sigma^n \in \tilde{S}^\Phi_p$ for all $n \in \mathbb{N}$. We restrict all $\boldsymbol{\mu}^n$ to be defined on $[0,\tilde{T}^\Phi(\mu_0)]$.

By Theorem 3.2.18, $\boldsymbol{\mu}^n$ w^*-converges up to subsequences to an admissible trajectory $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\tilde{T}^\Phi(\mu_0)]}$ starting from μ_0 driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,\tilde{T}^\Phi(\mu_0)]}$, and by w^*-closure of $\tilde{S}^\Phi$ we have $\sigma^n \rightharpoonup^* \mu|_{t=\tilde{T}^\Phi(\mu_0)} \in \tilde{S}^\Phi$. Applying again Lemma 3.2.7, we have that $\mu|_{t=\tilde{T}^\Phi(\mu_0)} \in \tilde{S}^\Phi_p$. Thus $\tilde{T}^\Phi(\mu_0) = \tilde{T}^\Phi_p(\mu_0) = J_F(\boldsymbol{\mu}, \boldsymbol{\nu})$. $\square$

The following result, which allows us to embed classical admissible trajectories into an admissible trajectory in the space of measures, will be the main tool used to prove the next comparison results (Corollaries 3.2.22 and 3.2.23) between the classical and the generalized minimum time function. We will see that these results allow us to justify the name of *generalized minimum time* given to functions $\tilde{T}^\Phi(\cdot)$ and $\tilde{T}^\Phi_p(\cdot)$.

Lemma 3.2.21 (Convexity property of the embedding of classical trajectories)**.** *Let $N \in \mathbb{N} \setminus \{0\}$, $T > 0$ be given. Assume (F_0) and (F_1). Consider a family of continuous curves and real numbers $\{(\gamma_i, \lambda_i)\}_{i=1,\dots,N} \subseteq \Gamma_T \times [0,1]$ such that $\gamma_i(\cdot)$ is a trajectory of $\dot{x}(t) \in F(x(t))$ for $i = 1, \dots, N$, and $\displaystyle\sum_{i=1}^N \lambda_i = 1$.*

For all $i = 1, \dots, N$ and $t \in [0,T]$, define the measures $\mu_t^{(i)} = \delta_{\gamma_i(t)}$, $\mu_t = \displaystyle\sum_{i=1}^N \lambda_i \mu_t^{(i)}$,

$$\nu_t^{(i)} = \begin{cases} \dot{\gamma}_i(t)\delta_{\gamma_i(t)}, & \text{if } \dot{\gamma}_i(t) \text{ exists,} \\ \\ 0, & \text{otherwise,} \end{cases}$$

and $\nu_t = \displaystyle\sum_{i=1}^N \lambda_i \nu_t^{(i)}$. Then $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ is an admissible trajectory driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$.

Proof. By linearity, clearly we have that

$$\partial_t \mu_t + \operatorname{div} \nu_t = 0$$

is satisfied in the sense of distributions, moreover $\mu_t(B) = 0$ implies $\nu_t(B) = 0$ for every Borel set $B \subseteq \mathbb{R}^d$, thus $|\nu_t| \ll \mu_t$. It remains only to prove that for a.e. $t \in [0,T]$ we have $\nu_t = v_t \mu_t$ for a vector-valued function $v_t \in L^1(\mathbb{R}^d; \mathbb{R}^d)$ satisfying $v_t(x) \in F(x)$ for μ_t-a.e. $x \in \mathbb{R}^d$. Set

$$\tau = \{t \in [0,T] : \dot{\gamma}_i(t) \text{ exists for all } i = 1, \dots, N \text{ and } \dot{\gamma}_i(t) \in F(\gamma_i(t))\},$$

and notice that τ has full measure in $[0,T]$.

Fix $t \in \tau$, $x \in \operatorname{supp} \mu_t$. By definition of μ_t, we have that there exists $I \subseteq \{1, \dots, N\}$ such that $\mu_t^{(i)} = \delta_x$ if and only if $i \in I$. So it is possible to find $\delta > 0$ such that for all $0 < \rho < \delta$ we have

$$\mu_t(B(x,\rho)) = \sum_{j\in I} \lambda_j, \qquad \nu_t(B(x,\rho)) = \sum_{i\in I} \lambda_i \int_{B(x,\rho)} \frac{\nu_t^{(i)}}{\mu_t^{(i)}}(y)\, d\mu_t^{(i)}(y) = \sum_{i\in I} \lambda_i \frac{\nu_t^{(i)}}{\mu_t^{(i)}}(x).$$

Thus for every $t \in \tau$ and $x \in \operatorname{supp} \mu_t$ we have

$$v_t(x) := \lim_{\rho\to 0^+} \frac{\nu_t(B(x,\rho))}{\mu_t(B(x,\rho))} = \sum_{i\in I} \frac{\lambda_i}{\sum_{j\in I} \lambda_j} \frac{\nu_t^{(i)}}{\mu_t^{(i)}}(x),$$

i.e., a convex combination of $\dot{\gamma}_i(t) = \dfrac{\nu_t^{(i)}}{\mu_t^{(i)}}(x) \in F(x)$ for μ_t-a.e. $x \in \mathbb{R}^d$. Thus $\dfrac{\nu_t}{\mu_t}(x) = v_t(x) \in F(x)$, and so $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ is an admissible trajectory driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$. □

Corollary 3.2.22. *Assume (F_0), (F_1). Let $\Phi \subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1.1, and assume that the generalized target $\tilde{S}^\Phi$ admits a classical counterpart $S \subseteq \mathbb{R}^d$ which is weakly invariant for the dynamics $\dot{x}(t) \in F(x(t))$. Let $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$ with $p > 1$. Then $\tilde{T}_p^\Phi(\mu_0) = \tilde{T}^\Phi(\mu_0) = \|T(\cdot)\|_{L^\infty_{\mu_0}}$.*

Proof. Since $\tilde{S}^\Phi$ admits classical counterpart S, we have that S is closed and we can always take $\Phi = \{d_S(\cdot)\}$. Thus in this proof we will simply write $\tilde{T}_p$ and $\tilde{S}_p$ in place of $\tilde{T}_p^\Phi$ and $\tilde{S}_p^\Phi$, respectively.

By Proposition 3.2.12, we have only to prove that $\tilde{T}_p(\mu_0) \le T := \|T(\cdot)\|_{L^\infty_{\mu_0}}$. Assume that $T < +\infty$, otherwise there is nothing to prove. For μ_0-a.e. point $x \in \mathbb{R}^d$ we have $T(x) \le T$, thus there exists a trajectory $\gamma_x(\cdot)$ such that $\gamma_x(T(x)) \in S$. By the weak invariance of S, we can extend this trajectory to be defined on $[0,T]$ with the constraint $\gamma_x(t) \in S$ for all $T(x) \le t \le T$, thus in particular $\gamma_x(T) \in S$. Fix $\varepsilon > 0$, then there exists $N = N_\varepsilon \in \mathbb{N}\setminus\{0\}$, and $\{(x_i,\lambda_i) : i = 1,\dots,N_\varepsilon\} \subseteq \operatorname{supp}\mu_0 \times [0,1]$ such that:

1. $\displaystyle\sum_{i=1}^{N_\varepsilon} \lambda_i = 1$;
2. $W_p\left(\mu_0, \displaystyle\sum_{i=1}^{N_\varepsilon} \lambda_i\delta_{x_i}\right) < \varepsilon$;
3. there exist classical admissible trajectories $\{\gamma_i : [0,T] \to \mathbb{R}^d : i = 1,\dots,N_\varepsilon\}$ satisfying $\gamma_i(0) = x_i$ and $\gamma_i(T) \in S$ for all $i = 1,\dots,N_\varepsilon$.

It is possible to find an admissible trajectory $\boldsymbol{\mu}^{(\varepsilon)} = \left\{\mu_t^{(\varepsilon)}\right\}_{t\in[0,T]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ such that $\mu_0^{(\varepsilon)} = \sum_{i=1}^{N_\varepsilon} \lambda_i\delta_{x_i}$ and $\mu_T^{(\varepsilon)} \in \tilde{S}_p$, indeed, we can set

$$\mu_t^{(\varepsilon)} = \sum_{i=1}^{N_\varepsilon} \lambda_i\delta_{\gamma_i(t)}, \qquad \nu_t^{(\varepsilon)} = \begin{cases} \displaystyle\sum_{i=1}^{N_\varepsilon} \lambda_i\dot{\gamma}_i(t)\delta_{\gamma_i(t)}, & \text{if } \dot{\gamma}_i(t) \text{ exists for all } i = 1,\dots,N_\varepsilon, \\ 0, & \text{otherwise,} \end{cases}$$

and then apply Lemma 3.2.21.

Since $\mu_0^{(\varepsilon)}$ converges in W_p to μ_0, we have that there exists $\bar{\varepsilon} > 0$ such that the set $\left\{\mathrm{m}_p\left(\mu_0^{(\varepsilon)}\right) : 0 < \varepsilon < \bar{\varepsilon}\right\}$ is uniformly bounded by $\mathrm{m}_p(\mu_0) + 1$. In particular, by taking a sequence $\varepsilon_k \to 0^+$, and the corresponding admissible trajectories $\boldsymbol{\mu}^{(\varepsilon_k)}$ driven by $\boldsymbol{\nu}^{(\varepsilon_k)}$, we can extract by Theorem 3.2.18 a subsequence converging to an admissible trajectory $\bar{\boldsymbol{\mu}}$ driven by $\bar{\boldsymbol{\nu}}$ satisfying $\bar{\mu}_0 = \mu_0$. Since $\mu_T^{(\varepsilon)} \in \tilde{S}_p$ for all $\varepsilon > 0$, by the closure of $\tilde{S}_p$ we have $\bar{\mu}_T \in \tilde{S}_p$, thus $\tilde{T}_p(\mu_0) \le T$. □

Corollary 3.2.23 (Second comparison result)**.** *Assume (F_0), (F_1). Let $\Phi \subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1.1, and assume that the generalized target $\tilde{S}^\Phi$ admits a classical counterpart S. Then, for every $x_0 \in \mathbb{R}^d$ we have $\tilde{T}^\Phi(\delta_{x_0}) = \tilde{T}^\Phi_p(\delta_{x_0}) = T(x_0)$ for all $p \geq 1$, where $T(\cdot)$ is the classical minimum time function for $\dot{x}(t) \in F(x(t))$ with target S.*

Proof. Apply Lemma 3.2.21 to the family $\{(\gamma, 1)\}$, where $\gamma(\cdot)$ is an admissible trajectory of $\dot{x}(t) \in F(x(t))$ satisfying $\gamma(0) = x_0$ and $\gamma(T(x_0)) \in S$. We obtain an admissible trajectory steering δ_{x_0} to $\tilde{S}_p$ for all $p \geq 1$ in time $T(x_0)$, thus $\tilde{T}_p(\delta_{x_0}) \leq T(x_0)$. By Proposition 3.2.12, since $\|T(\cdot)\|_{L^\infty_{\delta_{x_0}}} = T(x_0)$, equality holds. □

Remark 3.2.24. This means that if we have a precise knowledge of the initial state, we recover exactly the classical objects in finite-dimension.

The following is a generalization of a cardinal result in Optimal Control Theory recalled in Theorem 1.4.8. The proof is based on gluing results for solutions of the continuity equation.

Theorem 3.2.25 (Dynamic programming principle)**.** *Let $0 \leq s \leq \tau$, let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be a set-valued function, let $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\tau]}$ be an admissible curve for Σ_F. Then we have*

$$\tilde{T}^\Phi(\mu_0) \leq s + \tilde{T}^\Phi(\mu_s).$$

Moreover, if $\tilde{T}^\Phi(\mu_0) < +\infty$, equality holds for all $s \in [0, \tilde{T}^\Phi(\mu_0)]$ if and only if $\boldsymbol{\mu}$ is optimal for $\mu_0 = \mu_{|t=0}$. The same result holds for $\tilde{T}^\Phi_p$ in place of $\tilde{T}^\Phi$, $p \geq 1$.

Proof. Let $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,\tau]} \subseteq \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ be such that $\boldsymbol{\mu}$ is driven by $\boldsymbol{\nu}$. Fix $s \in [0,\tau]$, $\varepsilon > 0$. If $\tilde{T}^\Phi(\mu_s) = +\infty$ there is nothing to prove. Otherwise there exists an admissible curve $\boldsymbol{\mu}^\varepsilon := \{\mu^\varepsilon_t\}_{t\in[0,\tilde{T}^\Phi(\mu_s)+\varepsilon]} \subseteq \mathscr{P}(\mathbb{R}^d)$ driven by $\boldsymbol{\nu}^\varepsilon = \{\nu^\varepsilon_t\}_{t\in[0,\tilde{T}^\Phi(\mu_s)+\varepsilon]} \subseteq \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ such that $\mu^\varepsilon_{|t=0} = \mu_s$ and $\mu^\varepsilon_{|t=\tilde{T}^\Phi(\mu_s)+\varepsilon} \in \tilde{S}^\Phi$. We consider

$$\tilde{v}^\varepsilon_t(x) := \begin{cases} \dfrac{\nu_t}{\mu_t}(x), & \text{for } 0 \leq t \leq s,\\[2ex] \dfrac{\nu^\varepsilon_{t-s}}{\mu^\varepsilon_{t-s}}(x), & \text{for } s < t \leq \tilde{T}^\Phi(\mu_s) + s + \varepsilon.\end{cases}$$

$$\tilde{\mu}^\varepsilon_t := \begin{cases} \mu_t, & \text{for } 0 \leq t \leq s,\\[2ex] \mu^\varepsilon_{t-s}, & \text{for } s < t \leq \tilde{T}^\Phi(\mu_s) + s + \varepsilon.\end{cases}$$

It is clear that $\tilde{\mu}^\varepsilon_{|t=0} = \mu_0$, that $\tilde{\mu}^\varepsilon_{|t=\tilde{T}^\Phi(\mu_s)+s+\varepsilon} \in \tilde{S}^\Phi$, and that $\tilde{v}^\varepsilon_t(x) \in F(x)$ for $\tilde{\mu}^\varepsilon_t$–a.e. $x \in \mathbb{R}^d$ and a.e. $t \in [0, \tilde{T}^\Phi(\mu_s) + \varepsilon]$. Moreover, $t \mapsto \tilde{\mu}^\varepsilon_t$ is narrowly continuous. Since Lemma 1.3.2 ensures that $\tilde{\boldsymbol{\mu}}^{\boldsymbol{\varepsilon}} := \{\tilde{\mu}^\varepsilon_t\}_{t\in[0,\tilde{T}^\Phi(\mu_s)+s+\varepsilon]}$ is a solution of the continuity equation driven by $\tilde{\boldsymbol{\nu}}^{\boldsymbol{\varepsilon}} = \{\tilde{\nu}^\varepsilon_t = \tilde{v}^\varepsilon_t \tilde{\mu}^\varepsilon_t\}_{t\in[0,\tilde{T}^\Phi(\mu_s)+s+\varepsilon]}$, thus an admissible trajectory, we have that

$$\tilde{T}^\Phi(\mu_0) \leq J_F(\tilde{\boldsymbol{\mu}}^{\boldsymbol{\varepsilon}}, \tilde{\boldsymbol{\nu}}^{\boldsymbol{\varepsilon}}) = \tilde{T}^\Phi(\mu_s) + s + \varepsilon\,.$$

By arbitrariness of $\varepsilon > 0$, we conclude that $\tilde{T}^\Phi(\mu_0) \leq s + \tilde{T}^\Phi(\mu_s)$.

Assume now that $\tilde{T}^\Phi(\mu_0) < +\infty$ and equality holds for all $s \in [0, \tilde{T}^\Phi(\mu_0)]$. Then, in particular, when $s = \tilde{T}^\Phi(\mu_0)$ we get

$$\tilde{T}^\Phi(\mu_0) = \tilde{T}^\Phi(\mu_0) + \tilde{T}^\Phi(\mu_{\tilde{T}^\Phi(\mu_0)}) \quad \Rightarrow \quad \tilde{T}^\Phi(\mu_{\tilde{T}^\Phi(\mu_0)}) = 0\,.$$

In turn, this implies $\mu_{\tilde{T}^\Phi(\mu_0)} = \mu_{s+\tilde{T}^\Phi(\mu_s)} \in \tilde{S}^\Phi$, and so $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\tilde{T}^\Phi(\mu_0)]}$ joins μ_0 with the generalized target in the minimum time $\tilde{T}^\Phi(\mu_0)$, thus $\boldsymbol{\mu}$ is optimal for μ_0.

Finally, assume that $\boldsymbol{\mu}$, driven by $\boldsymbol{\nu}$, is optimal for μ_0 and $\tilde{T}^\Phi(\mu_0) < +\infty$. To have equality $\tilde{T}^\Phi(\mu_0) = s + \tilde{T}^\Phi(\mu_s)$, it is enough to show that $\tilde{T}^\Phi(\mu_0) \geq s + \tilde{T}^\Phi(\mu_s)$. If we define $\nu'_t := \nu_{t+s}$, we have that $\boldsymbol{\mu'} = \{\mu'_t\}_{t\in[0,\tilde{T}^\Phi(\mu_0)-s]} := \{\mu_{t+s}\}_{t\in[0,\tilde{T}^\Phi(\mu_0)-s]}$ is an admissible trajectory driven by $\boldsymbol{\nu'} = \{\nu'_t\}_{t\in[0,\tilde{T}^\Phi(\mu_0)-s]}$ and starting by μ_s. This implies that

$$\begin{aligned}\tilde{T}^\Phi(\mu_0) = J_F(\boldsymbol{\mu},\boldsymbol{\nu}) &= s + \int_s^{\tilde{T}^\Phi(\mu_0)} \int_{\mathbb{R}^d} \left(1 + I_{F(x)}\left(\frac{\nu_t}{\mu_t}(x)\right)\right) d\mu_t(x)\,dt \\ &= s + \int_0^{\tilde{T}^\Phi(\mu_0)-s} \int_{\mathbb{R}^d} \left(1 + I_{F(x)}\left(\frac{\nu'_t}{\mu'_t}(x)\right)\right) d\mu'_t(x)\,dt \geq s + \tilde{T}^\Phi(\mu_s),\end{aligned}$$

which concludes the proof. □

3.2.1 Attainability results

We are now interested in proving *sufficient* conditions on the set-valued function $F(\cdot)$ in order to have *attainability* of the generalized control system, i.e. to steer a probability measure on the generalized target by following an admissible trajectory in finite time.

In other words, we want to prove a generalization of the so called *Petrov's condition* that gives, in the classical case, an attainability property for the control system, i.e. a sufficient condition for continuity of the minimum time function at the boundary of the target.

Theorem 3.2.26 (Attainability in the smooth case)**.** *Assume (F_0), (F_1). Let $\Phi \subseteq C^1_b(\mathbb{R}^d;\mathbb{R}) \cap \mathrm{Lip}(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1.1 and let $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$, $p \geq 1$. Assume that:*

1. *for all $\phi \in \Phi$ there exists a $\mathscr{L}^1$-integrable map $k^\phi :]0,+\infty[\to]0,+\infty[$;*
2. *there exists $T \in [0,+\infty[$ such that*
$$T \geq \sup_{\phi\in\Phi} \inf \left\{ t \geq 0 \,:\, \int_{\mathbb{R}^d} \phi(x)\,d\mu_0(x) \leq \int_0^t k^\phi(s)\,ds \right\};$$
3. *there exist a Borel vector field $v : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and an admissible trajectory $\boldsymbol{\mu} := \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ driven by $\boldsymbol{\nu} = \{\nu_t := v_t\mu_t\}_{t\in[0,T]}$, and satisfying $\mu_{|t=0} = \mu_0$,*

such that the following condition holds:

(C_c) for all $\phi \in \Phi$ we have $\displaystyle\int_{\mathbb{R}^d} \langle \nabla\phi(x), v_t(x)\rangle\,d\mu_t(x) \leq -k^\phi(t)$ for a.e. $t \in]0,T]$.

Then we have

$$\tilde{T}_p^{\Phi}(\mu_0) \le \sup_{\phi\in\Phi} \inf \left\{ t \ge 0 \,:\, \int_{\mathbb{R}^d} \phi(x)\, d\mu_0(x) \le \int_0^t k^{\phi}(s)\, ds \right\}.$$

Proof. We notice that by Lemma 3.2.7, we have $\boldsymbol{\mu} \subseteq \mathscr{P}_p(\mathbb{R}^d)$.

Given $\phi \in \Phi$, we set $L_t^{\phi} := \displaystyle\int_{\mathbb{R}^d} \phi(x)\, d\mu_t(x)$. Take $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$ and notice that if $T = 0$ we have

$$\sup_{\phi\in\Phi} \inf \left\{ t \ge 0 \,:\, \int_{\mathbb{R}^d} \phi(x)\, d\mu_0(x) \le \int_0^t k^{\phi}(s)\, ds \right\} = 0,$$

so $\mu_0 \in \tilde{S}_p^{\Phi}$ and $\tilde{T}_p^{\Phi}(\mu_0) = 0$. We assume then $T > 0$.

From the continuity equation we have that in the distributional sense it holds (see Remark 8.1.1 in [9], allowing to use the functions of Φ as test functions)

$$\dot{L}_t^{\phi} = \frac{d}{dt} \int_{\mathbb{R}^d} \phi(x) d\mu_t(x) = \int_{\mathbb{R}^d} \langle \nabla\phi(x), v_t(x)\rangle\, d\mu_t(x) \le -k^{\phi}(t).$$

Then $L_t^{\phi} \le L_0^{\phi} - \int_0^t k^{\phi}(s)\, ds$ for $0 < t \le T$. Thus if we take $t \in]0, T]$ s.t. we have $\int_{\mathbb{R}^d} \phi(x)\, d\mu_0(x) \le \int_0^t k^{\phi}(s)\, ds$ for all $\phi \in \Phi$, then we have that $L_t^{\phi} \le 0$ for all $\phi \in \Phi$, hence $\mu_t \in \tilde{S}_p^{\Phi}$ for all such t, which ends the proof. $\square$

Remark 3.2.27. In particular, if in the condition (C_c) above we can choose $k^{\phi}(t) \equiv k^{\phi}$ for a.e. $t > 0$, for a constant $k^{\phi} > 0$, then we get $\tilde{T}_p^{\Phi}(\mu_0) \le \sup_{\phi\in\Phi} \left\{ \frac{1}{k^{\phi}} \int_{\mathbb{R}^d} \phi(x)\, d\mu_0(x) \right\}$.

In the next part, we will weaken the strong assumptions required in the previous result, dealing with the case $p = 2$, proving the attainability result in Theorem 3.2.32.

Throughout this and the next section we will use the following notation.

Definition 3.2.28. Given $Q, T, H, M, h > 0$ and $\Phi \subseteq C^0(\mathbb{R}^d; \mathbb{R})$, we define

$$SC_{M,H}(\mathbb{R}^d;\mathbb{R}) := \left\{ \phi : \mathbb{R}^d \to \mathbb{R} : \begin{array}{l} \phi \in SC(\mathbb{R}^d;\mathbb{R}) \text{ has semiconcavity constant} \\ \text{less or equal than } M \text{ and} \\ \mathrm{Lip}(\phi, B(0, 2R+1)) \le H(R+1), \text{ for all } R > 0 \end{array} \right\},$$

$$D_{Q,H,h}(s) := \frac{2\sqrt{3}}{h} H(s + Q + 1)^{\frac{1}{2}},$$

$$G_{M,H}(r, s) := H(Mr + 2s + 3) \cdot Mr,$$

$$\mathcal{A}_{Q,T,H}^{M,h,\Phi} := \left\{ \mu \in \mathscr{P}_2(\mathbb{R}^d) : \begin{array}{l} T \ge \dfrac{2}{h} \left(\displaystyle\sup_{\phi\in\Phi} \int_{\mathbb{R}^d} \phi(x)\, d\mu(x) + G_{M,H}(\tilde{T}_2^{\Phi}(\mu), \mathrm{m}_2(\mu)) \right) \\ \mathrm{m}_2(\mu) \le Q \end{array} \right\}.$$

Lemma 3.2.29. *Let $C > 0$, and consider the problem*

$$\begin{cases} \partial_t \mu_t(x) + \mathrm{div}(v(x)\mu_t(x)) = 0, & \text{for } t \in]0, T],\ x \in \mathbb{R}^d, \\ \mu_{|t=0} = \mu_0 \in \mathscr{P}_2(\mathbb{R}^d), \end{cases} \tag{3.7}$$

where $v : \mathbb{R}^d \to \mathbb{R}^d$ is a Borel map satisfying $|v(x)| \le C(|x|+1)$ for every $x \in \mathbb{R}^d$ and $t \mapsto \|v\|_{L^1_{\mu_t}} \in L^1$. Let $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ be such that $t \mapsto e_t \sharp \boldsymbol{\eta}$ is a solution *of* (3.7) *as in the Superposition Principle (Theorem 1.3.3). If $v \in C^0_b(\mathbb{R}^d;\mathbb{R}^d)$ and for all $x \in \mathbb{R}^d$ it admits a nondecreasing modulus of continuity $\omega_x(\cdot)$ at the point x, with $(x,r) \mapsto \omega_x(r)$ in $L^2_{\mu_0\otimes\mathscr{L}^1}(\mathbb{R}^d \times [0,T\|v\|_\infty])$, then*

$$\left\|\frac{e_t - e_0}{t} - v\circ e_0\right\|^2_{L^2_{\boldsymbol{\eta}}} \le \frac{1}{\|v\|_\infty}\int_{\mathbb{R}^d}\int_0^{\|v\|_\infty} \omega_x^2(rt)\,dr\,d\mu_0(x),$$

and the left hand side tends to zero for $t \to 0$.

Proof. We write $\boldsymbol{\eta} = \mu_0 \otimes \eta_x$, $\eta_x \in \mathscr{P}(\Gamma^x_T)$, thus for μ_0-a.e. $x \in \mathbb{R}^d$ and η_x-a.e. $\gamma \in \Gamma^x_T$ we have that γ is an absolutely continuous solution of

$$\begin{cases} \dot\gamma(t) = v(\gamma(t)), \text{ for } \mathscr{L}^1\text{-a.e. } t \in]0,T], \\ \gamma(0) = x. \end{cases}$$

Let $M := \|v\|_\infty$. By hypothesis we have

$$\begin{aligned}
\left\|\frac{e_t - e_0}{t} - v\circ e_0\right\|^2_{L^2_{\boldsymbol{\eta}}} &= \int_{\mathbb{R}^d}\int_{\Gamma^x_T}\left|\frac{\gamma(t)-\gamma(0)}{t} - v\circ\gamma(0)\right|^2 d\eta_x(\gamma)\,d\mu_0(x)\\
&= \int_{\mathbb{R}^d}\int_{\Gamma^x_T}\left|\frac{1}{t}\int_0^t \dot\gamma(s)ds - v\circ\gamma(0)\right|^2 d\eta_x(\gamma)\,d\mu_0(x)\\
&= \int_{\mathbb{R}^d}\int_{\Gamma^x_T}\left|\frac{1}{t}\int_0^t (v\circ\gamma(s) - v\circ\gamma(0))ds\right|^2 d\eta_x(\gamma)\,d\mu_0(x)\\
&\le \int_{\mathbb{R}^d}\int_{\Gamma^x_T}\left(\frac{1}{t}\int_0^t \omega_{\gamma(0)}(|\gamma(s)-\gamma(0)|)ds\right)^2 d\eta_x(\gamma)\,d\mu_0(x)\\
&\le \int_{\mathbb{R}^d}\left(\frac{1}{t}\int_0^t \omega_x(M\cdot s)\,ds\right)^2 d\mu_0(x)\\
&= \int_{\mathbb{R}^d}\left(\frac{1}{M}\int_0^M \omega_x(rt)\,dr\right)^2 d\mu_0(x)\\
&\le \frac{1}{M}\int_{\mathbb{R}^d}\int_0^M \omega_x^2(rt)\,dr\,d\mu_0(x),
\end{aligned}$$

where we used Jensen's inequality for the last passage.

Finally, recalling the assumptions on ω_x, we conclude by letting $t \to 0^+$ and using the Dominated Convergence Theorem. □

The following result gives an upper bound on the "observable measurements", involved in the definition of generalized target set, evaluated along an evolving admissible trajectory.

Lemma 3.2.30. *Assume (F_0), (F_4) and take M as in (F_4). Let $\tau > 0$, $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$. Let $\Phi \subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1.1. Suppose that there exists $H > 0$ s.t. for all $R > 0$, we have $\mathrm{Lip}(\phi, B(0,2R+1)) \le H(R+1)$ for all $\phi \in \Phi$. Then for any admissible trajectory $\boldsymbol{\mu} := \{\mu_t\}_{t\in[0,\tau]}$, $\mu_{|t=0} = \mu_0$, we have*

$$\sup_{\phi\in\Phi}\int_{\mathbb{R}^d}\phi(x)\,d\mu_t(x) \le \sup_{\phi\in\Phi}\int_{\mathbb{R}^d}\phi(x)\,d\mu_0(x) + G_{M,H}(\tau, \mathrm{m}_2(\mu_0)).$$

for all $0 \le t \le \tau$.

Proof. Let $\boldsymbol{\mu} := \{\mu_t\}_{t\in[0,\tau]}$, $\mu_{|t=0} = \mu_0$, be an admissible trajectory and $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_\tau)$ be such that $\mu_t = e_t\sharp\boldsymbol{\eta}$, for $0 \le t \le \tau$, as in the Superposition Principle (Theorem 1.3.3). We write $\boldsymbol{\eta} = \mu_0 \otimes \eta_x$, $\eta_x \in \mathscr{P}(\Gamma_\tau^x)$, thus for μ_0-a.e. $x \in \mathbb{R}^d$ and η_x-a.e. $\gamma \in \Gamma_\tau^x$ we have that γ is an absolutely continuous solution of

$$\begin{cases}\dot\gamma(t) \in F(\gamma(t)), \text{ for } \mathscr{L}^1\text{-a.e. } t \in]0,\tau],\\ \gamma(0) = x.\end{cases}$$

In particular, for all $t \in [0,\tau]$ we have that $|\gamma(t)-\gamma(0)| \le \int_0^t |\dot\gamma(s)|\,ds \le Mt$.

Notice that for all $\phi \in \Phi$ and $t \in [0,\tau]$, it holds

$$\begin{aligned}|\phi(\gamma(t)) - \phi(\gamma(0))| &\le H(|\gamma(t)| + |\gamma(0)| + 1)\cdot|\gamma(t)-\gamma(0)|\\ &\le H(|\gamma(t)-\gamma(0)| + 2|\gamma(0)| + 1)\cdot Mt\\ &\le H(Mt + 2|\gamma(0)| + 1)\cdot Mt =: P(t).\end{aligned}$$

Hence,

$$\int_{\mathbb{R}^d\times\Gamma_\tau} \phi(\gamma(t))\,d\boldsymbol{\eta}(x,\gamma) \le \int_{\mathbb{R}^d\times\Gamma_\tau} \phi(\gamma(0))\,d\boldsymbol{\eta}(x,\gamma) + \int_{\mathbb{R}^d\times\Gamma_\tau} P(t)\,d\boldsymbol{\eta}(x,\gamma) \tag{3.8}$$

$$\iff \int_{\mathbb{R}^d} \phi(x)\,d\mu_t(x) \le \int_{\mathbb{R}^d} \phi(x)\,d\mu_0(x) + \int_{\mathbb{R}^d\times\Gamma_\tau} P(t)\,d\boldsymbol{\eta}(x,\gamma), \tag{3.9}$$

for all $0 \le t \le \tau$, $\phi \in \Phi$.

Observe that

$$\int_{\mathbb{R}^d\times\Gamma_\tau} P(t)\,d\boldsymbol{\eta}(x,\gamma) \le H(M\tau + 2\mathrm{m}_1(\mu_0) + 1)\cdot M\tau \tag{3.10}$$

$$\le H(M\tau + 2\mathrm{m}_2(\mu_0) + 3)\cdot M\tau =: G_{M,H}(\tau, \mathrm{m}_2(\mu_0)), \tag{3.11}$$

for all $0 \le t \le \tau$, where we used the fact that $\mathrm{m}_1(\mu) \le \mathrm{m}_2(\mu)^{\frac{1}{2}} \le \mathrm{m}_2(\mu) + 1$, for any $\mu \in \mathscr{P}(\mathbb{R}^d)$ by Hölder inequality.

Hence the conclusion follows by passing to the supremum on $\phi \in \Phi$ in (3.9) and using the estimate (3.11). □

Remark 3.2.31. The simplest choice for Φ is to take $\Phi = \{d_S\}$, where d_S is the distance function from a given closed set $S \subseteq \mathbb{R}^d$. This case can be used to model the so called *evacuation problem*, i.e. situations that arise for example in pedestrian dynamics in which we want to steer a mass of people outside a room with one or more exits. In this kind of problems the set-valued function F, representing the admissible velocities of the pedestrians, takes into account the presence of possible obstacles modelling the geometry of the environment. In this case, the next result will bound the total time needed to evacuate the room by taking into account the initial distribution of the agents.

Theorem 3.2.32 (Attainability result)**.** *Assume* (F_0), (F_4) *and take* M *as in* (F_4)*. Let* $K, H > 0$, $\Phi \subseteq SC_{K,H}(\mathbb{R}^d;\mathbb{R})$ *such that* Φ *satisfies* (T_E) *in Definition 3.1.1.*

Assume that there exist $h, T > 0$ and a modulus of continuity $\tilde{\omega}(\cdot)$ such that for all $\mu \in \mathscr{P}_2(\mathbb{R}^d)\backslash\tilde{S}_2^\Phi$ there exist a continuous vector field $v = v_\mu \in C^0(\mathbb{R}^d;\mathbb{R}^d)$ and a function $(x,r) \mapsto \omega_x(r)$ in $L^2_{\mu\otimes\mathscr{L}^1}(\mathbb{R}^d \times [0, TM])$ satisfying:

1. *$v_\mu(x) \in F(x)$ for μ-a.e. $x \in \mathbb{R}^d$;*

2. *$\omega_x(\cdot)$ is a nondecreasing modulus of continuity at x for v_μ for μ-a.e. $x \in \mathbb{R}^d$, and*

$$\omega_\mu(t) := \left(\frac{1}{M}\int_{\mathbb{R}^d}\int_0^M \omega_x^2(rt)\,dr\,d\mu(x)\right)^{\frac{1}{2}} \leq \tilde{\omega}(t),$$

   for $0 \leq t \leq T$;

3. *for all $\phi \in \Phi$ there exists $\zeta^{\mu,\phi} \in \mathrm{Bor}(\mathbb{R}^d;\mathbb{R}^d)$ satisfying $\zeta^{\mu,\phi}(x) \in \partial^+\phi(x)$ for μ-a.e. $x \in \mathbb{R}^d$ and*

$$\int_{\mathbb{R}^d} \langle \zeta^{\mu,\phi}(x), v(x)\rangle d\mu(x) < -h.$$

Then we have

$$\tilde{T}_2^\Phi(\bar{\mu}) \leq \frac{2}{h}\sup_{\phi\in\Phi}\int_{\mathbb{R}^d}\phi(x)d\bar{\mu}(x),$$

for all $\bar{\mu} \in \mathscr{P}_2(\mathbb{R}^d)$ such that $\displaystyle\frac{2}{h}\sup_{\phi\in\Phi}\int_{\mathbb{R}^d}\phi(x)d\bar{\mu}(x) \leq T$.

Proof. We will adapt a method used in finite-dimensional case in Theorem 5.10 in [59].

First, notice that by hypothesis (F_4) we have $v_\mu \in C^0_b(\mathbb{R}^d;\mathbb{R}^d)$, for all $\mu \in \mathscr{P}_2(\mathbb{R}^d) \setminus \tilde{S}_2^\Phi$.

For all $\phi \in \Phi$, $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, set

$$L(\mu) := \sup_{\phi\in\Phi}\int_{\mathbb{R}^d}\phi(x)d\mu(x).$$

Take $\bar{\mu} \in \mathscr{P}_2(\mathbb{R}^d)$ and notice that if $L(\bar{\mu}) \leq 0$, then $\bar{\mu} \in \tilde{S}_2^\Phi$. We assume then $L(\bar{\mu}) > 0$ and $T \geq \frac{2}{h}L(\bar{\mu})$, otherwise there is nothing to prove.

We define by recurrence the sequences $\{t_i\}_{i\in\mathbb{N}} \subseteq [0,+\infty[$, $\{\mu^{(i)}\}_{i\in\mathbb{N}} \subseteq \mathscr{P}_2(\mathbb{R}^d)$, $\{v^{(i)}\}_{i\in\mathbb{N}} \subseteq C(\mathbb{R}^d;\mathbb{R}^d)$, $\{\zeta^{(i),\phi}\}_{i\in\mathbb{N}} \subseteq \mathrm{Bor}(\mathbb{R}^d;\mathbb{R}^d)$ and $\{L_i\}_{i\in\mathbb{N}} \subseteq [0,+\infty[$.

Define $t_0 = 0$, $\mu^{(0)} = \bar{\mu}$, and, for all $\phi \in \Phi$, let $v^{(0)} = v_{\bar{\mu}}$, $\zeta^{(0),\phi} = \zeta^{\bar{\mu},\phi}$ as in the statement with $\mu = \bar{\mu}$. Set $L_0 = L(\bar{\mu}) > 0$ as above.

Suppose to have defined for all $\phi \in \Phi$ the quantities $t_i, \mu^{(i)}$ and $v^{(i)} = v_{\mu^{(i)}}, \zeta^{(i),\phi} = \zeta^{\mu^{(i)},\phi}$, $L_i = L(\mu^{(i)}) \geq 0$, where $v^{(i)} = v_{\mu^{(i)}}, \zeta^{(i),\phi} = \zeta^{\mu^{(i)},\phi}$ are taken as in the statement with $\mu = \mu^{(i)}$ and $\displaystyle\sum_{k=0}^i t_k < T$, and where $\mu^{(i)}$ is joined to $\bar{\mu}$ by an admissible trajectory in time $\displaystyle\sum_{k=0}^i t_k$.

Consider the problem

$$\begin{cases}\partial_t\mu_t(x)+\mathrm{div}(v^{(i)}(x)\mu_t(x))=0, & \text{for } t\in]0,T-\sum_{k=0}^i t_k],\, x\in\mathbb{R}^d,\\ \\ \mu_{|t=0}=\mu^{(i)},\end{cases} \tag{3.12}$$

and let $\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_{T-\sum_{k=0}^i t_k})$ be such that $t\mapsto\mu_t=e_t\sharp\boldsymbol{\eta}$ solves (3.12) as in the Superposition Principle (Theorem 1.3.3) and μ_t is connected to $\bar\mu$ along an admissible trajectory in time $t+\displaystyle\sum_{k=0}^i t_k$. We recall that $e_0\sharp\boldsymbol{\eta}=\mu^{(i)}$.

Recalling the hypothesis on Φ, we have that for all $\phi\in\Phi$

$$\begin{aligned}\|\zeta^{(i),\phi}\|_{L^2_{\mu^{(i)}}}&\le\|\mathrm{Lip}(\phi,B(0,2|\cdot|+1))\|_{L^2_{\mu^{(i)}}}\le H\left(\int_{\mathbb{R}^d}(|x|+1)^2\,d\mu^{(i)}(x)\right)^{\frac12}\\&\le\sqrt{2}H(\mathrm{m}_2(\mu^{(i)})+1)^{\frac12}\le 2H(\mathrm{m}_2(\mu^{(i)})+1).\end{aligned}$$

Furthermore, by definition of $L(\cdot)$, for any $t\in\left]0,T-\sum_{k=0}^i t_k\right[$ there exists $\bar\phi=\phi^{t,i}\in\Phi$ such that $L(\mu_t)\le\displaystyle\int_{\mathbb{R}^d}\bar\phi(x)\,d\mu_t(x)+t^3$.

Thus, recalling the semiconcavity property of Φ, Lemma 3.2.29, and taking C'_T,C''_T as in Lemma 3.2.7, we have the existence of $\mathcal{C}',\mathcal{C}''>0$ depending only on H,K,T and $\bar\mu$ such that

$$\begin{aligned}L(\mu_t)-L_i&\le\int_{\mathbb{R}^d}\bar\phi(x)d\mu_t(x)-\int_{\mathbb{R}^d}\bar\phi(x)d\mu_0(x)+t^3\\&=\int_{\mathbb{R}^d\times\Gamma_T}(\bar\phi\circ e_t-\bar\phi\circ e_0)d\boldsymbol{\eta}(x,\gamma)+t^3\\&\le\int_{\mathbb{R}^d\times\Gamma_T}\langle\zeta^{(i),\bar\phi}\circ e_0,e_t-e_0\rangle d\boldsymbol{\eta}(x,\gamma)+K\|e_t-e_0\|^2_{L^2_{\boldsymbol{\eta}}}+t^3\\&=t\int_{\mathbb{R}^d\times\Gamma_T}\langle\zeta^{(i),\bar\phi}\circ e_0,v^{(i)}\circ e_0\rangle d\boldsymbol{\eta}+t\int_{\mathbb{R}^d\times\Gamma_T}\langle\zeta^{(i),\bar\phi}\circ e_0,\frac{e_t-e_0}{t}-v^{(i)}\circ e_0\rangle d\boldsymbol{\eta}+\\&\quad+t^2K\left\|\frac{e_t-e_0}{t}\right\|^2_{L^2_{\boldsymbol{\eta}}}+t^3\\&\le t\int_{\mathbb{R}^d}\langle\zeta^{(i),\bar\phi},v^{(i)}\rangle d\mu^{(i)}+t\|\zeta^{(i),\bar\phi}\|_{L^2_{\mu^{(i)}}}\cdot\left\|\frac{e_t-e_0}{t}-v^{(i)}\circ e_0\right\|_{L^2_{\boldsymbol{\eta}}}+\\&\quad+t^2KC'_T(\mathrm{m}_2(\mu^{(i)})+1)+t^3\\&\le-ht+2tH(\mathrm{m}_2(\mu^{(i)})+1)\,\omega_{\mu^{(i)}}(t)+t^2KC'_T(\mathrm{m}_2(\mu^{(i)})+1)+t^3\\&\le-ht+2tH\tilde\omega(t)\,[C''_T(\mathrm{m}_2(\bar\mu)+1)+1]+t^2KC'_T[C''_T(\mathrm{m}_2(\bar\mu)+1)+1]+t^3\\&\le-ht+\tilde\omega(t)\mathcal{C}'t+\mathcal{C}''t^2+t^3.\end{aligned}$$

Thus we have that there exists $\tau>0$ independent on i such that $L(\mu_t)-L_i\le-\frac{h}{2}t$ for $0<t\le\tau\wedge\left[T-\sum_{k=0}^i t_k\right]$, where we adopt the notation $a\wedge b=\min\{a,b\}$.

At this point we can define $t_{i+1}:=\tau\wedge\left[T-\sum_{k=0}^i t_k\right]\wedge\tilde T_2^\Phi(\mu^{(i)})$, $\mu^{(i+1)}=\mu_{t_{i+1}}\in\mathscr{P}_2(\mathbb{R}^d)$ and take, for all $\phi\in\Phi$, a velocity field $v^{(i+1)}=v_{\mu^{(i+1)}}$ and a

Borel map $\zeta^{(i+1),\phi} = \zeta^{\mu^{(i+1)},\phi}$ as in the statement. Define $L_{i+1} = L(\mu^{(i+1)}) \geq 0$. In this way, we have also provided that $\displaystyle\sum_{k=0}^{i+1} t_k \leq T$.

Thus we have
$$L_{i+1} - L_i \leq -\frac{h}{2} t_{i+1} \leq 0.$$
It follows that $\{L_j\}_{j\in\mathbb{N}}$ is a decreasing sequence bounded from below by 0, so it admits a limit value $L_\infty \geq 0$. From the above relation we have also
$$\frac{2}{h}(L_i - L_{i+1}) \geq t_{i+1},$$
and so
$$T \geq \frac{2}{h} L_0 \geq \frac{2}{h}(L_0 - L_\infty) = \frac{2}{h}\sum_{i=0}^{\infty}(L_i - L_{i+1}) \geq \sum_{i=0}^{\infty} t_{i+1} \geq 0.$$
Thus, in particular, we have also that $t_j \to 0$ as $j \to +\infty$.

We notice that
$$\begin{aligned} W_2^2(\mu^{(i+1)}, \mu^{(i)}) &\leq \|e_{t_{i+1}} - e_0\|^2_{L^2_{\boldsymbol{\eta}}} = \int_{\mathbb{R}^d}\int_{\Gamma_T^x} |\gamma(t_{i+1}) - \gamma(0)|^2 d\eta_x\, d\mu^{(i)} \\ &\leq \int_{\mathbb{R}^d}\int_{\Gamma_T^x} \left(\int_0^{t_{i+1}} |\dot\gamma(s)|^2 ds\right) d\eta_x\, d\mu^{(i)} \leq M^2 \cdot t_{i+1}, \end{aligned}$$
where for the first inequality we have used the property (7.1.6) in [9] ($\mu^{(i+1)} = e_{t_{i+1}}\sharp\boldsymbol{\eta}$, $\mu^{(i)} = e_0\sharp\boldsymbol{\eta}$). Then we used the disintegration Lemma, the property of absolute continuity of $\gamma \in \Gamma_T$, Jensen's inequality and hypothesis (F_4). Since the series $\displaystyle\sum_{i=0}^{\infty} t_{i+1}$ converges, we have that $\{\mu^{(i)}\}_{i\in\mathbb{N}}$ is a Cauchy sequence in the complete space $(\mathscr{P}_2(\mathbb{R}^d), W_2)$, and so there exists $\tilde\mu \in \mathscr{P}_2(\mathbb{R}^d)$ such that $\mu^{(i)} \to \tilde\mu$ in W_2 for $i \to +\infty$.

According to the definition of t_{i+1}, we have:
$$0 = \lim_{i\to\infty} t_{i+1} = \liminf_{i\to\infty}\left(\tau \wedge \left[T - \sum_{k=0}^{i} t_k\right] \wedge \tilde T_2^\Phi(\mu^{(i)})\right),$$
this implies
$$\liminf_{i\to\infty} \tilde T_2^\Phi(\mu^{(i)}) = 0,$$
and so, using l.s.c. property of the minimum time function proved in Theorem 3.2.19, we have that $\tilde T_2^\Phi(\tilde\mu) = 0$, i.e. $\tilde\mu \in \tilde S_2^\Phi$. Since we have constructed an admissible trajectory connecting $\bar\mu$ to $\tilde S_2^\Phi$ in time $\displaystyle\sum_{i=0}^{\infty} t_{i+1}$, we have $\displaystyle\sum_{i=0}^{\infty} t_{i+1} \geq \tilde T_2^\Phi(\bar\mu)$, and so
$$\tilde T_2^\Phi(\bar\mu) \leq \frac{2}{h}\sup_{\phi\in\Phi}\int_{\mathbb{R}^d}\phi(x)d\bar\mu(x).$$
□

Remark 3.2.33. In the special case of Remark 3.2.31, the above result yields $\tilde{T}_2^\Phi(\bar{\mu}) \leq \dfrac{2}{h}\|d_S\|_{L^1_{\bar{\mu}}}$, hence by Proposition 3.1.8, we obtain $\tilde{T}_2^\Phi(\bar{\mu}) \leq \dfrac{2}{h}\tilde{d}_{\tilde{S}_p^\Phi}(\bar{\mu})$.

The result of Theorem 3.2.32 can be applied also to the system described in the following example.

Example 3.2.34 (A model of optimal displacement of solar panels on a hill). Assume to have a certain amount of solar panels distributed in an initial configuration (for instance stored in some warehouses) near to a hill in a fixed region. Our aim is to steer the solar panels in suitable positions on the hill, such that the new configuration achieves a fixed minimum efficiency threshold (*target*) averaged in one year of solar exposition, and minimizing a cost depending on the "effort" required to move them from the initial position to the final configuration. This problem can be modelized as follows.

After a normalization, we represent by $\mu_0 \in \mathscr{P}_2(\mathbb{R}^2)$ the given initial distribution of solar panels, and by a map $h \in C^\infty_C(\mathbb{R}^2; [0, h_{\max}])$ the shape of the hill and the surrounding region, where $h(x, y)$ represents the altitude of the point (x, y). We are assuming that the region is quite small compared with the surface of the Earth, i.e., that the Earth's curvature effects are negligible w.r.t. the scale of the system. Furthermore, let $\hat{r}(s) = (r_1(s), r_2(s), r_3(s)) \in \mathbb{R}^3$ be the unit vector giving the direction joining an observer in the region with the position of the sun at time $s \in [0, T]$, where T is set to one year. Of course, the function $\hat{r}(\cdot)$ is given taking into account the latitude, and we have $\hat{r} \in C^\infty([0, T]; \mathbb{R}^3)$. If the scale of the system is not too large, we may assume that $\hat{r}(s)$ does not depend on the position of the observer in the region of interest.

Then, given $\varepsilon, \delta, \alpha > 0$, we can model the *instantaneous efficiency* $\psi^{\varepsilon,\delta,\alpha}(s, x, y)$ at time $s \in [0, T]$, for a panel lying at the position $(x, y) \in \mathbb{R}^2$, for example by the formula

$$\psi^{\varepsilon,\delta,\alpha}(s, x, y) = \psi_1^\delta(r_3(s))\ \psi_2^\varepsilon\left(\hat{r}(s) \cdot \frac{(-\nabla h(x, y), 1)}{|(-\nabla h(x, y), 1)|}\right)\ \psi_3^{\varepsilon,\alpha}(s, x, y),$$

where

- $\psi_1^\delta \in C^\infty([-1, 1]; [0, 1])$, represents the *presence* of solar light, hence $\psi_1^\delta(z)$ is set to 1 when $z \in [\delta, 1]$ (*day time*), it is set to 0 when $z \in [-\delta, -1]$ (*night time*), and it is strictly increasing for $z \in]-\delta, \delta[$ (*dawn* and *twilight*).
- $\psi_2^\varepsilon \in C^\infty([-1, 1]; [\varepsilon, 1])$, expresses the *instantanous performance* at time s for a solar panel lying on the ground in position (x, y), which depends on the angle of exposure to the sun light, i.e., on the angle between $\hat{r}(s)$ and the normal to the ground at (x, y) (which is the normal to the hypograph of h). The function ψ_2^ε is strictly increasing and we set $\psi_2^\varepsilon(-1) = \varepsilon$ (representing the *default background radiation* due to the diffusion effect of the atmosphere), and $\psi_2^\varepsilon(1) = 1$, hence the maximal instantaneous performance at (x, y) is achieved when the panel's surface is orthogonal to the direction of the sun light.
- $\psi_3^{\varepsilon,\alpha} \in C^\infty([0, T] \times \mathbb{R}^2; [\varepsilon, 1])$, takes into account the presence of bumps in the straight line between the panel and the sun. For any $s \in [0, T]$ we define the set of points directly exposed to the sun at time s by

  $$V(s) := \left\{(x, y) \in \mathbb{R}^2 \,:\, h(x, y) + \lambda r_3(s) \geq h(x + \lambda r_1(s), y + \lambda r_2(s)),\ \text{ for all } \lambda \geq 0\right\},$$

and we set $\psi_3^{\varepsilon,\alpha}(s,x,y)=1$ if $(x,y)\in V(s)$, $\psi_3^{\varepsilon,\alpha}(s,x,y)=\varepsilon$ if $d_{V(s)}(x,y)>\alpha$ (which defines the *shadow region*, where the only radiation is given by the default background radiation).

The *averaged efficiency* for a configuration $\mu\in\mathscr{P}_2(\mathbb{R}^2)$ is given by

$$E(\mu):=\int_{\mathbb{R}^2}\left(\int_0^T\psi^{\varepsilon,\delta,\alpha}(s,x,y)\,ds\right)d\mu(x,y)\in[0,T].$$

Given a target efficiency $\bar{c}>0$, our aim is to have $E(\mu)\geq\bar{c}$, hence the target set $\tilde{S}_2^\Phi$ is defined as in Definition 3.1.1 by taking $\Phi=\{\phi\}$, where

$$\phi(x,y):=\bar{c}-\int_0^T\psi^{\varepsilon,\delta,\alpha}(s,x,y)\,ds\in C^\infty(\mathbb{R}^2;\mathbb{R})\cap\mathrm{Lip}(\mathbb{R}^2)\subseteq SC(\mathbb{R}^2;\mathbb{R}).$$

We take into accout the "effort" (cost) to move the panels in the controlled dynamics by defining the set-valued map $F:\mathbb{R}^2\rightrightarrows\mathbb{R}^2$ as

$$F(x,y)=\overline{B\left(0,\frac{1}{|\nabla h(x,y)|^2+1}\right)},$$

which expresses the fact that the movements are much costly at the point of the hill where the slope is higher. The assumptions of Theorem 3.2.32 are thus satisfied.

We notice that the model can be refined by adding further cost terms, e.g., penalizing an excessive concentration or sparsity in the position of the panels. These effects can be included by considering instead of the usual Wasserstein distance, some variants of it (we refer e.g. to [56] for further details).

With much milder assumptions w.r.t. the previous attainability result, in the case of existence of a classical counterpart for the generalized target set, it is possible to prove a weaker controllability result, as showed below.

Indeed, representation formula for the generalized minimum time provided in Corollary 3.2.22 allows us to recover many results valid for the classical minimum time function also in the framework of generalized systems.

Theorem 3.2.35 (Controllability)**.** *Let $\Phi\subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1.1. Assume that the generalized target $\tilde{S}^\Phi$ admits a classical counterpart $S\subseteq\mathbb{R}^d$ which is weakly invariant for F. Assume (F_0), (F_1), (F_3) and that for every $R>0$ there exist η_R, $\sigma_R>0$ such that for a.e. $x\in B(0,R)\setminus S$ with $d_S(x)\leq\sigma_R$ there holds*

$$\sigma_{F(x)}(-\nabla d_S(x))>\eta_R\,,\tag{3.13}$$

where $\sigma_{F(x)}$ is the support function of the set $F(x)$ as in (1.1).
Then, if we set for $p>1$

$$\mathscr{P}_p(\mathbb{R}^d)_{|R}:=\{\mu\in\mathscr{P}_p(\mathbb{R}^d):\,\|d_S\|_{L^\infty_\mu}\leq R\text{ and }\operatorname{supp}\mu\subseteq\overline{B(0,\sigma_R)}\},$$

there exists $c_R>0$ such that for every $\mu_0\in\mathscr{P}_p(\mathbb{R}^d)_{|R}$ we have

$$\tilde{T}_p^\Phi(\mu_0)\leq\frac{1}{c_R}\|d_S\|_{L^\infty_{\mu_0}}\leq\frac{R}{c_R}.$$

Proof. According to Proposition 2.2 in [21], the present assumptions imply that there exists a constant $c_R > 0$ such that the classical minimum time function satisfies

$$T(x) \leq \frac{1}{c_R} d_S(x)\,, \tag{3.14}$$

for every $x \in B(0,R) \setminus S$ with $d_S(x) \leq \sigma_R$. Moreover, $T(\cdot)$ is Lipschitz continuous in such set.

Now, the result follows immediately from (3.14) and Corollary 3.2.22. $\square$

Remark 3.2.36. For other controllability conditions generalizing (3.13), the reader may refer e.g. to [37, 58].

Remark 3.2.37. Notice that the result above is, in a certain sense, sharp for $\tilde{T}_p(\mu_0)$ in such mild hypothesis. In particular, although the assumptions of Theorem 3.2.35 imply that the classical minimum time function satisfies $T(x) \leq \dfrac{1}{c_R} d_S(x)$, the natural conjecture $\tilde{T}_p^{\Phi}(\mu_0) \leq \dfrac{1}{c_R}\tilde{d}_{\tilde{S}_p^{\Phi}}(\mu_0)$ in general fails for the generalized minimum time function, as the following example shows.

Example 3.2.38. In $\mathbb{R}^2$, let $S = \{0\}$, $\tilde{S}_p = \tilde{S} = \{\delta_0\}$, $x_0 \in \mathbb{R}^2 \setminus \{0\}$. Define $\mu_0^\lambda := \lambda\delta_0 + (1-\lambda)\delta_{x_0}$, and set $F(x) \equiv \overline{B(0,1)}$ for all $x \in \mathbb{R}^d$. We have that (3.13) is satisfied, since S is convex, and by setting $v_t(x) := -\frac{x}{|x|}$ for $x \neq 0$ and $v_t(0) := 0$, we obtain that $\tilde{T}_p(\mu_0^\lambda) = T(x_0)$ for every $\lambda \in [0,1]$. On the other hand, $\lim\limits_{\lambda\to 1} W_p(\mu_0^\lambda, \delta_0) = 0$, hence the quotient $\tilde{T}_p(\mu_0^\lambda)/\tilde{d}_{\tilde{S}_p}(\mu_0^\lambda)$ is unbounded as $\lambda \to 1$.

3.2.2 Lipschitz continuity of $\tilde{T}_2^{\Phi}$

This section is devoted to the study of sufficient conditions yielding Lipschitz continuity property for the generalized minimum time function once we have the estimate of attainability previously proved in Theorem 3.2.32.

We stress the fact that the lack of a result of continuous dependence on initial data for the continuity equation with no strong regularity hypothesis on the optimal velocity field makes hard to have a property of Lipschitz continuity of the generalized minimum time function. Indeed, in this case this property is not a direct consequence of an attainability result as it is for the classical case with smooth dynamics.

Next result states a relation between the generalized minimum time function, $\tilde{T}_2^{\Phi}$, and the distance from the generalized target set, $\tilde{d}_{\tilde{S}_2^{\Phi}}$. This will be used to prove Lipschitz continuity of $\tilde{T}_2^{\Phi}$ in Theorem 3.2.42. A similar result, called *Petrov's condition*, holds for the correspondent classical objects.

Corollary 3.2.39. *Assume the same hypothesis and notation of Theorem 3.2.32 and that there exists $C > 0$ such that $\mathrm{m}_2(\mu) \leq C$ for all $\mu \in \tilde{S}_2^{\Phi}$. Then $\tilde{T}_2^{\Phi}(\bar{\mu}) \leq D_{C,H,h}(\mathrm{m}_2(\bar{\mu}))\cdot\tilde{d}_{\tilde{S}_2^{\Phi}}(\bar{\mu})$, for all $\bar{\mu} \in \mathscr{P}_2(\mathbb{R}^d)$ such that $\dfrac{2}{h}\sup\limits_{\phi\in\Phi}\displaystyle\int_{\mathbb{R}^d}\phi(x)\,d\bar{\mu}(x) \leq T$.*

Proof. Note that for all $\phi \in \Phi$ it holds

$$|\phi(x) - \phi(y)| \leq H(|x| + |y| + 1)\,|x - y|.$$

Thus for all $\phi \in \Phi$, $\mu' \in \tilde{S}_2^\Phi$ and $\pi \in \Pi(\bar{\mu}, \mu')$, by Hölder's inequality and using the fact that $(a+b+c)^2 \le 3(a^2+b^2+c^2)$ for any $a, b, c \ge 0$, we have

$$\int_{\mathbb{R}^d} \phi(x) d\bar{\mu}(x) - \int_{\mathbb{R}^d} \phi(y) d\mu'(y) \le \iint_{\mathbb{R}^d \times \mathbb{R}^d} H(|x|+|y|+1)\,|x-y|\,d\pi(x,y)$$
$$\le \sqrt{3} H \left[\iint_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2\,d\pi(x,y) \right]^{\frac{1}{2}} \cdot [\mathrm{m}_2(\bar{\mu}) + \mathrm{m}_2(\mu') + 1]^{\frac{1}{2}}.$$

Note that the left hand side is greater than $\displaystyle\int_{\mathbb{R}^d} \phi(x) d\bar{\mu}(x)$, since $\mu' \in \tilde{S}_2^\Phi$. By passing to the infimum w.r.t. $\pi \in \Pi(\bar{\mu}, \mu')$, we get

$$\frac{2}{h} \int_{\mathbb{R}^d} \phi(x) d\bar{\mu}(x) \le \frac{2\sqrt{3}}{h} H [\mathrm{m}_2(\bar{\mu}) + \mathrm{m}_2(\mu') + 1]^{\frac{1}{2}} \cdot W_2(\bar{\mu}, \mu')$$
$$\le \frac{2\sqrt{3}}{h} H [\mathrm{m}_2(\bar{\mu}) + C + 1]^{\frac{1}{2}} \cdot W_2(\bar{\mu}, \mu')$$

Recalling Theorem 3.2.32, the thesis now follows by passing to the supremum w.r.t. $\phi \in \Phi$ and to the infimum w.r.t. $\mu' \in \tilde{S}_2^\Phi$. □

Next two propositions will lead to the Lipschitz continuity result proved in Theorem 3.2.42 through various degrees of generality, giving more relaxed estimates under weaker assumptions.

Proposition 3.2.40. *Assume the same hypothesis and notation of Theorem 3.2.32 and that there exists $C > 0$ such that $\mathrm{m}_2(\bar{\mu}) \le C$ for all $\bar{\mu} \in \tilde{S}_2^\Phi$. Then, for any $Q > 0$ and any $\mu_0^1, \mu_0^2 \in \mathcal{A}_{Q,T,H}^{M,h,\Phi}$, there exists a constant $\mathcal{C}_{H,h,C}(Q) > 0$ such that we have*

$$|\tilde{T}_2^\Phi(\mu_0^1) - \tilde{T}_2^\Phi(\mu_0^2)| \le \mathcal{C}_{H,h,C}(Q) \cdot W_2(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2),$$

for every $\boldsymbol{\eta}^i := \mu_0^i \otimes \eta_x^i \in \mathscr{P}(\mathbb{R}^d \times \Gamma_{\bar{t}})$, $i = 1, 2$, $\bar{t} := \tilde{T}_2^\Phi(\mu_0^1) \wedge \tilde{T}_2^\Phi(\mu_0^2)$, such that $\eta_x^i \in \mathscr{P}(\Gamma_{\bar{t}}^x)$ is concentrated on absolutely continuous solutions of

$$\begin{cases} \dot{\gamma}(t) \in F(\gamma(t)), & \text{for } \mathscr{L}^1\text{-a.e. } 0 < t \le \bar{t} \\ \gamma(0) = x, \end{cases}$$

for μ_0^i-a.e. $x \in \mathbb{R}^d$ and such that if $\tilde{T}_2^\Phi(\mu_0^i) = \bar{t}$, then $\{e_t \sharp \boldsymbol{\eta}^i\}_{t \in [0, \tilde{T}_2^\Phi(\mu_0^i)]} \subseteq \mathscr{P}(\mathbb{R}^d)$ is an optimal trajectory for μ_0^i.

Proof. Fix any $Q > 0$ and set $\mathcal{A} := \mathcal{A}_{Q,T,H}^{M,h,\Phi}$. Let $\mu_0^i \in \mathcal{A}$, $i = 1, 2$, and notice that if μ_0^1 or μ_0^2 belongs to $\tilde{S}_2^\Phi$, the conclusion immediately follows from Corollary 3.2.39. From now on we suppose $\mu_0^i \notin \tilde{S}_2^\Phi$ for $i = 1, 2$. Assume that $t_2 := \tilde{T}_2^\Phi(\mu_0^2) \ge t_1 := \tilde{T}_2^\Phi(\mu_0^1)$.

Notice that $T \ge \dfrac{2}{h} \sup_{\phi \in \Phi} \displaystyle\int_{\mathbb{R}^d} \phi(x)\,d\mu_{t_1}^2(x)$, for every admissible trajectory $t \mapsto \mu_t^2$, $\mu_{|t=0}^2 = \mu_0^2$. Indeed, by Lemma 3.2.30 with $\tau = t_2$, we have

$$\frac{2}{h} \sup_{\phi \in \Phi} \int_{\mathbb{R}^d} \phi(x)\,d\mu_{t_1}^2(x) \le \frac{2}{h} \left(\sup_{\phi \in \Phi} \int_{\mathbb{R}^d} \phi(x)\,d\mu_0^2(x) + G_{M,H}(t_2, \mathrm{m}_2(\mu_0^2)) \right) \le T,$$

where the last inequality comes from the fact that we took $\mu_0^2 \in \mathcal{A}$.

Hence, we can apply Corollary 3.2.39 along with the Dynamic Programming Principle (Theorem 3.2.25) to obtain

$$t_2 \leq t_1 + \tilde{T}_2^{\Phi}(\mu_{t_1}^2) \leq t_1 + D_{C,H,h}(\mathrm{m}_2(\mu_{t_1}^2)) \cdot W_2(\mu_{t_1}^2, \mu_{t_1}^1),$$

for every admissible trajectory $t \mapsto \mu_t^2$, $\mu_{|t=0}^2 = \mu_0^2$, and for every optimal trajectory $t \mapsto \mu_t^1$, $\mu_{|t=0}^1 = \mu_0^1$, since $\mu_{t_1}^1 \in \tilde{S}_2^{\Phi}$.

Let $\bar{t} := \tilde{T}_2^{\Phi}(\mu_0^1) \wedge \tilde{T}_2^{\Phi}(\mu_0^2)$. Let $\boldsymbol{\mu}^i := \{\mu_t^i\}_{t\in[0,\bar{t}]}$, and $\boldsymbol{\eta}^i \in \mathscr{P}(\mathbb{R}^d \times \Gamma_{\bar{t}})$, $i = 1, 2$, be such that $\mu_t^i = e_t \sharp \boldsymbol{\eta}^i$ for $0 \leq t \leq \bar{t}$ as in the Superposition Principle (Theorem 1.3.3). Since the evaluation map e_t is 1-Lipschitz continuous, we have

$$t_2 \leq t_1 + D_{C,H,h}(\mathrm{m}_2(\mu_{t_1}^2)) \cdot W_2(e_{t_1}\sharp\boldsymbol{\eta}^2, e_{t_1}\sharp\boldsymbol{\eta}^1) \leq t_1 + D_{C,H,h}(\mathrm{m}_2(\mu_{t_1}^2)) \cdot W_2(\boldsymbol{\eta}^2, \boldsymbol{\eta}^1),$$

for every $\boldsymbol{\eta}^2$ such that $\boldsymbol{\mu}^2$ is an admissible trajectory and for every $\boldsymbol{\eta}^1$ such that $\boldsymbol{\mu}^1$ is an optimal trajectory. By reversing the roles of μ_0^1 and μ_0^2, we obtain

$$\begin{aligned} |\tilde{T}_2^{\Phi}(\mu_0^2) - \tilde{T}_2^{\Phi}(\mu_0^1)| &\leq \max\{D_{C,H,h}(\mathrm{m}_2(\mu_{t_1}^2)), D_{C,H,h}(\mathrm{m}_2(\mu_{t_2}^1))\} W_2(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2) \\ &\leq \mathcal{C}'_{H,h,C}\left(\mathrm{m}_2(\mu_0^1), \mathrm{m}_2(\mu_0^2)\right) W_2(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2), \end{aligned}$$

for every $\boldsymbol{\eta}^i$ such that $\boldsymbol{\mu}^i$ is an admissible trajectory, $i = 1, 2$, with $\boldsymbol{\mu}^i$ an optimal trajectory if $\tilde{T}_2^{\Phi}(\mu_0^i) = \bar{t}$, and with $\mathcal{C}'_{H,h,C}(\cdot,\cdot)$ coming from estimates in Lemma 3.2.7. Note that $\mathcal{C}'_{H,h,C}(\cdot,\cdot)$ is increasing w.r.t. all the arguments by construction. Hence, the result follows. □

Proposition 3.2.41. *Assume the same hypothesis and notation of Theorem 3.2.32 and that there exists $C > 0$ such that $\mathrm{m}_2(\bar{\mu}) \leq C$ for all $\bar{\mu} \in \tilde{S}_2^{\Phi}$. Furthermore, assume the following:*

(OC) *: there exists a strictly increasing modulus of continuity $\omega^0(\cdot)$ such that for all $\mu \in \mathscr{P}_2(\mathbb{R}^d) \setminus \tilde{S}_2^{\Phi}$ there exists a uniformly continuous vector field $\bar{v}_\mu : \mathbb{R}^d \to \mathbb{R}^d$ with modulus of continuity ω^0, such that the trajectory $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\tilde{T}_2^{\Phi}(\mu)]}$, $\mu_{|t=0} = \mu$, driven by $\boldsymbol{\nu} = \{\bar{v}_\mu \mu_t\}_{t\in[0,\tilde{T}_2^{\Phi}(\mu)]}$, is optimal.*

Then, for any $Q > 0$ and any $\mu_0^1, \mu_0^2 \in \mathcal{A}_{Q,T,H}^{M,h,\Phi}$, there exists a constant $\mathcal{C}_{H,h,C}(Q) > 0$ such that we have

$$|\tilde{T}_2^{\Phi}(\mu_0^1) - \tilde{T}_2^{\Phi}(\mu_0^2)| \leq \mathcal{C}_{H,h,C}(Q) \cdot \left\{ \iint_{\mathbb{R}^d\times\mathbb{R}^d} \left(|x-y|^2 + \left[\psi^{-1}\left(\psi(|x-y|) + \bar{t}\right)\right]^2 \right) d\tilde{\pi}(x,y) \right\}^{\frac{1}{2}},$$

for any $\tilde{\pi} \in \Pi(\mu_0^1, \mu_0^2)$, where $\psi : [0, +\infty] \to [0, +\infty]$ is such that $\dfrac{d\psi}{dr}(r) = \dfrac{1}{\omega^0(r)}$, for all $r \in]0, +\infty[$, and $\bar{t} := \tilde{T}_2^{\Phi}(\mu_0^1) \wedge \tilde{T}_2^{\Phi}(\mu_0^2)$.

Proof. By Proposition 3.2.40, for any $Q > 0$ and any $\mu_0^1, \mu_0^2 \in \mathcal{A}_{Q,T,H}^{M,h,\Phi}$, there exists a constant $\mathcal{C}_{H,h,C}(Q) > 0$ such that we have that the following esimate

$$|\tilde{T}_2^{\Phi}(\mu_0^1) - \tilde{T}_2^{\Phi}(\mu_0^2)| \leq \mathcal{C}_{H,h,C}(Q) \cdot W_2(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2), \tag{3.15}$$

holds in particular for every $\boldsymbol{\eta}^i := \mu_0^i \otimes \eta_x^i \in \mathscr{P}(\mathbb{R}^d \times \Gamma_{\bar{t}})$, $i = 1, 2$, $\bar{t} := \tilde{T}_2^{\Phi}(\mu_0^1) \wedge \tilde{T}_2^{\Phi}(\mu_0^2)$, such that $\eta_x^i \in \mathscr{P}(\Gamma_{\bar{t}}^x)$ is concentrated on absolutely continuous solutions of

$$\begin{cases} \dot{\gamma}(t) = \bar{v}_{\mu_0^i}(\gamma(t)), & \text{for } \mathscr{L}^1\text{-a.e. } 0 < t \leq \bar{t} \\ \gamma(0) = x, \end{cases}$$

for μ_0^j-a.e. $x \in \mathbb{R}^d$ and where $j \in \{1,2\}$ is such that $\tilde{T}_2^\Phi(\mu_0^j) = \bar{t}$, and $\bar{v}_{\mu_0^j}$ is taken as in the current statement, satisfying (OC) with $\mu = \mu_0^j$.

Hence, by (3.15) we have

$$|\tilde{T}_2^\Phi(\mu_0^1) - \tilde{T}_2^\Phi(\mu_0^2)| \leq$$

$$\leq \mathcal{C}_{H,h,C}(Q) \cdot \left\{ \int_{(\mathbb{R}^d \times \Gamma_{\bar{t}}) \times (\mathbb{R}^d \times \Gamma_{\bar{t}})} \left[|x-y|^2 + \|\gamma_x - \gamma_y\|^2 \right] d\pi((x,\gamma_x),(y,\gamma_y)) \right\}^{\frac{1}{2}},$$

for every $\pi \in \Pi(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)$. Notice that for $\boldsymbol{\eta}^1$-a.e. (x, γ_x) and for $\boldsymbol{\eta}^2$-a.e. (y, γ_y) we have

$$z(t) := |\gamma_x(t) - \gamma_y(t)| \leq |x-y| + \int_0^t \left| \bar{v}_{\mu_0^j}(\gamma_x(s)) - \bar{v}_{\mu_0^j}(\gamma_y(s)) \right| ds,$$

for all $t \in [0, \bar{t}]$. Thus, by hypothesis we have that $z(t) \leq z(0) + \displaystyle\int_0^t \omega^0(z(s))\, ds$, and so $\dot{z}(t) \leq \omega^0(z(t))$. By solving $\dot{x}(t) = \omega^0(x(t))$, we get $\psi(x(t)) - \psi(x(0)) = t$, where $\psi : [0,+\infty] \to [0,+\infty]$ is such that $\dfrac{d\psi}{dr}(r) = \dfrac{1}{\omega^0}(r)$. Notice that $\psi(\cdot)$ is invertible since ω^0 is strictly increasing, hence we get $z(t) \leq \psi^{-1}(\psi(z(0)) + t) \leq \psi^{-1}(\psi(z(0)) + \bar{t})$ for all $t \in [0, \bar{t}]$.

By the previous estimate, we obtain

$$|\tilde{T}_2^\Phi(\mu_0^1) - \tilde{T}_2^\Phi(\mu_0^2)| \leq$$

$$\leq \mathcal{C}_{H,h,C}(Q) \cdot \left\{ \int_{(\mathbb{R}^d \times \Gamma_{\bar{t}}) \times (\mathbb{R}^d \times \Gamma_{\bar{t}})} \left[|x-y|^2 + \left[\psi^{-1}(\psi(|x-y|) + \bar{t})\right]^2 \right] d\pi((x,\gamma_x),(y,\gamma_y)) \right\}^{\frac{1}{2}},$$

for every $\pi \in \Pi(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)$. Defining $\tilde{\pi} := (e_0, e_0) \sharp \pi \in \mathscr{P}(\mathbb{R}^d \times \mathbb{R}^d)$, we can easily prove that $\tilde{\pi} \in \Pi(\mu_0^1, \mu_0^2)$. Hence, we conclude that

$$|\tilde{T}_2^\Phi(\mu_0^1) - \tilde{T}_2^\Phi(\mu_0^2)| \leq$$

$$\leq \mathcal{C}_{H,h,C}(Q) \cdot \left\{ \iint_{\mathbb{R}^d \times \mathbb{R}^d} \left[|x-y|^2 + \left[\psi^{-1}(\psi(|x-y|) + \bar{t})\right]^2 \right] d\tilde{\pi}((x,y) \right\}^{\frac{1}{2}},$$

for every $\tilde{\pi} \in \Pi(\mu_0^1, \mu_0^2)$. □

Theorem 3.2.42 (Lipschitz continuity). *Assume (F_0), (F_4) and take M as in (F_4). Let $K, H > 0$, $\Phi \subseteq SC_{K,H}(\mathbb{R}^d;\mathbb{R})$ such that Φ satisfies (T_E) in Definition 3.1.1. Suppose that there exists $C > 0$ such that $\mathrm{m}_2(\bar{\mu}) \leq C$ for all $\bar{\mu} \in \tilde{S}_2^\Phi$.*

Assume that there exist $h, T > 0$ and a modulus of continuity $\tilde{\omega}(\cdot)$ such that for all $\mu \in \mathscr{P}_2(\mathbb{R}^d) \setminus \tilde{S}_2^\Phi$ there exist a continuous vector field $v_\mu \in C^0(\mathbb{R}^d;\mathbb{R}^d)$ and a function $(x,r) \mapsto \omega_x(r)$ in $L^2_{\mu \otimes \mathscr{L}^1}(\mathbb{R}^d \times [0, TM])$ satisfying:

1. *$v_\mu(x) \in F(x)$ for μ-a.e. $x \in \mathbb{R}^d$;*
2. *$\omega_x(\cdot)$ is a nondecreasing modulus of continuity at x for v_μ for μ-a.e. $x \in \mathbb{R}^d$, and*

$$\left(\frac{1}{M} \int_{\mathbb{R}^d} \int_0^M \omega_x^2(rt)\, dr\, d\mu(x) \right)^{\frac{1}{2}} \leq \tilde{\omega}(t),$$

for $0 \leq t \leq T$;

3. *for all* $\phi \in \Phi$ *there exists* $\zeta^{\mu,\phi} \in \mathrm{Bor}(\mathbb{R}^d;\mathbb{R}^d)$ *satisfying* $\zeta^{\mu,\phi}(x) \in \partial^+\phi(x)$ *for* μ*-a.e.* $x \in \mathbb{R}^d$ *and*

$$\int_{\mathbb{R}^d} \langle \zeta^{\mu,\phi}(x), v(x)\rangle d\mu(x) < -h.$$

Furthermore, assume the following

$(OC+)$ *: as in* (OC) *with* $\omega^0(s) := Ls$ *for all* $s \in [0,+\infty]$.

Then $\tilde{T}_2^\Phi(\cdot)$ *is Lipschitz continuous in the set* $\mathcal{A}^{M,h,\Phi}_{Q,T,H} \cap \{\mu \in \mathscr{P}_2(\mathbb{R}^d) : \tilde{d}_{\tilde{S}_2^\Phi}(\mu) \leq R\}$ *for any* $Q, R > 0$.

Proof. The proof follows from Proposition 3.2.41. More precisely, by Proposition 3.2.41, for any $Q > 0$ and any $\mu_0^1, \mu_0^2 \in \mathcal{A}^{M,h,\Phi}_{Q,T,H}$, there exists a constant $\mathcal{C}_{H,h,C}(Q) > 0$ such that we have

$$|\tilde{T}_2^\Phi(\mu_0^1)-\tilde{T}_2^\Phi(\mu_0^2)| \leq \mathcal{C}_{H,h,C}(Q)\cdot\left\{\iint_{\mathbb{R}^d\times\mathbb{R}^d} \left(|x-y|^2 + \left[\psi^{-1}\left(\psi(|x-y|)+\bar{t}\right)\right]^2\right) d\tilde{\pi}(x,y)\right\}^{\frac{1}{2}}, \tag{3.16}$$

for any $\tilde{\pi} \in \Pi(\mu_0^1, \mu_0^2)$, where $\psi : [0,+\infty] \to [0,+\infty]$ is such that $\dfrac{d\psi}{dr}(r) = \dfrac{1}{\omega^0(r)}$, for all $r \in]0,+\infty[$, and $\bar{t} := \tilde{T}_2^\Phi(\mu_0^1) \wedge \tilde{T}_2^\Phi(\mu_0^2)$.

Hence, we can take $\psi(r) = \log r^{\frac{1}{L}}$, $r \in]0,+\infty]$. Then, $\psi^{-1}\left(\psi(|x-y|)+\bar{t}\right) = e^{L\bar{t}}|x-y|$, and by (3.16) we get

$$|\tilde{T}_2^\Phi(\mu_0^1) - \tilde{T}_2^\Phi(\mu_0^2)| \leq (e^{2L\bar{t}}+1)\,\mathcal{C}_{H,h,C}(Q)\cdot\left\{\iint_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2\, d\tilde{\pi}(x,y)\right\}^{\frac{1}{2}},$$

for any $\tilde{\pi} \in \Pi(\mu_0^1, \mu_0^2)$. Thus, by passing to the infimum on $\tilde{\pi} \in \Pi(\mu_0^1, \mu_0^2)$, we have

$$|\tilde{T}_2^\Phi(\mu_0^1) - \tilde{T}_2^\Phi(\mu_0^2)| \leq (e^{2L\bar{t}}+1)\,\mathcal{C}_{H,h,C}(Q)\cdot W_2(\mu_0^1,\mu_0^2).$$

Recalling Corollary 3.2.39, we have $\bar{t} \leq D_{C,H,h}(Q)\cdot R$, in the set $\mathcal{A}^{M,h,\Phi}_{Q,T,H} \cap \{\mu \in \mathscr{P}_2(\mathbb{R}^d) : \tilde{d}_{\tilde{S}_2^\Phi}(\mu) \leq R\}$, for any $Q, R > 0$. This fact yields

$$|\tilde{T}_2^\Phi(\mu_0^2) - \tilde{T}_2^\Phi(\mu_0^1)| \leq \mathcal{C}'_{H,h,C,L}(Q,R)\, W_2(\mu_0^2,\mu_0^1),$$

for a constant $\mathcal{C}'_{H,h,C,L}(Q,R) > 0$, hence Lipschitz continuity of $\tilde{T}_2^\Phi(\cdot)$ in the set $\mathcal{A}^{M,h,\Phi}_{Q,T,H} \cap \{\mu \in \mathscr{P}_2(\mathbb{R}^d) : \tilde{d}_{\tilde{S}_2^\Phi}(\mu) \leq R\}$. □

Remark 3.2.43. Note that requiring assumption $(OC+)$ in the previous theorem is equivalent to ask that the vector field $\bar{v}_\mu$ is globally Lipschitz continuous with $\mathrm{Lip}(\bar{v}_\mu) \leq L$, hence $\mu_t = T_t \sharp \mu$, where $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\tilde{T}_2^\Phi(\mu)]}$ is the optimal trajectory driven by $\boldsymbol{\nu} = \{\bar{v}_\mu \mu_t\}_{t\in[0,\tilde{T}_2^\Phi(\mu)]}$, and $\dot{T}_t(x) = \bar{v}_\mu \circ T_t(x)$, $T_0(x) = x$ for all $x \in \mathbb{R}^d$ and $0 < t \leq \tilde{T}_2^\Phi(\mu)$.

Notice that assumption $(OC+)$ of the previous theorem, which was required in order to get Lipschitz continuity of the generalized minimum time function, is quite demanding. In the following example we show a situation where it is fullfilled.

Example 3.2.44. Let $A \in \mathrm{Mat}_{d\times d}(\mathbb{R})$ be a symmetric matrix satisfying $\lambda_{\max}, |\lambda_{\min}| < 1$, where $\lambda_{\max}$ and $\lambda_{\min}$, are its maximum and minimum eigenvalues, respectively. We study a minimum time problem in the case where the underlying time-optimal control problem in $\mathbb{R}^d$ has the dynamics $\dot{x}(t) \in F(x) := \{Ax+u : u \in \overline{B(0,1)}\}$ and target set $S = \overline{B(0,1)}$. We notice that the classical Petrov's condition holds (see for instance Definition 8.2.2 in [22]), since for all $x \in \partial S$, we have

$$\min_{u\in\overline{B(0,1)}} \langle Ax+u, x\rangle \le \lambda_{\max} + \min_{u\in\overline{B(0,1)}} \langle u, x\rangle = \lambda_{\max} - 1 < 0.$$

Recalling the linearity of the dynamics, by Theorem 5.2 in [21] and Theorem 8.3.4 in [22], we have that the classical minimum time function $T(\cdot)$ is $C^{1,1}$ on every compact set of $\mathbb{R}^d \setminus S$, in particular it is a solution of the Hamilton-Jacobi-Bellman equation

$$-\langle Ax, \nabla T(x)\rangle + |\nabla T(x)| = 1, \qquad \text{in } \mathbb{R}^d \setminus S, \tag{3.17}$$

which implies also

$$\begin{aligned}\liminf_{d_S(x)\to 0^+} |\nabla T(x)| &\ge \liminf_{d_S(x)\to 0^+} \frac{1}{1+|Ax|} \ge \frac{1}{1+\|A\|} > 0,\\ \limsup_{d_S(x)\to 0^+} |\nabla T(x)| &\le \limsup_{d_S(x)\to 0^+} \frac{1}{1-|Ax|} \le \frac{1}{1-\|A\|}.\end{aligned}$$

It can be seen that

$$u^*(x) := \begin{cases} -\dfrac{\nabla T(x)}{|\nabla T(x)|}, & \text{for all } x \in \mathbb{R}^d \setminus S,\\[2ex] \lim\limits_{\substack{\bar{x}\to x\\ \bar{x}\in\mathbb{R}^d\setminus S}} u^*(\bar{x}), & \text{for all } x \in \partial S,\\[2ex] |x| \cdot u^*\left(\dfrac{x}{|x|}\right), & \text{for all } x \in \overset{\circ}{S}\setminus\{0\},\\[2ex] 0, & \text{for } x = 0.\end{cases}$$

is a locally Lipschitz continuous map defined on the whole of $\mathbb{R}^d$ (since $T(\cdot)$ can be extended to a $C^{1,1}$ map defined on $\overline{\mathbb{R}^d \setminus S}$). Set $v(x) := Ax+u^*(x)$, we obtain a locally Lipschitz vector field, which is optimal for the classical problem in $\mathbb{R}^d$. Hence, by taking $\Phi := \{d_S\}$, v is optimal also for the generalized problem by invariance of S w.r.t. v. Indeed, we have $\tilde{T}_2^\Phi(\mu) = \|T(\cdot)\|_{L^\infty_\mu}$ for all $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ (see [31]) and the assumptions of Theorem 3.2.42 are satisfied.

We emphasize again that proving Lipschitz continuity for the generalized minimum time function without requiring strong assumptions yielding Gronwall-like estimates is a difficult task.

For this reason, an interesting open problem would be to investigate regularity of $\tilde{T}_2^\Phi$ with milder assumptions on the dynamics, stating the problem in a suitable *smaller class* of probability measures, for example for measures that are absolutely continuous w.r.t. Lebesgue's measure. In [1–5, 25, 38, 40] there are many results concerning the *Lagrangian flow problem*, i.e. the study of existence, uniqueness and stability properties for the continuity equation restricted to suitable subclasses of $\mathscr{P}(\mathbb{R}^d)$ under very mild regularity assumptions on the driving vector field.

Another possible issue is due to the fact that for several reasons we can be interested in *restricting* the regularity class of the vector field governing the evolution from L^2-selection of F to smoother selections (e.g. Lipschitz or C^1) In particular, this may be a critical issue when we are interested in constructing numerical approximations of the solutions enyoing some stability properties.

A possible way to face these problems is to incorporate such constraints directly in the definition of admissible trajectories, for example by redefining the functional J_F as follows

$$J_F(\boldsymbol{\mu},\boldsymbol{\nu}) := \begin{cases} \displaystyle\int_a^b \int_{\mathbb{R}^d} \left[1 + I_{F(x)}\left(\frac{\nu_t}{\mu_t}(x)\right) + I_{\mathscr{S}}(\mu_t) + I_{[0,M]}\left(\left\|\nabla \frac{\nu_t}{\mu_t}\right\|_{L^\infty}\right)\right] d\mu_t(x)\,dt, & \\ & \text{for } \mathscr{L}^1\text{-a.e. } t \in I, \\ +\infty, & \text{otherwise,} \end{cases} \tag{3.18}$$

where $\mathscr{S} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ is a given class of measures. In this way the finiteness of J_F implies that the evolution occurs only inside a class $\mathscr{S}$ of measures with Lipschitz continuous driving vector fields v_t, with Lipschitz constant less or equal than M.

Many of the results (Theorem 3.2.32, Lemma 3.2.30, Corollary 3.2.39, Proposition 3.2.40, Proposition 3.2.41, Theorem 3.2.42) can be reformulated in this way, with almost identical proofs, but requiring less restrictive assumptions in the statement. For instance, in Theorem 3.2.42 we can drop assumption $(OC+)$ and require that the others hold for all $\mu \in \mathscr{S}\setminus\tilde{S}_2^\Phi$ instead of all $\mu \in \mathscr{P}_2(\mathbb{R}^d)\setminus\tilde{S}_2^\Phi$.

Many other constraints, more related to the nature of the model, can be treated in this way, e.g. penalizing concentration or rarefaction of the agents, or other effects due to the global distribution of the agents.

3.3 Hamilton-Jacobi-Bellman equation

In this section we will prove that under suitable assumptions the generalized minimum time function solves a natural Hamilton-Jacobi-Bellman equation on $\mathscr{P}_2(\mathbb{R}^d)$ in the viscosity sense (Theorem 3.3.9). The notion of viscosity sub-/superdifferential that we are going to use is different from other currently available in literature (e.g. [9, 26, 46, 47]), being modeled on this particular problem.

Throughout this section we will mainly use the alternative definition of admissible curve and the notation provided by Definition 1.0.6 and 3.2.4.

Definition 3.3.1 (Averaged speed set)**.** Assume (F_0) and (F_1), $T > 0$. For any

$\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$, $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$, we set

$$\mathscr{V}(\boldsymbol{\eta}) := \Bigg\{ w_{\boldsymbol{\eta}} \in L^2_{\boldsymbol{\eta}}(\mathbb{R}^d \times \Gamma_T) : \exists \{t_i\}_{i\in\mathbb{N}} \subseteq]0,T[, \text{ with } t_i \to 0^+ \text{ and } \frac{e_{t_i} - e_0}{t_i} \rightharpoonup w_{\boldsymbol{\eta}} \text{ weakly in } L^2_{\boldsymbol{\eta}}(\mathbb{R}^d \times \Gamma_T; \mathbb{R}^d) \Bigg\}.$$

We notice that, according to the boundedness result of Lemma 3.2.7 (iii), for any sequence $\{t_i\}_{i\in\mathbb{N}} \subseteq]0,T[$ with $t_i \to 0^+$, there exists a subsequence $\tau = \{t_{i_k}\}_{k\in\mathbb{N}}$ and $w_{\boldsymbol{\eta}} \in L^2_{\boldsymbol{\eta}}(\mathbb{R}^d \times \Gamma_T; \mathbb{R}^d)$ such that $\dfrac{e_{t_{i_k}} - e_0}{t_{i_k}}$ weakly converges to an element of $L^2_{\boldsymbol{\eta}}(\mathbb{R}^d \times \Gamma_T; \mathbb{R}^d)$, thus $\mathscr{V}(\boldsymbol{\eta}) \neq \emptyset$.

Lemma 3.3.2 (Properties of the averaged speed set)**.** *Assume* (F_0) *and* (F_1)*,* $T > 0$*. For any* $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$*,* $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$ *and every* $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$ *we have that*

(i) $w_{\boldsymbol{\eta}}(x,\gamma) \in F(\gamma(0))$ *for* $\boldsymbol{\eta}$*-a.e* $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$*.*

(ii) if we denote by $\{\eta_x\}_{x\in\mathbb{R}^d}$ *the disintegration of* $\boldsymbol{\eta}$ *w.r.t. the map* e_0*, the map*

$$x \mapsto \int_{\Gamma^x_T} w_{\boldsymbol{\eta}}(x,\gamma)\, d\eta_x(\gamma),$$

belongs to $L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)$*.*

Proof. We prove (i). Fix $\varepsilon > 0$ and $(x,\gamma) \in \operatorname{supp}\boldsymbol{\eta}$. Since $\gamma(\cdot)$ and $F(\cdot)$ are continuous, there exists $t^*_{\varepsilon,\gamma} > 0$ such that for all $0 < t < t^*_{\varepsilon,\gamma}$ we have $F(\gamma(t)) \subseteq F(\gamma(0)) + \varepsilon B(0,1)$. In particular, for all $0 < t < t^*_{\varepsilon,\gamma}$ and $v \in \mathbb{R}^d$ we have

$$\begin{aligned}\langle v, \varphi_t(x,\gamma)\rangle = \langle v, \frac{\gamma(t)-\gamma(0)}{t}\rangle &= \frac{1}{t}\int_0^t \langle v, \dot\gamma(s)\rangle\, ds \\ &\leq \frac{1}{t}\int_0^t \sigma_{F(\gamma(s))}(v)\, ds \leq \sigma_{F(\gamma(0))+\varepsilon B(0,1)}(v),\end{aligned}$$

where $\varphi_t(x,\gamma) = \dfrac{e_t(x,\gamma) - e_0(x,\gamma)}{t}$.

Thus

$$\overline{\text{co}}\{\varphi_t(x,\gamma) :\, 0 < t < t^*_{\varepsilon,\gamma}\} \subseteq F(\gamma(0)) + \varepsilon\overline{B(0,1)}$$

Given $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$, let $\{t_i\}_{i\in\mathbb{N}} \subseteq]0,1]$ be a sequence such that $t_i \to 0^+$ and $\varphi_{t_i} \rightharpoonup w_{\boldsymbol{\eta}}$ weakly in $L^2_{\boldsymbol{\eta}}$. In particular, by Mazur's Lemma, there is a sequence in $\text{co}\{\varphi_{t_i} :\, i \in \mathbb{N}\}$ strongly convergent to $w_{\boldsymbol{\eta}}$. In particular, for (x,γ)-a.e. point of $\mathbb{R}^d \times \Gamma_T$ we have pointwise convergence, i.e.

$$w_{\boldsymbol{\eta}}(x,\gamma) \in \overline{\text{co}}\{\varphi_{t_i}(x,\gamma) :\, i \in \mathbb{N}\}.$$

Given a point (x,γ) where above pointwise convergence occurs, we can consider a subsequence $\{t_{i_k}\}_{k\in\mathbb{N}}$ of t_i satisfying $0 < t_{i_k} < t^*_{\varepsilon,\gamma}$, obtaining that

$$\begin{aligned}w_{\boldsymbol{\eta}}(x,\gamma) \in \overline{\text{co}}\{\varphi_{t_{i_k}}(x,\gamma) :\, k \in \mathbb{N}\}) &\subseteq \overline{\text{co}}\{\varphi_t(x,\gamma) :\, 0 < t < t^*_{\varepsilon,\gamma}\} \\ &\subseteq F(\gamma(0)) + \varepsilon\overline{B(0,1)}.\end{aligned}$$

By letting $\varepsilon \to 0^+$ we have that $w_{\boldsymbol{\eta}}(x,\gamma) \in F(\gamma(0))$ for $\boldsymbol{\eta}$-a.e. $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$.

We prove now (ii). By definition, the disintegration of $\boldsymbol{\eta}$ w.r.t. the evaluation map e_0 is a family of measures $\{\eta_x\}_{x\in\mathbb{R}^d}$ satisfying (recall that $e_0\sharp\boldsymbol{\eta} = \mu_0$)

$$\iint_{\mathbb{R}^d\times\Gamma_T} f(x,\gamma)w_{\boldsymbol{\eta}}(x,\gamma)\,d\boldsymbol{\eta}(x,\gamma) = \int_{\mathbb{R}^d}\left(\int_{\Gamma_T^x}\langle f(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle\,d\eta_x(\gamma)\right)d\mu_0(x),$$

for all Borel map $f : \mathbb{R}^d \times \Gamma_T \to \mathbb{R}^d$. Moreover the family $\{\eta_x\}_{x\in\mathbb{R}^d}$ is uniquely determined for μ_0-a.e. $x \in \mathbb{R}^d$ (see e.g. Theorem 5.3.1 in [9]).

For any $\psi \in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)$, clearly we have $\psi\circ e_0 \in L^2_{\boldsymbol{\eta}}(\mathbb{R}^d\times\Gamma_T;\mathbb{R}^d)$, since $e_0\sharp\boldsymbol{\eta} = \mu_0$. Recalling that $w_{\boldsymbol{\eta}} \in L^2_{\boldsymbol{\eta}}$, we obtain

$$\begin{aligned}\int_{\mathbb{R}^d}\langle\psi(x),\int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(x,\gamma)\,d\eta_x(\gamma)\rangle\,d\mu_0(x) &= \int_{\mathbb{R}^d}\int_{\Gamma_T^x}\langle\psi(x), w_{\boldsymbol{\eta}}(x,\gamma)\rangle\,d\eta_x(\gamma)\,d\mu_0(x)\\ &= \int_{\mathbb{R}^d}\int_{\Gamma_T^x}\langle\psi\circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle\,d\eta_x(\gamma)\,d\mu_0(x)\\ &= \iint_{\mathbb{R}^d\times\Gamma_T}\langle\psi\circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle\,d\boldsymbol{\eta}(x,\gamma) < +\infty.\end{aligned}$$

By the arbitrariness of $\psi \in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)$, we obtain that the map

$$x \mapsto \int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(x,\gamma)\,d\eta_x(\gamma),$$

belongs to $L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)$, moreover for μ_0-a.e. $x \in \mathbb{R}^d$, we have from (i) that

$$\int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(\gamma)\,d\eta_x(\gamma) \in F(x).$$

□

Remark 3.3.3. We can interpret each $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$ as a sort of averaged vector field of initial velocity in the sense of measure (we recall that in general an admissible trajectory γ may fail to possess a tangent vector at $t = 0$). The map

$$x \mapsto \int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(\gamma)\,d\eta_x(\gamma),$$

can be interpreted as a *initial barycentric speed* of all the (weighted) trajectories emanating from x in the support of $\boldsymbol{\eta}$. This approach is quite related to Theorem 5.4.4. in [9].

In the case in which the trajectory $t \mapsto e_t\sharp\boldsymbol{\eta}$ is driven by a sufficient smooth vector field, we recover exactly as averaged vector field and initial barycentric speed the expected objects, as shown below.

Lemma 3.3.4 (Regular driving vector fields)**.** *Assume* (F_0), (F_1) *and let* $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$. *Let* $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ *be an absolutely continuous solution of*

$$\begin{cases}\partial_t\mu_t + \operatorname{div}(v_t\mu_t) = 0,\ t \in]0,T[\\ \\ \mu_{|t=0} = \mu_0,\end{cases}$$

where $v \in C^0([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ *satisfies* $v_0(x) \in F(x)$ *for all* $x \in \mathbb{R}^d$. *Then if* $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$ *satisfies* $\mu_t = e_t \sharp \boldsymbol{\eta}$ *for all* $t \in [0,T]$, *we have that*

$$\lim_{t\to 0} \left\| \frac{e_t - e_0}{t} - v_0 \circ e_0 \right\|_{L^2_{\boldsymbol{\eta}}} = 0,$$

and so $\mathscr{V}(\boldsymbol{\eta}) = \{v_0 \circ e_0\}$, *thus we have*

$$\left\{ x \mapsto \int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(x,\gamma)\, d\eta_x : w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta}) \right\} = \{v_0(\cdot)\}.$$

Proof. We have

$$\left\| \frac{e_t - e_0}{t} - v_0 \circ e_0 \right\|^2_{L^2_{\boldsymbol{\eta}}} = \iint_{\mathbb{R}^d \times \Gamma_T} \left| \frac{\gamma(t) - \gamma(0)}{t} - v_0(\gamma(0)) \right|^2 d\boldsymbol{\eta}(x,\gamma),$$

For $\boldsymbol{\eta}$-a.e. $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$, by continuity of v we have $\gamma \in C^1$ and $\dot\gamma(t) = v_t(\gamma(t))$, hence for t small enough we get

$$\begin{aligned}\left| \frac{\gamma(t) - \gamma(0)}{t} - v_0(\gamma(0)) \right| &\le \frac{1}{t}\int_0^t |\dot\gamma(s)|\, ds + |v_0(\gamma(0))| = \frac{1}{t}\int_0^t |v_s(\gamma(s))|\, ds + |v_0(\gamma(0))| \\ &\le 2|v_0(\gamma(0))| + 1 \in L^2_{\boldsymbol{\eta}},\end{aligned}$$

indeed by (F_1) we have

$$\begin{aligned}\iint_{\mathbb{R}^d \times \Gamma_T} |v_0(\gamma(0))|^2\, d\boldsymbol{\eta}(x,\gamma) &= \int_{\mathbb{R}^d} |v_0(x)|^2\, d\mu_0(x) \le C^2 \int_{\mathbb{R}^d} (|x|+1)^2\, d\mu_0(x) \\ &\le 2C^2\left(\mathrm{m}_2(\mu_0) + 1\right),\end{aligned}$$

with $C > 0$ as in (F_1). Thus, for $\boldsymbol{\eta}$-a.e. $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$,

$$\lim_{t\to 0^+} \left| \frac{\gamma(t) - \gamma(0)}{t} - v_0(\gamma(0)) \right| = 0.$$

Thus applying Lebesgue's Dominated Convergence Theorem we obtain

$$\lim_{t\to 0} \left\| \frac{e_t - e_0}{t} - v_0 \circ e_0 \right\|^2_{L^2_{\boldsymbol{\eta}}} = 0,$$

hence $w_{\boldsymbol{\eta}} = v_0 \circ e_0$. The last assertion now follows. □

We have already proved that the set

$$\left\{ x \mapsto \int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(x,\gamma)\, d\eta_x : \boldsymbol{\eta} \in \mathscr{T}_F(\mu_0),\ w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta}) \right\}$$

is contained in the set of all $L^2_{\mu_0}(\mathbb{R}^d; \mathbb{R}^d)$-selections of $F(\cdot)$. The next density result shows that, indeed, equality holds: since allows to approximate every $L^2_{\mu_0}$-selections by C^0-selections, and then use Lemma 3.3.4. This will be the main ingredient used to prove Theorem 3.3.9, i.e. that the generalized minimum time function is a solution of an Hamilton-Jacobi-Bellman equation in a suitable viscosity sense.

Lemma 3.3.5 (Approximation). *Let $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$. Assume (F_0) and (F_1). Then given any $v \in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)$ satisfying $v(x) \in F(x)$ for μ_0-a.e. $x \in \mathbb{R}^d$, there exists a sequence of continuous maps $\{g_n\}_{n\in\mathbb{N}} \subseteq C^0(\mathbb{R}^d;\mathbb{R}^d)$ such that*

1. $\lim\limits_{n\to\infty} \|g_n - v\|_{L^2_{\mu_0}} = 0$;
2. $g_n(x) \in F(x)$ *for all* $x \in \mathbb{R}^d$.

In particular, we have

$$\{v \in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d) : v(x) \in F(x) \text{ for } \mu_0\text{-a.e. } x \in \mathbb{R}^d\} =$$
$$= \left\{ x \mapsto \int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(x,\gamma)\, d\eta_x : \boldsymbol{\eta} \in \mathscr{T}_F(\mu_0),\, w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta}) \right\}.$$

Proof. By Lusin's Theorem (see e.g. Theorem 1.45 in [6]), we can construct a sequence of compact sets $\{K_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^d$ and of continuous maps $\{v_n\}_{n\in\mathbb{N}} \subseteq C^0_c(\mathbb{R}^d;\mathbb{R}^d)$ such that $v_n = v$ on K_n and $\mu_0(\mathbb{R}^d \setminus K_n) \le 1/n$. For all $n \in \mathbb{N}$ define the set valued maps

$$G_n(x) := \begin{cases} F(x), & \text{for } x \in \mathbb{R}^d \setminus K_n, \\ \\ \{v_n(x)\}, & \text{for } x \in K_n. \end{cases}$$

We prove that $G_n(\cdot)$ is lower semicontinuous. If $x \in \mathbb{R}^d \setminus K_n$, then in a neighborhood of x we have $G_n = F$, thus G_n is lower semicontinuous. Let $x \in K_n$ and V be an open set such that $V \cap G_n(x) \neq \emptyset$. In particular, we have that V is an open neighborhood of $v_n(x)$. Without loss of generality, we may assume that $V = B(v_n(x),\varepsilon)$ for $\varepsilon > 0$, thus there exists $\delta > 0$ such that if $y \in B(x,\delta) \cap K_n$ we have $v_n(y) \in V$, and so $G_n(y) \cap V \neq \emptyset$. On the other hand, by continuity of F, there exists an open neighborhood U of x such that $V \cap F(y) \neq \emptyset$ for all $y \in U$. Thus, if we set $U' = U \cap B(x,\delta) \setminus K_n$, we have that U' is an open neighborhood of x satisfying:

(*a*) for all $y \in U' \setminus K_n$ we have $F(y) = G_n(y)$ and so $G_n(y) \cap V \neq \emptyset$;

(*b*) for all $y \in U' \cap K_n$ we have $v_n(y) \in V$, and so $G_n(y) \cap V \neq \emptyset$;

and so given V for all $y \in U'$ we have $G_n(y) \cap V \neq \emptyset$, which proves lower semicontinuity. Since $G_n(\cdot)$ is lower semicontinuous with compact convex values, by Michael's Selection Theorem (see e.g. Theorem 9.1.2 in [13]) we can find a continuous selection $g_n \in C^0(\mathbb{R}^d;\mathbb{R}^d)$ which by construction agrees with v and v_n on K_n and satisfies $g_n(x) \in G_n(x) \subseteq F(x)$ for all $x \in \mathbb{R}^d$. Finally, we have

$$\begin{aligned}\int_{\mathbb{R}^d} |v(x) - g_n(x)|^2\, d\mu_0(x) &= \int_{\mathbb{R}^d\setminus K_n} |v(x) - g_n(x)|^2\, d\mu_0(x) \\ &\le \int_{\mathbb{R}^d\setminus K_n} 4C^2(|x|+1)^2\, d\mu_0(x) \le 8C^2\left(\mathrm{m}_2(\mu_0) + 1\right),\end{aligned}$$

with $C > 0$ as in (F_1), hence (1) follows. The last assertion comes from Lemma 3.3.4 with $v = v_0$. ☐

We introduce now the following definition of viscosity sub-/superdifferential. For other concepts of viscosity sub-/superdifferential, we refer the reader to [9, 26].

Definition 3.3.6 (Sub-/Super-differential in $\mathscr{P}_2(\mathbb{R}^d)$)**.** Let $V : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be a function. Fix $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\delta > 0$. We say that $p_\mu \in L^2_\mu(\mathbb{R}^d; \mathbb{R}^d)$ belongs to the δ-superdifferential $D^+_\delta V(\mu)$ at μ if for all $T > 0$ and $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ such that $t \mapsto e_t \sharp \boldsymbol{\eta}$ is an absolutely continuous curve in $\mathscr{P}_2(\mathbb{R}^d)$ defined in $[0, T]$ with $e_0 \sharp \boldsymbol{\eta} = \mu$ we have

$$\limsup_{t\to 0^+} \frac{V(e_t\sharp\boldsymbol{\eta}) - V(e_0\sharp\boldsymbol{\eta}) - \displaystyle\iint_{\mathbb{R}^d\times\Gamma_T} \langle p_\mu \circ e_0(x,\gamma), e_t(x,\gamma) - e_0(x,\gamma)\rangle \, d\boldsymbol{\eta}(x,\gamma)}{\|e_t - e_0\|_{L^2_{\boldsymbol{\eta}}}} \leq \delta. \tag{3.19}$$

In the same way, $q_\mu \in L^2_\mu(\mathbb{R}^d; \mathbb{R}^d)$ belongs to the δ-subdifferential $D^-_\delta V(\mu)$ at μ if $-q_\mu \in D^+_\delta[-V](\mu)$. Moreover, $D^\pm_\delta[V](\mu)$ is the closure in L^2_μ of $D^\pm_\delta[V](\mu) \cap C^0_b(\mathbb{R}^d; \mathbb{R}^d)$.

Definition 3.3.7 (Viscosity solutions)**.** Let $V : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be a function and $\mathscr{H} : T^*\mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$. We say that V is a

1. *viscosity supersolution* of $\mathscr{H}(\mu, DV(\mu)) = 0$ if V is l.s.c. and there exists $C > 0$ depending only on $\mathscr{H}$ such that $\mathscr{H}(\mu, q_\mu) \geq -C\delta$ for all $q_\mu \in D^-_\delta V(\mu)$, $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, and for all $\delta > 0$.

2. *viscosity subsolution* of $\mathscr{H}(\mu, DV(\mu)) = 0$ if V is u.s.c. and there exists $C > 0$ depending only on $\mathscr{H}$ such that $\mathscr{H}(\mu, p_\mu) \leq C\delta$ for all $p_\mu \in D^+_\delta V(\mu)$, $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, and for all $\delta > 0$.

3. *viscosity solution* of $\mathscr{H}(\mu, DV(\mu)) = 0$ if it is both a viscosity subsolution and a viscosity supersolution.

Definition 3.3.8 (Hamiltonian Function)**.** Given $\mu \in \mathscr{P}(\mathbb{R}^d)$, define

$$\mathscr{D}(\mu) := \left\{ \nu \in \mathscr{M}(\mathbb{R}^d; \mathbb{R}^d) : |\nu| \ll \mu \text{ and } \int_{\mathbb{R}^d} \left(\left| \frac{\nu}{\mu} \right|^2 + I_{F(x)}\left(\frac{\nu}{\mu}(x) \right) \right) d\mu < +\infty \right\}.$$

Since the tangent space $T_\mu \mathscr{P}_2(\mathbb{R}^d)$ to $\mathscr{P}_2(\mathbb{R}^d)$ at $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ is $L^2_\mu(\mathbb{R}^d; \mathbb{R}^d)$, which coincides with its dual, we can define a map $\mathscr{H}_F : T^*\mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ by setting

$$\begin{aligned}
\mathscr{H}_F(\mu, \psi) &:= -\left[1 + \inf_{\nu \in \mathscr{D}(\mu)} \int_{\mathbb{R}^d} \langle \psi(x), \frac{\nu}{\mu}(x) \rangle \, d\mu \right], \\
&= -\left[1 + \inf_{\substack{v \in L^2_\mu(\mathbb{R}^d;\mathbb{R}^d) \\ v(x) \in F(x) \text{ for } \mu\text{-a.e. } x}} \int_{\mathbb{R}^d} \langle \psi(x), v(x) \rangle \, d\mu \right],
\end{aligned}$$

where $(\mu, \psi) \in T^*\mathscr{P}_2(\mathbb{R}^d)$, i.e., $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\psi \in L^2_\mu(\mathbb{R}^d; \mathbb{R}^d)$.

If we assume (F_4), or more generally that F possesses a Borel selection uniformly bounded, we have

$$\mathscr{H}_F(\mu,\psi) := -1 + \int_{\mathbb{R}^d} \sigma_{-F(x)}(\psi(x))\, d\mu,$$

by using a consequence of classical Measurable Selection Lemma (see e.g. Theorem 6.31 p. 119 in [35]).

Now, we can prove the main result of this chapter.

Theorem 3.3.9 (Viscosity solution)**.** *Let $\mathcal{A}$ be any open subset of $\mathscr{P}_2(\mathbb{R}^d)$ with uniformly bounded $2-$moments. Assume (F_0) and (F_1) and that $\tilde{T}_2^{\Phi}(\cdot)$ is continuous on $\mathcal{A}$. Then $\tilde{T}_2^{\Phi}(\cdot)$ is a viscosity solution of $\mathscr{H}_F(\mu, D\tilde{T}_2^{\Phi}(\mu)) = 0$ on $\mathcal{A}$, with $\mathscr{H}_F$ defined as in Definition 3.3.8.*

Proof. The proof is splitted in two claims.

Claim 1: $\tilde{T}_2^{\Phi}(\cdot)$ is a subsolution of $\mathscr{H}_F(\mu, D\tilde{T}_2^{\Phi}(\mu)) = 0$ on $\mathcal{A}$.

Proof of Claim 1. Let $\mu_0 \in \mathcal{A}$. Given $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$ and set $\mu_t = e_t \sharp \boldsymbol{\eta}$ for all t, by the Dynamic Programming Principle (Theorem 3.2.25) we have $\tilde{T}_2^{\Phi}(\mu_0) \le \tilde{T}_2^{\Phi}(\mu_s) + s$ for all $0 < s \le \tilde{T}_2^{\Phi}(\mu_0)$. Without loss of generality, we can assume $0 < s < 1$. Given any $p_{\mu_0} \in D_\delta^+ \tilde{T}_2^{\Phi}(\mu_0)$, and set

$$\begin{aligned}
A(s, p_{\mu_0}, \boldsymbol{\eta}) &:= -\, s - \iint_{\mathbb{R}^d\times\Gamma_T} \langle p_{\mu_0} \circ e_0(x,\gamma), e_s(x,\gamma) - e_0(x,\gamma)\rangle \, d\boldsymbol{\eta},\\
B(s, p_{\mu_0}, \boldsymbol{\eta}) &:= \tilde{T}_2^{\Phi}(\mu_s) - \tilde{T}_2^{\Phi}(\mu_0) - \iint_{\mathbb{R}^d\times\Gamma_T} \langle p_{\mu_0} \circ e_0(x,\gamma), e_s(x,\gamma) - e_0(x,\gamma)\rangle d\boldsymbol{\eta},
\end{aligned}$$

we have $A(s, p_{\mu_0}, \boldsymbol{\eta}) \le B(s, p_{\mu_0}, \boldsymbol{\eta})$.

We recall that since by definition $p_{\mu_0} \in L^2_{\mu_0}$, we have that $p_{\mu_0} \circ e_0 \in L^2_{\boldsymbol{\eta}}$. Dividing by $s > 0$, we obtain that

$$\limsup_{s\to 0^+} \frac{A(s, p_{\mu_0}, \boldsymbol{\eta})}{s} \ge -1 - \iint_{\mathbb{R}^d\times\Gamma_T} \langle p_{\mu_0} \circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\boldsymbol{\eta}(x,\gamma),$$

for all $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$.

Recalling the choice of p_{μ_0}, we have

$$\limsup_{s\to 0^+} \frac{B(s, p_{\mu_0}, \boldsymbol{\eta})}{s} = \limsup_{s\to 0^+} \frac{B(s, p_{\mu_0}, \boldsymbol{\eta})}{\|e_s - e_0\|_{L^2_{\boldsymbol{\eta}}}} \cdot \left\| \frac{e_s - e_0}{s} \right\|_{L^2_{\boldsymbol{\eta}}} \le K\delta,$$

where $K > 0$ is a suitable constant coming from Lemma 3.2.7 and from hypothesis.

We thus obtain for all $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$ and all $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$, that

$$1 + \iint_{\mathbb{R}^d\times\Gamma_T} \langle p_{\mu_0} \circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\boldsymbol{\eta}(x,\gamma) \ge -K\delta.$$

By passing to the infimum on $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$ and $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$, and recalling Lemma 3.3.5, we have

$$\begin{aligned}
-K\delta &\leq 1 + \inf_{\substack{\boldsymbol{\eta}\in\mathscr{T}_F(\mu_0)\\ w_{\boldsymbol{\eta}}\in\mathscr{V}(\boldsymbol{\eta})}} \iint_{\mathbb{R}^d\times\Gamma_T} \langle p_{\mu_0}\circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\boldsymbol{\eta}(x,\gamma)\\
&= 1 + \inf_{\substack{\boldsymbol{\eta}\in\mathscr{T}_F(\mu_0)\\ w_{\boldsymbol{\eta}}\in\mathscr{V}(\boldsymbol{\eta})}} \int_{\mathbb{R}^d}\int_{\Gamma_T^x} \langle p_{\mu_0}\circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\eta_x\, d\mu_0\\
&= 1 + \inf_{\substack{\boldsymbol{\eta}\in\mathscr{T}_F(\mu_0)\\ w_{\boldsymbol{\eta}}\in\mathscr{V}(\boldsymbol{\eta})}} \int_{\mathbb{R}^d} \langle p_{\mu_0}\circ e_0(x,\gamma), \int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(x,\gamma)\, d\eta_x\rangle \, d\mu_0\\
&= 1 + \inf_{\substack{v\in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)\\ v(x)\in F(x)\ \mu_0\text{-a.e. } x}} \int_{\mathbb{R}^d} \langle p_{\mu_0}, v\rangle \, d\mu_0 = -\mathscr{H}_F(\mu_0, p_{\mu_0}),
\end{aligned}$$

so $\tilde{T}_2^\Phi(\cdot)$ is a subsolution, thus confirming Claim 1. $\diamond$

Claim 2: $\tilde{T}_2^\Phi(\cdot)$ is a supersolution of $\mathscr{H}_F(\mu, D\tilde{T}_2^\Phi(\mu)) = 0$ on $\mathcal{A}$.

Proof of Claim 2. Let $\mu_0 \in \mathcal{A}$. Given $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$ and defined the admissible trajectory $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} = \{e_t\sharp\boldsymbol{\eta}\}_{t\in[0,T]}$, and $q_{\mu_0} \in D^-_\delta \tilde{T}_2^\Phi(\mu_0)$, there is a sequence $\{s_i\}_{i\in\mathbb{N}} \subseteq]0,T[$ and $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$ such that $s_i \to 0^+$, $\dfrac{e_{s_i}-e_0}{s_i}$ weakly converges to $w_{\boldsymbol{\eta}}$ in $L^2_{\boldsymbol{\eta}}$, and for all $i \in \mathbb{N}$

$$\begin{aligned}
\iint_{\mathbb{R}^d\times\Gamma_T} &\langle q_{\mu_0}\circ e_0(x,\gamma), \frac{e_{s_i}(x,\gamma)-e_0(x,\gamma)}{s_i}\rangle \, d\boldsymbol{\eta}(x,\gamma)\\
&\leq 2\delta \left\| \frac{e_{s_i}-e_0}{s_i} \right\|_{L^2_{\boldsymbol{\eta}}} - \frac{\tilde{T}_2^\Phi(\mu_0)-\tilde{T}_2^\Phi(\mu_{s_i})}{s_i}.
\end{aligned}$$

By taking i sufficiently large we thus obtain

$$\iint_{\mathbb{R}^d\times\Gamma_T} \langle q_{\mu_0}\circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\boldsymbol{\eta}(x,\gamma) \leq 3K\delta - \frac{\tilde{T}_2^\Phi(\mu_0)-\tilde{T}_2^\Phi(\mu_{s_i})}{s_i}.$$

By using Lemma 3.3.5 and arguing as in Claim 1, we have

$$\inf_{\substack{\boldsymbol{\eta}\in\mathscr{T}_F(\mu_0)\\ w_{\boldsymbol{\eta}}\in\mathscr{V}(\boldsymbol{\eta})}} \iint_{\mathbb{R}^d\times\Gamma_T} \langle q_{\mu_0}\circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\boldsymbol{\eta}(x,\gamma) = -\mathscr{H}_F(\mu_0, q_{\mu_0}) - 1,$$

and so

$$\mathscr{H}_F(\mu_0, q_{\mu_0}) \geq -3K\delta + \frac{\tilde{T}_2^\Phi(\mu_0)-\tilde{T}_2^\Phi(\mu_{s_i})}{s_i} - 1.$$

By the Dynamic Programming Principle, passing to the infimum on all admissible curves and recalling that $\dfrac{\tilde{T}_2^\Phi(\mu_0)-\tilde{T}_2^\Phi(\mu_s)}{s} - 1 \leq 0$ with equality holding if and only if $\boldsymbol{\mu}$ is optimal, we obtain $\mathscr{H}_F(\mu_0, q_{\mu_0}) \geq -C'\delta$, which proves that $\tilde{T}_2^\Phi(\cdot)$ is a supersolution, thus confirming Claim 2. $\square$

Remark 3.3.10. Unfortunately, we have that $\tilde{T}_2^\Phi(\cdot)$ in general fails to be continuous, being just lower semicontinuous. Moreover, it seems to be quite a difficult problem to provide general necessary and sufficient conditions on problem data granting this continuity property. Thus, an open problem is the extension of the definition of viscosity solutions and the subsequent result on Hamilton-Jacobi-Bellman equation, to the case where we have only lower semicontinuity of the minimum time function, instead of continuity, in spirit of Barron-Jensen's approach to viscosity solutions.

Anyway, regarding the present result, as seen in Theorem 3.2.42, we can give sufficient conditions for local Lipschitz continuity of $\tilde{T}_2^\Phi(\cdot)$. In the following we will provide simple examples in which this sufficient conditions are not satisfied, but it is still possible to have continuity of the minimum time function.

Example 3.3.11. In $\mathbb{R}^2$, take $\Phi = \{\phi\}$, where $\phi(x,y) = 1 - \displaystyle\int_{-\infty}^{x} e^{-|s|}ds \in C_b^1 \cap \mathrm{Lip}(\mathbb{R}^2;\mathbb{R})$ and denote with L the Lipschitz constant of ϕ. Observe that $\partial_x\phi(x,y) = -e^{-|x|} < 0$ and $\partial_x\phi \in C_b^0$. Let $F(x,y) := \{(\alpha,0) : \alpha \in [0,1]\}$, $\mu_0 \in \mathscr{P}_2(\mathbb{R}^2)$. If we denote with $t \mapsto \gamma(t)$ an absolutely continuous solution of the characteristic system

$$\begin{cases} \dot{\gamma}(t) \in F(\gamma(t)), & t > 0, \\ \gamma(0) = (x,y), \end{cases}$$

we have $\phi \circ \gamma(t) = \phi(x + \int_0^t \alpha(s)\,ds, y) \geq \phi(x+t,y)$.

Thus, every trajectory $\boldsymbol{\mu} = \{\mu_t\}_{t>0}$, starting with $\mu|_{t=0} = \mu_0$, and defined by by $\mu_t = (\mathrm{Id} + tv)\sharp\mu_0$ for $v = (1,0)$ is optimal for μ_0.

Moreover, if we define $G : [0,+\infty[\times\mathscr{P}_2(\mathbb{R}^2) \to \mathbb{R}$ by setting

$$G(t,\mu_0) := \int_{\mathbb{R}^2} \phi((x,y)+tv)\,d\mu_0 = \int_{\mathbb{R}^2} \phi(x,y)\,d\mu_t(x,y),$$

we have that $\mu_t \in \tilde{S}_2^\Phi$ if and only if $G(t,\mu_0) \leq 0$, thus

$$\tilde{T}_2^\Phi(\mu_0) = \inf\{t \geq 0 : G(t,\mu_0) = 0\},$$

due to the strictly decreasing property of $G(t,\mu_0)$ w.r.t. t (due to the sign of $\partial_x\phi$).

In order to prove continuity of $\tilde{T}_2^\Phi(\cdot)$ we use the same procedure of Dini's theorem.

First, observe that for $G : [0,+\infty[\times\mathscr{P}_2(\mathbb{R}^2) \to \mathbb{R}$ we have

$$\frac{\partial}{\partial t}G(t,\mu) = \iint_{\mathbb{R}^2} \partial_x\phi(x+t,y)\,d\mu = \iint_{\mathbb{R}^2} -e^{-|x+t|}d\mu < 0.$$

Furthermore the map $t \mapsto G(t,\mu)$ is continuous $\forall \mu \in \mathscr{P}_2(\mathbb{R}^2)$ by dominated convergence theorem, and $\mu \mapsto G(t,\mu)$ is continuous $\forall t \geq 0$ since $\phi \in C_b^0$. The function G is also jointly continuous w.r.t. both variables, indeed

$$|G(t_n,\mu_n) - G(t,\mu)| \leq |G(t_n,\mu) - G(t,\mu)| + |G(t_n,\mu_n) - G(t_n,\mu)|,$$

where the first term tends to zero for $n \to +\infty$ by continuity of G w.r.t. t. Focusing on the second term, by Kantorovich duality and Hölder inequality, we get

$$\begin{aligned}|G(t_n,\mu_n)-G(t_n,\mu)| &= \left|\int_{\mathbb{R}^2}\phi(x_1+t_n,y_1)d\mu_n(x_1,y_1)-\int_{\mathbb{R}^2}\phi(x_2+t_n,y_2)d\mu(x_2,y_2)\right|\\ &\le L\,W_1(\mu_n,\mu)\\ &\le L\,W_2(\mu_n,\mu),\end{aligned}$$

that goes to zero for $n\to+\infty$. Hence jointly continuity of G.

Moreover $\partial_t G$ is continuous w.r.t. t (and w.r.t. μ) for the same reasons.

Since $\partial_t G<0$ everywhere, if we fix $\mu\in\mathscr{P}_2(\mathbb{R}^2)$ there exists at most a unique t such that $G(t,\mu)=0$. Note that $\lim_{t\to+\infty}G(t,\mu)=-1$, hence $\tilde{T}_2^\Phi(\mu)<+\infty$ for all $\mu\in\mathscr{P}_2(\mathbb{R}^2)$. Let $\mu\in\mathscr{P}_2(\mathbb{R}^2)\setminus\tilde{S}_2^\Phi$ (otherwise there is nothing to prove), then there exists a unique t such that $G(t,\mu)=0$ and so $t=\tilde{T}_2^\Phi(\mu)$.

Take a sequence $\{\mu_n\}_{n\in\mathbb{N}}\subseteq\mathscr{P}_2(\mathbb{R}^2)\setminus\tilde{S}_2^\Phi$, such that $\mu_n\rightharpoonup^*\mu$, then $G(\tilde{T}_2^\Phi(\mu_n),\mu_n)$ $=0$ for all $n\in\mathbb{N}$, hence by jointly continuity of G we have that $G(\limsup\limits_{n\to+\infty}\tilde{T}_2^\Phi(\mu_n),\mu)$ $=0$, thus $\tilde{T}_2^\Phi(\mu)=\limsup_{n\to+\infty}\tilde{T}_2^\Phi(\mu_n)$.

So we have proved upper semicontinuity of $\tilde{T}_2^\Phi$, hence continuity by Theorem 3.2.19.

Example 3.3.12. In $\mathbb{R}^2$, set $\phi(x,y):=\arctan(x(1+\arctan^2 y))$. We have that ϕ is bounded, continuous and since

$$\nabla\phi(x,y)=\left(\frac{1+\arctan^2 y}{x^2\left(1+\arctan^2 y\right)^2+1},\frac{2x\arctan y}{(y^2+1)\left(x^2\left(1+\arctan^2 y\right)^2+1\right)}\right),$$

we have $\phi\in C_b^1$, $\nabla\phi\in\mathrm{Lip}_{\mathrm{loc}}(\mathbb{R}^2,\mathbb{R}^2)$ and $\nabla\phi(x,y)\neq(0,0)$ for all $(x,y)\in\mathbb{R}^2$. Take $\Phi=\{\phi\}$, and notice that $\phi(-1,0)<0$, and so assumptions of Definition 3.1.1 are satisfied. Notice also that the gradient of ϕ is uniformly bounded on $\mathbb{R}^2$, and let $L>0$ such that $|\nabla\phi(x,y)|\le L$ for all $(x,y)\in\mathbb{R}^2$.

Let $F:\mathbb{R}^2\rightrightarrows\mathbb{R}^2$ defined as $F(x,y)=[-\nabla\phi(x,y),\nabla\phi(x,y)]$. We have that F is Lipschitz continuous and bounded, moreover, for all $(x,y)\in\mathbb{R}^2$ we have that

$$\inf_{\xi\in F(x,y)}\langle\xi,\nabla\phi(x,y)\rangle=-|\nabla\phi(x,y)|^2.$$

Let now $T_t(\cdot)$ be the solution of $\dot{T}_t(x,y)=-\nabla\phi\circ T_t(x,y)$, $T_0(x,y)=(x,y)$. Given $\mu_0\in\mathscr{P}_2(\mathbb{R}^2)$, and set $\mu_t=T_t\sharp\mu_0$, we have that $\boldsymbol{\mu}=\{\mu_t\}_{t\ge0}$ is an optimal trajectory starting from μ_0.

Notice that $\mu_t\in\tilde{S}_2^\Phi$ if and only if $G(t,\mu_0)\le0$ where $G:[0,+\infty[\times\mathscr{P}_2(\mathbb{R}^2)\to\mathbb{R}$ is defined by

$$G(t,\mu):=\int_{\mathbb{R}^2}\phi\circ T_t(x,y)\,d\mu_0(x,y),$$

and so

$$\tilde{T}_2^\Phi(\mu_0)=\inf\{t\ge0\,:\,G(t,\mu_0)\le0\}.$$

The function G is jointly continuous w.r.t. both variables, indeed

$$|G(t_n,\mu_n)-G(t,\mu)|\le|G(t_n,\mu)-G(t,\mu)|+|G(t_n,\mu_n)-G(t_n,\mu)|,$$

where the first term tends to zero for $n\to+\infty$ by Dominated Convergence Theorem. Focusing on the second term, recalling that ϕ and T_t are Lipschitz continuous with constant L, by Kantorovich duality and Hölder inequality we get

$$\begin{aligned}|G(t_n,\mu_n)-G(t_n,\mu)| &= \left|\int_{\mathbb{R}^2}\phi\circ T_{t_n}(x_1,y_1)\,d\mu_n(x_1,y_1)-\int_{\mathbb{R}^2}\phi\circ T_{t_n}(x_2,y_2)\,d\mu(x_2,y_2)\right|\\ &\leq L^2\,W_1(\mu_n,\mu)\leq L^2\,W_2(\mu_n,\mu),\end{aligned}$$

that goes to zero for $n\to+\infty$.

Since

$$\frac{\partial}{\partial t}G(t,\mu)=\int_{\mathbb{R}^2}\langle\nabla\phi\circ T_t(x,y),\dot{T}_t(x,y)\rangle\,d\mu(x,y)=-\int_{\mathbb{R}^2}|\nabla\phi\circ T_t(x,y)|^2\,d\mu(x,y)<0,$$

if we fix $\mu_0\in\mathscr{P}_2(\mathbb{R}^2)$ there exists at most a unique $t\geq 0$ such that $G(t,\mu_0)=0$, and in this case we have $t=\tilde{T}_2^\Phi(\mu_0)$.

Notice that we have $|\partial_x\phi(x,y)|\geq\frac{1}{Kx^2+1}$ for a suitable constant $K>0$ independent on (x,y). Set $z(t)=\langle T_t(x,y),(1,0)\rangle$, then $\dot{z}(t)\leq-\frac{1}{Kz(t)^2+1}$, and so $\lim_{t\to+\infty}\langle T_t(x,y),(1,0)\rangle=-\infty$. By Dominated Convergence Theorem, this implies

$$\lim_{t\to+\infty}G(t,\mu)=-\frac{\pi}{2},$$

and so we have that for all $\mu\notin\tilde{S}_2^\Phi$ there exists $\bar{t}\geq 0$ such that $G(\bar{t},\mu)\leq 0$, hence $\tilde{T}_2^\Phi(\mu)<+\infty$ for all $\mu\in\mathscr{P}_2(\mathbb{R}^2)$.

Take now a sequence $\{\mu_n\}_{n\in\mathbb{N}}\subseteq\mathscr{P}_2(\mathbb{R}^2)\setminus\tilde{S}_2^\Phi$, such that $\mu_n\rightharpoonup^*\mu$, then $G(\tilde{T}_2^\Phi(\mu_n),\mu_n)=0$ for all $n\in\mathbb{N}$, hence by jointly continuity of G we have that $G\left(\limsup_{n\to+\infty}\tilde{T}_2^\Phi(\mu_n),\mu\right)=0$, thus $\tilde{T}_2^\Phi(\mu)=\limsup_{n\to+\infty}\tilde{T}_2^\Phi(\mu_n)$.

Applying the same procedure with lim inf, we get continuity of $\tilde{T}_2^\Phi$.

3.4 Measure-theoretic Lie bracket for nonsmooth vector fields

In this section we prove a generalization of the classical notion of commutators of vector fields in our framework of measure theory (see [29]), providing an extension of the set-valued Lie bracket introduced in [68, 69] for Lipschitz continuous vector fields.

Indded, in [68, 69] the authors give a generalization of the classical notion of *Lie bracket* (or *commutator*) of two smooth vector fields X, Y, in order to study the commutativity of the flows of two vector fields basically just assuming that the flows are well-defined (e.g., the two vector fields are locally Lipschitz continuous). In this framework, the classical Lie bracket $[X,Y](\cdot)$ appears to be defined only a.e. w.r.t. Lebesgue measure, moreover, as showed with many examples in [68], even at the point where it can be defined, it does not catch all the local features of the two flows.

By mean of a suitable construction, in [68] the authors define an object, called *set-valued Lie bracket*, which associates to *every point* of the space a suitable *set* $[X,Y]_{\mathrm{set}}(\cdot)$, which in the classical smooth case reduces to the usual Lie bracket, and turns out to be the convex hull of the upper Kuratowski limit of the classical Lie bracket (which are defined in a Lebesgue full measure subset, in particular in a dense subset).

They also prove that the basic properties enjoyed by the classical Lie bracket (asymptotic formula, commutativity of the flows, simultaneous flow-box theorem), have their natural counterparts.

The main ingredient to prove the results of [68] is an *exact* integral formula expressing the difference $\phi^Y_{-t} \circ \phi^X_{-s} \circ \phi^Y_t \circ \phi^X_s(q) - q$ (proved in Lemma 4.5 of [68]), where X, Y are locally Lipschitz vector fields and ϕ^X_t, ϕ^Y_t their flows at time t. In this context, the term *exact* is used in opposition to *asymptotic*. This integral formula turns out very useful to be handled and, together with a regularization argument, yields all the main results of the paper.

In [67], these results are applied to give a nonsmooth version of the Frobenius theorem for Lipschitz distributions of vector fields on a manifold. The generalization of the construction of [68] to higher order Lie bracket is not straigthforward, as pointed out in Section 7 of [68], and has been recently proved in the two papers [42], which generalized the exact formula for the single Lie bracket to general nested brackets, and the forthcoming [43].

It is well known that, in the classical framework, the vector space Lie$(\mathscr{F})$ generated by all the vector fields built from a given set $\mathscr{F}$ of vector fields by mean of possibly nested Lie bracket, is deeply related to *controllability* properties of the finite-dimensional driftless control-affine systems where the controlled vector fields are the element of $\mathscr{F}$. Roughly speaking, Lie bracket operations enlarge the set of admissible *displacements* that a particle can reach in a given amount of time by following the admissible trajectories of the system, even if, in general, a Lie bracket does not give an admissible *direction* for the system.

The study of *higher order conditions for attainability* plays an important role also in the classical finite-dimensional setting. Petrov's condition represents a first order requirement on the trajectory and can be interpreted as the request that for each point sufficiently near to the target there exists an admissible trajectory which points *sufficiently towards* the target at the first order, indeed it involves the first order term of at least one admissible trajectory, i.e. an admissible velocity. Since it is a strong condition to be satisfied, it is natural to look for higher order conditions when the first one doesn't hold, by involving higher order terms of the expansion of the trajectory. It has been studied (see [52]) that these conditions involves Lie bracket of admissible vector fields and can be viewed as Petrov's conditions of higher order.

Hence, in order to give higher order conditions for controllability in our framework, it turns out to be a natural problem to define some correspondent quantity for the Lie bracket in a measure-theoretic setting by using tools of transport theory. The study of controllability conditions involving measure-theoretic Lie bracket is still an open problem in this setting. We refer the reader to [59, 60] for the study of sufficient conditions granting *small time-local attainability* in finite-dimension.

Our strategy can be summarized as follows: by exploiting the main idea of the *Agrachev–Gamkrelidze formalism* (AGF) formalism (see for example [68]), we consider probability measures on $\mathbb{R}^d$, and define our object as limit (in a suitable topology) of an asymptotic formula like the one considered in [68], but instead of the *evaluation* at a point q, corresponding to the choice of δ_q, we consider the *push forward* of a probability measure μ along the flow. Under suitable assumptions, we are able to consider the convexified Kuratowski upper limit of this construction as in [68], thus defining a set-valued measure theoretic Lie bracket, which - by construction - satisfies the asymptotic formula and the commutativity property. We notice that this object, being a set of vector-valued measures absolutely continuous w.r.t. μ, has no longer a pointwise meaning, unless the starting measure is purely atomic.

We give also some representation formula, which allows to compare our results with the results of [68], showing that in the case of Dirac deltas, the two constructions agrees and, slightly more generally, under the Lipschitz assumptions of [68], the density of each element w.r.t. a general probability measure μ is an L^p_μ-selection of the set-valued Lie bracket defined in [68].

This Section is structured as follows: in Subsection 3.4.1 we review some preliminaries of differential geometry, in Subsection 3.4.2 we introduce the main objects of our study and formulate the main results, in Subsection 3.4.3 we compare our result with the construction in [68]. We conclude providing an example illustrating our construction in Subsection 3.4.4.

3.4.1 Preliminaries on differential geometry

Definition 3.4.1 (Formal bracket)**.** We denote by $\mathrm{Diffeo}(\mathbb{R}^d)$ the set of all diffeomorphisms of $\mathbb{R}^d$. Let $\psi, \varphi \in \mathrm{Diffeo}(\mathbb{R}^d)$ be two diffeomorphisms. We define their *formal bracket* by setting:

$$[\psi, \varphi](x) := \psi \circ \varphi \circ \psi^{-1} \circ \varphi^{-1}(x).$$

Since for every $\psi, \varphi \in \mathrm{Diffeo}(\mathbb{R}^d)$ we have that $[\psi, \varphi] \in \mathrm{Diffeo}(\mathbb{R}^d)$, by iterating the procedure we can construct *formal bracket expressions* by nesting formal brackets of diffeomorphisms. Given a subset $\mathscr{S} \subseteq \mathrm{Diffeo}(\mathbb{R}^d)$, we define the *length* (also *order* or *depth*) of nested formal brackets of elements of $\mathscr{S}$ by induction. If $\varphi \in \mathscr{S}$ is a single diffeomorphism, then $\mathrm{ord}\,(\varphi) = 1$. Otherwise, if A and B are formal bracket expressions of elements of $\mathscr{S}$, we set $\mathrm{ord}\,[A, B] = \mathrm{ord}\, A + \mathrm{ord}\, B$.

Definition 3.4.2. Let $X : \mathbb{R}^d \to \mathbb{R}^d$ be a locally Lipschitz vector field. Given $x \in \mathbb{R}^d$, we denote by $\phi^X_t(x)$ or $\phi^X(t, x)$ the *flow of* X *starting from* x, i.e. the (unique) solution of $\dot{x}(s) = X(x(s))$, $x(0) = x$ evaluated at $s = t$. We have $\phi^X(0, x) = x$ and $\dfrac{\partial}{\partial t}\phi^X(t, x) = X(\phi^X(t, x))$.

For t sufficiently small, it is well known that $\phi^X_t(\cdot)$ is a diffeomorphism. Given two C^1-smooth vector fields X, Y, we have that

$$\begin{cases} \dfrac{d}{dt}[\phi^X_t, \phi^Y_t](x)_{|t=0} = 0, \\[2ex] \dfrac{d^2}{dt^2}[\phi^X_t, \phi^Y_t](x)_{|t=0} = 2[X, Y](x), \end{cases}$$

where on the right hand side we have the usual *Lie bracket* of vector fields defined in local coordinates by:

$$[X, Y](x) = \langle \nabla Y(x), X(x)\rangle - \langle \nabla X(x), Y(x)\rangle.$$

The correspondence between the first nonvanishing derivative at 0 of flows generating the bracket and the order of the Lie bracket is explained in the following classical result (see e.g., Theorem 1 in [61]).

Theorem 3.4.3. *Let $k \in \mathbb{N} \setminus \{0,1\}$, M be a manifold of class C^k, and for $i = 1, \dots, k$ let $\phi^i : \mathbb{R} \times M \supset U_{\phi^i} \to M$ be a smooth map of class C^k such that*

1. *U_{ϕ^i} is an open neighborhood of $\{0\} \times M$ in $\mathbb{R} \times M$,*
2. *ϕ^i_t is a diffeomorphism of class C^k on its domain,*
3. *$\phi^i_0 = \mathrm{Id}_M$ and $\dfrac{\partial}{\partial t}\phi^i_t\big|_{t=0} = X_i \in \mathrm{Vec}_{k-1}(M)$,*

where $\mathrm{Vec}_k(M)$ is the set of vector fields on M of class C^k. Then for each formal bracket expression B of order k (w.r.t. $\mathscr{S} = \{\phi^i : i = 1, \dots, k\}$) we have

$$\begin{aligned} \frac{\partial^j}{\partial t^j} B(\phi^1_t, \dots, \phi^k_t)\Big|_{t=0} &= 0 \quad \forall 1 \le j < k, \\ \frac{1}{k!} \cdot \frac{\partial^k}{\partial t^k} B(\phi^1_t, \dots, \phi^k_t)\Big|_{t=0} &= B(X_1, \dots, X_k), \end{aligned}$$

where the last expression is computed substituting each ϕ^i_t with X_i in $B(\phi^1_t, \dots, \phi^k_t)$, and then computing the nested Lie bracket of vector fields.

3.4.2 Measure-theoretic Lie bracket

Here, we introduce the basic objects of our analysis, proving also the main results of this section.

We will adopt the following notations. Given a family of Banach spaces $\{X_i\}_{i \in I}$, we define the Borel maps $r_i : \prod_{j \in I} X_j \to X_i$, $r_i(x_I) = x_i$ for all $i \in I$.

We will denote with $d_{\mathscr{P}}$ any metric on $\mathscr{P}(X)$ inducing the w^*-topology on $\mathscr{P}(X)$.

Definition 3.4.4 (Measures associated to a family of transformations)**.** Let $T > 0$, $\mathcal{K} \subseteq \mathscr{P}(\mathbb{R}^d)$, $\mu \in \mathrm{cl}_{d_{\mathscr{P}}}\mathcal{K}$ and let $\mathbf{\Psi}_{\mathcal{K}} = \{\Psi_t(\cdot)\}_{t \in [0,T]}$ be a family of maps such that

(D_1) $\Psi_t : \mathbb{R}^d \to \mathbb{R}^d$ is a Borel map for all $t \in [0,T]$;

(D_2) $t \mapsto \Psi_t(x)$ is continuous from $[0,T]$ to $\mathbb{R}^d$;

(D_3) $\Psi_0 = \mathrm{Id}_{\mathbb{R}^d}$;

(D_4) $\Psi_t \sharp \mu \in \mathcal{K}$ for all $t \in]0,T]$,

where $\mathrm{cl}_{d_{\mathscr{P}}}$ denotes the closure in the w^*-topology. If $\mathcal{K} = \mathscr{P}(\mathbb{R}^d)$ we will omit the subscript $\mathcal{K}$.
Define the measures $\boldsymbol{\eta}^{\mathbf{\Psi}_{\mathcal{K}}}_\mu \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ and $\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu,t} \in \mathscr{P}(\mathbb{R}^d \times \mathbb{R}^d)$ by setting for any $t \in]0,T]$, $m \in \mathbb{N} \setminus \{0\}$, $\varphi \in \mathrm{Bor}_b(\mathbb{R}^d \times \Gamma_T)$, $\psi \in \mathrm{Bor}_b(\mathbb{R}^d \times \mathbb{R}^d)$

$$\begin{aligned} \iint_{\mathbb{R}^d \times \Gamma_T} \varphi(x,\gamma)\, d\boldsymbol{\eta}^{\mathbf{\Psi}_{\mathcal{K}}}_\mu(x,\gamma) &:= \int_{\mathbb{R}^d} \varphi(x,\gamma_x)\, d\mu(x), \\ \iint_{\mathbb{R}^d \times \mathbb{R}^d} \psi(x,y)\, d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu,t}(x,y) &:= \int_{\mathbb{R}^d} \psi\left(x, \frac{\Psi_t(x) - x}{t^m}\right) d\mu(x), \end{aligned}$$

where $\gamma_x(\cdot) \in \Gamma_T$ is defined by $\gamma_x(t) = \Psi_t(x)$.

Defined the map $Q_t^m : \mathbb{R}^d \times \Gamma_T \to \mathbb{R}^d$ by

$$Q_t^m(x, \gamma) := \frac{e_t(x, \gamma) - e_0(x, \gamma)}{t^m},$$

we have $\boldsymbol{\eta}_\mu^{\boldsymbol{\Psi}_{\mathcal{K}}} = \mu \otimes \delta_{\gamma_x}$, $\pi_{\mu,t}^{\boldsymbol{\Psi}_{\mathcal{K}},m} = (e_0 \times Q_t^m)\sharp\boldsymbol{\eta}_\mu^{\boldsymbol{\Psi}_{\mathcal{K}}} = \left(\mathrm{Id}_{\mathbb{R}^d}, \dfrac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m}\right)\sharp\mu$, where for $t \neq 0$ the map $e_0 \times Q_t^m : \mathbb{R}^d \times \Gamma_T \to \mathbb{R}^d \times \mathbb{R}^d$ is defined as

$$(e_0 \times Q_t^m)(x, \gamma) = \left(\gamma(0), \frac{\gamma(t) - \gamma(0)}{t^m}\right).$$

Remark 3.4.5. The main motivation for considering a general subset $\mathcal{K}$ of $\mathscr{P}(\mathbb{R}^d)$ comes from applications, where for example we are able to measure only averaged quantities w.r.t. Lebesgue's measure.

We will now provide some estimates on the p-moments of the measures $\boldsymbol{\eta}_\mu^{\boldsymbol{\Psi}_{\mathcal{K}}}$ and $\pi_{\mu,t}^{\boldsymbol{\Psi}_{\mathcal{K}},m}$ associated to $\boldsymbol{\Psi}_{\mathcal{K}}$.

Lemma 3.4.6 (Estimates on moments)**.** *Let $T > 0$, $p \geq 1$, $\mathcal{K} \subseteq \mathscr{P}_p(\mathbb{R}^d)$, $\mu \in \mathrm{cl}_{d_{\mathscr{P}}}\mathcal{K}$ and let $\boldsymbol{\Psi}_{\mathcal{K}} = \{\Psi_t(\cdot)\}_{t\in[0,T]}$ be a family of maps satisfying assumptions (D_1), (D_2), (D_3), (D_4).*

1. *If $\dfrac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m} \in L^p_\mu(\mathbb{R}^d)$, we have*
$$\mathrm{m}_p(\pi_{\mu,t}^{\boldsymbol{\Psi}_{\mathcal{K}},m}) \leq \left(\left\|\frac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m}\right\|_{L^p_\mu} + \mathrm{m}_p^{1/p}(\mu)\right)^p.$$

2. *If there exists a Borel map $f : \mathbb{R}^d \to [0, +\infty]$ with $|\Psi_t(x) - x| \leq f(x)$ for all $t \in [0, T]$ and $x \in \mathbb{R}^d$, we have*
$$\mathrm{m}_p(\boldsymbol{\eta}_\mu^{\boldsymbol{\Psi}_{\mathcal{K}}}) \leq \mathrm{m}_p(\mu) + \left(\|f\|_{L^p_\mu} + \mathrm{m}_p^{1/p}(\mu)\right)^p.$$

Proof.

1. If $\dfrac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m} \in L^p_\mu(\mathbb{R}^d)$, we have
$$\begin{aligned}\mathrm{m}_p(\pi_{\mu,t}^{\boldsymbol{\Psi}_{\mathcal{K}},m}) &\leq \iint_{\mathbb{R}^d\times\mathbb{R}^d} (|x| + |y|)^p \, d\pi_{\mu,t}^{\boldsymbol{\Psi}_{\mathcal{K}},m}(x, y)\\ &\leq \left(\left(\iint_{\mathbb{R}^d\times\mathbb{R}^d} |x|^p \, d\pi_{\mu,t}^{\boldsymbol{\Psi}_{\mathcal{K}},m}(x, y)\right)^{1/p} + \left(\iint_{\mathbb{R}^d\times\mathbb{R}^d} |y|^p \, d\pi_{\mu,t}^{\boldsymbol{\Psi}_{\mathcal{K}},m}(x, y)\right)^{1/p}\right)^p\\ &= \left(\left\|\frac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m}\right\|_{L^p_\mu} + \mathrm{m}_p^{1/p}(\mu)\right)^p.\end{aligned}$$

2. If there exists a Borel map $f : \mathbb{R}^d \to [0, +\infty]$ with $|\Psi_t(x) - x| \leq f(x)$ for all $t \in [0, T]$ and $x \in \mathbb{R}^d$, we have by Monotone Convergence Theorem
$$\begin{aligned}\mathrm{m}_p(\boldsymbol{\eta}_\mu^{\boldsymbol{\Psi}_{\mathcal{K}}}) &= \iint_{\mathbb{R}^d\times\Gamma_T} (|x|^p + \|\gamma\|^p_\infty) \, d\boldsymbol{\eta}_\mu^{\boldsymbol{\Psi}_{\mathcal{K}}}(x, \gamma) = \int_{\mathbb{R}^d} (|x|^p + \|\gamma_x\|^p_\infty) \, d\mu(x)\\ &\leq \mathrm{m}_p(\mu) + \int_{\mathbb{R}^d} (\|\gamma_x - x\|_\infty + |x|)^p \, d\mu(x)\\ &\leq \mathrm{m}_p(\mu) + \left(\|f\|_{L^p_\mu} + \mathrm{m}_p^{1/p}(\mu)\right)^p.\end{aligned}$$

□

We define now a measure-theoretic object related to the limit of $\dfrac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m}$ as $t \to 0^+$.

Definition 3.4.7 (Measure-theoretic expansion)**.** Let $T > 0$, $m \in \mathbb{N}$, $m \geq 1$, $p \geq 1$, $\mathcal{K} \subseteq \mathscr{P}_p(\mathbb{R}^d)$, $\mu \in \mathrm{cl}_{d_{\mathscr{P}}}\mathcal{K}$ and let $\mathbf{\Psi}_{\mathcal{K}} = \{\Psi_t(\cdot)\}_{t\in[0,T]}$ be a family of maps satisfying assumptions (D_1), (D_2), (D_3), (D_4). Define the following set

$$P^p_m(\mu, \mathbf{\Psi}_{\mathcal{K}}) := \bigcap_{\substack{\delta>0\\ 0<\sigma<T}} \mathrm{cl}_{W_p}\left\{\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu',t} \in \mathscr{P}_p(\mathbb{R}^d \times \mathbb{R}^d) : 0 < t \leq \sigma,\, 0 < d_{\mathscr{P}}(\mu', \mu) \leq \delta,\, \mu' \in \mathcal{K}\right\},$$

where cl_{W_p} denotes the closure in the W_p-topology, and $\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu',t}$ is defined as in Definition 3.4.4.

We notice that

1. $P^p_m(\mu, \mathbf{\Psi}_{\mathcal{K}})$ is W_p-closed.
2. $\pi \in P^p_m(\mu, \mathbf{\Psi}_{\mathcal{K}})$ if and only if there exist $\{t_i\}_{i\in\mathbb{N}} \subseteq]0, T]$ and $\{\mu^{(i)}\}_{i\in\mathbb{N}} \subseteq \mathcal{K}$ such that $t_i \to 0$, $\mu^{(i)} \rightharpoonup^* \mu$, and $W_p\left(\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}, \pi\right) \to 0$ as $i \to +\infty$.
3. For any $\pi \in P^p_m(\mu, \mathbf{\Psi}_{\mathcal{K}})$ we have that $r_1\sharp\pi = \mu$, indeed, given $t_i \to 0^+$, $\{\mu^{(i)}\}_{i\in\mathbb{N}} \subseteq \mathcal{K}$, $\mu^{(i)} \rightharpoonup^* \mu$ such that $W_p(\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}, \pi) \to 0$, we have in particular $r_1\sharp\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i} \rightharpoonup^* r_1\sharp\pi$, since convergence in W_p implies w^*-convergence, and $r_1\sharp\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i} = \mu^{(i)} \rightharpoonup^* \mu$.

We can disintegrate each element $\pi \in P^p_m(\mu, \mathbf{\Psi}_{\mathcal{K}})$ with respect to r_1 obtaining a family of probability measures $\{\sigma^\pi_x\}_{x\in\mathbb{R}^d}$ which is μ-a.e. uniquely defined and satisfies $\pi = \mu \otimes \sigma^\pi_x$. Thus we can define the set

$$V^p_m(\mu, \mathbf{\Psi}_{\mathcal{K}}) := \left\{V \in L^p_\mu(\mathbb{R}^d; \mathbb{R}^d) : V(x) = \int_{\mathbb{R}^d} y\, d\sigma^\pi_x(y),\, \pi = \mu \otimes \sigma^\pi_x \in P^p_m(\mu, \mathbf{\Psi}_{\mathcal{K}})\right\}.$$

Remark 3.4.8. Roughly speaking, the second marginal of each element $\pi \in P^p_m(\mu, \mathbf{\Psi}_{\mathcal{K}})$ represents a limit point of the vector valued measure $\dfrac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m}\mu'$ for $\mu' \in \mathcal{K}$ converging to μ and $t \to 0^+$. To recover an object defined pointwise μ-a.e., we take its barycenter, obtaining the map V.

The set of vector-valued measures $\{V\mu : V \in V^p_m(\mu, \mathbf{\Psi}_{\mathcal{K}})\}$ will be the object generalizing the asymptotic behaviour of the vector-valued measure $\dfrac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m}\mu'$, in the sense precised below.

Lemma 3.4.9 (Interpretation)**.** *Let $T > 0$, $m \in \mathbb{N}$, $m \geq 1$, $p \geq 2$, $\mathcal{K} \subseteq \mathscr{P}_p(\mathbb{R}^d)$, $\mu \in \mathrm{cl}_{d_{\mathscr{P}}}\mathcal{K}$ and let $\mathbf{\Psi}_{\mathcal{K}} = \{\Psi_t(\cdot)\}_{t\in[0,T]}$ be a family of maps satisfying assumptions (D_1), (D_2), (D_3), (D_4). Then if $V \in V^p_m(\mu, \mathbf{\Psi}_{\mathcal{K}})$ there exist $\{\mu^{(i)}\}_{i\in\mathbb{N}} \subseteq \mathcal{K}$ and $\{t_i\}_{i\in\mathbb{N}} \subseteq]0, T]$ such that $\mu^{(i)} \rightharpoonup^* \mu$, $t_i \to 0^+$ and*

$$\lim_{i\to+\infty} \frac{\Psi_{t_i}\sharp\mu^{(i)} - \mu^{(i)}}{t_i^m} = -\mathrm{div}(V\mu),$$

in the sense of distributions.

Proof. Let $V \in V_m^p(\mu, \mathbf{\Psi}_{\mathcal{K}})$. There exist sequences $\{t_i\}_{i\in\mathbb{N}} \subseteq]0,T]$ and $\{\mu^{(i)}\}_{i\in\mathbb{N}} \subseteq \mathcal{K}$, and a family of probability measures $\{\sigma_x\}_{x\in\mathbb{R}^d}$ uniquely defined for μ-a.e. $x \in \mathbb{R}^d$ such that $\mu^{(i)} \rightharpoonup^* \mu$, $t_i \to 0^+$ and, set $\pi := \mu \otimes \sigma_x$, we have $W_p(\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}, \pi) \to 0^+$ and

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} \varphi(x) y \, d\pi(x,y) = \int_{\mathbb{R}^d} \varphi(x) V(x)\, d\mu,$$

for any $\varphi \in C_C^\infty(\mathbb{R}^d)$.

For any $\varphi \in C_C^\infty(\mathbb{R}^d)$ we set $R_\varphi : [0,+\infty[\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}$,

$$R_\varphi(t,x,y) := \varphi\left(x + t^m y\right) - \varphi(x) - \langle \nabla\varphi(x), t^m y\rangle,$$

and, recalling the smoothness of φ, we have

$$\frac{|R_\varphi(t,x,y)|}{t^m} \le t^m \|D^2\varphi\|_\infty |y|^2 \chi_{\operatorname{supp}\varphi}(x).$$

In particular, for i sufficiently large we obtain

$$\begin{aligned}\iint_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|R_\varphi(t_i,x,y)|}{t_i^m}\, d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y) &\le t_i^m \|D^2\varphi\|_\infty \iint_{\mathbb{R}^d\times\mathbb{R}^d} |y|^2\, d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y)\\ &\le t_i^m \|D^2\varphi\|_\infty \mathrm{m}_2(\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i})\\ &\le t_i^m \|D^2\varphi\|_\infty (1 + \mathrm{m}_p(\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i})).\end{aligned}$$

Hence we have

$$\begin{aligned}\langle \varphi, \frac{\Psi_{t_i}\sharp\mu^{(i)} - \mu^{(i)}}{t_i^m}\rangle &= \frac{1}{t_i^m}\left[\int_{\mathbb{R}^d} \varphi(x)\, d\Psi_{t_i}\sharp\mu^{(i)}(x) - \int_{\mathbb{R}^d} \varphi(x)\, d\mu^{(i)}(x)\right]\\ &= \frac{1}{t_i^m}\int_{\mathbb{R}^d}\left[\varphi\left(x + t_i^m \frac{\Psi_{t_i}(x) - x}{t_i^m}\right) - \varphi(x)\right] d\mu^{(i)}(x)\\ &= \frac{1}{t_i^m}\iint_{\mathbb{R}^d\times\mathbb{R}^d} \left[\varphi\left(x + t_i^m y\right) - \varphi(x)\right] d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y)\\ &= \iint_{\mathbb{R}^d\times\mathbb{R}^d} \langle\nabla\varphi(x), y\rangle\, d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y) + \iint_{\mathbb{R}^d\times\mathbb{R}^d} \frac{R_\varphi(t_i,x,y)}{t_i^m}\, d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y)\\ &\le \iint_{\mathbb{R}^d\times\mathbb{R}^d} \langle\nabla\varphi(x), y\rangle\, d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y) + t_i^m \|D^2\varphi\|_\infty (1 + \mathrm{m}_p(\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i})).\end{aligned}$$

Taking the limit for $i \to +\infty$, and recalling that $\mathrm{m}_p(\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{t_i,\mu^{(i)}})$ is uniformly bounded since $W_p(\pi, \pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}) \to 0$, we have

$$\lim_{i\to+\infty} \langle \varphi, \frac{\Psi_{t_i}\sharp\mu^{(i)} - \mu^{(i)}}{t_i^m}\rangle \le \int_{\mathbb{R}^d} \langle\nabla\varphi(x), V(x)\rangle\, d\mu(x) = -\langle \varphi, \operatorname{div}(V\mu)\rangle,$$

which concludes the proof by the arbitrariness of $\varphi \in C_C^\infty(\mathbb{R}^d)$. $\square$

Corollary 3.4.10. *In the same assumptions of Lemma 3.4.9, assume that*

$$\lim_{t\to 0}\left\| \frac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m}\right\|_{L^p_\mu} = 0.$$

Then

1. $\displaystyle\lim_{t\to 0}\frac{W_p(\Psi_t\sharp\mu,\mu)}{t^m}=0$;
2. for every $\varphi\in\mathrm{Lip}(\mathbb{R}^d)$ we have

$$\lim_{t\to 0}\int_{\mathbb{R}^d}\frac{\varphi\circ\Psi_t(x)-\varphi(x)}{t^m}\,d\mu(x)=0.$$

Proof. The result comes immediately, since we have

$$\begin{aligned}\left(\frac{W_p(\Psi_t\sharp\mu,\mu)}{t^m}\right)^p&\le\iint_{\mathbb{R}^d\times\mathbb{R}^d}\frac{|\Psi_t(x)-x|^p}{t^{pm}}\,d\mu(x)\\&=\iint_{\mathbb{R}^d\times\mathbb{R}^d}|y|^p\,d\pi^{\mathbf{\Psi}_\mathcal{K},m}_{\mu,t}(x,y),\\ \left|\int_{\mathbb{R}^d}\frac{\varphi\circ\Psi_t(x)-\varphi(x)}{t^m}\,d\mu(x)\right|^p&=\left|\iint_{\mathbb{R}^d\times\mathbb{R}^d}\frac{\varphi(x+t^my)-\varphi(x)}{t^m}\,d\pi^{\mathbf{\Psi}_\mathcal{K},m}_{\mu,t}(x,y)\right|^p\\&\le\mathrm{Lip}^p(\varphi)\iint_{\mathbb{R}^d\times\mathbb{R}^d}|y|^p\,d\pi^{\mathbf{\Psi}_\mathcal{K},m}_{\mu,t}(x,y),\end{aligned}$$

and in both cases the right hand side tends to 0 by assumption. □

We are going to provide now a sufficent condition ensuring that the above defined sets are nonempty.

Lemma 3.4.11 (Nontriviality). *Let $T>0$, $m\in\mathbb{N}$, $m\ge 1$, $p\ge 1$, $\mathcal{K}\subseteq\mathscr{P}_p(\mathbb{R}^d)$, $\mu\in\mathrm{cl}_{d_\mathscr{P}}\mathcal{K}$ and let $\mathbf{\Psi}_\mathcal{K}=\{\Psi_t(\cdot)\}_{t\in[0,T]}$ be a family of maps satisfying assumptions (D_1), (D_2), (D_3), (D_4).*

1. *$P^p_m(\mu,\mathbf{\Psi}_\mathcal{K})\neq\emptyset$ if and only if $V^p_m(\mu,\mathbf{\Psi}_\mathcal{K})\neq\emptyset$. More precisely, if $\pi=\mu\otimes\sigma^\pi_x\in P^p_m(\mu,\mathbf{\Psi}_\mathcal{K})$ then the map defined as*
$$V(x)=\int_{\mathbb{R}^d}y\,d\sigma^\pi_x(y)$$
belongs to $L^p_\mu(\mathbb{R}^d;\mathbb{R}^d)$.
2. *Assume that*
$$\liminf_{\substack{W_p(\mu',\mu)\to 0\\ \mu'\in\mathcal{K}\\ t\to 0^+}}\frac{\|\Psi_t-\mathrm{Id}_{\mathbb{R}^d}\|_{L^p_{\mu'}}}{t^m}=:C<+\infty,$$
then $P^p_m(\mu,\mathbf{\Psi}_\mathcal{K})\neq\emptyset$, which implies also $V^p_m(\mu,\mathbf{\Psi}_\mathcal{K})\neq\emptyset$.

Proof.

1. Given $\pi\in P^p_m(\mu,\mathbf{\Psi}_\mathcal{K})$ as in the statement, we estimate the L^p_μ-norm of $V(\cdot)$ by applying Jensen's inequality
$$\begin{aligned}\|V\|^p_{L^p_\mu}&=\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}y\,d\sigma^\pi_x(y)\right|^p\,d\mu(x)\le\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|y|^p\,d\sigma^\pi_x(y)\right)\,d\mu(x)\\&=\iint_{\mathbb{R}^d\times\mathbb{R}^d}|y|^p\,d\pi(x,y)\le\mathrm{m}_p(\pi)<+\infty.\end{aligned}$$
Then we have that $V\in V^p_m(\mu,\mathbf{\Psi}_\mathcal{K})$, which turns out to be nonempty. The converse is trivial.

2. Let $\{\mu^{(i)}\}_{i\in\mathbb{N}}$ be a sequence in $\mathcal{K}$, $\{t_i\}_{i\in\mathbb{N}} \subseteq]0,T]$ be such that

$$W_p(\mu^{(i)},\mu)\to 0, \qquad t_i\to 0^+, \qquad \lim_{i\to+\infty}\frac{\|\Psi_{t_i}-\mathrm{Id}_{\mathbb{R}^d}\|_{L^p_{\mu^{(i)}}}}{t_i^m}=C.$$

Since $W_p(\mu^{(i)},\mu)\to 0$, we have that there exists $C'>0$ such that $\mathrm{m}_p^{1/p}(\mu^{(i)})\le C'$ for all $i\in\mathbb{N}$. Define $\pi^{\mathbf{\Psi}_\mathcal{K},m}_{\mu^{(i)},t_i}$ as in Definition 3.4.4, and notice that, by assumption, for i sufficiently large we have $\left\|\dfrac{\Psi_{t_i}-\mathrm{Id}_{\mathbb{R}^d}}{t_i^m}\right\|_{L^p_{\mu^{(i)}}}\le$ $C+1$. Thus, according to Lemma 3.4.6 item (1),

$$\mathrm{m}_p(\pi^{\mathbf{\Psi}_\mathcal{K},m}_{\mu^{(i)},t_i})\le\left(\left\|\frac{\Psi_{t_i}-\mathrm{Id}_{\mathbb{R}^d}}{t_i^m}\right\|_{L^p_{\mu^{(i)}}}+\mathrm{m}_p^{1/p}(\mu^{(i)})\right)^p\le(C+C'+1)^p.$$

In particular, according to Remark 5.1.5 in [9], up to passing to a subsequence, we can assume that there exists $\pi_\infty\in\mathscr{P}_p(\mathbb{R}^d\times\mathbb{R}^d)$ such that $W_p(\pi^{\mathbf{\Psi}_\mathcal{K},m}_{\mu^{(i)},t_i},\pi_\infty)\to 0$, yielding $\pi_\infty\in P^p_m(\mu,\mathbf{\Psi}_\mathcal{K})$ and $\mathrm{m}_p(\pi_\infty)\le(C+C'+1)^p$. To conclude, it is enough to apply the previous item.

□

The following localization result allows us to restrict our attention in the computation of $P^p_m(\mu,\mathbf{\Psi}_\mathcal{K})$ just on the measures supported in a neighborhood of $\operatorname{supp}\mu$.

Lemma 3.4.12 (Localization)**.** *Let $T>0$, $m\in\mathbb{N}$, $m\ge 1$, $p\ge 1$, $\mathcal{K}\subseteq\mathscr{P}_p(\mathbb{R}^d)$ such that if $\mu_1\in\mathcal{K}$ and $\mu_2\ll\mu_1$, then also $\mu_2\in\mathcal{K}$. Let $\mu\in\mathrm{cl}_{d_\mathscr{P}}\mathcal{K}$ and $\mathbf{\Psi}_\mathcal{K}=\{\Psi_t(\cdot)\}_{t\in[0,T]}$ be a family of maps satisfying assumptions (D_1), (D_2), (D_3), (D_4). Then we have*

$$P^p_m(\mu,\mathbf{\Psi}_\mathcal{K})=\bigcap_{\substack{0<\delta<T\\ W\subseteq\mathbb{R}^d\text{ open}\\ \operatorname{supp}\mu\subseteq W}}\mathrm{cl}_{W_p}\left\{\pi^{\mathbf{\Psi}_\mathcal{K},m}_{\mu',t}\in\mathscr{P}_p(\mathbb{R}^d\times\mathbb{R}^d):\ \begin{array}{l}0<d_\mathscr{P}(\mu',\mu)\le\delta,\ \mu'\in\mathcal{K}\\ 0<t\le\delta,\ \operatorname{supp}\mu'\subseteq\overline{W}\end{array}\right\},$$

Proof. The inclusion $\supseteq$ holds trivially true. We prove the converse inclusion. Let $\pi\in P^p_m(\mu,\mathbf{\Psi}_\mathcal{K})$, in particular there exists $\{\mu^{(i)}\}_{i\in\mathbb{N}}\subseteq\mathcal{K}$, $\mu^{(i)}\rightharpoonup^*\mu$, $t_i\to 0^+$ such that $W_p(\pi^{\mathbf{\Psi}_\mathcal{K},m}_{\mu^{(i)},t_i},\pi)\to 0$. Let $W\subseteq\mathbb{R}^d$ be open and such that $\operatorname{supp}\mu\subseteq W$. Define $\varphi_W\in C^\infty_C(\mathbb{R}^d)$ such that $0\le\varphi_W(\mathbb{R}^d)\le 1$, $\varphi_W(x)\equiv 1$ for all $x\in\operatorname{supp}(\mu)$ and $\operatorname{supp}\varphi_W\subseteq W$. Set

$$\mu^{(i)}_W:=\frac{\varphi_W\mu^{(i)}}{\displaystyle\int_{\mathbb{R}^d}\varphi_W(x)\,d\mu^{(i)}(x)}\in\mathcal{K},$$

by hypothesis. Let $\psi\in C^0_b(\mathbb{R}^d)$. Then, since $\psi\varphi_W\in C^0_b(\mathbb{R}^d)$, we have

$$\lim_{i\to+\infty}\int_{\mathbb{R}^d}\psi(x)\,d\mu^{(i)}_W(x)=\lim_{i\to+\infty}\frac{\displaystyle\int_{\mathbb{R}^d}\psi(x)\varphi_W(x)\,d\mu^{(i)}(x)}{\displaystyle\int_{\mathbb{R}^d}\varphi_W(x)\,d\mu^{(i)}(x)}=\frac{\displaystyle\int_{\mathbb{R}^d}\psi(x)\varphi_W(x)\,d\mu(x)}{\displaystyle\int_{\mathbb{R}^d}\varphi_W(x)\,d\mu(x)}$$
$$=\int_{\mathbb{R}^d}\psi(x)\,d\mu(x),$$

since $\varphi_W \equiv 1$ on $\operatorname{supp}\mu$. Thus we have $\mu_W^{(i)} \rightharpoonup^* \mu$ for all $0<\delta<T$. For any $0<\delta<T$ we have

$$\lim_{i\to+\infty}\int_{\mathbb{R}^d} \varphi_W(x)\,d\mu^{(i)}(x) = 1,$$

thus there exists $i_\delta\in\mathbb{N}$ such that $\displaystyle\int_{\mathbb{R}^d} \varphi_W(x)\,d\mu^{(i)}(x) \geq \frac{1}{2}$, for all $i\geq i_\delta$.

This implies $\mathrm{m}_p(\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu_W^{(i)},t_i}) \leq 2\mathrm{m}_p(\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i})$, for all $i\geq i_\delta$, by Monotone Convergence Theorem. Since by assumption $W_p(\pi,\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i})\to 0$, we have that $\mathrm{m}_p(\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i})$ is uniformly bounded, and so, up to passing to a non relabeled subsequence, we have that there exists $\pi'\in\mathscr{P}_p(\mathbb{R}^d\times\mathbb{R}^d)$ such that $W_p(\pi',\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu_W^{(i)},t_i})\to 0$ as $i\to+\infty$. To prove that $\pi=\pi'$, which will conclude the proof by the arbitrariness of W and δ, it is enough to show that $d_{\mathscr{P}}(\pi,\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu_W^{(i)},t_i})\to 0$. Indeed, for any $\psi\in C_b^0(\mathbb{R}^d\times\mathbb{R}^d)$ we have

$$\begin{aligned}
\lim_{i\to+\infty}\iint_{\mathbb{R}^d\times\mathbb{R}^d}\psi(x,y)\,d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu_W^{(i)},t_i}(x,y) &= \lim_{i\to+\infty}\int_{\mathbb{R}^d}\psi\left(x,\frac{\Psi_{t_i}(x)-x}{t_i^m}\right)d\mu_W^{(i)}(x)\\
&= \lim_{i\to+\infty}\frac{\displaystyle\int_{\mathbb{R}^d}\varphi_W(x)\psi\left(x,\frac{\Psi_{t_i}(x)-x}{t_i^m}\right)d\mu^{(i)}(x)}{\displaystyle\int_{\mathbb{R}^d}\varphi_W(x)\,d\mu^{(i)}}\\
&= \lim_{i\to+\infty}\int_{\mathbb{R}^d}\varphi_W(x)\psi(x,y)\,d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y)\\
&= \int_{\mathbb{R}^d}\varphi_W(x)\psi(x,y)\,d\pi(x,y)\\
&= \int_{\mathbb{R}^d}\psi(x,y)\,d\pi(x,y),
\end{aligned}$$

and so $W_p(\pi,\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{t_i,\mu_W^{(i)}})\to 0$ as $i\to+\infty$, $\{\mu^{(i)}\}_{i\in\mathbb{N}}\subseteq\mathcal{K}$ and $\operatorname{supp}\mu_W^{(i)}\subseteq\overline{W}$ for all $i\in\mathbb{N}$. □

We will now provide some representation formulas for the function on $V_m^p(\mu,\mathbf{\Psi}_{\mathcal{K}})$, proving also some refinement under additional assumptions. These will be used to establish a comparison with the set-valued Lie bracket defined by Rampazzo-Sussmann in [68].

Definition 3.4.13. Let $T>0$, $m\in\mathbb{N}$, $m\geq 1$, $\mathcal{K}\subseteq\mathscr{P}(\mathbb{R}^d)$, $\mu\in\mathrm{cl}_{d_{\mathscr{P}}}\mathcal{K}$, $D\subseteq\mathbb{R}^d$, and let $\mathbf{\Psi}_{\mathcal{K}}=\{\Psi_t(\cdot)\}_{t\in[0,T]}$ be a family of maps satisfying assumptions (D_1), (D_2), (D_3), (D_4). For every $\delta>0$, $0<\sigma<T$, and $z\in\mathbb{R}^d$, define the sets

$$\begin{aligned}
S^{\sigma,\delta}_{m,D}(z) &:= \left\{\frac{\Psi_t(y)-y}{t^m}\,:\,0<t<\sigma,\,y\in B(z,\delta)\cap D\right\},\\
K^{\sigma,\delta}_{m,D}(z) &:= \begin{cases}\overline{\mathrm{co}}\,S^{\sigma,\delta}_{m,D}(z), & \text{if } S^{\sigma,\delta}_{m,D}(z)\neq\emptyset,\\[2ex] \emptyset, & \text{otherwise,}\end{cases}\\
E_{m,D} &:= \{z\in D:\ \text{there exists } \sigma_z,\delta_z>0 \text{ such that } S^{\sigma_z,\delta_z}_{m,D}(z) \text{ is bounded}\}.
\end{aligned}$$

If $D=\mathbb{R}^d$ we will write $S^{\sigma,\delta}_m(z)$, $K^{\sigma,\delta}_m(z)$, thus omitting D.

Theorem 3.4.14 (Representation formula)**.** *Let $T > 0$, $m \in \mathbb{N}$, $m \geq 1$, $p \geq 1$, $\mathcal{K} \subseteq \mathscr{P}_p(\mathbb{R}^d)$, $\mu \in \mathrm{cl}_{d_\mathscr{P}}\mathcal{K}$ and let $\mathbf{\Psi}_\mathcal{K} = \{\Psi_t(\cdot)\}_{t\in[0,T]}$ be a family of maps satisfying assumptions (D_1), (D_2), (D_3), (D_4). Let $D \subseteq \mathbb{R}^d$ and assume that the following condition holds*

(H_1) *$\mu'(D) = 1$ for all $\mu' \in \mathcal{K}$.*

Then if $V \in V^p_m(\mu, \mathbf{\Psi}_\mathcal{K})$ we have

$$V(z) \in \bigcap_{\sigma,\delta>0} K^{\sigma,\delta}_{m,D}(z), \qquad \text{for } \mu\text{-a.e. } z \in \mathbb{R}^d, \tag{3.20}$$

$$V(z) \in \mathrm{co} \bigcap_{\sigma,\delta>0} \overline{S^{\sigma,\delta}_{m,D}(z)}, \qquad \text{for } \mu\text{-a.e. } z \in E_{m,D}. \tag{3.21}$$

Proof. Let $V \in V^p_m(\mu, \mathbf{\Psi}_\mathcal{K})$. There exist sequences $\{t_i\}_{i\in\mathbb{N}} \subseteq]0, T]$, $t_i \to 0^+$ and $\{\mu^{(i)}\}_{i\in\mathbb{N}} \subseteq \mathcal{K}$, $\mu^{(i)} \rightharpoonup^* \mu$, and a family of probability measures $\{\xi_x\}_{x\in\mathbb{R}^d}$ uniquely defined for μ-a.e. $x \in \mathbb{R}^d$ such that denoted by $\pi := \mu \otimes \xi_x$, we have $W_p(\pi^{\mathbf{\Psi}_\mathcal{K},m}_{\mu^{(i)},t_i}, \pi) \to 0$ and $V(x) = \displaystyle\int_{\mathbb{R}^d} y d\xi_x(y)$ for μ-a.e. $x \in \mathbb{R}^d$.

For any $\sigma \in]0, T]$ we define a set-valued map $G_\sigma : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ by taking

$$G_\sigma(x) := \bigcap_{\delta>0} K^{\sigma,\delta}_{m,D}(x).$$

Notice that $\mathrm{dom}\, G_\sigma \supseteq D$. This set-valued map has closed graph, indeed, let $\{x_n\}_{n\in\mathbb{N}}$, $\{y_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^d$, $x, y \in \mathbb{R}^d$ be such that $x_n \to x$, $y_n \to y$, $y_n \in G_\sigma(x_n)$ for all $n \in \mathbb{N}$. Fix $\delta > 0$ and let $n_\delta > 0$ be such that $|x_n - x| < \delta$ for all $n \geq n_\delta$. For every $\delta' > 0$ and $n \geq n_\delta$ we have that

$$y_n \in \overline{\mathrm{co}}\, S^{\sigma,\delta'}_{m,D}(x_n) \subseteq \overline{\mathrm{co}}\, S^{\sigma,\delta'+|x_n-x|}_{m,D}(x) \subseteq \overline{\mathrm{co}}\, S^{\sigma,\delta'+\delta}_{m,D}(x).$$

By passing to the limit as $n \to +\infty$ we have $y \in \overline{\mathrm{co}}\, S^{\sigma,\delta'+\delta}_{m,D}(x)$ for all $\delta', \delta > 0$, and then by taking the intersection on $\delta, \delta' > 0$ we have $y \in G_\sigma(x)$.

Since G_σ has closed graph, the map $g_\sigma(x, y) := I_{G_\sigma(x)}(y)$ is l.s.c. and non-negative (set $I_\emptyset \equiv +\infty$), moreover $g_\sigma(x, \cdot)$ is convex for all $x \in \mathbb{R}^d$.

By Jensen's inequality we have

$$\begin{aligned}\int_{\mathbb{R}^d} g_\sigma(x, V(x))\, d\mu(x) &= \int_{\mathbb{R}^d} g_\sigma\left(x, \int_{\mathbb{R}^d} y\, d\xi_x(y)\right) d\mu(x)\\ &\leq \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} g_\sigma(x, y)\, d\xi_x(x)\, d\mu(x)\\ &= \iint_{\mathbb{R}^d\times\mathbb{R}^d} g_\sigma(x, y)\, d\pi(x, y).\end{aligned}$$

Recalling Lemma 5.1.7 in [9], by l.s.c. of $g_\sigma(\cdot, \cdot)$ we have

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} g_\sigma(x, y)\, d\pi(x, y) \leq \liminf_{i\to+\infty} \iint_{\mathbb{R}^d\times\mathbb{R}^d} g_\sigma(x, y)\, d\pi^{\mathbf{\Psi}_\mathcal{K},m}_{\mu^{(i)},t_i}(x, y).$$

We obtain

$$\begin{aligned}\int_{\mathbb{R}^d} g_\sigma(x,V(x))\,d\mu(x) &\le \iint_{\mathbb{R}^d\times\mathbb{R}^d} g_\sigma\,(x,y)\;d\pi(x,y)\\ &\le \liminf_{i\to+\infty}\iint_{\mathbb{R}^d\times\mathbb{R}^d} g_\sigma(x,y)\,d\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y)\\ &= \liminf_{i\to+\infty}\int_{\mathbb{R}^d} g_\sigma\left(x,\frac{\Psi_{t_i}(x)-x}{t_i^m}\right)\,d\mu^{(i)}(x).\end{aligned}$$

Since there exists $i_\sigma \ge 0$ such that $t_i \le \sigma$ for all $i \ge i_\sigma$, then for any $x \in D$ we have

$$g_\sigma\left(x,\frac{\Psi_{t_i}(x)-x}{t_i^m}\right)=0, \qquad \text{for all } i\ge i_\sigma. \tag{3.22}$$

This implies

$$\int_{\mathbb{R}^d} g_\sigma(x,V(x))\,d\mu(x) \le \liminf_{i\to+\infty}\int_{\mathbb{R}^d\setminus D} g_\sigma\left(x,\frac{\Psi_{t_i}(x)-x}{t_i^m}\right)\,d\mu^{(i)}(x).$$

Thus, since by hypothesis $\mu^{(i)}(D)=1$ for all $i\in\mathbb{N}$, we have $g_\sigma(x,V(x))=0$ for μ-a.e. $x\in\mathbb{R}^d$. Recalling the arbitrariness of $\sigma>0$, for μ-a.e. $x\in\mathbb{R}^d$

$$V(x)\in\bigcap_{\sigma>0}\bigcap_{\delta>0}K^{\sigma,\delta}_{m,D}(x)=\bigcap_{\sigma,\delta>0}K^{\sigma,\delta}_{m,D}(x),$$

which proves (3.20).

Since

$$\bigcap_{\sigma,\delta>0}K^{\sigma,\delta}_{m,D}(z)\supseteq \operatorname{co}\bigcap_{\sigma,\delta>0}\overline{S^{\sigma,\delta}_{m,D}(z)}$$

for all $z\in\mathbb{R}^d$, to prove (3.21) we must show that equality holds when $z\in E_{m,D}$. By definition of $E_{m,D}$, there exist $\delta_z>0$ and $0<\sigma_z<T$ such that $S^{\sigma,\delta}_{m,D}(z)$ is bounded for all $0<\sigma<\sigma_z$ and $0<\delta<\delta_z$, so we can find a sequence $t_i\to 0^+$, a sequence $y_i\to z$, and a vector $\xi(z)\in\mathbb{R}^d$ such that

$$\lim_{i\to\infty}\frac{\Psi_{t_i}(y_i)-y_i}{t_i^m}=\xi(z),$$

and, by construction, we have $\xi(z)\in\overline{S^{\sigma,\delta}_{m,D}(z)}$ for all $\sigma,\delta>0$.

Thus $\xi(z)\in\bigcap_{\sigma,\delta>0}\overline{S^{\sigma,\delta}_{m,D}(z)}$, and so the set $\operatorname{co}\bigcap_{\sigma,\delta>0}\overline{S^{\sigma,\delta}_{m,D}(z)}$ is closed, convex, and nonempty.

Assume by contradiction that $w\in\bigcap_{\sigma,\delta>0}K^{\sigma,\delta}_{m,D}(z)\setminus\operatorname{co}\bigcap_{\sigma,\delta>0}\overline{S^{\sigma,\delta}_{m,D}(z)}$. By Hahn-Banach separation theorem, there exist $\varepsilon>0$ and $\bar v\in\mathbb{R}^d$ such that

$$\langle\bar v,w\rangle\ge\langle\bar v,\xi\rangle+\varepsilon,\ \text{ for all } \xi\in\operatorname{co}\bigcap_{\sigma,\delta>0}\overline{S^{\sigma,\delta}_{m,D}(z)},$$

in particular we have

$$\langle\bar v,w\rangle\ge\langle\bar v,\xi\rangle+\varepsilon,\ \text{ for all } \xi\in\bigcap_{\sigma,\delta>0}\overline{S^{\sigma,\delta}_{m,D}(z)}.$$

On the other hand, we have that

$$w \in \bigcap_{\sigma,\delta>0} \overline{\mathrm{co}}\, S^{\sigma,\delta}_{m,D}(z)$$

implies that for all $v \in \mathbb{R}^d$, $\sigma, \delta > 0$ we have

$$\langle v, w\rangle \le \sup_{p\in\overline{\mathrm{co}}\, S^{\sigma,\delta}_{m,D}(z)} \langle v, p\rangle = \sup_{p\in S^{\sigma,\delta}_{m,D}(z)} \langle v, p\rangle,$$

so for every sequence $\sigma_i \to 0^+$ and $\delta_i \to 0$ we choose $\xi_i \in S^{\sigma_i,\delta_i}_{m,D}(z)$ such that

$$\sup_{p\in S^{\sigma_i,\delta_i}_{m,D}(z)} \langle v, p\rangle \le \langle v, \xi_i\rangle + \frac{1}{i}.$$

Up to passing to a subsequence, we can assume that $\xi_i \to \bar{\xi}$. By construction, we have that $\bar{\xi} \in \overline{S^{\sigma,\delta}_{m,D}(z)}$ for all $\sigma, \delta > 0$, and

$$\langle v, w\rangle \le \langle v, \bar{\xi}\rangle,$$

contradicting the fact that $\langle \bar{v}, w\rangle \ge \langle \bar{v}, \xi\rangle + \varepsilon$ for all $\xi \in \bigcap_{\sigma,\delta>0} \overline{S^{\sigma,\delta}_{m,D}(z)}$. $\square$

Remark 3.4.15. In the case in which the maps $\mathbf{\Psi}_{\mathcal{K}} \ni \Psi_t : \mathbb{R}^d \to \mathbb{R}^d$ are continuous for all $t \in [0, T]$, then Theorem 3.4.14 holds also if instead of condition (H_1) we assume

(H_2) $\mu'(\overline{D}) = 1$ for all $\mu' \in \mathcal{K}$.

Indeed, in this case property (3.22) holds for all $x \in \overline{D}$ and not only for all $x \in D$, thanks to lower semicontinuity of g_σ. Furthermore, property (3.21) holds for μ-a.e. $z \in \tilde{E}_{m,D}$, where

$$\tilde{E}_{m,D} := \{z \in \overline{D} : \text{ there exists } \sigma_z, \delta_z > 0 \text{ such that } S^{\sigma_z,\delta_z}_{m,D}(z) \text{ is bounded}\}.$$

We notice also that if D is a dense subset of $\mathbb{R}^d$, condition (H_2) is trivially satisfied.

3.4.3 Application to the composition of flows of vector fields

As seen in the introduction of this Section 3.4, in [68] the authors extended the definition of a Lie bracket of two C^1 vector fields to the case of two Lipschitz continuous vector fields X, Y, that is an assumption implying continuity of $\Psi_t(\cdot) := [\phi^X_t, \phi^Y_t](\cdot)$. In this case, the Lie bracket of the vector fields at every point turns out to be a set. Moreover, they provided in this framework an asymptotic formula for the flows and the generalization of other classical results holding for the Lie bracket of vector fields.

A natural question is to compare our construction with the one in [68] when the starting measure reduces to a Dirac delta, in the spirit of the AGF formalism. The aim of this section is to perform such a comparison, showing that - roughly speaking - the density V of the measure theoretic bracket $V\mu$ is a L^p_μ selection of the Rampazzo-Sussmann set-valued Lie bracket. In particular, when $\mu = \delta_q$, the two constructions reduces to the same object.

We will take $\mathcal{K} = \mathscr{P}(\mathbb{R}^d)$ throughout the section, hence we will omit the condition (D_4) in Definition 3.4.4 since it follows from (D_1).

We recall the following definition from [68].

Definition 3.4.16 (Set-valued Lie bracket)**.** Let f, g be locally Lipschitz vector fields on $\mathbb{R}^d$. The (set-valued) Lie bracket of f and g at $x \in \mathbb{R}^d$ is

$$[f,g]_{\mathrm{set}}(x) := \mathrm{co}\Big\{v \in \mathbb{R}^d : \text{ there exists a sequence } \{x_j\}_{j\in\mathbb{N}} \subseteq \mathrm{dom}(Df) \cap \mathrm{dom}(Dg), \\ \text{such that } x_j \to x \text{ and } v = \lim_{j\to\infty} [f,g](x_j)\Big\}$$

where $\mathrm{dom}(Df)$ and $\mathrm{dom}(Dg)$ denotes the set of differentiability points of f and g, respectively. Recalling Rademacher's Theorem, when f is Lipschitz continuous it is differentiable at a.e. $x \in \mathbb{R}^d$, thus $\mathrm{dom}(Df) \cap \mathrm{dom}(Dg)$ has full measure in $\mathbb{R}^d$.

According to Remark 3.6 in [68], the following equivalent definition can be given

$$[f,g]_{\mathrm{set}}(x) = \{Bf(x) - Ag(x) : (A,B) \in \partial(f \times g)(x)\},$$

where $f \times g$ is the map defined as $(f \times g)(x) = (f(x), g(x))$, and ∂ denotes the Clarke's generalized Jacobian, which for a Lipschitz continuous map $h : \mathbb{R}^k \to \mathbb{R}^m$ is defined as

$$\begin{aligned}\partial h(x) &:= \mathrm{co}\Big\{L : \mathbb{R}^k \to \mathbb{R}^m : \text{there exists } \{x_j\}_{j\in\mathbb{N}} \subseteq \mathrm{dom}(Dh) \text{ s.t. } L = \lim_{j\to\infty} Dh(x_j)\Big\} \\ &= \mathrm{co} \bigcap_{\delta>0} \overline{\{Dh(y) : y \in \mathrm{dom}(Dh) \cap B(x,\delta)\}}.\end{aligned}$$

Recall that in general $\partial(f \times g)(x) \subseteq \partial f(x) \times \partial g(x)$, and the inclusion may be strict.

We can recast the above definition by

$$[f,g]_{\mathrm{set}}(x) = \mathrm{co} \bigcap_{\delta>0} \overline{\{Dg(y)f(y) - Df(y)g(y) : y \in \mathrm{dom}(Df) \cap \mathrm{dom}(Dg) \cap B(x,\delta)\}}.$$

Remark 3.4.17. Let v be a Lipschitz continuous vector field with Lipschitz constant $L > 0$. Fix a set of smooth mollifiers $\{s_\rho\}_{\rho>0}$ and set $v_\rho = v * s_\rho$. For any $\varepsilon > 0$ there exists $\rho > 0$ such that for all $0 \le t \le T$

$$\begin{aligned}|\phi^v_t(x) - \phi^{v_\rho}_t(y)| &\le |x-y| + \int_0^t |v(\phi^v_s(x)) - v_\rho(\phi^{v_\rho}_s(y))|\, ds \\ &\le |x-y| + \int_0^t |v(\phi^v_s(x)) - v(\phi^{v_\rho}_s(y))| + \int_0^t |v(\phi^{v_\rho}_s(y)) - v_\rho(\phi^{v_\rho}_s(y))|\, ds \\ &\le |x-y| + L\int_0^t |\phi^v_s(x) - \phi^{v_\rho}_s(y)| + \varepsilon T.\end{aligned}$$

By Gronwall's inequality,

$$|\phi_t^v(x) - \phi_t^{v_\rho}(y)| \leq (|x-y| + \varepsilon T)\, e^{LT}$$

and so if $|x-y| \leq C'\varepsilon$, there exists $C'' > 0$ such that $|\phi_t^v(x) - \phi_t^{v_\rho}(y)| \leq C''\varepsilon$. The argument can be iterated for concatenation of flows of Lipschitz continuous vector fields.

Remark 3.4.18. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz continuous map. Then, if f is differentiable at $x \in \mathbb{R}^d$, we have $\nabla f_\rho(x) \to \nabla f(x)$, where $f_\rho(x) = (f * s_\rho)(x)$, and $\{s_\rho\}_{\rho>0}$ is any family of smooth mollifiers. It is enough to check the assertion for the directional derivatives of f, so let $v \in \mathbb{R}^d$, $\|v\| = 1$. Recalling that f_ρ converges uniformly to f on compact sets, we have

$$\begin{aligned}
\{\partial_v f(x)\} &= \bigcap_{\sigma>0} \overline{\left\{\frac{f(x+tv)-f(x)}{t} : 0 < t < \sigma\right\}} \\
&= \bigcap_{\sigma>0} \bigcap_{\rho>0} \overline{\left\{\frac{f_\tau(x+tv)-f_\tau(x)}{t} : 0 < t < \sigma,\, 0 < \tau < \rho\right\}} \\
&= \bigcap_{\rho>0} \bigcap_{\sigma>0} \overline{\left\{\frac{f_\tau(x+tv)-f_\tau(x)}{t} : 0 < t < \sigma,\, 0 < \tau < \rho\right\}} \\
&= \bigcap_{\rho>0} \overline{\{\partial_v f_\tau(x) : 0 < \tau < \rho\}} = \left\{\lim_{\rho\to 0} \partial_v f_\rho(x)\right\}.
\end{aligned}$$

We will show now a result stating the main connection between our construction and [68]. Indeed, we prove that in the same framework of [68], the two constructions agree.

Proposition 3.4.19. *Let now X, Y be locally Lipschitz continuous vector fields, set $\Psi_t(x) = [\phi_t^X, \phi_t^Y](x)$, then $\mathbf{\Psi} = \{\Psi_t(\cdot)\}_{t\in[0,T]}$ satisfies assumptions (D_1), (D_2), (D_3). For any $z \in \mathbb{R}^d$ and $V \in V_2^p(\delta_z, \mathbf{\Psi})$ we have*

$$V(z) \in [X, Y]_{\mathrm{set}}(z).$$

Proof. Let D be the set of differentiability points of X and Y, in particular it is dense in $\mathbb{R}^d$. Fix $z \in \mathbb{R}^d$. By Lemma 3.4.12, we can restrict ourselves to measures supported on a compact neighborhood of z, thus without loss of generality we can assume that X, Y are globally Lipschitz continuous.

Fix a smooth family of mollifiers $\{s_\rho\}_{\rho>0}$, and let $X^\rho = X * s_\rho$ and $Y^\rho = Y * s_\rho$. We set $\Psi_t^\rho(x) = [\phi_t^{X^\rho}, \phi_t^{Y^\rho}]$ and notice that $\mathbf{\Psi}^\rho$ converges uniformly to $\mathbf{\Psi}$ on every compact subset of $[0,T] \times \mathbb{R}^d$. Moreover, if $x \in D$ we have

$\nabla X^\rho(x) \to \nabla X(x)$ as $\rho \to 0^+$ by Remark 3.4.18. These two facts implies that

$$\begin{aligned}
\operatorname{co}\bigcap_{\sigma,\delta>0}\overline{S^{\sigma,\delta}_{m,D}(z)} &= \operatorname{co}\bigcap_{\sigma,\delta>0}\bigcap_{\rho>0}\overline{\left\{\frac{\Psi^\tau_t(x)-x}{t^2}\,:\,0<\tau<\rho,\,x\in B(z,\delta)\cap D,\,0<t<\sigma\right\}}\\
&= \operatorname{co}\bigcap_{\delta>0}\bigcap_{\rho>0}\bigcap_{\sigma>0}\overline{\left\{\frac{\Psi^\tau_t(x)-x}{t^2}\,:\,0<\tau<\rho,\,x\in B(z,\delta)\cap D,\,0<t<\sigma\right\}}\\
&= \operatorname{co}\bigcap_{\delta>0}\bigcap_{\rho>0}\overline{\{[X^\tau,Y^\tau](x)\,:\,0<\tau<\rho,\,x\in B(z,\delta)\cap D\}}\\
&= \operatorname{co}\bigcap_{\delta>0}\overline{\{\nabla Y(x)\cdot X(x)-\nabla X(x)\cdot Y(x)\,:\,x\in B(z,\delta)\cap D\}}\\
&= [X,Y]_{\mathrm{set}}(z).
\end{aligned}$$

Hence we can conclude, thanks to Remark 3.4.15 and noticing that we have $\tilde{E}_{2,D} = \mathbb{R}^d$ by density of D in $\mathbb{R}^d$. □

Exploiting this representation formula, and the results of [68] (see in particular Theorem 5.3 for commutativity), the asymptotic result given by Corollary 3.4.10 can be refined as follows.

Corollary 3.4.20. *Let $T>0$, $m\in\mathbb{N}$, $m\ge 1$, $p\ge 1$, and let X, Y be locally Lipschitz continuous vector fields. Set $\Psi_t(x) = [\phi^X_t,\phi^Y_t](x)$, $\boldsymbol{\Psi}_t = \{\Psi_t(\cdot)\}_{t\in[0,T]}$. Then, if $V^p_2(\mu,\boldsymbol{\Psi}) = \{0\}$ for all $\mu\in\mathscr{P}_p(\mathbb{R}^d)$ we have $(\phi^X_t\circ\phi^Y_t)\sharp\mu = (\phi^Y_t\circ\phi^X_t)\sharp\mu$ for all $\mu\in\mathscr{P}_p(\mathbb{R}^d)$, $t\in[0,T]$.*

Apparently, the construction of Proposition 3.4.19 can be extended to any formal bracket by using Theorem 3.4.3. However, it has been pointed out in [68] that the step between the definition of the *single* set-valued bracket, and the definition of *higher order* bracket is quite nontrivial. Indeed, we can give just a partial answer to this issue.

Definition 3.4.21. Let $k\in\mathbb{N}\setminus\{0,1\}$, and $X_1,\dots,X_k$ be vector fields of class $C^{k-2,1}(\mathbb{R}^d)$. Let $\mathscr{S} := \{\phi^{X_i}_t\,:\,i=1,\dots,k\}$ and consider a formal bracket $B(\phi^{X_1}_t,\dots,\phi^{X_k}_t)$ of order k w.r.t. $\mathscr{S}$. Let $D\subseteq\mathbb{R}^d$. We define for any $z\in\mathbb{R}^d$

$$B_{\mathrm{set}}(X_1,\dots,X_k)(z) = \operatorname{co}\bigcap_{\delta>0}\bigcap_{\rho>0}\overline{\{B(X^\tau_1,\dots,X^\tau_k)(x)\,:\,x\in B(z,\delta)\cap D,\,0<\tau<\rho\}}. \tag{3.23}$$

The motivation for such a definition is the following.

Remark 3.4.22. Set $\Psi_t(x) = B(\phi^{X_1}_t,\dots,\phi^{X_k}_t)(x)$ and let D be the set of differentiability points for all the vector fields involved and for their derivatives up to the order appearing in the bracket B. In particular, D is dense in $\mathbb{R}^d$. By Theorem 3.4.14, for all $z\in\mathbb{R}^d$ we have

$$V(z)\in\operatorname{co}\bigcap_{\sigma,\delta>0}\overline{S^{\sigma,\delta}_{k,D}(z)},$$

for all $V\in V^p_k(\delta_z,\boldsymbol{\Psi})$. Thus it make sense to define

$$B_{\mathrm{set}}(X_1,\dots,X_k)(z) = \operatorname{co}\bigcap_{\sigma,\delta>0}\overline{S^{\sigma,\delta}_{k,D}(z)},$$

indeed, equality follows by the very same argument of Proposition 3.4.19.

When z is a differentiability point for all the vector fields involved and for their derivatives up to the order appearing in the bracket B, we can refine (3.23), in the spirit of Proposition 3.4.19, i.e., we set D as the set of common differentiability points for all the vector fields and their derivatives, and we have for all $z \in D$

$$B_{\mathrm{set}}(X_1, \ldots, X_k)(z) = \mathrm{co} \bigcap_{\delta>0} \overline{\{B(X_1, \ldots, X_k)(x) : x \in B(z, \delta) \cap D\}}. \tag{3.24}$$

However, in general, the definition given in (3.24) is *not consistent with the asymptotic formula when $z \notin D$*, in the following sense: to have

$$\mathrm{co} \bigcap_{\delta>0} \overline{\{B(X_1, \ldots, X_k)(x) : x \in B(y, \delta) \cap D\}} = 0$$

for all y in a neighborhood of z, in general *does not imply* that $\lim_{t \to 0} \dfrac{\Psi_t(z) - z}{t^m} = 0$, as showed with a counterexample in Section 7.1 and Section 7.2 of [68], where the possibility to extend the construction of [68] to higher order bracket respecting the asymptotic formulas is extensively studied.

On the other hand, (3.23) is coherent with the asymptotic formula at all $z \in \mathbb{R}^d$, by construction, but lacks of a simpler representation.

The problem for the pointwise set-valued bracket has been partially treated in [42], and will be concluded in [43], by using different techniques w.r.t. this paper. We just point out here that a useful tool to study the cluster points of $B(X_1^\tau, \ldots, X_k^\tau)(x)$ as $\tau \to 0$ is provided by the following result, which is a simplified version of Theorem 9.67 in [70].

Proposition 3.4.23. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally Lipschitz function, and let $\{s_\rho\}_{\rho>0}$ be a sequence of smooth mollifiers. Set $f_\rho = f * s_\rho$. Then*

$$\mathrm{co} \bigcap_{\delta>0} \bigcap_{\rho>0} \overline{\{\nabla f_\tau(x') : x' \in B(x, \delta),\ 0 < \tau < \rho\}} = \partial_C f(x).$$

3.4.4 An Example

In this section we provide an example illustrating our approach.

In the example below, we first consider the case in which the measure μ is blind w.r.t. the singularity set H of the vector fields, i.e. the singularities of the vector fields are contained in a μ-negligible closed set. In this case, roughly speaking, we can neglect them and perform the computations exactly as in the classical case. In the same setting, we then analyze the behaviour of the system on the singular set H. To this aim, we will set $D = \mathbb{R}^d \setminus H$.

Example 3.4.24. In $\mathbb{R}^2$, set $H := \{(x, y) \in \mathbb{R}^2 : xy = 0\}$ and consider two Borel vector fields safifying for $(x, y) \in D$

$$X(x, y) := \sqrt{\frac{3}{5}} \cdot \frac{x}{y^{2/3}} \cdot (1, 1), \qquad Y(x, y) := X(y, x).$$

Since in the open set D these vector fields are smooth, we can set $\Psi_t(x, y) = [\phi_t^X, \phi_t^Y](x, y)$ for $(x, y) \in D$ and t small enough, thus for all $(x, y) \in D$ we have

$$\lim_{\substack{(u,w)\to(x,y)\\ t\to 0^+}} \frac{\Psi_t(u, v) - (u, v)}{t^2} = [X, Y](x, y) = \frac{x - y}{x^{2/3} y^{2/3}}(1, 1).$$

According to the representation formula, we have that if $V_2^p(\mu, \mathbf{\Psi}) \neq \emptyset$, we must have

$$V(x, y) = \frac{x-y}{x^{2/3}y^{2/3}}(1,1), \text{ for } \mu\text{-a.e. } (x, y) \in D \text{ and all } V \in V_2^p(\mu, \mathbf{\Psi}).$$

Thus if the map $(x, y) \mapsto \dfrac{x-y}{x^{2/3}y^{2/3}}(1,1) \in L^p_\mu(\mathbb{R}^d)$ and $\mu(D) = 1$, we obtain that $V_2^p(\mu, \mathbf{\Psi})$ reduces to the singleton $(x, y) \mapsto \dfrac{x-y}{x^{2/3}y^{2/3}}(1,1)$. For istance, this holds for $1 \leq p < 3/2$ and any $\mu \ll \mathscr{L}$ with compact support.

Fix $x_0 \neq 0$. For every $\delta, \sigma > 0$ the set $\overline{S^{\sigma,\delta}_{2,D}(x_0, 0)}$ is unbounded, since

$$\overline{S^{\sigma,\delta}_{2,D}(x_0, 0)} \supseteq \bigcap_{\sigma>0} \overline{S^{\sigma,\delta}_{2,D}(x_0, 0)} = \overline{\left\{ \frac{x-y}{x^{2/3}y^{2/3}}(1,1) : \, (x, y) \in B((x_0, 0), \delta) \cap D \right\}}.$$

According to the representation formula, we have that if $V^p_{2,D}(\mu, \mathbf{\Psi}) \neq \emptyset$, we must have for μ-a.e. $(x_0, 0) \in \mathbb{R}^2$

$$V(x_0, 0) \in \bigcap_{\sigma,\delta>0} \text{co}\overline{S^{\sigma,\delta}_{2,D}(x_0, 0)},$$

but this set is empty. Thus if $\mu(\{(x_0, 0) \, : \, x_0 > 0\}) > 0$ we have that $V^p_{2,D}(\mu, \mathbf{\Psi}) = \emptyset$. However, it is easy to show that for $1 < m < 2$ we have

$$\bigcap_{\sigma,\delta>0} \text{co}\overline{S^{\sigma,\delta}_{m,D}(x_0, 0)} = \{\lambda(1, 1) : \, \lambda \geq 0\}.$$

We can reason in a similar way on all the points of $H \setminus \{(0, 0)\}$.

Concerning the origin, we notice that

$$\bigcap_{\sigma,\delta>0} \text{co}\overline{S^{\sigma,\delta}_{2,D}(0, 0)} = \mathbb{R}^2,$$

thus in the case that $\mu(H \setminus \{(0, 0)\}) = 0$, we are able to define again $V(\cdot) \in V^p_{2,D}(\mu, \mathbf{\Psi})$ provided that $(x, y) \mapsto \dfrac{x-y}{x^{2/3}y^{2/3}}(1,1) \in L^p_\mu(\mathbb{R}^d \setminus \{(0, 0)\})$ (we can simply set $V(0, 0) = 0$).

Chapter 4

Time-optimal control problem in a non-isolated case

The formulation of the problem we are going to study in the present chapter (see [33]) is strictly related to the theory presented in the previous one for the mass-preserving case (cfr. the related papers [28, 30–32]), where a time-optimal control problem in the space of probability measures is investigated. As already discussed, in the mass-preserving case, the admissible trajectories are time-depending Borel probability measures on $\mathbb{R}^d$ solving an homogeneous (controlled) continuity equation

$$\partial_t \mu_t + \mathrm{div}(v_t \mu_t) = 0$$

in the distributional sense, thus granting the preservation of the total mass during the evolution. The Borel velocity field v_t is the control parameter, and ranges among $L^1_{\mu_t}$-selections of the multifunction F driving the underling ODE.

Given a set $S \subseteq \mathbb{R}^d$ closed and nonempty, we can choose as target set $\tilde{S} \subseteq \mathscr{P}(\mathbb{R}^d)$ the set of all the probability measures supported in S (recalling the concept of *classical counterpart*), and so the aim is to steer an initial state $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$ towards $\tilde{S}$ along a mass-preserving trajectory driven by an admissible velocity field v_t. The cost-functional associated to this trajectory driven by v_t is chosen to be the final time T for which $\mathrm{supp}(\mu_{|t=T}) \subseteq S$.

In that setting, the natural definition of minimum time function starting by a probability measure μ_0 is the infimum of the cost-functionals associated to admissible trajectories with initial state μ_0, as usual. We adopt the notation $\tilde{T} : \mathscr{P}(\mathbb{R}^d) \to [0, +\infty]$ to refer to the generalized minimum time function associated to this problem without the superscript Φ in Definition 3.2.10 since we are considering the case of existence of a classical counterpart for the target.

In this chapter we face a different problem, more related to the study of the *evacuation problem*, i.e. the problem to find the minimum time for a crowd to completely leave a region under some constraints on the trajectory of each pedestrian.

The problem we are going to introduce can be seen also as a *logistic problem involving non renewable resources.* More precisely, we consider again an initial state $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$, representing for example the initial statistic distribution of agents. At the initial time, to each agent of the system is given an amount of supplies depending on his/her initial position, represented by a funcion $f_0 : \mathbb{R}^d \to [0, +\infty]$ called *clock-function.* The aim for each agent is to reach a fixed region $S \subseteq \mathbb{R}^d$ (common for all the agents) before the full consumption of the provided supplies, which decrease linearly in time during the evolution. The goal is to find the minimum amount of supplies which must be assigned at the beginning to each agent to comply the task, together with the macroscopic description of the trajectories of the agents allowing them to reach S with this minimum amount of supplies.

Notice that we ask the target set S to be strongly invariant for F in order to remove the agents once they have achieved their own task.

Another possible way to interpret this problem as a *time-optimal control problem*, is to associate to each admissible mass-preserving trajectory $\boldsymbol{\mu}$ starting by μ_0, a function $f_0 : \mathbb{R}^d \to [0, +\infty]$ whose aim is to bound from above the time needed by the particles in the support of μ_0 to reach the target S following the trajectory $\boldsymbol{\mu}$.

This formulation gives us the possibility to study a new class of trajectories $\tilde{\boldsymbol{\mu}}$ for μ_0, called *clock-trajectories*, which are no longer mass-preserving, but time-depending positive measures which loose their mass linearly in time, as prescribed by the *clock-function* f_0. At this point, an upper bound on the time weighted on the initial agents' distribution to reach the target is given by $\displaystyle\int_{\mathbb{R}^d} f_0(x)\, d\mu_0(x)$, and we look for the least of these upper bounds.

We notice that such a formulation is different from the problem of instantaneous annihilation of the mass discussed in Section 2.3.

The main results obtained in this Chapter can be summarized as follows:

1. a theorem of existence of a solution for the problem, with a characterization of the value function in this case (Corollary 4.3.7);

2. a Dynamic Programming Principle (Corollary 4.3.8) and some regularity results on the value function (Corollaries 4.3.10 and 4.3.11);

3. the introduction of a natural Hamilton-Jacobi-Bellman equation for the value function of this problem, which turns out to be a solution in a suitable infinite-dimensional viscosity sense (Theorem 4.4.3).

4.1 Statement of the problem and preliminary results

We formalize now the objects involved in the present study recalling also the ones defined in Chapter 3 for the mass-preserving case as done below.

Definition 4.1.1. Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be a set-valued map, $\bar{\mu} \in \mathscr{P}(\mathbb{R}^d)$.

1. Let $T > 0$. We say that $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ is an *admissible mass-preserving trajectory defined on* $[0, T]$ *and starting from* $\bar{\mu}$ if there exists

$\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]} \subseteq \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ such that $\nu_t \ll \mu_t$ for a.e. $t \in [0,T]$, $\mu_0 = \bar{\mu}$, $\partial_t\mu_t + \operatorname{div}\nu_t = 0$ in the sense of distributions and $v_t(x) := \dfrac{\nu_t}{\mu_t}(x) \in F(x)$ for a.e. $t \in [0,T]$ and μ_t-a.e. $x \in \mathbb{R}^d$. In this case, we will say also that the admissible mass-preserving trajectory $\boldsymbol{\mu}$ is *driven* by $\boldsymbol{\nu}$.

2. Let $T > 0$, $\boldsymbol{\mu}$ be an admissible mass-preserving trajectory defined on $[0,T]$ starting from $\bar{\mu}$ and driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$. We will say that $\boldsymbol{\mu}$ *is represented by* $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ if we have $e_t\sharp\boldsymbol{\eta} = \mu_t$ for all $t \in [0,T]$ and $\boldsymbol{\eta}$ is concentrated on the pairs $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$ where γ is an absolutely continuous solution of
$$\begin{cases} \dot{\gamma}(t) = v_t(\gamma(t)), & \text{for a.e. } 0 < t \le T \\ \gamma(0) = x, \end{cases}$$
where $v_t(x) := \dfrac{\nu_t}{\mu_t}(x)$.

We now define the concept of clock-trajectory and clock-function. The fact that the clock is ticking downward is recapitulated by condition 4 of the following definition.

Notice that, since we want to define the admissible clock-trajectory for possible infinite times, we need to have a sequence of mass-preserving trajectories, each extending the previous one, defined in increasing finite time intervals. In this way, we can use resuts valid for separable metric spaces as Γ_T for every $0 < T < +\infty$.

Definition 4.1.2. Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be a set-valued map, $S \subseteq \mathbb{R}^d$ be closed, nonempty and strongly invariant for F, $\bar{\mu} \in \mathscr{P}(\mathbb{R}^d)$ with $\operatorname{supp}(\bar{\mu}) \subseteq \mathbb{R}^d \setminus S$.
A Borel family of positive and finite measures $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[} \subseteq \mathscr{M}^+(\mathbb{R}^d)$ is an *admissible clock-trajectory (curve)* for $\bar{\mu}$ with target S if there exist a Borel map $f_0 : \mathbb{R}^d \to [0,+\infty]$ called *clock-function*, and sequences $\{T_n\}_{n\in\mathbb{N}} \subseteq [0,+\infty[$, $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$, $\{\boldsymbol{\nu}^n\}_{n\in\mathbb{N}}$, and $\{\boldsymbol{\eta}_n\}_{n\in\mathbb{N}}$ such that

1. $T_n \to +\infty$;

2. for any $n \in \mathbb{N}$ we have that $\boldsymbol{\mu}^n = \{\mu^n_t\}_{t\in[0,T_n]}$ is an admissible mass-preserving trajectory defined on $[0,T_n]$, starting from $\bar{\mu}$, driven by $\boldsymbol{\nu}^n := \{\nu^n_t\}_{t\in[0,T_n]}$, and represented by $\boldsymbol{\eta}_n$;

3. given $n_1, n_2 \in \mathbb{N}$ with $T_{n_1} \le T_{n_2}$, we have $(\mathrm{Id}_{\mathbb{R}^d} \times r_{n_2,n_1})\sharp\boldsymbol{\eta}_{n_2} = \boldsymbol{\eta}_{n_1}$, where $r_{n_2,n_1} : \Gamma_{T_{n_2}} \to \Gamma_{T_{n_1}}$ is the linear and continuous operator defined by setting $r_{n_2,n_1}\gamma(t) = \gamma(t)$ for all $t \in [0,T_{n_1}]$. Clearly, $r_{n_2,n_1}\gamma \in \Gamma_{T_{n_1}}$ for all $\gamma \in \Gamma_{T_{n_2}}$. In particular, this implies $\mu^{n_1}_t = \mu^{n_2}_t$ for all $t \in [0,T_{n_1}]$.

4. for any $n \in \mathbb{N}$, $t \in [0,T_n]$, $\varphi \in C^0_C(\mathbb{R}^d)$, we have
$$\int_{\mathbb{R}^d} \varphi(x)\,d\tilde{\mu}_t = \iint_{\mathbb{R}^d\times\Gamma_{T_n}} \varphi(\gamma(t))\chi_{S^c}(\gamma(t))(f_0(x) - t)\,d\boldsymbol{\eta}_n(x,\gamma)$$

In this case we will say that $\tilde{\boldsymbol{\mu}}$ *follows the family of mass-preserving trajectories* $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$. Notice that, since we ask $\tilde{\mu}_0(\mathbb{R}^d) < +\infty$, then we can identify f_0 with $\dfrac{\tilde{\mu}_0}{\bar{\mu}} \in L^1_{\mu_0}$.

Remark 4.1.3. We recall that if the time-dependent vector field $v_t(x) := \dfrac{\nu_t}{\mu_t}(x)$ satisfies the assumption of the Superposition Principle (Theorem 1.3.3) then there exists $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ representing $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$.

Necessarily, since by Definition 4.1.2, $\tilde{\mu}_t$ is a positive measure, then we have the following first comparison result between an admissible clock-function and the classical minimum time function for the underlying finite-dimensional differential inclusion with target S.

Lemma 4.1.4 (Lower bound on the clock function)**.** *Let $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$, $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[}$ be an admissible clock-trajectory for μ_0 with clock-function f_0. Then we have $f_0(x) \geq T(x)$ for μ_0-a.e. $x \in \mathbb{R}^d$, where $T : \mathbb{R}^d \to [0,+\infty]$ is the classical minimum time function for the same target set $S \subseteq \mathbb{R}^d$.*

Proof. By assumption, let $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$ be a family of admissible mass-preserving trajectories starting from μ_0 represented by $\{\boldsymbol{\eta}_n\}_{n\in\mathbb{N}}$, and $\{T_n\}_{n\in\mathbb{N}} \subseteq [0,+\infty[$ such that $\tilde{\boldsymbol{\mu}}$ follows $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$, $T_n \to +\infty$ and $\boldsymbol{\mu}^n$ is defined on $[0,T_n]$. For any $t \geq 0$, chosen $T_n \geq t$, we have

$$\int_{\mathbb{R}^d} \varphi(x)\, d\tilde{\mu}_t(x) = \iint_{\mathbb{R}^d\times\Gamma_{T_n}} \varphi(\gamma(t))\chi_{S^c}(\gamma(t))(f_0(x)-t)\, d\boldsymbol{\eta}_n(x,\gamma).$$

In particular, since $\tilde{\mu}_t$ is a positive measure, we must have $f_0(x) \geq t$ for $\boldsymbol{\eta}_n$-a.e. (x,γ) such that $\gamma(t) \notin S$. Thus we must have $f_0(x) \geq t$ for $\boldsymbol{\eta}_n$-a.e. (x,γ) such that $t \leq \min\{T_n, T(x)\}$, i.e., $f_0(x) \geq t$ for μ_0-a.e. x with $0 < t \leq \min\{T_n, T(x)\}$, so $f_0(x) \geq \min\{T_n, T(x)\}$ for μ_0-a.e. $x \in \mathbb{R}^d$ and for all $n \in \mathbb{N}$. We conclude that $f_0(x) \geq T(x)$ for μ_0-a.e. $x \in \mathbb{R}^d$. $\square$

Proposition 4.1.5 (Clock trajectory and mass-preserving trajectory)**.** *Definition 4.1.2 is well-posed in the sense that it defines a Radon measure $\tilde{\mu}_t$ for all $t \geq 0$.*

Moreover, let $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[} \subseteq \mathscr{M}^+(\mathbb{R}^d)$ be an admissible clock-trajectory with clock-function f_0, and let us call with $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}} := \{\{\mu^n_t\}_{t\in[0,T_n]}\}_n$ the family of mass-preserving trajectories followed by $\tilde{\boldsymbol{\mu}}$. Then for all $n \in \mathbb{N}$ we have $\tilde{\mu}_t \ll \mu^n_t$ for all $t \in [0,T_n]$.

Proof. Let $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[} \subseteq \mathscr{M}^+(\mathbb{R}^d)$ be an admissible clock-trajectory for μ_0 with clock-function f_0 following the family of mass-preserving trajectories $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}} := \{\{\mu^n_t\}_{t\in[0,T_n]}\}_n$ represented by $\{\boldsymbol{\eta}_n\}_{n\in\mathbb{N}}$.

For any $n \in \mathbb{N}$ let us fix any $t \in [0,T_n]$. We disintegrate $\boldsymbol{\eta}_n$ with respect to the continuous map $e_0 : \mathbb{R}^d \times \Gamma_{T_n} \to \mathbb{R}^d$. This yields a family of probability measures $\{\eta^n_x\}_{x\in\mathbb{R}^d}$ which is uniquely defined $e_0\sharp\boldsymbol{\eta}_n$-a.e. such that $\boldsymbol{\eta}_n = \mu_0 \otimes \eta^n_x$ and so the right-hand side of Definition 4.1.2(4) can be written as

$$\int_{\mathbb{R}^d}\int_{\Gamma^x_{T_n}} \varphi(\gamma(t))\, \chi_{\mathbb{R}^d\setminus S}(\gamma(t)) \cdot (f_0(x)-t)\, d\eta^n_x(\gamma)\, d\mu_0(x),$$

where $\varphi(\gamma(t))\,\chi_{\mathbb{R}^d\setminus S}(\gamma(t))$ is l.s.c. in γ and $(f_0(x)-t)$ is Borel measurable in x, hence the integrand is Borel measurable w.r.t. η^n_x. Thus the whole expression is well-posed in terms of measurability.

Let us consider the operator $L : C^0_C(\mathbb{R}^d) \to [0,+\infty[$ defined as follows

$$L(\varphi) := \iint_{\mathbb{R}^d\times\Gamma_{T_n}} \varphi(\gamma(t))\,\chi_{\mathbb{R}^d\setminus S}(\gamma(t))\cdot(f_0(x)-t)\,d\boldsymbol{\eta}_n(x,\gamma).$$

In order to prove that Definition 4.1.2(4) gives a Radon measure $\tilde{\mu}_t \in \mathscr{M}^+(\mathbb{R}^d)$ we should prove that the operator L is linear and continuous w.r.t. the sup norm, hence $\tilde{\mu}_t \in [C^0_C(\mathbb{R}^d)]'$.

Claim 1: linearity. Immediate by definition of L.

Claim 2: continuity. Immediate by boundedness of $\varphi \in C^0_C$ and by the fact that $f_0 \in L^1_{\mu_0}$, indeed

$$\begin{aligned}|L(\varphi)| &= \left|\int_{\mathbb{R}^d}\int_{\Gamma^x_{T_n}} \varphi(\gamma(t))\,\chi_{\mathbb{R}^d\setminus S}(\gamma(t))\cdot(f_0(x)-t)\,d\eta^n_x(\gamma)\,d\mu_0(x)\right|\\ &\le \|\varphi\|_\infty\cdot\int_{\mathbb{R}^d}|f_0(x)-t|\,d\mu_0(x) < +\infty,\end{aligned}$$

Thus, recalling linearity property, we conclude continuity of the operator L.

For the last assertion, let us consider again any $n\in\mathbb{N}$ and any $t\in[0,T_n]$. We disintegrate $\boldsymbol{\eta}_n$ with respect to the continuous map $e_t : \mathbb{R}^d\times\Gamma_{T_n}\to\mathbb{R}^d$. This yields a family of probability measures $\{\eta^n_y\}_{y\in\mathbb{R}^d}$ which is uniquely defined $e_t\sharp\boldsymbol{\eta}_n$-a.e. such that

$$\begin{aligned}\int_{\mathbb{R}^d}\varphi(x)\,d\tilde{\mu}_t(x) &= \int_{\mathbb{R}^d}\int_{e_t^{-1}(y)}\varphi(y)\chi_{S^c}(y)(f_0(\gamma(0))-t)\,d\eta^n_y(x,\gamma)\,d\mu^n_t(y)\\ &= \int_{\mathbb{R}^d}\varphi(y)\chi_{S^c}(y)\left(\int_{e_t^{-1}(y)}f_0(\gamma(0))\,d\eta^n_y(x,\gamma)-t\right)d\mu^n_t(y),\end{aligned}$$

We define the Borel map (see Section 5.3 in [9])

$$\Psi^{\boldsymbol{\eta}_n}(t,y) := \int_{e_t^{-1}(y)} f_0(\gamma(0))\,d\eta^n_y(x,\gamma),$$

and we notice that $\tilde{\mu}_t \ll \mu^n_t$ for all $t\in[0,T_n]$ and for all $n\in\mathbb{N}$, with

$$\frac{\tilde{\mu}_t}{\mu^n_t}(y) = \chi_{S^c}(y)(\Psi^{\boldsymbol{\eta}_n}(t,y)-t),$$

in particular, for μ_0-a.e. $y\in\mathbb{R}^d$ we have $f_0(y)=\chi_{S^c}(y)\Psi^{\boldsymbol{\eta}_n}(0,y)$ for all $n\in\mathbb{N}$. □

4.2 Some results in a mass-preserving setting

In this section, we prove some approximation results on the mass-preserving trajectories on which our objects are built.

Given $N \in \mathbb{N}$, $N > 0$, consider a set of N agents moving along admissible trajectories $\gamma_i(\cdot)$, $i = 1, \dots, N$ of the differential inclusion $\dot{x}(t) \in F(x(t))$. We want to associate to the evolution of such a system an admissible mass-preserving trajectory.

Proposition 4.2.1 (Finite embedding of classical admissible trajectories)**.** *Assume hypothesis (F_0). Let $N \in \mathbb{N} \setminus \{0\}$, and consider a set of N admissible trajectories $\{\gamma_i(\cdot),\ i = 1, \dots, N\} \subseteq \Gamma_T$ of the differential inclusion $\dot{x}(t) \in F(x(t))$. For any $t \in [0, T]$, we define the empirical measures*

$$\boldsymbol{\eta}^N(x, \gamma) = \frac{1}{N} \sum_{i=1}^{N} \delta_{\gamma_i(0)} \otimes \delta_{\gamma_i} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T),$$

$$\mu_t^N = e_t \sharp \boldsymbol{\eta}^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\gamma_i(t)} \in \mathscr{P}(\mathbb{R}^d).$$

Then $\boldsymbol{\mu}^N = \{\mu_t^N\}_{t \in [0,T]}$ is an admissible mass-preserving trajectory driven by $\boldsymbol{\nu}^N = \{\nu_t^N\}_{t \in [0,T]}$ and represented by $\boldsymbol{\eta}^N$ for every $N \in \mathbb{N}$, where $\nu_t^N \in \mathscr{M}(\mathbb{R}^d; \mathbb{R}^d)$ is defined for a.e. $t \in [0, T]$ by

$$\nu_t^N = \frac{1}{N} \sum_{i=1}^{N} \dot{\gamma}_i(t) \delta_{\gamma_i(t)} \in \mathscr{M}(\mathbb{R}^d; \mathbb{R}^d).$$

Proof. For any $\varphi \in C_C^\infty(\mathbb{R}^d)$ and for a.e. $t \in [0, T]$ we have

$$\begin{aligned}\frac{d}{dt} \int_{\mathbb{R}^d} \varphi(x)\, d\mu_t^N &= \frac{1}{N} \sum_{i=1}^{N} \frac{d}{dt} \varphi(\gamma_i(t)) = \frac{1}{N} \sum_{i=1}^{N} \langle \nabla \varphi(\gamma_i(t)), \dot{\gamma}_i(t) \rangle \\ &= \int_{\mathbb{R}^d} \nabla \varphi(x)\, d \left(\frac{1}{N} \sum_{i=1}^{N} \dot{\gamma}_i(t) \delta_{\gamma_i(t)} \right),\end{aligned}$$

since the set in which $\dot{\gamma}_i(t)$ exists for all $i = 1, \dots, N$ has full measure in $[0, T]$.

Defining

$$\nu_t^N = \frac{1}{N} \sum_{i=1}^{N} \dot{\gamma}_i(t) \delta_{\gamma_i(t)} \in \mathscr{M}(\mathbb{R}^d; \mathbb{R}^d),$$

we obtain that $\boldsymbol{\mu}^N = \{\mu_t^N\}_{t \in [0,T]}$ and $\boldsymbol{\nu}^N = \{\nu_t^N\}_{t \in [0,T]}$ satisfy the continuity equation

$$\partial_t \mu_t + \operatorname{div} \nu_t = 0,$$

and $\nu_t^N \ll \mu_t^N$ for a.e. $t \in [0, T]$. We adopt now an Eulerian point of view: for any Borel set B we are interested in the average speed of the agents which at time t are inside B, i.e., for a.e. $t \in [0, T]$ we set

$$I_{B,t}^N := \{i \in \{1, \dots, N\} : \gamma_i(t) \in B\},$$

and so if $I_{B,t}^N \neq \emptyset$, we have

$$\frac{\nu_t^N(B)}{\mu_t^N(B)} = \frac{\dfrac{1}{N} \displaystyle\sum_{i \in I_{B,t}^N} \dot{\gamma}_i(t)}{\dfrac{1}{N} \displaystyle\sum_{i \in I_{B,t}^N} 1} = \frac{1}{|I_{B,t}^N|} \sum_{i \in I_{B,t}^N} \dot{\gamma}_i(t).$$

Fix now $x \in \mathbb{R}^d$ and $\varepsilon > 0$. Recalling that the set-valued map F is convex valued and upper semicontinuous, there exists $\delta > 0$ such that $F(y) \subseteq F(x) + \varepsilon B(0,1)$ for all $y \in B(x,\delta)$. In particular, if $I^N_{B(x,\delta),t} \neq \emptyset$

$$\frac{\nu^N_t(B(x,\delta))}{\mu^N_t(B(x,\delta))} = \frac{1}{|I^N_{B(x,\delta),t}|} \sum_{i \in I^N_{B(x,\delta),t}} \dot{\gamma}_i(t) \in F(x) + \varepsilon B(0,1).$$

We have that $I^N_{B(x,\delta),t} \neq \emptyset$ for all $\delta > 0$ if and only if $x \in \{\gamma_i(t) : i = 1, \ldots, N\}$, i.e., if and only if $x \in \operatorname{supp} \mu^N_t$. So for any $x \in \operatorname{supp} \mu^N_t$, by taking the limit for $\delta \to 0^+$ and then letting $\varepsilon \to 0^+$, we have

$$\frac{\nu^N_t}{\mu^N_t}(x) = \lim_{\delta \to 0^+} \frac{\nu^N_t(B(x,\delta))}{\mu^N_t(B(x,\delta))} \in F(x).$$

We thus obtain that $\boldsymbol{\mu}^N = \{\mu^N_t\}_{t\in[0,T]}$ is an admissible mass-preserving trajectory driven by $\boldsymbol{\nu}^N = \{\nu^N_t\}_{t\in[0,T]}$ and represented by $\boldsymbol{\eta}^N$ for every $N \in \mathbb{N}$. $\square$

We consider now the limit of the above construction as $N \to +\infty$ in the case $p > 1$.

Proposition 4.2.2 (Mean Field Limit)**.** *Assume hypothesis* (F_0) *and* (F_1)*. Let* $\{\gamma_i\}_{i\in\mathbb{N}} \subseteq \Gamma_T$ *be a sequence of admissible trajectories of the differential inclusion* $\dot{x}(t) \in F(x(t))$*,* $p > 1$*. For any* $N \in \mathbb{N} \setminus \{0\}$*, we define*

$$\begin{aligned}
\boldsymbol{\eta}^N &= \frac{1}{N}\sum_{i=1}^N \delta_{\gamma_i(0)} \otimes \delta_{\gamma_i} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T),\\
\mu^N_t &= e_t \sharp \boldsymbol{\eta}^N = \frac{1}{N}\sum_{i=1}^N \delta_{\gamma_i(t)} \in \mathscr{P}(\mathbb{R}^d) \text{ for all } t \in [0,T],\\
\nu^N_t &= \frac{1}{N}\sum_{i=1}^N \dot{\gamma}(t)\delta_{\gamma_i(t)} \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d) \text{ for a.e. } t \in [0,T].
\end{aligned}$$

Assume that there exists $C_1 > 0$ *such that*

$$\lim_{N\to+\infty} \mathrm{m}_p(\mu^N_0) = \sup_{N\to+\infty} \mathrm{m}_p(\mu^N_0) < C_1.$$

Then there exist a sequence $\{N_k\}_{k\in\mathbb{N}}$ *such that* $N_k \to +\infty$*,* $\boldsymbol{\eta}^\infty \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$*,* $\boldsymbol{\mu}^\infty = \{\mu^\infty_t\}_{t\in[0,T]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$*,* $\boldsymbol{\nu}^\infty = \{\nu^\infty_t\}_{t\in[0,T]}$*, such that*

a. $\boldsymbol{\eta}^{N_k} \rightharpoonup^* \boldsymbol{\eta}^\infty$*;*

b. $W_p(\mu^{N_k}_t, \mu^\infty_t) \to 0$ *for all* $t \in [0,T]$*;*

c. $\nu^{N_k}_t \rightharpoonup^* \nu^\infty_t$ *for a.e.* $t \in [0,T]$*;*

d. $\boldsymbol{\mu}^\infty$ *is an admissible trajectory driven by* $\boldsymbol{\nu}^\infty$ *and represented by* $\boldsymbol{\eta}^\infty$*;*

e. for any closed set $\mathcal{K} \subseteq \Gamma_T$ *such that* $\{\gamma_i\}_{i\in\mathbb{N}} \subseteq \mathcal{K}$*, we have that*

$$\operatorname{supp} \boldsymbol{\eta}^\infty \subseteq \{(\gamma(0),\gamma) \in \mathbb{R}^d \times \Gamma_T : \gamma \in \mathcal{K}\}.$$

We will say also that $\boldsymbol{\mu}^\infty$ *is a* mean field limit *associated to* $\{\gamma_i\}_{i\in\mathbb{N}} \subseteq \Gamma_T$.

Proof. Thanks to Proposition 4.2.1, we can apply Lemma 3.2.7 to $\boldsymbol{\mu}^N = \{\mu_t^N\}_{t\in[0,T]}$ and $\boldsymbol{\nu}^N = \{\nu_t^N\}_{t\in[0,T]}$, and we have that there exist $D', D'' > 0$ such that

$$\begin{aligned} \mathrm{m}_p(\mu_t^N) &\le D'\left(\mathrm{m}_p(\mu_0^N)+1\right) \le D'(C_1+1),\\ \mathrm{m}_{p-1}(|\nu_t^N|) &\le D''(C_1+1). \end{aligned} \tag{4.1}$$

Claim 1: The sequence $\{\boldsymbol{\eta}^N\}_{N\in\mathbb{N}}$ is tight, thus there exists a subsequence $\{\boldsymbol{\eta}^{N_k}\}_{k\in\mathbb{N}}$ and $\boldsymbol{\eta}^\infty \in \mathscr{P}(\mathbb{R}^d\times\Gamma_T)$ such that $\boldsymbol{\eta}^{N_k} \rightharpoonup^* \boldsymbol{\eta}^\infty$.

It is enough to prove that $\{r_k\sharp\boldsymbol{\eta}^N\}_{N\in\mathbb{N}}$, $k=1,2$, are tight, where $r_1 : \mathbb{R}^d\times\Gamma_T \to \mathbb{R}^d$ and $r_2 : \mathbb{R}^d\times\Gamma_T\to\Gamma_T$ are defined by $r_1(x,\gamma)=x$ and $r_2(x,\gamma)=\gamma$. Recalling Remark 5.1.5 in [9], it is enough to prove that there are two Borel functions $\psi_1 : \mathbb{R}^d \to [0,+\infty]$ and $\psi_2 : \Gamma_T \to [0,+\infty]$ with compact sublevels such that

$$\sup_{N\in\mathbb{N}}\int_{\mathbb{R}^d}\psi_1(y)\,d(r_1\sharp\boldsymbol{\eta}^N)(y) < +\infty, \qquad \sup_{N\in\mathbb{N}}\int_{\Gamma_T}\psi_2(\gamma)\,d(r_2\sharp\boldsymbol{\eta}^N)(\gamma) < +\infty.$$

We set

$$\psi_1(y) = |y|^p, \qquad\qquad \psi_2(\gamma) = \begin{cases} \displaystyle\int_0^T |\dot\gamma(t)|^p\,dt, & \text{if } \gamma\in AC^p([0,T]),\\[2ex] +\infty, & \text{otherwise.}\end{cases}$$

We have that $\psi_2(\cdot)$ has compact sublevels if $p>1$. We recall that if $\dot\gamma(t) \in F\circ\gamma(t)$ for a.e. t, then by (F_1) we can apply Lemma 1.4.3 to have

$$|\gamma(t)| \le (|\gamma(0)| + Ct)e^{Ct}.$$

Indeed, for all $N\in\mathbb{N}$ we have

$$\begin{aligned} \int_{\mathbb{R}^d}\psi_1(y)\,d(r_1\sharp\boldsymbol{\eta}^N)(y) &= \iint_{\mathbb{R}^d\times\Gamma_T}|x|^p\,d\boldsymbol{\eta}^N(x,\gamma) = \mathrm{m}_p(\mu_0^N) \le C_1,\\ \int_{\Gamma_T}\psi_2(\gamma)\,d(r_2\sharp\boldsymbol{\eta}^N)(\gamma) &\le \iint_{\mathbb{R}^d\times\Gamma_T}\left(\int_0^T C^p(|\gamma(t)|+1)^p\,dt\right)d\boldsymbol{\eta}^N(x,\gamma)\\ &\le TC^p\int_{\mathbb{R}^d}((|x|+CT)e^{CT}+1)^p\,d\mu_0^N(x)\\ &\le TC^p(e^{CT}\mathrm{m}_p^{1/p}(\mu_0^N)+CTe^{CT}+1)^p\\ &\le TC^p(e^{CT}C_1^{1/p}+CTe^{CT}+1)^p, \end{aligned}$$

which confirms Claim 1. ◇

Claim 2: Set $\mu_t^\infty = e_t\sharp\boldsymbol{\eta}^\infty$. Then $\boldsymbol{\mu}^\infty = \{\mu_t^\infty\}_{t\in[0,T]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ and $W_p(\mu_t^{N_k},\mu_t^\infty)\to 0$ as $k\to+\infty$ for all $t\in[0,T]$. Moreover, for a.e. $t\in[0,T]$ the sequence $\{\nu_t^N\}_{N\in\mathbb{N}}$ is tight, thus up to a non relabeled subsequence, it weakly* converges to a measure $\nu_t^\infty \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$.

Since the map $e_t : \mathbb{R}^d \times \Gamma_T \to \mathbb{R}^d$ is continuous, we have that

$$\mu_t^{N_k} = e_t \sharp \boldsymbol{\eta}^{N_k} \rightharpoonup^* e_t \sharp \boldsymbol{\eta}^\infty = \mu_t^\infty, \text{ for all } t \in [0,T].$$

All the other assertions follow from the fact that the moments of μ_t^N are uniformly bounded for $N \in \mathbb{N}$ and $t \in [0,T]$, also the tightness of $\{\nu_t^N\}_{N\in\mathbb{N}}$ follows from (4.1). $\diamond$

Claim 3: $\boldsymbol{\mu}^\infty$ is an admissible trajectory driven by $\boldsymbol{\nu}^\infty = \{\nu_t^\infty\}_{t\in[0,T]}$.

Notice that the functionals

$$(\boldsymbol{\mu},\boldsymbol{\nu}) \mapsto \begin{cases} \displaystyle\int_0^T \int_{\mathbb{R}^d} \left[\left| \frac{\nu_t}{\mu_t}(x) \right|^p + I_{F(x)} \left(\frac{\nu_t}{\mu_t}(x) \right) \right] d\mu_t(x)\, dt, & \text{if } \nu_t \ll \mu_t \text{ for a.e. } t, \\ +\infty, \text{ otherwise,} \end{cases}$$

$$(\boldsymbol{\mu},\boldsymbol{\nu}) \mapsto \sup_{\varphi \in C^1_C([0,T]\times\mathbb{R}^d)} \int_0^T \left(\int_{\mathbb{R}^d} \partial_t \varphi d\mu_t + \int_{\mathbb{R}^d} \nabla\varphi d\nu_t \right) dt,$$

are l.s.c. w.r.t. a.e. pointwise weak* convergence of measures (see Lemma 2.2.3, p. 39, Theorem 3.4.1, p.115, and Corollary 3.4.2 in [18] or Theorem 2.34 in [6]). Then we have that the equation

$$\partial_t \mu_t^\infty + \operatorname{div} \nu_t^\infty = 0$$

holds in the sense of distributions, and for a.e. $t \in [0,T]$ we have $\nu_t^\infty \ll \mu_t^\infty$, $\dfrac{\nu_t^\infty}{\mu_t^\infty}(x) \in F(x)$ for μ_t^∞-a.e. $x \in \mathbb{R}^d$ with $\dfrac{\nu_t^\infty}{\mu_t^\infty}(x) \in L^p_{\mu_t^\infty}$. $\diamond$

Consider now the last assertion to be proved. Let $(x,\gamma) \in \operatorname{supp} \boldsymbol{\eta}^\infty$. By Proposition 5.1.8 in [9] there exists a sequence $\{\hat{\gamma}_k\}_{k\in\mathbb{N}} \in \Gamma_T$ such that $(\hat{\gamma}_k(0),\hat{\gamma}_k) \in \operatorname{supp} \boldsymbol{\eta}^N$ for all $N \in \mathbb{N}$ and $\|\hat{\gamma}_k - \gamma\|_\infty \to 0$. By definition of $\boldsymbol{\eta}^N$ we have $\hat{\gamma}_k = \gamma_{j_k}$ for an index $0 < j_k \le N$, and so $\{\hat{\gamma}_k\}_{k\in\mathbb{N}} \subseteq \mathcal{K}$, thus $\gamma \in \mathcal{K}$.

$\square$

Remark 4.2.3. We notice that the tightness of $\{\mu_t^N\}_{N\in\mathbb{N}}$ holds also in the case $p = 1$ by (4.1).

The following result provides us with the possibility to construct an admissible mass-preserving trajectory $\boldsymbol{\mu} := \{\mu_t\}_{t\in[0,T]}$, i.e., a curve in $\mathscr{P}(\mathbb{R}^d)$ that satisfies a continuity equation with velocity field that is an $L^p_{\mu_t}$-selection of the multifunction F, by constructing it on admissible trajectories of the finite-dimensional system of characteristics in a consistent way.

Corollary 4.2.4. *Assume hypothesis* (F_0) *and* (F_1). *Let* $p > 1$, $K \subseteq \mathbb{R}^d$ *be closed,* $f \in C^0(\mathbb{R}^d; [0,T])$.

1. *For any sequence* $\{\gamma_i\}_{i\in\mathbb{N}}$ *of admissible trajectories of the differential inclusion* $\dot{x}(t) \in F(x(t))$ *satisfying*

$$\lim_{N\to+\infty} \frac{1}{N} \sum_{i=1}^N |\gamma_i(0)|^p < +\infty, \qquad \gamma_i(f(\gamma_i(0))) \in K \textit{ for all } i \in \mathbb{N},$$

we have that all the corresponding mean field limits $\boldsymbol{\mu}^\infty$ are represented by measures $\boldsymbol{\eta}^\infty$ such that γ is an admissible trajectory of the differential inclusion satisfying $\gamma(f(\gamma(0))) = \gamma(f(x)) \in K$ and $\gamma(0) = x$, for $\boldsymbol{\eta}^\infty$-a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_T$.

2. *For any $\mu \in \mathscr{P}_p(\mathbb{R}^d)$ such that for μ-a.e. $x \in \mathbb{R}^d$ there exists an admissible trajectory for the finite-dimensional system $\dot\gamma(t) \in F(\gamma(t))$ satisfying $\gamma(0) = x$ and $\gamma \circ f(x) \in K$, there exist $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ and $\boldsymbol{\eta}$ such that $\boldsymbol{\mu}$ is an admissible mass-preserving trajectory represented by $\boldsymbol{\eta}$ with $\mu_0 = \mu$, and γ is an admissible trajectory of the differential inclusion satisfying $\gamma(f(\gamma(0))) = \gamma(f(x)) \in K$ and $\gamma(0) = x$, for $\boldsymbol{\eta}$-a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_T$.*

Proof. For the first assertion is enough to notice that the set

$$\mathcal{K} := \{\gamma \in \Gamma_T : \gamma(f(\gamma(0))) \in K\}$$

is closed in Γ_T and then apply Proposition 4.2.2. In the second case, we have that there exists a sequence of compact sets $\{C_j\}_{j\in\mathbb{N}}$ such that $\mu(\mathbb{R}^d \setminus C_j) \le \dfrac{1}{j}$ for all $j \in \mathbb{N} \setminus \{0\}$. Set

$$\mu_j(B) = \frac{1}{\mu(C_j)} \mu(B \cap C_j) \in \mathscr{P}(\mathbb{R}^d),$$

clearly $\mu_j \rightharpoonup^* \mu$ and $\mathrm{m}_p(\mu_j) \to m_p(\mu)$ as $j \to +\infty$ by Dominated Convergence Theorem, thus $W_p(\mu_j, \mu) \to 0$. There exists $\{x_{i,j}\}_{i,j\in\mathbb{N}}$ such that $x_{i,j} \in C_j$ for all $i, j \in \mathbb{N}$ and

$$\mu_0^{k,j} = \frac{1}{k}\sum_{i=1}^{k} \delta_{x_{i,j}} \rightharpoonup^* \mu_j, \text{ as } k \to +\infty.$$

Since $\operatorname{supp} \mu_0^{k,j} \subseteq C_j$ and $\operatorname{supp} \mu_j \subseteq C_j$, we have also $\mathrm{m}_p(\mu_0^{k,j}) \to \mathrm{m}_p(\mu_j)$ as $k \to +\infty$. For any $j \in \mathbb{N}$, let $k_j \in \mathbb{N}$ be such that

$$\mathrm{m}_p(\mu_0^{k_j,j}) \le \mathrm{m}_p(\mu_j) + \frac{1}{j} \text{ and } W_p\left(\mu_0^{k_j,j}, \mu_j\right) \le \frac{1}{j}.$$

In particular, we have $W_p\left(\mu_0^{k_j,j}, \mu\right) \le \dfrac{1}{j} + W_p(\mu, \mu_j) \to 0^+$ as $j \to +\infty$, and so

$$\sup_{j\in\mathbb{N}} \mathrm{m}_p(\mu_0^{k_j,j}) < +\infty.$$

Consider the countable set of points $\{x_{i,j} : i = 1, \dots, k_j,\ j = 1, \dots, \infty\}$. We can order it by stating that $(i, j) < (i', j')$ if either $j < j'$ or $j = j'$ and $i < i'$, thus we obtain the sequence of points $\{x_h\}_{h\in\mathbb{N}}$. By assumption, for each $h \in \mathbb{N}$ there exists $\gamma_h \in \Gamma_T$ admissible trajectory of the differential inclusion satisfying $\gamma_h(0) = x_h$ and $\gamma_h \circ f(x_h) \in K$. We then apply item (1) to this sequence to conclude the proof. □

Remark 4.2.5. The assumption $f \in C^0(\mathbb{R}^d)$ of the previous corollary can be weakened by assuming that $f(\cdot)$ is continuous at x for μ_0-a.e. $x \in \mathbb{R}^d$ or, equivalently, that the set of discontinuities of $f(\cdot)$ are contained in a μ_0-negligible closed set.

4.3 A Dynamic Programming Principle

This section is devoted to state a time-optimal control problem in the space of positive finite Borel measures for a non-isolated case with mass loss using the definition of clock-trajectory given in Definition 4.1.2 and then prove a Dynamic Programming Principle related to such a minimization problem.

From now on, we will consider only closed, nonempty and strongly invariant target sets for our dynamics.

Definition 4.3.1 (Clock-generalized minimum time)**.** Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be a set-valued function, $S \subseteq \mathbb{R}^d$ be a target set for F. In analogy with the classical case, we define the *clock-generalized minimum time function* $\tau : \mathscr{P}(\mathbb{R}^d) \to [0, +\infty]$ by setting

$$\tau(\mu_0) := \inf \big\{ \tilde{\mu}_0(\mathbb{R}^d) : \tilde{\boldsymbol{\mu}} := \{\tilde{\mu}_t\}_{t\in[0,+\infty[} \subseteq \mathscr{M}^+(\mathbb{R}^d) \text{ is an admissible clock-} \tag{4.2}$$
$$\text{-trajectory for the measure } \mu_0,\ \tilde{\mu}_{|t=0} = \tilde{\mu}_0 \big\},$$

where, by convention, $\inf \emptyset = +\infty$.

Given $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$ with $\tau(\mu_0) < +\infty$, an admissible clock-curve $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[} \subseteq \mathscr{M}^+(\mathbb{R}^d)$ for μ_0 is *optimal for* μ_0 if

$$\tau(\mu_0) = \tilde{\mu}_{|t=0}(\mathbb{R}^d).$$

Given $p \geq 1$, we define also a clock-generalized minimum time function $\tau_p : \mathscr{P}_p(\mathbb{R}^d) \to [0, +\infty]$ by replacing in the above definitions $\mathscr{P}(\mathbb{R}^d)$ by $\mathscr{P}_p(\mathbb{R}^d)$ and $\mathscr{M}^+(\mathbb{R}^d)$ by $\mathscr{M}^+_p(\mathbb{R}^d)$. Since $\mathscr{M}^+_p(\mathbb{R}^d) \subseteq \mathscr{M}^+(\mathbb{R}^d)$, it is clear that $\tau_p(\mu_0) \geq \tau(\mu_0)$.

The main task of this section is to prove a Dynamic Programming Principle for our minimization problem. To this end we will prove a representation result expressing $\tau(\mu)$ as an average of the classical minimum-time function $T(\cdot)$, and then applying the well-known Dynamic Programming Principle (Theorem 1.4.8) holding for $T(\cdot)$.

Before treating the case with milder assumptions in Section 4.3.2, we will see a result yielding the Dynamic Programming Principle in a more regular case (Section 4.3.1).

4.3.1 Regular case

Lemma 4.3.2 (Extension)**.** *Assume hypothesis* (F_0) *and* (F_1)*. Let* $p > 1$ *and* $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$*. Let* $T > 0$ *and* $\bar{\boldsymbol{\mu}} = \{\bar{\mu}_t\}_{t\in[0,T]}$ *be an admissible mass-preserving trajectory driven by* $\bar{\boldsymbol{\nu}} = \{\bar{\nu}_t\}_{t\in[0,T]}$ *and represented by* $\bar{\boldsymbol{\eta}} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$*, with* $\bar{\mu}_{|t=0} = \mu_0$*. Then there exist a sequence* $\{T_n\}_{n\in\mathbb{N}} \subseteq [0, +\infty[$*,* $T_n \geq T$ *for all* $n \in \mathbb{N}$*,* $T_n \to +\infty$ *and a family of admissible mass-preserving trajectories* $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$*,* $\boldsymbol{\mu^n} = \{\mu^n_t\}_{t\in[0,T_n]}$*, driven by* $\{\boldsymbol{\nu}^n\}_{n\in\mathbb{N}}$*, such that given* $n_1, n_2 \in \mathbb{N}$ *with* $T_{n_1} \leq T_{n_2}$*, we have* $\mu^{n_1}_t = \mu^{n_2}_t$ *for all* $t \in [0, T_{n_1}]$*, and there exists a sequence* $\{\boldsymbol{\eta}_n\}_{n\in\mathbb{N}}$ *such that* $\boldsymbol{\eta}_n$ *represents* $\{\mu^n_t\}_{t\in[0,T_n]}$*.*

Proof. For any $\varepsilon > 0$, let us define by induction an increasing sequence $\{T_n\}_{n\in\mathbb{N}}$ such that $T_n \to +\infty$. Take $T_0 := T$, and suppose to have defined T_i, $i \geq 0$. Then $T_{i+1} := T_i + \varepsilon$, for all $i \in \mathbb{N}$.

We can define by induction the family $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$, $\boldsymbol{\mu}^n := \{\mu^n_t\}_{t\in[0,T_n]}$, in the following way. We take $\boldsymbol{\mu}^0 = \bar{\boldsymbol{\mu}}$. Let us suppose to have defined $\boldsymbol{\mu}^i$, $i \geq 0$. Then, for any $i \in \mathbb{N}$ we define $\boldsymbol{\mu}^{i+1}$ as follows. Consider a continuous selection v^{i+1} of F and the solution $\{\hat{\mu}^{i+1}_t\}_{t\in[0,\varepsilon]}$ of

$$\begin{cases} \partial_t \mu_t + \operatorname{div} v^{i+1}\mu_t = 0, \\ \mu_{|t=0} = \mu^i_{T_i} \end{cases}$$

By setting

$$\mu^{i+1}_t := \begin{cases} \mu^i_t, & \text{for } 0 \leq t < T_i, \\ \\ \hat{\mu}^{i+1}_{t-T_i}, & \text{for } T_i \leq t \leq T_i + \varepsilon = T_{i+1}, \end{cases}$$

$$\nu^{i+1}_t := \begin{cases} \nu^i_t, & \text{for } 0 \leq t < T_i, \\ \\ v^{i+1}\hat{\mu}^{i+1}_{t-T_i}, & \text{for } T_i \leq t \leq T_i + \varepsilon = T_{i+1}, \end{cases}$$

then by gluing results (see Lemma 4.4 in [41]) we obtain an admissibile trajectory $\boldsymbol{\mu}^{i+1} = \{\mu^{i+1}_t\}_t$ driven by $\boldsymbol{\nu}^{i+1} = \{\nu^{i+1}_t\}_t$ which is defined on $[0, T_{i+1}]$ and agrees with $\boldsymbol{\mu}^i$ on $[0, T_i]$. The last assertion follows from the Superposition Principle on the family of admissible trajectories $\{\mu^n_t\}_{n\in\mathbb{N}}$. □

In the following result we prove the existence of optimal trajectories in the case in which $T(\cdot)$ is continuous. As we can imagine, the classical minimum time function turns out to be the optimal clock function.

Lemma 4.3.3. *Assume that $T(\cdot)$ is continuous, $p > 1$ and that (F_0) and (F_1) hold true. Let $S \subseteq \mathbb{R}^d$ be a target set for F. Given $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$ with $\operatorname{supp}\mu_0 \subseteq \mathbb{R}^d \setminus S$, such that $\bar{T} := \|T\|_{L^\infty_{\mu_0}} < +\infty$, then there exists an admissible clock-trajectory $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[}$ with target S for μ_0 with clock-function $T(\cdot)$.*

Proof. We take $f(\cdot) = T(\cdot)$ in Corollary 4.2.4 with $T = \bar{T}$ and with $K = S$, obtaining an admissible mass-preserving trajectory represented by $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_{\bar{T}})$ satisfying $\gamma(T(x)) \in S$ for a.e. $(x, \gamma) \in \boldsymbol{\eta}$.

We can use Lemma 4.3.2 to construct a sequence $\{T_n\}_{n\in\mathbb{N}}$, $T_n \geq \bar{T}$, $T_n \to +\infty$, and an extended family of admissible mass preserving trajectories represented by $\{\boldsymbol{\eta}_n\}_{n\in\mathbb{N}} \subseteq \mathscr{P}(\mathbb{R}^d \times \Gamma_{T_n})$ satisfying $\gamma(T(x)) \in S$ for a.e. $(x, \gamma) \in \boldsymbol{\eta}_n$. In particular, by the strongly invariance of S, we have that if $T(x) < t \leq T_n$ then $\chi_{S^c}(\gamma(t)) = 0$. Thus $\chi_{S^c}(\gamma(t))(T(x) - t) \geq 0$ for all $t \in [0, T_n]$ and a.e. $(x, \gamma) \in \boldsymbol{\eta}_n$. Then we can construct by definition an admissible clock-trajectory following the family of admissible mass-preserving trajectories represented by $\{\boldsymbol{\eta}_n\}_{n\in\mathbb{N}}$ and with clock-function $T(\cdot)$. □

Remark 4.3.4. As remarked for Corollary 4.2.4, we can weaken the assumption of continuity on $T(\cdot)$ by requiring that T is continuous at μ_0-a.e. $x \in \mathbb{R}^d$.

The Dynamic Programming Principle is then a direct consequence of Lemma 4.1.4 and Lemma 4.3.3 which together say that, under regularity hypothesis, it is possible to construct an admissible clock-trajectory with clock-function $T(\cdot)$ and this turns out to be an optimal trajectory for the system. Hence we conclude by applying the result holding for the classical minimum-time function.

4.3.2 L^1 case

In this section we will see how to prove a Dynamic Programming Principle (Corollary 4.3.8) requiring a natural assumption, i.e. boundedness of the L^1-norm of $T(\cdot)$ w.r.t. a given initial measure $\mu \in \mathscr{P}_p(\mathbb{R}^d)$. The proof is based on a result of optimality of the classical minimum time function among the admissible clock-functions for a given initial measure (Corollary 4.3.7). The main tools used are selection and disintegration results.

It is possible to note that we can actually construct such optimal trajectories by approximation techniques, in particular by Lusin's Theorem and Corollary 4.2.4 (see the forthcoming paper [34]), however we will not present this construction here since it is not necessary for present purposes.

Lemma 4.3.5 (Borel selection of optimal trajectories)**.** *Let $T > 0$, $\mathscr{R} = T^{-1}([0, +\infty[)$, and $\mu \in \mathscr{P}(\mathbb{R}^d)$ be such that $\mu(\mathbb{R}^d \setminus \mathscr{R}) = 0$. Then there exist*

1. *a Borel map $\psi : \mathscr{R} \to \Gamma_T$ such that $\gamma_x := \psi(x)$ is an admissible trajectory starting from x,*
2. *an optimal trajectory $\hat{\gamma}_x : [0, T(x)] \to \mathbb{R}^d$ such that $\hat{\gamma}_x(t) = \gamma_x(t)$ for all $t \in [0, T]$,*
3. *an admissible mass-preserving trajectory $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ with $\mu_0 = \mu$, driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$, and represented by $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ with*

$$\boldsymbol{\eta} = \mu \otimes \delta_{\gamma_x}.$$

Proof. Define the set of admissible trajectories defined in $[0, T]$ for the finite-dimensional system, $\mathcal{A}_T \subseteq \Gamma_T$, and the set-valued map $G_T : \mathscr{R} \rightrightarrows \Gamma_T$ by

$$\mathcal{A}_T := \{\gamma \in \Gamma_T : \dot{\gamma} \in F \circ \gamma(t) \text{ for a.e. } 0 < t < T\},$$
$$G_T(x) := \begin{cases} \{\gamma \in \mathcal{A}_T : \gamma(0) = x, \text{ and } T(\gamma(0)) = T(\gamma(T)) + T\}, & \text{for } T < T(x), \\ \\ \{\gamma \in \mathcal{A}_T : \gamma(0) = x, \text{ and } \gamma(T(x)) \in S\}, & \text{for } T \geq T(x). \end{cases}$$

We notice that $G_T(x)$ is closed and nonempty for every $x \in \mathscr{R}$. Given $(x, \gamma) \in \mathscr{R} \times \Gamma_T$, we have that $\gamma \in G(x)$ if and only if there exists an optimal trajectory $\hat{\gamma}$ defined on $[0, T(x)]$ starting from x such that $\hat{\gamma}(t) = \gamma(t)$ for all $0 \leq t \leq \min\{T, T(x)\}$. Define the map

$$g(x, \gamma) := \begin{cases} I_x(\gamma(0)) + I_{\mathcal{A}_T}(\gamma) + I_S(\gamma(T(x))), & \text{if } T \geq T(x), \\ \\ I_x(\gamma(0)) + I_{\mathcal{A}_T}(\gamma) + I_{\{0\}}(T(x) - T(\gamma(T)) - T), & \text{if } T < T(x), \end{cases}$$

and notice that $(x, \gamma) \in \mathrm{Graph}(G_T)$ if and only if $g(x, \gamma) = 0$. Since we have

$$g(x, \gamma) = \sup_{\substack{q_1, q_2 \in \mathbb{R}^d \\ q_3 \in \mathbb{R}}} \Big\{ q_1(x - \gamma(0)) + I_{\mathcal{A}_T}(\gamma) + \chi_{[0,T]}(T(x))[\langle q_2, \gamma(T(x))\rangle - \sigma_S(q_2)] +$$
$$+ (1 - \chi_{[0,T]}(T(x)))q_3(T(x) - T(\gamma(T)) - T) \Big\},$$

we have that g is the pointwise supremum of Borel maps, and so it is Borel (we recall that $\gamma \mapsto I_{\mathcal{A}_T}(\gamma)$ is l.s.c. since $\mathcal{A}_T$ is closed, and $\gamma \mapsto T(\gamma(T))$ is l.s.c.).

Hence $\operatorname{Graph} G_T = g^{-1}(0)$ is a Borel set. By Theorem 8.1.4 p. 310 in [13], we have that the set-valued map $G_T : \mathscr{R} \rightrightarrows \Gamma_T$ is Borel measurable, and so by Theorem 8.1.3 p. 308 in [13] it admits a Borel selection $\psi : \mathscr{R} \to \Gamma_T$.

Since $\mu(\mathbb{R}^d \setminus \mathscr{R}) = 0$ we can define the probability measure

$$\boldsymbol{\eta} = \mu \otimes \delta_{\psi(x)} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T),$$

which is concentrated on (x, γ) such that γ is an admissible curve of the finite-dimensional system satisfying $\gamma(0) = x$, and $\gamma(T(x)) \in S$ if $T \geq T(x)$, or $T(\gamma(0)) = T(\gamma(T)) + T$, if $T(x) > T$, i.e., there exists an optimal trajectory $\hat{\gamma}$ defined on $[0, T(x)]$ such that $\hat{\gamma}(t) = \gamma(t)$ on $[0, T]$. This definition of $\boldsymbol{\eta}$ induces a curve $\boldsymbol{\mu} = \{\mu_t\}_{t \in [0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ defined by

$$\int_{\mathbb{R}^d} \varphi(x)\, d\mu_t(x) = \iint_{\mathbb{R}^d \times \Gamma_T} \varphi(\gamma(t))\, d\boldsymbol{\eta}(x, \gamma),$$

for all $\varphi \in C^0_C(\mathbb{R}^d)$, with $\mu_{|t=0} = \mu$. We want to show that $\boldsymbol{\mu}$ is an admissible mass-preserving trajectory.

The set $\mathcal{N}$ of $(t, x, \gamma) \in [0, T] \times \mathbb{R}^d \times \Gamma_T$ for which $\gamma(0) \neq x$ or $\dot{\gamma}(t)$ does not exists or $\dot{\gamma}(t) \notin F(\gamma(t))$ is $\mathscr{L}^1 \otimes \boldsymbol{\eta}$-negligible as seen at the beginning of Section 2.3, thus by projection on the first component, we have that $\dot{\gamma}(t) \in F(\gamma(t))$ for $\boldsymbol{\eta}$-a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_T$ and a.e. $t \in [0, T]$. For a.e. $t \in [0, T]$ we disintegrate $\boldsymbol{\eta}$ w.r.t. $e_t : \mathbb{R}^d \times \Gamma_T \to \mathbb{R}^d$, obtaining $\boldsymbol{\eta} = \mu_t \otimes \boldsymbol{\eta}_{t,y}$

$$\begin{aligned}
\frac{d}{dt} \int_{\mathbb{R}^d} \varphi(x)\, d\mu_t(x) &= \iint_{\mathbb{R}^d \times \Gamma_T} \nabla\varphi(\gamma(t)) \cdot \dot{\gamma}(t)\, d\boldsymbol{\eta}(x, \gamma) \\
&= \int_{\mathbb{R}^d} \int_{e_t^{-1}(y)} \nabla\varphi(\gamma(t)) \cdot \dot{\gamma}(t)\, d\boldsymbol{\eta}_{t,y}(x, \gamma)\, d\mu_t(y) \\
&= \int_{\mathbb{R}^d} \nabla\varphi(y) \cdot \int_{e_t^{-1}(y)} \dot{\gamma}(t)\, d\boldsymbol{\eta}_{t,y}(x, \gamma)\, d\mu_t(y),
\end{aligned}$$

We define $\boldsymbol{\nu} = \{v_t \mu_t\}_{t \in [0,T]}$ by setting for a.e. $t \in [0, T]$

$$v_t(y) = \int_{e_t^{-1}(y)} \dot{\gamma}(t)\, d\boldsymbol{\eta}_{t,y}(x, \gamma).$$

In order to conclude that $\boldsymbol{\mu}$ is an admissible trajectory driven by $\boldsymbol{\nu}$, it is enough to show that

$$\int_{e_t^{-1}(y)} \dot{\gamma}(t)\, d\boldsymbol{\eta}_{t,y}(x, \gamma) \in F(y)$$

for μ_t-a.e. $y \in \mathbb{R}^d$ and a.e. $t \in [0, T]$. This follows from Jensen's inequality, since

$$I_{F(y)}\left(\int_{e_t^{-1}(y)} \dot{\gamma}(t)\, d\boldsymbol{\eta}_{t,y}(x, \gamma) \right) \leq \int_{e_t^{-1}(y)} I_{F(y)}(\dot{\gamma}(t))\, d\boldsymbol{\eta}_{t,y}(x, \gamma) = 0.$$

□

Definition 4.3.6 (Movements along time-optimal trajectories)**.** Let $T > 0$, $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$. We say that $(\{\mu_t\}_{t\in[0,T[}, \{\nu_t\}_{t\in[0,T[})$ is a *movement along time-optimal curves from* μ_0 (μ_0-MATOC) if

a. there exists $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ such that for $\boldsymbol{\eta}$-a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_T$ we have $\gamma \in AC([0, T]; \mathbb{R}^d)$ and $\gamma(0) = x$, $\dot\gamma(t) \in F(\gamma(t))$ for a.e. $t \in [0, T]$, and either $\gamma(T(x)) \in S$ if $T(x) \le T$ or $T(x) = T(\gamma(T)) + T$ if $T(x) > T$;

b. $\mu_{|t=0} = \mu_0$, $\mu_t = e_t\sharp\boldsymbol{\eta}$ for all $t \in [0, T[$, and we set $\mu_T = e_T\sharp\boldsymbol{\eta}$;

c. $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ is an admissible mass-preserving trajectory driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$.

Corollary 4.3.7 (Optimal clock)**.** *Assume hypothesis* (F_0) *and* (F_1)*. Let* $S \subseteq \mathbb{R}^d$ *be a target set for* F*. Let* $p > 1$ *and* $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$*, with* $\mathrm{supp}\mu_0 \subseteq \mathbb{R}^d \setminus S$*, be such that* $\|T(\cdot)\|_{L^1_{\mu_0}} < +\infty$*. Then* $T(\cdot)$ *is the optimal clock function for* μ_0*.*

Proof. By assumption, we have that $\mu_0(\mathbb{R}^d \setminus \mathscr{R}) = 0$.

We consider the set (see Definition 4.3.6)

$$\mathcal{X} := \left\{ \left(\{\mu_t\}_{t\in[0,T[}, \{\nu_t\}_{t\in[0,T[}\right) : T > 0,\ (\{\mu_t\}_{t\in[0,T[}, \{\nu_t\}_{t\in[0,T[}) \text{ is a } \mu_0\text{-MATOC} \right\}.$$

By Lemma 4.3.5, we have $\mathcal{X} \neq \emptyset$. We endow $\mathcal{X}$ with the partial order relation defined by

$$(\boldsymbol{\mu}^1, \boldsymbol{\nu}^1) \preceq (\boldsymbol{\mu}^2, \boldsymbol{\nu}^2) \text{ iff } \tau_1 \le \tau_2, \text{ and } \mu^1_t = \mu^2_t,\ \nu^1_t = \nu^2_t \text{ for all } t \in [0, \tau_1[,$$

where $\boldsymbol{\mu}^i = \{\mu^i_t\}_{t\in[0,\tau_i[}$, $\boldsymbol{\nu}^i = \{\nu^i_t\}_{t\in[0,\tau_i[}$, $i = 1, 2$. Consider a total ordered chain

$$\mathcal{C} = \{(\boldsymbol{\mu}^\alpha = \{\mu^\alpha_t\}_{t\in[0,\tau_\alpha[}, \boldsymbol{\nu}^\alpha = \{\nu^\alpha_t\}_{t\in[0,\tau_\alpha[})\}_{\alpha\in A} \subseteq \mathcal{X}.$$

We define $\left(\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\sup\tau_\alpha[}, \boldsymbol{\nu} = \{\nu_t\}_{t\in[0,\sup\tau_\alpha[}\right)$ by setting $\mu_t = \mu^\alpha_t$ and $\nu_t = \nu^\alpha_t$ for all $\alpha \in A$ such that $t \in [0, \tau_\alpha[$. The definition is well-posed since all the elements of $\mathcal{C}$ agree on the set where they are defined, moreover given $0 \le t < \sup\tau_\alpha$ there exists $t \le \tau_\alpha < \sup\tau_\alpha$, and so we can define $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ on the whole of $[0, \sup\tau_\alpha[$.

Finally, we prove that $\boldsymbol{\mu}$ is an admissible trajectory driven by $\boldsymbol{\nu}$. Given any $\varphi \in C^1_C([0, \sup\tau_\alpha[\times\mathbb{R}^d)$ we have that $\mathrm{supp}\,\varphi \subseteq [0, \tau_{\bar\alpha}[\times\mathbb{R}^d$ for a certain $\bar\alpha \in A$, and, since $\boldsymbol{\mu}$ agrees with an admissible trajectory on $[0, \tau_{\bar\alpha}[$, we have that

$$\begin{aligned}\iint_{[0,\sup\tau_\alpha[\times\mathbb{R}^d} \partial_t\varphi(t,x)\,d\mu_t\,dt &= \iint_{[0,\tau_{\bar\alpha}[\times\mathbb{R}^d} \partial_t\varphi(t,x)\,d\mu^\alpha_t\,dt \\ &= -\iint_{[0,\tau_{\bar\alpha}[\times\mathbb{R}^d} \nabla\varphi(t,x)\,d\nu^\alpha_t\,dt = -\iint_{[0,\sup\tau_\alpha[\times\mathbb{R}^d} \nabla\varphi(t,x)\,d\nu_t\,dt,\end{aligned}$$

and so $\boldsymbol{\mu}$ is an admissible trajectory driven by $\boldsymbol{\nu}$. In particular, we have $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathcal{X}$ and $(\boldsymbol{\mu}^\alpha, \boldsymbol{\nu}^\alpha) \preceq (\boldsymbol{\mu}, \boldsymbol{\nu})$ for all $\alpha \in A$. By Zorn's Lemma there exist maximal elements in $\mathcal{X}$.

Let $(\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\tau[}, \boldsymbol{\nu} = \{\nu_t\}_{t\in[0,\tau[})$ be one of these maximal elements. We want to prove that $\tau = +\infty$. By contradiction, assume that $\tau < +\infty$. By Lemma 3.2.7, there exist $D', D'' > 0$ such that for all $t \in [0, \tau]$ we have

$$\begin{aligned}\mathrm{m}_p(\mu_t) &\le D'(\mathrm{m}_p(\mu_0) + 1),\\ \mathrm{m}_{p-1}(|\nu_t|) &\le D''(\mathrm{m}_p(\mu_0) + 1).\end{aligned}$$

Thus, according to Remark 5.1.5 in [9], there exist $\mu_\tau \in \mathscr{P}(\mathbb{R}^d)$ and $\nu_\tau \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ such that $\mu_t \rightharpoonup^* \mu_\tau$ and $\nu_t \rightharpoonup^* \nu_\tau$ as $t \to \tau^-$. Consider now $\varepsilon > 0$, a continuous selection v of F and the solution $\{\mu'_t\}_{t\in[0,\varepsilon]}$ of

$$\begin{cases}\partial_t\mu_t + \operatorname{div} v\mu_t = 0,\\ \mu_{|t=0} = \mu_\tau\end{cases}$$

By setting

$$\mu_t^\circ := \begin{cases}\mu_t, & \text{for } 0 \le t < \tau,\\ \\ \mu'_{t-\tau}, & \text{for } \tau \le t \le \tau+\varepsilon,\end{cases}$$

$$\nu_t^\circ := \begin{cases}\nu_t, & \text{for } 0 \le t < \tau,\\ \\ v\mu'_{t-\tau}, & \text{for } \tau \le t \le \tau+\varepsilon,\end{cases}$$

we obtain an admissibile trajectory $\boldsymbol{\mu}^\circ = \{\mu_t^\circ\}_t$ driven by $\boldsymbol{\nu}^\circ = \{\nu_t^\circ\}_t$ which is defined on $[0,\tau+\varepsilon[$ and agrees with $\boldsymbol{\mu}$ on $[0,\tau[$, thus contradicting the maximality of $(\boldsymbol{\mu},\boldsymbol{\nu})$. Thus $\tau = +\infty$.

Let $\{T_n\}_{n\in\mathbb{N}}$ be a sequence with $T_n \to +\infty$ and $(\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,+\infty[}, \boldsymbol{\nu} = \{\nu_t\}_{t\in[0,+\infty[})$ be a maximal element in $\mathcal{X}$. Then $\{(\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T_n[}, \boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T_n[}) : n \in \mathbb{N}\}$ is a totally ordered chain in $\mathcal{X}$ whose upper bound is $(\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,+\infty[}, \boldsymbol{\nu} = \{\nu_t\}_{t\in[0,+\infty[})$. Then, by Definition 4.3.6, we have a sequence of probability measures $\{\boldsymbol{\eta}_n\}_{n\in\mathbb{N}} \subseteq \mathscr{P}(\mathbb{R}^d\times\Gamma_{T_n})$ such that $\{\mu_t\}_{t\in[0,T_n]}$ is represented by $\boldsymbol{\eta}_n$. We notice that by construction if $n_1 \le n_2$ then for all $t \in [0,T_{n_1}]$ we have

$$\iint_{\mathbb{R}^d\times\Gamma_{T_{n_1}}} \varphi(\gamma(t))\chi_{S^c}(\gamma(t))(T(x)-t)\,d\boldsymbol{\eta}_{n_1} = \iint_{\mathbb{R}^d\times\Gamma_{T_{n_2}}} \varphi(\gamma(t))\chi_{S^c}(\gamma(t))(T(x)-t)\,d\boldsymbol{\eta}_{n_2},$$

thus we can define $\tilde{\boldsymbol{\mu}} = \{\tilde\mu_t\}_{t\in[0,+\infty[}$ by setting for all $n \in \mathbb{N}$ and for all $t \in [0,T_n[$

$$\int_{\mathbb{R}^d} \varphi(x)\tilde\mu_t(x) = \iint_{\mathbb{R}^d\times\Gamma_{T_n}} \varphi(\gamma(t))\chi_{S^c}(\gamma(t))(T(x)-t)\,d\boldsymbol{\eta}_n(x,\gamma).$$

Since $\boldsymbol{\eta}_n$ is concentrated on (restriction to $[0,T_n]$ of) optimal trajectories and S is strongly invariant, we have that $t \ge T(x)$ if and only if $\gamma(t) \in S$, and so $\tilde\mu_t \in \mathscr{M}^+(\mathbb{R}^d)$ for all $t \ge 0$. Thus $T(\cdot) = \dfrac{\tilde\mu_0}{\mu_0}(\cdot)$ is an admissible clock for μ_0. Moreover, since for μ_0-a.e. $x \in \mathbb{R}^d$ and for every admissible clock $f_0(\cdot)$ for μ_0 we must have $f_0(x) \ge T(x)$ by Lemma 4.1.4, we conclude that $T(\cdot)$ is the optimal clock for μ_0. □

Now we can deduce the following Dynamic Programming Principle.

Corollary 4.3.8 (DPP for the clock problem). *Assume hypothesis (F_0) and (F_1). Let $S \subseteq \mathbb{R}^d$ be a target set for F. Let $p > 1$ and $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$, with* $\operatorname{supp}\mu_0 \subseteq \mathbb{R}^d\setminus S$, *be such that $\|T(\cdot)\|_{L^1_{\mu_0}} < +\infty$. We have*

$$\tau_p(\mu_0) = \int_{\mathbb{R}^d} T(x)\,d\mu_0(x).$$

Let $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[}$ be an admissible clock-trajectory for μ_0 following a family of admissible mass-preserving trajectories $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$ starting from μ_0. For any $s \geq 0$ we choose $n > 0$ such that $\boldsymbol{\mu}^n$ is defined on an interval $[0, T_n]$ containing s and it is represented by $\boldsymbol{\eta}_n \in \mathscr{P}(\mathbb{R}^d \times \Gamma_{T_n})$. Then we have

$$\tau_p(\mu_0) = \iint_{\mathbb{R}^d\times\Gamma_{T_n}} T(\gamma(0))\, d\boldsymbol{\eta}_n \leq \iint_{\mathbb{R}^d\times\Gamma_{T_n}} [T(\gamma(s)) + s]\, d\boldsymbol{\eta}_n \leq s + \tau_p(\mu^n_s).$$

Moreover, if $\boldsymbol{\eta}_n$ is concentrated on (restriction to $[0, T_n]$ of) time-optimal trajectories, then for all $s \geq 0$ such that $\mathrm{supp}\mu^n_s \subseteq \mathbb{R}^d \setminus S$, we have

$$\tau_p(\mu_0) = s + \tau_p(\mu^n_s),$$

and so for such $s \geq 0$ we have

$$\tau_p(\mu_0) = \inf_{\boldsymbol{\mu}} \{s + \tau_p(\mu_s)\},$$

where the infimum is taken on admissible mass-preserving trajectories $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,s]}$ satisfying $\mu_{t=0} = \mu_0$.

The proof is a direct consequence of Corollary 4.3.7, Theorem 1.4.8 and Lemma 4.1.4.

Remark 4.3.9. We notice that, in the same hypothesis of Corollary 4.3.7, if $\mu \in \mathscr{P}_p(\mathbb{R}^d)$ we have that $\tau_p(\mu) = \|T(\cdot)\|_{L^1_\mu} \leq \|T(\cdot)\|_{L^\infty_\mu} = \tilde{T}_p(\mu)$, where $\tilde{T}_p(\cdot)$ is the generalized minimum time function studied in the previous chapter for the mass-preserving case (see Definition 3.2.10) with generalized target set $\tilde{S} := \{\sigma \in \mathscr{P}(\mathbb{R}^d) : \mathrm{supp}\,\sigma \subseteq S\}$ (i.e. we are requiring the existence of a classical counterpart for the target set which coincides with S). In particular, we refer to Corollary 3.2.22 in the previous chapter for the last equivalence.

4.3.3 Regularity results

Thanks to Corollary 4.3.7, under suitable assumptions, the clock-generalized minimum time function inherits regularity results from the classical one as shown in the next corollaries. For the following result, we refer to [51] for conditions under which the classical minimum time function $T(\cdot)$ is l.s.c..

Corollary 4.3.10 (L.s.c. of the clock-generalized minimum time function)**.** *Assume that $T(\cdot)$ is l.s.c.. Assume hypothesis (F_0) and (F_1). Let $S \subseteq \mathbb{R}^d$ be a target set for F. Let $p > 1$ and $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$, with $\mathrm{supp}\mu_0 \subseteq \mathbb{R}^d \setminus S$, be such that $\|T(\cdot)\|_{L^1_{\mu_0}} < +\infty$. Then $\tau_p : \mathscr{P}_p(\mathbb{R}^d) \to [0, +\infty]$ is l.s.c..*

Proof. We have to prove that $\tau_p(\mu_0) \leq \liminf\limits_{W_p(\mu,\mu_0)\to 0} \tau_p(\mu)$. Taken a sequence $\{\mu^n_0\}_{n\in\mathbb{N}} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ s.t. $W_p(\mu^n_0, \mu_0) \to 0$ for $n \to +\infty$, and $\liminf\limits_{W_p(\mu,\mu_0)\to 0} \tau_p(\mu) = \liminf\limits_{n\to+\infty} \tau_p(\mu^n_0)$, we want to prove that $\tau_p(\mu_0) \leq \liminf\limits_{n\to+\infty} \tau_p(\mu^n_0)$.

By Lemma 4.1.4, Lemma 5.1.7. in [9] and Corollary 4.3.7, we conclude immediately that

$$\liminf_{n\to+\infty} \tau_p(\mu^n_0) \geq \liminf_{n\to+\infty} \int_{\mathbb{R}^d} T(x)\, d\mu^n_0(x) \geq \int_{\mathbb{R}^d} T(x)\, d\mu_0(x) = \tau_p(\mu_0).$$

□

We are now intersted in proving *sufficient* conditions on the set-valued function $F(\cdot)$ in order to have *controllability* of the generalized control system, i.e. to steer a probability measure on the generalized target by an admissible trajectory in finite time.

Representation formula for the generalized minimum time provided in Corollary 4.3.7 allows us to recover many results valid for the classical minimum time function also in the framework of the generalized systems. We refer the reader to Chapter 2 in [21] and Sections 2 and 3 in [21] for a definition and classical results about semiconcave functions, in particular regarding the classical minimum time function.

Corollary 4.3.11 (Controllability)**.** *Assume (F_0), (F_1), (F_3). Let $S \subseteq \mathbb{R}^d$ be a target set for F. Assume furthermore that for every $R > 0$ there exist $\eta_R, \sigma_R > 0$ such that for a.e. $x \in B(0,R) \setminus S$ with $d_S(x) \leq \sigma_R$ there holds*

$$\sigma_{F(x)}(-\nabla d_S(x)) > \eta_R\,, \tag{4.3}$$

where $\sigma_{F(x)}$ is the support function of the set $F(x)$ as in (1.1).
Then, if we set for $p > 1$

$$\mathscr{P}_p(\mathbb{R}^d)_{|R} := \{\mu \in \mathscr{P}_p(\mathbb{R}^d) : \|T(\cdot)\|_{L^1_\mu} < +\infty \text{ and } \operatorname{supp}\mu \subseteq \overline{B(0,\sigma_R)} \setminus S\},$$

there exists $c_R > 0$ such that for every $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)_{|R}$ the following properties hold.

1. *$\tau_p(\mu_0) \leq \dfrac{1}{c_R}\|d_S\|_{L^1_{\mu_0}}$.*
2. *The function $\tau_p : \mathscr{P}_p(\mathbb{R}^d) \to [0,+\infty]$ is Lipschitz continuous on $\mathscr{P}_p(\mathbb{R}^d)_{|R}$ with respect to the metric W_p^p.*
3. *If $\partial S \in C^{1,1}$, then the function $\tau_p : \mathscr{P}_p(\mathbb{R}^d) \to [0,+\infty]$ is semiconcave on*
$$\{\mu \in \mathscr{P}_p(\mathbb{R}^d)_{|R} : \operatorname{supp}\mu \cap S = \emptyset\}$$
with respect to the metric W_2.

Proof. According to Proposition 2.2 in [21], the present assumptions imply that there exists a constant $c_R > 0$ such that the classical minimum time function satisfies

$$T(x) \leq \frac{1}{c_R} d_S(x)\,, \tag{4.4}$$

for every $x \in B(0,R) \setminus S$ with $d_S(x) \leq \sigma_R$. Moreover, $T(\cdot)$ is Lipschitz continuous in such set. We denote by $k_R > 0$ its Lipschitz constant.
Now, property (1) follows from (4.4) and Corollary 4.3.7, since

$$\tau_p(\mu_0) = \int_{\mathbb{R}^d} T(x)\,d\mu_0 \leq \frac{1}{c_R}\int_{\mathbb{R}^d} d_S(x)\,d\mu_0 = \frac{1}{c_R}\|d_S\|_{L^1_{\mu_0}}.$$

To prove (2), fix $\mu_1, \mu_2 \in \mathscr{P}_p(\mathbb{R}^d)_{|R}$. By setting

$$c'_R := \frac{c_R}{(1+c_R)(1+k_R)}\,,$$

we have that the function $c'_R T(\cdot)$ is Lipschitz continuous with constant less than 1 and that $c'_R T(\cdot) \leq R$. Hence, it can be extended to a continuous bounded function on the whole $\mathbb{R}^d$, and $|c'_R T(x) - c'_R T(y)| \leq |x-y|^p$ for all $x, y \in \mathbb{R}^d$. According to Kantorovich duality (1.3) and Corollary 4.3.7 we then have

$$W_p^p(\mu_1, \mu_2) \geq \int_{\mathbb{R}^d} c'_R T(x)\, d\mu_1(x) - \int_{\mathbb{R}^d} c'_R T(y)\, d\mu_2(y) = c'_R(\tau_p(\mu_1) - \tau_p(\mu_2))\,.$$

By switching the roles of μ_1 and μ_2, we obtain (2).

Finally, according to Theorem 3.1 in [21], when $\partial S \in C^{1,1}$ we have that the classical minimum time function is semiconcave in $\{x : T(x) < +\infty\} \setminus S$. In particular, there exists $D_R > 0$ such that

$$T(tx_1 + (1-t)x_2) \geq tT(x_1) + (1-t)T(x_2) - D_R\, t(1-t)\, |x_1 - x_2|^2, \tag{4.5}$$

for every $x_1, x_2 \in \{x : T(x) < +\infty\} \setminus S$.

Let $K := \overline{B(0, \sigma_R)}$. For any Borel sets $A, B \subseteq \mathbb{R}^d$ and $\pi \in \Pi(\mu_1, \mu_2)$, we now have

$$A \times B \subseteq [(A \times B) \cap (K \times K)] \cup [(A \setminus K) \times \mathbb{R}^d] \cup [\mathbb{R}^d \times (B \setminus K)],$$

so that

$$\begin{aligned} \pi(A \times B) &\leq \pi((A \times B) \cap (K \times K)) + \mu_0(A \setminus K) + \mu_1(B \setminus K) \\ &= \pi((A \times B) \cap (K \times K)), \end{aligned}$$

because μ_1 and μ_2 are concentrated on K. In particular, $\operatorname{supp}(\pi) \subseteq K \times K$. Therefore, we choose an optimal transport plan $\pi \in \Pi(\mu_1, \mu_2)$ realizing the p-Wasserstein distance between μ_1 and μ_2, so that $\mu_t := t\mu_1 + (1-t)\mu_2 = \big(t\mathrm{pr}^1 + (1-t)\mathrm{pr}^2\big)\sharp\pi$, where $\mathrm{pr}^i \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, $i = 1, 2$, is the projection on the i–th component, i.e., $\mathrm{pr}^i(x_1, x_2) = x_i$. We integrate the estimate (4.5) to find that, by using Lemma 4.1.4 and Corollary 4.3.7,

$$\begin{aligned} \tau_p(\mu_t) &\geq \int_{\mathbb{R}^d} T(x)\, d\mu_t(x) \\ &= \iint_{\mathbb{R}^d \times \mathbb{R}^d} T(tx_1 + (1-t)x_2)\, d\pi(x_1, x_2) \\ &\geq t \int_{\mathbb{R}^d} T(x_1)\, d\mu_1 + (1-t) \int_{\mathbb{R}^d} T(x_2)\, d\mu_2 \\ &\qquad - D_R\, t\,(1-t) \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x_1 - x_2|^2\, d\pi(x_1, x_2) \\ &= t\tau_p(\mu_1) + (1-t)\tau_p(\mu_2) - D_R\, t(1-t)\, W_2^2(\mu_1, \mu_2). \end{aligned}$$

□

Remark 4.3.12. For other controllability conditions generalizing (4.3), the reader may refer e.g. to [37, 58].

4.4 Hamilton-Jacobi-Bellman equation

In this section we will prove that under the assumptions granting the validity of the Dynamic Programming Principle and of a result which aims to recover the initial velocity of admissible trajectories, the clock-generalized minimum time function solves a natural Hamilton-Jacobi-Bellman equation on $\mathscr{P}_2(\mathbb{R}^d)$ in a suitable viscosity sense (Theorem 4.4.3).

We observe also that once we have the Dynamic Programming Principle and once the problem is modeled on the same notion of admissible mass-preserving trajectories, then the Hamilton-Jacobi-Bellman equation related to the present problem is the same considered in Section 3.3 for the mass-preserving case. We then follow a very similar approach as the one discussed in Section 3.3.

First, let us point out that in the following we will use Lemma 3.2.7 about properties of the evaluation operator already seen in the previous Chapter.

The following proposition allows to construct an admissible mass-preserving trajectory concentrated on characteristics of class C^1 with initial velocity the given one.

Proposition 4.4.1. *Assume hypothesis* (F_0), (F_1). *Let* $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, *and* $x \mapsto v_x$ *be a Borel selection of* F *belonging to* L^2_μ. *Then for any* $T > 0$ *there exists an admissible mass-preserving curve* $\boldsymbol{\mu}$ *defined on* $[0,T]$ *starting from* μ *and represented by* $\boldsymbol{\eta}$ *such that for* $\boldsymbol{\eta}$*-a.e.* $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$ *we have that* $\gamma \in C^1([0,T])$, $\dot\gamma(t) \in F(\gamma(t))$ *for all* $t \in [0,T]$, $\gamma(0) = x$ *and* $\dot\gamma(0) = v_x$.

Proof. Let $T > 0$ be fixed. Consider the set-valued map $G : \mathbb{R}^d \rightrightarrows C^0(\mathbb{R}^d;\mathbb{R}^d)$ defined by

$$G(x) := \{v \in C^0(\mathbb{R}^d;\mathbb{R}^d) : v(x) = v_x,\ v(y) \in F(y) \text{ for all } y \in \mathbb{R}^d\},$$

and notice that, recalling the assumptions on F, we have that $G(x)$ is nonempty, convex and closed. Indeed, for every $x \in \mathbb{R}^d$ and $v_x \in F(x)$ there exists by Michael's continuous selection Theorem a continuous selection v of F such that $v(x) = v_x$.

Define the map $g : \mathbb{R}^d \times C^0(\mathbb{R}^d;\mathbb{R}^d) \to [0,+\infty]$ by setting

$$g(x,v) := \sup_{q,y\in\mathbb{R}^d} \left\{ I_{F(y)}(v(y)) + \langle q, v_x - v(x)\rangle \right\},$$

noticing that $v \in G(x)$ if and only if $g(x,v) = 0$.

To prove that g is a Borel map, it is enough to show that $(v,y) \mapsto I_{F(y)}(v(y))$ is a Borel map from $C^0(\mathbb{R}^d;\mathbb{R}^d) \times \mathbb{R}^d$ to $\{0,+\infty\}$.

Indeed, consider any sequence $\{v_n\}_{n\in\mathbb{N}} \subseteq C^0(\mathbb{R}^d;\mathbb{R}^d)$ uniformly convergent to $v \in C^0(\mathbb{R}^d;\mathbb{R}^d)$ on compact sets, and $\{y_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^d$ converging to y.

Then, $v_n(y_n) \to v(y)$, $n \to +\infty$. Indeed, denoted with $\omega_y(\cdot)$ a modulus of continuity for v at the point y, we have

$$\begin{aligned} |v_n(y_n) - v(y)| &\le |v_n(y_n) - v(y_n)| + |v(y_n) - v(y)| \\ &\le \|v_n - v\|_{L^\infty(B(y,s))} + \omega_y(|y_n - y|), \end{aligned}$$

for a suitable $s > 0$. Hence, we deduce that

$$\liminf_{n\to+\infty} I_{F(y_n)}(v_n(y_n)) \ge I_{F(y)}(v(y)),$$

where we used the fact that the map $f:\mathbb{R}^d\times\mathbb{R}^d\to\{0,+\infty\}$, $f(z,w):=I_{F(z)}(w)$, is l.s.c. due to u.s.c. of F.

Thus we have just proved that $(v,y)\mapsto I_{F(y)}(v(y))$ is l.s.c. and hence a Borel map. Hence $\mathrm{Graph}\,G=g^{-1}(0)$ is a Borel set. By Theorem 8.1.4 p. 310 in [13], we have that the set-valued map $G:\mathbb{R}^d\rightrightarrows C^0(\mathbb{R}^d;\mathbb{R}^d)$ is Borel measurable, and so by Theorem 8.1.3 p. 308 in [13] it admits a Borel selection $V:\mathbb{R}^d\to C^0(\mathbb{R}^d;\mathbb{R}^d)$. We denote $V(x)\in C^0(\mathbb{R}^d;\mathbb{R}^d)$ by V_x.

We fix a family of smooth mollifiers $\{\rho_\varepsilon\}_{\varepsilon>0}\subseteq C^\infty_C(\mathbb{R}^d)$ such that $\mathrm{supp}\,\rho_\varepsilon\subseteq\overline{B(0,\varepsilon)}$, and denote by $H^T_{x,\varepsilon}$ the (unique) $\gamma\in\Gamma_T$ satisfying $\dot\gamma(t)=(V_x*\rho_\varepsilon)\circ\gamma(t)$, $\gamma(0)=x$. We want to prove that $H^T_{x,\varepsilon}$ is a Borel map in x.

For any $x\in\mathbb{R}^d$ and $W\in\mathrm{Lip}(\mathbb{R}^d;\mathbb{R}^d)$ denote by $h_{x,W}(t)$ the solution of $\dot x(t)=W\circ x(t)$, $x(0)=x$. The map $h:\mathbb{R}^d\times\mathrm{Lip}(\mathbb{R}^d;\mathbb{R}^d)\to\Gamma_T$ is continuous, hence Borel, since for all $x,y\in\mathbb{R}^d$, $W_1,W_2\in\mathrm{Lip}(\mathbb{R}^d;\mathbb{R}^d)$, we have

$$\begin{aligned}|h_{x,W_1}(t)-h_{y,W_2}(t)|&\le|x-y|+\int_0^t|W_1(h_{x,W_1}(s))-W_2(h_{y,W_2}(s))|\,ds\\&\le|x-y|+\int_0^t|W_1(h_{x,W_1}(s))-W_1(h_{y,W_2}(s))|\,ds+\\&+\int_0^t|W_1(h_{y,W_2}(s))-W_2(h_{y,W_2}(s))|\,ds\\&\le|x-y|+\mathrm{Lip}(W_1)\int_0^t|h_{x,W_1}(s))-h_{y,W_2}(s)|\,ds+t\|W_1-W_2\|_\infty\end{aligned}$$

and so by Gronwall's inequality

$$|h_{x,W_1}(t)-h_{y,W_2}(t)|\le(|x-y|+t\|W_1-W_2\|_\infty)e^{t\mathrm{Lip}(W_1)},$$

which implies

$$\|h_{x,W_1}-h_{y,W_2}\|_\infty\le(|x-y|+T\|W_1-W_2\|_\infty)e^{T\mathrm{Lip}(W_1)}.$$

Since $H^T_{x,\varepsilon}$ can be written as the composition of the Borel maps $x\mapsto(x,V_x)$, $(x,Z)\mapsto(x,Z*\rho_\varepsilon)$, and $(x,W)\mapsto h_{x,W}$, we have that it is a Borel map.

Finally, we define the Kuratowski upper limit of $H^T_{x,\varepsilon}$ by

$$H^T(x):=\{\gamma\in\Gamma_T:\ \text{there exists }\{\varepsilon_n\}_{n\in\mathbb{N}}\text{ s.t. }\varepsilon_n\to0^+,H^T_{x,\varepsilon_n}\to\gamma,\ \text{as }n\to+\infty\}.$$

Thanks to Theorem 8.2.5 in [13], this is a Borel set-valued map from $\mathbb{R}^d$ to Γ_T, thus possesses a Borel selection $\psi:\mathbb{R}^d\to\Gamma_T$.

Given $x\in\mathbb{R}^d$, let $\{\varepsilon_n\}_{n\in\mathbb{N}}$ be such that $\varepsilon_n\to0^+$ and $H^T_{x,\varepsilon_n}\to\gamma_x:=\psi(x)$. In particular, we have that $H^T_{x,\varepsilon_n}(0)=x$ for all $n\in\mathbb{N}$, and so $\gamma_x(0)=x$. Since there exists a compact K containing $H^T_{x,\varepsilon_n}(\tau)$ for all $n\in\mathbb{N}$ sufficiently large and all $\tau\in[0,T]$, and moreover $V_x*\rho_{\varepsilon_n}$ converges to V_x in $C^0(\mathbb{R}^d)$ on all the compact sets of $\mathbb{R}^d$, we can pass to the limit by Dominated Convergence Theorem in

$$\frac{H^T_{x,\varepsilon_n}(s)-H^T_{x,\varepsilon_n}(t)}{s-t}=\frac{1}{s-t}\int_t^s V_x*\rho_{\varepsilon_n}(H^T_{x,\varepsilon_n}(\tau))\,d\tau,$$

obtaining

$$\frac{\gamma_x(s)-\gamma_x(t)}{s-t}=\frac{1}{s-t}\int_t^s V_x(\gamma_x(\tau))\,d\tau, \tag{4.6}$$

thus $\gamma_x \in C^1$ is an admissible curve satisfying $\dot\gamma_x(0)=v_x$.

We define the probability measure

$$\boldsymbol{\eta} := \mu\otimes\delta_{\gamma_x}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T),$$

which, as already seen in the last part of the proof of Lemma 4.3.5, induces an admissible trajectory $\boldsymbol{\mu}=\{\mu_t=e_t\sharp\boldsymbol{\eta}\}_{t\in[0,T]}$. Moreover, we prove that

$$\lim_{t\to 0}\left\|\frac{e_t-e_0}{t}-v_x\right\|_{L^2_{\boldsymbol{\eta}}}=0.$$

Indeed,

$$\begin{aligned}\left\|\frac{e_t-e_0}{t}-v_x\right\|^2_{L^2_{\boldsymbol{\eta}}} &= \int_{\mathbb{R}^d}\int_{\Gamma_T^x}\left|\frac{\gamma(t)-\gamma(0)}{t}-v_x\right|^2 d\delta_{\gamma_x}(\gamma)\,d\mu(x)\\ &= \int_{\mathbb{R}^d}\left|\frac{\gamma_x(t)-\gamma_x(0)}{t}-v_x\right|^2 d\mu(x),\end{aligned}$$

and for μ-a.e. $x\in\mathbb{R}^d$, recalling (4.6), continuity of $V_x(\cdot)$ and that $\gamma\in C^1$ and $\dot\gamma(0)=v_x$, we have

$$\begin{aligned}\left|\frac{\gamma_x(t)-\gamma_x(0)}{t}-v_x\right| &= \left|\frac{1}{t}\int_0^t V_x(\gamma_x(\tau))\,d\tau - v_x\right|\\ &\le \frac{1}{t}\int_0^t |V_x(\gamma_x(\tau))|\,d\tau+|v_x|\\ &\le \max_{t\in[0,T]}|V_x(\gamma_x(t))|+|v_x|,\\ \lim_{t\to 0^+}\left|\frac{\gamma_x(t)-\gamma_x(0)}{t}-v_x\right| &= 0.\end{aligned}$$

Thus we conclude applying Lebesgue's Dominated Convergence Theorem. □

Corollary 4.4.2. *Assume hypothesis (F_0), (F_1). Let $\mu\in\mathscr{P}_2(\mathbb{R}^d)$, $T>0$. Define the set $A_T(\mu)$ of the maps $w_{\boldsymbol{\eta}}\in L^2_{\boldsymbol{\eta}}$ satisfying the following*

1. *there exists an admissible mass-preserving trajectory $\boldsymbol{\mu}$ defined on $[0,T]$ and represented by $\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$ with $e_0\sharp\boldsymbol{\eta}=\mu$,*

2. *there exists a sequence $\{t_i\}_{i\in\mathbb{N}}\subseteq]0,T]$ such that $t_i\to 0$ and*

$$\begin{aligned}\lim_{i\to+\infty}\frac{1}{t_i}\iint_{\mathbb{R}^d\times\Gamma_T}\langle p\circ e_0(x,\gamma),e_{t_i}(x,\gamma)-e_0(x,\gamma)\rangle\,d\boldsymbol{\eta} &=\\ =\iint_{\mathbb{R}^d\times\Gamma_T}\langle p\circ e_0(x,\gamma),w_{\boldsymbol{\eta}}(x,\gamma)\rangle\,d\boldsymbol{\eta},\end{aligned}$$

for all $p\in L^2_\mu(\mathbb{R}^d;\mathbb{R}^d)$.

Then $A_T(\mu)=\{v\circ e_0 : v\in L^2_\mu,\ v(x)\in F(x) \text{ for } \mu\text{-a.e. } x\in\mathbb{R}^d\}$.

Proof. It is trivial that $A_T(\mu)$ is contained in the right hand side. The opposite inclusion follows from the previous Proposition with $v(x)=v_x$, noticing also that since $v\in L^2_\mu$, then $v\circ e_0\in L^2_{\boldsymbol{\eta}}$ with $\boldsymbol{\eta}$ as in 1 by Lemma 3.2.7.

Indeed, in Proposition 4.4.1 we proved strong convergence in $L^2_{\boldsymbol{\eta}}$ of $\dfrac{e_t-e_0}{t}$ to v_x for $t\to 0$. Hence we have weak convergence, in particular since $p\circ e_0\in L^2_{\boldsymbol{\eta}}$ for every $p\in L^2_\mu$ by Lemma 3.2.7, then there exists a sequence $\{t_i\}_{i\in\mathbb{N}}\subseteq]0,T]$ such that $t_i\to 0$ and

$$\lim_{i\to+\infty}\frac{1}{t_i}\iint_{\mathbb{R}^d\times\Gamma_T}\langle p\circ e_0(x,\gamma), e_{t_i}(x,\gamma)-e_0(x,\gamma)\rangle\,d\boldsymbol{\eta}=\iint_{\mathbb{R}^d\times\Gamma_T}\langle p\circ e_0(x,\gamma), v\circ e_0(x,\gamma)\rangle\,d\boldsymbol{\eta},$$

thus item 2 is satisfied with $w_{\boldsymbol{\eta}}=v\circ e_0$, and item 1 follows directly by the previous Proposition. □

We are now ready to prove the following theorem in which we adopt the same notion of sub-/super-differential defined in Definition 3.3.6 for the mass-preserving problem, and the corresponding notion of viscosity solutions as well as the same hamiltonian function of Definition 3.3.8.

The procedure used for the proof of the following result is like the one adopted in Theorem 3.3.9 for the generalized minimum time function of the mass-preserving case.

Theorem 4.4.3 (Viscosity solution)**.** *Let $S\subseteq\mathbb{R}^d$ be a target set for F. Let $\mathcal{A}$ be any open subset of $\mathscr{P}_2(\mathbb{R}^d)$ with uniformly bounded $2-$moments and such that if $\mu\in\mathcal{A}$ then $\mathrm{supp}\mu\subseteq\mathbb{R}^d\setminus S$. Assume hypothesis (F_0), (F_1). Assume that $\|T(\cdot)\|_{L^1_\mu}<+\infty$ for all $\mu\in\mathcal{A}$ and that $\tau:\mathscr{P}_2(\mathbb{R}^d)\to[0,+\infty]$ is continuous on $\mathcal{A}$. Then $\tau(\cdot)$ is a viscosity solution of $\mathscr{H}_F(\mu,D\tau(\mu))=0$ on $\mathcal{A}$, with $\mathscr{H}_F$ defined as in Definition 3.3.8.*

Proof. The proof is splitted in two claims.

Claim 1: $\tau(\cdot)$ is a subsolution of $\mathscr{H}_F(\mu,D\tau(\mu))=0$ on $\mathcal{A}$.

Proof of Claim 1. Let $\mu_0\in\mathcal{A}$. Let $\tilde{\boldsymbol{\mu}}=\{\tilde\mu_t\}_{t\in[0,+\infty[}$ be an admissible clock-trajectory for μ_0 following a family of admissible mass-preserving trajectories $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$ starting from μ_0. For any $s\geq 0$ we choose $n>0$ such that $\boldsymbol{\mu}^n$ is defined on an interval $[0,T_n]$ containing s and it is represented by $\boldsymbol{\eta}_n\in\mathscr{P}(\mathbb{R}^d\times\Gamma_{T_n})$. Then by the Dynamic Programming Principle we have $\tau(\mu_0)\leq\tau(\mu^n_s)+s$ for all $s>0$. Without loss of generality, we can assume $0<s<1$. Given any $p_{\mu_0}\in D^+_\delta\tau(\mu_0)$, and set

$$A(s,p_{\mu_0},\boldsymbol{\eta}_n):=-\,s-\iint_{\mathbb{R}^d\times\Gamma_{T_n}}\langle p_{\mu_0}\circ e_0(x,\gamma), e_s(x,\gamma)-e_0(x,\gamma)\rangle\,d\boldsymbol{\eta}_n,$$

$$B(s,p_{\mu_0},\boldsymbol{\eta}_n):=\tau(\mu^n_s)-\tau(\mu_0)-\iint_{\mathbb{R}^d\times\Gamma_{T_n}}\langle p_{\mu_0}\circ e_0(x,\gamma), e_s(x,\gamma)-e_0(x,\gamma)\rangle d\boldsymbol{\eta}_n,$$

we have $A(s,p_{\mu_0},\boldsymbol{\eta}_n)\leq B(s,p_{\mu_0},\boldsymbol{\eta}_n)$.

We recall that since by definition $p_{\mu_0}\in L^2_{\mu_0}$, we have that $p_{\mu_0}\circ e_0\in L^2_{\boldsymbol{\eta}_n}$ by Lemma 3.2.7. Dividing by $s>0$, we obtain that

$$\limsup_{s\to 0^+}\frac{A(s,p_{\mu_0},\boldsymbol{\eta}_n)}{s}\geq -1-\iint_{\mathbb{R}^d\times\Gamma_{T_n}}\langle p_{\mu_0}\circ e_0(x,\gamma), w_{\boldsymbol{\eta}_n}(x,\gamma)\rangle\,d\boldsymbol{\eta}_n(x,\gamma),$$

for all $w_{\boldsymbol{\eta}_n} \in A_{T_n}(\mu_0)$, with $A_{T_n}(\mu_0)$ defined as in Corollary 4.4.2.

Recalling the choice of p_{μ_0}, we have

$$\limsup_{s\to 0^+} \frac{B(s,p_{\mu_0},\boldsymbol{\eta}_n)}{s} = \limsup_{s\to 0^+} \frac{B(s,p_{\mu_0},\boldsymbol{\eta}_n)}{\|e_s - e_0\|_{L^2_{\boldsymbol{\eta}_n}}} \cdot \left\|\frac{e_s-e_0}{s}\right\|_{L^2_{\boldsymbol{\eta}_n}} \leq K\delta,$$

where $K > 0$ is a suitable constant coming from Lemma 3.2.7 and from hypothesis.

We thus obtain for all $\boldsymbol{\eta}_n$ as above and all $w_{\boldsymbol{\eta}_n} \in A_{T_n}(\mu_0)$, that

$$1 + \iint_{\mathbb{R}^d\times\Gamma_{T_n}} \langle p_{\mu_0}\circ e_0(x,\gamma), w_{\boldsymbol{\eta}_n}(x,\gamma)\rangle \, d\boldsymbol{\eta}_n(x,\gamma) \geq -K\delta.$$

By passing to the infimum on $\boldsymbol{\eta}_n$ and $w_{\boldsymbol{\eta}_n} \in A_{T_n}(\mu_0)$, and recalling Corollary 4.4.2, we have

$$\begin{aligned}
-K\delta &\leq 1 + \inf_{\substack{v\in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)\\ v(x)\in F(x)\ \mu_0\text{-a.e }x}} \iint_{\mathbb{R}^d\times\Gamma_{T_n}} \langle p_{\mu_0}\circ e_0(x,\gamma), v\circ e_0(x,\gamma)\rangle \, d\boldsymbol{\eta}_n(x,\gamma)\\
&= 1 + \inf_{\substack{v\in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)\\ v(x)\in F(x)\ \mu_0\text{-a.e }x}} \int_{\mathbb{R}^d}\int_{\Gamma^x_{T_n}} \langle p_{\mu_0}\circ e_0(x,\gamma), v\circ e_0(x,\gamma)\rangle \, d\eta^n_x(\gamma)\, d\mu_0(x)\\
&= 1 + \inf_{\substack{v\in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)\\ v(x)\in F(x)\ \mu_0\text{-a.e }x}} \int_{\mathbb{R}^d} \langle p_{\mu_0}, v\rangle \, d\mu_0 = -\mathscr{H}_F(\mu_0, p_{\mu_0}),
\end{aligned}$$

so $\tau(\cdot)$ is a subsolution, thus confirming Claim 1. ◇

Claim 2: $\tau(\cdot)$ is a supersolution of $\mathscr{H}_F(\mu, D\tau(\mu)) = 0$ on $\mathcal{A}$.

Proof of Claim 2. Let $\mu_0 \in \mathcal{A}$. Let $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[}$ be an admissible clock-trajectory for μ_0 following a family of admissible mass-preserving trajectories $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$ starting from μ_0. For any $s \geq 0$ we choose $n > 0$ such that $\boldsymbol{\mu}^n$ is defined on an interval $[0, T_n]$ containing s and it is represented by $\boldsymbol{\eta}_n \in \mathscr{P}(\mathbb{R}^d \times \Gamma_{T_n})$. Taken $q_{\mu_0} \in D^-_\delta \tau(\mu_0)$, there is a sequence $\{s_i\}_{i\in\mathbb{N}} \subseteq]0, T_n[$, $s_i \to 0^+$ and $w_{\boldsymbol{\eta}_n} \in A_{T_n}(\mu_0)$ as in Corollary 4.4.2 such that for all $i \in \mathbb{N}$

$$\begin{aligned}
\iint_{\mathbb{R}^d\times\Gamma_{T_n}} \langle q_{\mu_0}\circ e_0(x,\gamma), \frac{e_{s_i}(x,\gamma)-e_0(x,\gamma)}{s_i}\rangle \, d\boldsymbol{\eta}_n(x,\gamma)&\\
\leq 2\delta \left\|\frac{e_{s_i}-e_0}{s_i}\right\|_{L^2_{\boldsymbol{\eta}_n}} - \frac{\tau(\mu_0)-\tau(\mu^n_{s_i})}{s_i}&.
\end{aligned}$$

By taking i sufficiently large we thus obtain

$$\iint_{\mathbb{R}^d\times\Gamma_{T_n}} \langle q_{\mu_0}\circ e_0(x,\gamma), w_{\boldsymbol{\eta}_n}(x,\gamma)\rangle \, d\boldsymbol{\eta}_n(x,\gamma) \leq 3K\delta - \frac{\tau(\mu_0)-\tau(\mu^n_{s_i})}{s_i}.$$

By using Corollary 4.4.2 and arguing as in Claim 1, we have

$$\inf_{\substack{v\in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)\\ v(x)\in F(x)\ \mu_0\text{-a.e }x}} \iint_{\mathbb{R}^d\times\Gamma_{T_n}} \langle q_{\mu_0}\circ e_0(x,\gamma), v\circ e_0(x,\gamma)\rangle \, d\boldsymbol{\eta}_n(x,\gamma) = -\mathscr{H}_F(\mu_0, q_{\mu_0}) - 1,$$

and so
$$\mathscr{H}_F(\mu_0, q_{\mu_0}) \geq -3K\delta + \frac{\tau(\mu_0) - \tau(\mu^n_{s_i})}{s_i} - 1.$$
By the Dynamic Programming Principle, passing to the infimum on all admissible curves and recalling that $\dfrac{\tau(\mu_0) - \tau(\mu^n_s)}{s} - 1 \leq 0$ with equality holding if and only if $\boldsymbol{\eta}_n$ is concentrated on time-optimal trajectories, we obtain $\mathscr{H}_F(\mu_0, q_{\mu_0}) \geq -C'\delta$, which proves that $\tau(\cdot)$ is a supersolution, thus confirming Claim 2. □

Chapter 5

Open Problems

In order to conclude the discussion, we list below the main open issues.

1. Regarding the general treatment discussed in Chapter 2, the open problems are
   - to prove a result of existence of optimal trajectories (the idea is to use l.s.c. of the cost functional $J(T, \boldsymbol{\mu}, \boldsymbol{\nu})$ together with relative compactness of the set of admissible trajectories for the finite-dimensional underlying problem);
   - to find the corresponding HJB equation in a very general form, under further smoothness assumptions;
   - to prove some estimates for the value function (maybe related to the generalized distance from the target).
2. In Section 3.2.1, we discussed sufficient conditions on the dynamics granting attainability in the mass-preserving case and then, in Section 3.2.2 we strengthen this hypothesis in order to have Lipschitz continuity of the generalized minimum time function. In this line, an open problem consists in the study of further regularity properties of the minimum time function with milder assumptions on the dynamics, stating the problem in a suitable *smaller* class of probability measures, following the so called *Lagrangian flow problem.*
3. As pointed out in Section 3.4, in which a correspondent quantity for the Lie bracket in a measure-theoretic setting is presented, an interesting study will be related to the proof of higher order controllability conditions for the time-optimal control problem presented in Chapter 3.
4. The most important open problem of this thesis regards the framework of Chapters 3 and 4 which lack a *Comparison Principle* result that would lead to a characterization of the generalized minimum time function as the unique viscosity solution of an Hamilton-Jacobi-Bellman equation. Furthermore, as remarked in Section 3.3, another open problem is the extension of the definition of viscosity solutions and the related result on

HJB equation to the case where we have only lower semicontinuity of the minimum time function, instead of continuity, following a Barron-Jensen's approach to viscosity solutions.

5. Another open problem regarding Chapters 3 and 4 is to provide an analogous of the *Pontryagin maximum principle*, in order to formulate necessary conditions for an admissible trajectory to be optimal.

6. Finally, from an applicative point of view and in purpose of possible applications in *multi-agents systems*, it would be interesting to implement numerical symulations for the theory presented in Chapters 3 and 4.

Acknowledgments

I wish to conclude the thesis by thanking all the people who have been near to me during these years of PhD and who made this an interesting, motivating and nice period, supporting me expecially in harder periods and choices.

First of all, thanks mum, dad and brothers for your deep comprehension, faith, patience and constant support. Also thanks to all my relatives, expecially grandmother and "ziet".

My heartfelt thanks to the other author of this thesis, my supervisor Dr. Antonio Marigonda, who accompanied me throughout this period with great humanity and professionality. I am really proud to have had you as a supervisor, because of the work-style, the nice research topic, but also because of the suggestions, helpful discussions, opportunities and for your interest for my future.

A special thank to all my friends, who shared and will share with me meaningful moments during the years of study at University of Verona, at the AC group, in the sport, at University of Trento and in my village.

I wish to thank also the colleagues met at the University of Padova and during the conferences attended worldwide, together with all other researchers and professors with whom I had the opportunity to talk in these challenging events and courses attended.

I cannot forget Prof. Maxim V. Balashov, his kind invitation and teachings at MIPT (Dolgoprudny (Moscow), Russia) and the friends met during this nice and special period abroad.

I want to sincerely thank also Prof. Benedetto Piccoli for the interest shown for this research work, his invitation at Rutgers University-Camden (New Jersey, USA) and the study done there (which is part of the present thesis). Thanks also to all the friends met there, you made me feel at home from the very beginning.

Last but not least, I am grateful to the professors met at University of Verona, some of them now are elsewhere, who accompanied my university education, and Prof. Giandomenico Orlandi who has tutored expecially the starting period of my PhD program.

Bibliography

[1] Luigi Ambrosio, *The flow associated to weakly differentiable vector fields: recent results and open problems*, Nonlinear conservation laws and applications, IMA Vol. Math. Appl., vol. 153, Springer, New York, 2011, pp. 181–193, DOI 10.1007/978-1-4419-9554-4_7. MR2856995 (2012i:35224)

[2] ———, *Transport equation and Cauchy problem for non-smooth vector fields*, Calculus of Variations and Nonlinear Partial Differential Equations, Lecture Notes in Mathematics, vol. 1927, Springer-Verlag Berlin Heidelberg, Heidelberg, 2008, pp. 1–41, DOI 10.1007/978-3-540-75914-0_1.

[3] ———, *Transport equation and Cauchy problem for BV vector fields*, Invent. Math. **158** (2004), no. 2, 227–260, DOI 10.1007/s00222-004-0367-2. MR2096794

[4] Luigi Ambrosio and Gianluca Crippa, *Continuity equations and ODE flows with non-smooth velocity*, Proceedings of the Royal Society of Edinburgh: Section A **144** (2014), no. 6, 1191–1244, DOI 10.1017/S0308210513000085.

[5] Luigi Ambrosio and Alessio Figalli, *Almost everywhere well-posedness of continuity equations with measure initial data*, Comptes Rendus Mathematique **348** (2010), no. 5, 249–252, DOI 10.1016/j.crma.2010.01.018.

[6] Luigi Ambrosio, Nicola Fusco, and Diego Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2000. MR1857292 (2003a:49002)

[7] Luigi Ambrosio and Wilfred Gangbo, *Hamiltonian ODEs in the Wasserstein space of probability measures*, Comm. Pure Appl. Math. **61** (2008), no. 1, 18–53, DOI 10.1002/cpa.20188. MR2361303 (2009b:37101)

[8] Luigi Ambrosio and Nicola Gigli, *A User's Guide to Optimal Transport*, Lecture Notes in Mathematics, vol. 2062, Springer Berlin Heidelberg, 2013.

[9] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Gradient flows in metric spaces and in the space of probability measures*, 2nd ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008. MR2401600 (2009h:49002)

[10] Fabio Ancona and Alberto Bressan, *Nearly time optimal stabilizing patchy feedbacks*, Ann. Inst. Henri Poincaré, Analyse Non Linéaire **24** (2007), 279–310.

[11] ———, *Patchy vector fields and asymptotic stabilization*, ESAIM Control Optim. Calc. Var. **4** (1999), 445–471.

[12] Jean-Pierre Aubin and Arrigo Cellina, *Differential Inclusions*: *Set-Valued Maps and Viability Theory*, Springer-Verlag Berlin Heidelberg, Germany, 1984.

[13] Jean-Pierre Aubin and Hélène Frankowska, *Set-valued analysis*, Modern Birkhäuser Classics, Birkhäuser Boston Inc., Boston, MA, 2009. Reprint of the 1990 edition [MR1048347]. MR2458436

[14] Martino Bardi and Italo 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. MR1484411

[15] Patrick Bernard, *Young measures, superpositions and transport*, Indiana Univ. Math. J. **57** (2008), no. 1, 247–276.

[16] Alberto Bressan and Fabio S. Priuli, *Nearly optimal patchy feedbacks*, Discr. Cont. Dyn. Systems – Series A **21** (2008), 687–701.

[17] Roger W. Brockett and Daniel Liberzon, *On explicit steady-state solutions of Fokker-Planck equations for a class of nonlinear feedback systems*, Proceedings of the american Control Conference **1** (1998), 264–268.

[18] Giuseppe Buttazzo, *Semicontinuity, relaxation and integral representation in the calculus of variations*, Pitman Research Notes in Mathematics Series, vol. 207, Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1989. MR1020296

[19] Giuseppe Buttazzo, Chloé Jimenez, and Edouard Oudet, *An optimization problem for mass transportation with congested dynamics*, SIAM J. Control Optim. **48** (2009), no. 3, 1961–1976, DOI 10.1137/07070543X. MR2516195

[20] Piermarco Cannarsa, Antonio Marigonda, and Khai T. Nguyen, *Optimality conditions and regularity results for time optimal control problems with differential inclusions*, J. Math. Anal. Appl. **427** (2015), no. 1, 202–228, DOI 10.1016/j.jmaa.2015.02.027. MR3318195

[21] Piermarco Cannarsa and Carlo Sinestrari, *Convexity properties of the minimum time function*, Calc. Var. Partial Differential Equations **3** (1995), no. 3, 273–298, DOI 10.1007/BF01189393. MR1385289

[22] ———, *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*, Progress in Nonlinear Differential Equations and their Applications, 58, Birkhäuser Boston, Inc., Boston, MA, 2004. MR2041617

[23] Piermarco Cannarsa and Peter R. Wolenski, *Semiconcavity of the value function for a class of differential inclusions*, Discrete Contin. Dyn. Syst. **29** (2011), no. 2, 453–466, DOI 10.3934/dcds.2011.29.453. MR2728465 (2012a:49046)

[24] Marco Caponigro, Massimo Fornasier, Benedetto Piccoli, and Emmanuel Trélat, *Sparse stabilization and control of alignment models*, Math. Models and Methods in Appl. Sci. **25** (2015), no. 03, 521–564, DOI 10.1142/S0218202515400059.

[25] Laura Caravenna and Gianluca Crippa, *Uniqueness and Lagrangianity for solutions with low integrability of the continuity equation* (2016). arXiv:1608.04324v1 [math.AP].

[26] Pierre Cardaliaguet and Marc Quincampoix, *Deterministic differential games under probability knowledge of initial condition*, Int. Game Theory Rev. **10** (2008), no. 1, 1–16, DOI 10.1142/S021919890800173X. MR2423798

[27] José Antonio Carrillo, Young-Pil Choi, and Maxime Hauray, *The derivation of swarming models: mean-field limit and Wasserstein distances*, Collective dynamics from bacteria to crowds, CISM Courses and Lectures, vol. 553, Springer, Vienna, 2014, pp. 1–46, DOI 10.1007/978-3-7091-1785-9_1. MR3331178

[28] Giulia Cavagnari, *Regularity results for a time-optimal control problem in the space of probability measures*, to appear in Mathematical Control and Related Fields.

[29] Giulia Cavagnari and Antonio Marigonda, *Measure-theoretic Lie Brackets for nonsmooth vector fields*, preprint.

[30] ———, *Time-optimal control problem in the space of probability measures*, Large-scale scientific computing, Lecture Notes in Computer Science, vol. 9374, Springer, Cham, 2015, pp. 109–116, DOI 10.1007/978-3-319-26520-9. MR3480817

[31] Giulia Cavagnari, Antonio Marigonda, Khai T. Nguyen, and Fabio S. Priuli, *Generalized control systems in the space of probability measures*, submitted.

[32] Giulia Cavagnari, Antonio Marigonda, and Giandomenico Orlandi, *Hamilton-Jacobi-Bellman equation for a time-optimal control problem in the space of probability measures*, submitted.

[33] Giulia Cavagnari, Antonio Marigonda, and Benedetto Piccoli, *Averaged time-optimal control problem in the space of positive Borel measures*, preprint.

[34] ———, *Optimal syncronization problem for a multi-agent system*, submitted.

[35] Francis H. Clarke, *Functional analysis, calculus of variations and optimal control*, Graduate Texts in Mathematics, vol. 264, Springer, London, 2013. MR3026831

[36] Francis H. Clarke, Yuri S. Ledyaev, Ronald J. Stern, and Peter R. Wolenski, *Nonsmooth analysis and control theory*, Graduate Texts in Mathematics, vol. 178, Springer-Verlag, New York, 1998. MR1488695 (99a:49001)

[37] Giovanni Colombo, Antonio Marigonda, and Peter R. Wolenski, *Some new regularity properties for the minimal time function*, SIAM J. Control Optim. **44** (2006), no. 6, 2285–2299 (electronic), DOI 10.1137/050630076. MR2248184 (2008d:49021)

[38] Gianluca Crippa and Camillo De Lellis, *Estimates and regularity results for the DiPerna-Lions flow*, J. Reine Angew. Math. **616** (2008), 15–46, DOI 10.1515/CRELLE.2008.016. MR2369485

[39] Emiliano Cristiani, Benedetto Piccoli, and Andrea Tosin, *Multiscale modeling of pedestrian dynamics*, MS&A. Modeling, Simulation and Applications, vol. 12, Springer, Cham, 2014. MR3308728

[40] Ronald J. DiPerna and Pierre Louis Lions, *Ordinary differential equations, transport theory and Sobolev spaces*, Invent. Math. **98** (1989), no. 3, 511–547, DOI 10.1007/BF01393835. MR1022305

[41] Jean Dolbeault, Bruno Nazaret, and Giuseppe Savaré, *A new class of transport distances between measures*, Calc. Var. Partial Differential Equations **34** (2009), no. 2, 193–231, DOI 10.1007/s00526-008-0182-5. MR2448650 (2009g:49019)

[42] Ermal Feleqi and Franco Rampazzo, *Integral representations for bracket-generating multi-flows*, Discrete Contin. Dyn. Syst. **35** (2015), no. 9, 4345–4366, DOI 10.3934/dcds.2015.35.4345. MR3392629

[43] ———, *Iterated Lie brackets for nonsmooth vector fields*, preprint.

[44] Massimo Fornasier, Benedetto Piccoli, and Francesco Rossi, *Mean-field sparse optimal control*, Philosophical Transactions of the Royal Society A, posted on 2014, DOI 10.1098/rsta.2013.0400.

[45] Massimo Fornasier and Francesco Solombrino, *Mean-Field Optimal Control*, ESAIM: Control, Optimization and Calc. of Var. **20** (2014), no. 4, 1123–1152, DOI 10.1051/cocv/2014009.

[46] Wilfrid Gangbo, Truyen Nguyen, and Adrian Tudorascu, *Hamilton-Jacobi equations in the Wasserstein space*, Methods Appl. Anal. **15** (2008), no. 2, 155–183, DOI 10.4310/MAA.2008.v15.n2.a4. MR2481677 (2010f:49061)

[47] Wilfrid Gangbo and Andrzej Święch, *Optimal transport and large number of particles*, Discrete Contin. Dyn. Syst. **34** (2014), no. 4, 1397–1441, DOI 10.3934/dcds.2014.34.1397. MR3117847

[48] Yoshikazu Giga, Nao Hamamuki, and Atsushi Nakayasu, *Eikonal equations in metric spaces*, Trans. Amer. Math. Soc. **367** (2015), no. 1, 49–66, DOI 10.1090/S0002-9947-2014-05893-5.

[49] Diogo Gomes and Levon Nurbekyan, *An infinite-dimensional Weak KAM theory via random variables* (2015). arXiv:1508.00154v1 [math.DS].

[50] Seung-Yeal Ha and Eitan Tadmor, *From particle to kinetic and hydrodynamic descriptions of flocking*, Kinet. Relat. Models **1** (2008), no. 3, 415–435, DOI 10.3934/krm.2008.1.415. MR2425606

[51] Henry Hermes and Joseph P. LaSalle, *Functional analysis and time optimal control*, Academic Press, New York-London, 1969. Mathematics in Science and Engineering, Vol. 56. MR0420366

[52] Velimir Jurdjevic, *Geometric control theory*, Cambridge Studies in Advanced Mathematics, vol. 52, Cambridge University Press, Cambridge, 1997. MR1425878

[53] Leonid V. Kantorovich, *On the transfer of masses*, Dokl. Acad. Nauk **37** (1942), no. 7–8.

[54] Jean-Michel Lasry and Pierre Louis Lions, *Jeux à champ moyen. I. Le cas stationnaire*, C. R. Acad. Sci. Paris **343** (2006), 619–625.

[55] ———, *Jeux à champ moyen. II. Horizon fini et contrôle optimal*, C. R. Acad. Sci. Paris **343** (2006), 679–684.

[56] Stefano Lisini and Antonio Marigonda, *On a class of modified Wasserstein distances induced by concave mobility functions defined on bounded intervals*, Manuscripta Math. **133** (2010), no. 1-2, 197–224, DOI 10.1007/s00229-010-0371-3. MR2672546

[57] Stefania Maniglia, *Probabilistic representation and uniqueness results for measure-valued solutions of transport equations*, J. Math. Pures Appl. (9) **87** (2007), no. 6, 601–626, DOI 10.1016/j.matpur.2007.04.001 (English, with English and French summaries). MR2335089

[58] Antonio Marigonda, *Second order conditions for the controllability of nonlinear systems with drift*, Comm. Pure and Applied Analysis **5** (2006), no. 4, 861–885.

[59] Antonio Marigonda and Silvia Rigo, *Controllability of some nonlinear systems with drift via generalized curvature properties*, SIAM J. Control Optim. **53** (2015), no. 1, 434–474, DOI 10.1137/130920691. MR3310970

[60] Antonio Marigonda and Thuy T.T. Le, *Small-time local attainability for a class of control systems with state constraints*, ESAIM: Control, Optimization and Calc. of Var., posted on 2016, to appear, DOI 10.1051/cocv/2016022.

[61] Markus Mauhart and Peter W. Michor, *Commutators of flows and fields*, Arch. Math. (Brno) **28** (1992), no. 3-4, 229–236. MR1222291 (94e:58117)

[62] Gaspard Monge, *Mémoire sur la théorie des déblais at des remblais*, Histoire de l'Académie Royale des Sciences de Paris (1781), 666-704.

[63] Sebastien Motsch and Eitan Tadmor, *Heterophilious dynamics enhances consensus*, SIAM Rev. **56** (2014), no. 4, 577–621, DOI 10.1137/120901866. MR3274797

[64] Benedetto Piccoli and Francesco Rossi, *Generalized Wasserstein distance and its application to transport equations with source*, Arch. Ration. Mech. Anal. **211** (2014), no. 1, 335–358, DOI 10.1007/s00205-013-0669-x. MR3182483

[65] ———, *On properties of the Generalized Wasserstein distance* (2014). arXiv:1304.7014v3 [math.AP].

[66] Benedetto Piccoli and Andrea Tosin, *Time-evolving measures and macroscopic modeling of pedestrian flow*, Arch. Ration. Mech. Anal. **199** (2011), no. 3, 707–738, DOI 10.1007/s00205-010-0366-y. MR2771664

[67] Franco Rampazzo, *Frobenius-type theorems for Lipschitz distributions*, J. Differential Equations **243** (2007), no. 2, 270–300, DOI 10.1016/j.jde.2007.05.040. MR2371789

[68] Franco Rampazzo and Héctor J. Sussmann, *Commutators of flow maps of nonsmooth vector fields*, J. Differential Equations **232** (2007), no. 1, 134–175, DOI 10.1016/j.jde.2006.04.016. MR2281192 (2007j:49021)

[69] ———, *Set-valued differentials and a nonsmooth version of Chow's theorem*, Proc. of the 40th IEEE Conf. on Decision and Control, Orlando, FL, December 2001, Vol. 3, IEEE Publications, New York, 2001, pp. 2613–2618.

[70] R. Tyrrell Rockafellar and Roger J.-B. Wets, *Variational analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 317, Springer-Verlag, Berlin, 1998. MR1491362

[71] Filippo Santambrogio, *Optimal Transport for Applied Mathematicians*: *Calculus of Variations, PDEs, and Modeling*, 1st ed., Progress in Nonlinear Differential Equations and Their Applications, vol. 87, Birkhäuser Basel, Basel, 2015.

[72] Héctor J. Sussmann, *Subanalytic sets and feedback control*, J. Differential Equations **31** (1979), 31–52.

[73] Cédric Villani, *Optimal transport: Old and New*, Grundlehren der mathematischen Wissenschaften, vol. 338, Springer Science & Business Media, Heidelberg, 2008.

[74] ———, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003. MR1964483 (2004e:90003)

[75] Laurence C. Young, *Lecture on the Calculus of Variations and Optimal Control Theory*, AMS Chelsea Publishing, vol. 304, American Mathematical Society, Providence, RI, 1980.