# PARABOLIC SYSTEMS AND AN UNDERLYING LAGRANGIAN

A Thesis
Presented to
The Academic Faculty

by

Türkay Yolcu

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics

Georgia Institute of Technology
August 2009

# PARABOLIC SYSTEMS AND AN UNDERLYING LAGRANGIAN

Approved by:

Professor Wilfrid Gangbo, Advisor
School of Mathematics
*Georgia Institute of Technology*

Professor Shui-Nee Chow
School of Mathematics
*Georgia Institute of Technology*

Professor Evans Harrell
School of Mathematics
*Georgia Institute of Technology*

Professor Andrzej Swiech
School of Mathematics
*Georgia Institute of Technology*

Professor Anthony Joseph Yezzi
School of Electrical and Computer Engineering
*Georgia Institute of Technology*

Date Approved: 2 July 2009

*To Selma, my love, everything,*

# ACKNOWLEDGEMENTS

I would like to take this opportunity to express my deep and sincere gratitude to many people who contributed to my graduate experience at the School of Mathematics. This dissertation is possible because of their help, support and guidance.

I owe my deepest gratitude to my advisor, Wilfrid Gangbo, whose insightful comments and constructive criticisms have significantly strengthened this thesis. He taught me how to focus my ideas and impressed me with his knowledge in many areas. He was always there to listen and to give advice. Without his wisdom, knowledge, guidance and determination this work would never be accomplished. It has been an honor to be his student.

I wish to thank Alessio Figalli for his suggestions and comments. I am grateful to him for helping me to sort out the technical details of this work.

I would like to thank the committee members Shui-Nee Chow, Evans Harrell, Andrzej Swiech, and Anthony Joseph Yezzi for agreeing to serve on my committee and providing helpful comments and suggestions.

I am indebted to many of my friends for their kindness and camaraderie, in particular to Ramazan Tinaztepe, Jorge Viveros, and Suleyman Ulusoy.

I am also grateful to the following staff at the School of Mathematics, for their various forms of support during my graduate study Genola Turner, Sharon McDowell and Annette Rohrs. I also want to thank to William Green, Evans Harrell and Luca Dieci for providing advice on my graduate studies.

During the last year of this dissertation, I have been working as an instructor at Georgia College and State University. I am very thankful to all faculty members in the Department of Mathematics for making Georgia College a wonderful workplace and for providing a friendly and warm environment. I especially thank to Jason Huffman, Martha Allen, George Cazacu, Rodica Cazacu for their interesting discussions and friendship.

I wish to express my warm and sincere thanks to Rena Brakebill and Alfred Andrew for providing valuable advice and feedback on my teaching and helping me in my job search. Also, I am really lucky to meet Ms. Cathy Jacobson who helped a lot to improve my teaching as well as my english. I am grateful for their support and assistance during all the years at Georgia Tech.

I thank to my master thesis advisors, Ağacik Zafer and Marat Akhmet, for their encouragement to come to Georgia Tech.

Last but not least, I would like to extend my heartfelt gratitude to my family, especially my wife, Selma. Without their love, encouragement and constant support I would not have been able to finish any of these.

I offer my regards and blessings to all of those who supported me in any respect during the completion of my thesis. Thank you.

# TABLE OF CONTENTS

DEDICATION . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . iii

ACKNOWLEDGEMENTS . . . . . . . . . . . . . . . . . . . . . . . . . . iv

SUMMARY . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . viii

I INTRODUCTION . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

II PRELIMINARIES . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8

2.1 Notation, Definitions and Main Assumptions . . . . . . . . . . . . 8

2.2 Properties of Enthalpy and Pressure Functionals . . . . . . . . . . 14

2.3 Properties of $H$ and the Cost Functions . . . . . . . . . . . . . . 16

2.4 Total Works and Their Properties . . . . . . . . . . . . . . . . . . 19

III EXISTENCE OF SOLUTIONS IN A BOUNDED DOMAIN . . . . . . . 23

3.1 The Discrete Variational Problem . . . . . . . . . . . . . . . . . . 23

3.2 Characterization of Minimizers Via a Dual Problem . . . . . . . . . 27

3.3 Augmented Actions . . . . . . . . . . . . . . . . . . . . . . . . . 37

3.4 De Giorgi's Variational and Geodesic Interpolations . . . . . . . . . 40

3.5 Stability Property and Existence of Solutions . . . . . . . . . . . . 43

IV EXISTENCE OF SOLUTIONS IN UNBOUNDED DOMAINS FOR WEAK TONELLI LAGRANGIANS . . . . . . . . . . . . . . . . . . . . . . 54

4.1 Lower-semicontinuity of $\mathcal{U}$ . . . . . . . . . . . . . . . . . . . . . 54

4.2 Properties of Moments in the Unbounded Case . . . . . . . . . . . 56

4.3 Existence of Solutions . . . . . . . . . . . . . . . . . . . . . . . . 58

V UNIQUENESS OF DISTRIBUTIONAL SOLUTIONS . . . . . . . . . . 62

5.1 Assumptions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 62

5.2 A new Hamiltonian . . . . . . . . . . . . . . . . . . . . . . . . . . 62

5.3 Additional Properties Satisfied by Bounded Solutions . . . . . . . . 64

5.4 Uniqueness of Bounded Solutions . . . . . . . . . . . . . . . . . . 66

APPENDIX A SOME SUPPLEMENTARY RESULTS . . . . . . . . . . . 70

REFERENCES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 81

# SUMMARY

In this thesis, we extend De Giorgi's interpolation method to a class of parabolic equations which are not gradient flows but possess an entropy functional and an underlying Lagrangian. The new fact in the study is that not only the Lagrangian may depend on spatial variables, but also it does not induce a metric. Assuming the initial condition is a density function, not necessarily smooth, but solely of bounded first moments and finite "entropy", we use a variational scheme to discretize the equation in time and construct approximate solutions. Moreover, De Giorgi's interpolation method is revealed to be a powerful tool for proving convergence of our algorithm. Finally, we analyze uniqueness and stability of our solution in $L^1$.

# CHAPTER I

# INTRODUCTION

In the theory of existence of solutions of ordinary differential equations on a metric space, curves of maximal slope and minimizing movements play a crucial role. The minimizing movements in general results from a discrete scheme. They have the advantage of providing an approximate solution of the differential equation by discretizing in time while not requiring the initial condition to be smooth. Then a neat interpolation method introduced by De Giorgi [7, 8] ensures the compactness for the family of approximate solutions. Many recent works [3, 17] have used minimizing movement methods as a powerful tool for proving existence of solution for some classes of partial differential equations (PDEs). So far, most of these studies have been concerned with PDEs which can be interpreted as gradient flows of an entropy functional with respect to a metric on the space of probability measures. In this thesis, we extend the minimizing movements and De Giorgi's interpolation method to include PDEs which are not gradient flows, but possess an entropy functional and an underlying Lagrangian which may depend on the spatial variables. The main part of this work is studied in the joint work (cf. [11]).

In what follows $X \subset \mathbb{R}^d$ represents an open set whose boundary is of zero measure. We denote by $\mathcal{P}_\alpha(X)$ the set of Borel probability measures on $\mathbb{R}^d$ of bounded $\alpha$-moments, equipped with the $\alpha$-Wasserstein distance $W_\alpha$ (cf. Equation (2.1.9)). Let $\mathcal{P}_\alpha^{ac}(X)$ be the set of probability densities $\varrho$ such that $\varrho\mathcal{L}^d$ belongs to $\mathcal{P}_\alpha(X)$. We consider distributional solutions of a type of PDEs of the form

$$\partial_t \varrho_t + \operatorname{div}(\varrho_t V_t) = 0, \quad \text{in} \qquad \mathcal{D}'((0,T) \times X) \tag{1.0.1}$$

(this implicitly means that we have imposed Neumann boundary conditions), with

$$\varrho_t V_t := \varrho_t \nabla_{\mathbf{p}} H\big(\mathbf{x}, -\varrho_t^{-1} \nabla[P(\varrho_t)]\big) \quad \text{on} \quad (0,T) \times X$$

and

$$t \mapsto \varrho_t \in AC_1(0,T; \mathcal{P}_1^{ac}(X)) \subset C([0,T]; \mathcal{P}_1(X)).$$

The space to which the curve $t \mapsto \varrho_t$ belongs ensures that $\varrho_t$ converges to $\varrho_0$ in $\mathcal{P}_1^{ac}(X)$ as $t$ tends to 0. By abuse of notation, $\varrho_t$ will denote at the same time the solution at time $t$ and the function $(t, \mathbf{x}) \mapsto \varrho_t(\mathbf{x}) = \varrho(t, \mathbf{x})$ defined over $(0,T) \times X$. We only consider solutions such that $\nabla[P(\varrho_t)] \in L^1((0,T) \times X)$, and is absolutely continuous with respect to $\varrho_t$. If $\varrho_t$ satisfies additional conditions, then

$$t \mapsto \mathcal{U}(\varrho_t) := \int_X U(\varrho_t(\mathbf{x}))\, \mathrm{d}\mathbf{x}$$

is absolutely continuous, monotone nonincreasing, and

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{U}(\varrho_t) = \int_X \langle \nabla[P(\varrho_t)], V_t \rangle\, \mathrm{d}\mathbf{x}. \tag{1.0.2}$$

We recall that the unknown $\varrho_t$ is nonnegative, and can be interpreted as the density of a fluid, whose pressure is $P(\varrho_t)$. Here, the data $H$, $U$ and $P$ satisfy specific properties, which are stated in section 2.1.

Solutions of our equation can be regarded as curves of maximal slope on a metric space contained in $\mathcal{P}_1(X)$. They include the so-called minimizing movements (cf. [3]) obtained by many authors in case the Lagrangian does not depend on spatial variables (e.g. [16] when $H(\mathbf{p}) = 1/2|\mathbf{p}|^2$, [1, 3] when $H(\mathbf{x}, \mathbf{p}) \equiv H(\mathbf{p})$). These studies have been very recently extended to a special class of Lagrangian depending on spatial variables where the Hamiltonian assumes the form $H(\mathbf{x}, \mathbf{p}) = \langle A^*(\mathbf{x})\mathbf{p}, \mathbf{p} \rangle$ [17]. In their pioneering work Alt and Luckhaus [2] considered a system of quasilinear elliptic-parabolic differential equations of the form

$$\partial_t b^j(u) - \operatorname{div}(a^j(b(u), \nabla u)) = f^j(b(u)) \quad \text{on} \;\; (0,T) \times \Omega, \quad j = 1, ..., m.$$

$$b^j(u) = b^0 \text{ on } \{0\} \times \Omega, \quad u = u^D \text{ on } \{(0,T)\} \times \Gamma,$$

$$a^j(b(u), \nabla u) \cdot \nu = 0 \text{ on } (0,T) \times (\partial\Omega \setminus \Gamma), \ \ j = 1, ..., m.$$

similar to (1.0.1), requiring some assumptions not very comparable to ours. Their method of proof is very different from the ones practiced in the cited references above and is based on a Galerkin type approximation method. As an example, let us also remark that if

$$L(\mathbf{x}, \mathbf{v}) = \frac{|\mathbf{v}|^q}{q}, \quad 1 < q \neq 2 \quad \text{and} \quad U(t) = \frac{t^m}{m(m-1)}, \quad m = \frac{2p-3}{p-1}, \quad \frac{1}{p} + \frac{1}{q} = 1,$$

then Equation (1.0.1) becomes the (gradient flow) $p-$Laplacian Equation:

$$\partial_t \varrho_t = \operatorname{div}(|\nabla \varrho_t|^{p-2} \nabla \varrho_t).$$

Moreover, in the case of $q = 2$, taking $U(t) = t \log t$ together with $L(\mathbf{x}, \mathbf{v}) = \frac{|\mathbf{v}|^2}{2}$ turns Equation (1.0.1) into the Diffusion (Heat) Equation:

$$\partial_t \varrho_t = \Delta \varrho_t.$$

The strategy of the proof of our results is described as follows. As a first step, we show the existence of the solution. Let $L(\mathbf{x}, \cdot)$ be the Legendre transform of $H(\mathbf{x}, \cdot)$, which we refer to as a Lagrangian. For a time step $h > 0$, let $c_h(\mathbf{x}, \mathbf{y})$, the cost for moving a unit mass from a point $\mathbf{x}$ to a point $\mathbf{y}$, be the minimal action $\min_\sigma \int_0^h L(\sigma, \dot{\sigma}) \mathrm{d}t$. Here, the minimum is performed over the set of all paths (not necessarily contained in $X$) such that $\sigma(0) = \mathbf{x}$ and $\sigma(h) = \mathbf{y}$. The cost $c_h$ provides a way of defining the minimal total work $\mathcal{C}_h(\varrho_0, \varrho)$ (cf. (2.1.7)) for moving a mass of distribution $\varrho_0$ to another mass of distribution $\varrho$ in $X$. For absolutely continuous measures, the recent papers [5, 9, 10] give the uniqueness of minimizers in (2.1.7), which is concentrated on the graph of a function $T_h : \mathbb{R}^d \to \mathbb{R}^d$. Furthermore, $\mathcal{C}_h$ provides a natural way of interpolating between these measures: there exists a unique density $\bar{\varrho}_s$ such that

$$\mathcal{C}_h(\varrho_0, \varrho_h) = \mathcal{C}_s(\varrho_0, \bar{\varrho}_s) + \mathcal{C}_{h-s}(\bar{\varrho}_s, \varrho_h), \qquad s \in (0, h).$$

Assume for the moment that $X$ is bounded. For a given initial condition $\varrho_0 \in \mathcal{P}_1^{ac}(X)$ we inductively construct $\{\varrho_{nh}^h\}_n$ as follows: $\varrho_{(n+1)h}$ is the unique minimizer of

$$\mathcal{C}_h(\varrho_{nh}^h, \varrho) + \int_X U(\varrho)\,\mathrm{d}\mathbf{x}$$

over $\mathcal{P}_1^{ac}(X)$. We refer to this minimization problem as the primal problem. Under the additional condition that $L(\mathbf{x}, \mathbf{v}) > L(\mathbf{x}, \mathbf{0}) \equiv 0$ for all $\mathbf{x}, \mathbf{v} \in \mathbb{R}^d$ such that $\mathbf{v} \neq 0$, one has $c_h(\mathbf{x}, \mathbf{x}) < c_h(\mathbf{x}, \mathbf{y})$ for $\mathbf{x} \neq \mathbf{y}$. As a consequence, under that condition the following maximum principle (cf. Theorem 3.1.1) holds: if $\varrho_0 \leq M$ then $\varrho_{nh}^h \leq M$ for all $n \geq 0$. We then study a problem, dual to the primal one, which provides us with a characterization and some important regularity properties of the minimizer $\varrho_{(n+1)h}$. These properties would have been harder to obtain by studying only the primal problem. Having determined $\{\varrho_{nh}^h\}_{n \in \mathbb{N}}$, we consider two interpolating paths. The first one is the path $t \to \bar{\varrho}_t^h$ such that

$$\mathcal{C}_h(\varrho_{nh}^h, \varrho_{(n+1)h}^h) = \mathcal{C}_s(\varrho_{nh}^h, \bar{\varrho}_{nh+s}^h) + \mathcal{C}_{h-s}(\bar{\varrho}_{nh+s}^h, \varrho_{(n+1)h}^h), \qquad 0 < s < h.$$

The second path $t \to \varrho_t^h$ is defined by

$$\varrho_{nh+s}^h := \arg\min\left\{\mathcal{C}_s(\varrho_{nh}^h, \varrho) + \int_X U(\varrho)\,\mathrm{d}\mathbf{x}\right\}, \qquad 0 < s < h.$$

This interpolation was introduced by De Giorgi in the study of curves of maximal slopes when $\sqrt{\mathcal{C}_h}$ defines a metric. The path $\{\bar{\varrho}_t^h\}$ satisfies Equation (3.5.11), which is a discrete analogue of the differential quation in (1.0.1). Then we write a discrete energy inequality in terms of both paths $\{\bar{\varrho}_t^h\}$ and $\{\varrho_t^h\}$, and we prove that up to a subsequence both paths converge (in a sense to be made precise) to the same path $\varrho_t$. Furthermore, $\varrho_t$ satisfies the energy inequality

$$\mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_T) \geq \int_0^T \mathrm{d}t \int_X \Big[L\big(\mathbf{x}, V_t\big) + H\big(\mathbf{x}, -\varrho_t^{-1}\nabla[P(\varrho_t)]\big)\Big]\varrho_t\,\mathrm{d}\mathbf{x}, \tag{1.0.3}$$

which, by the assumptions on $H$ (cf. section 2.1), implies for instance that $\nabla[P(\varrho_t)] \in L^1((0,T) \times X)$. The above inequality corresponds to what can be regarded as one half of the chain rule:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{U}(\varrho_t) \leq \int_X \langle V_t, \nabla[P(\varrho_t)]\rangle \ \mathrm{d}\mathbf{x}.$$

Here $V_t$ is a velocity associated with the path $t \mapsto \varrho_t$, in the sense that Equation (1.0.1) holds without knowing that $\varrho_t V_t := \varrho_t \nabla_{\mathbf{p}} H\big(\mathbf{x}, -\varrho_t^{-1}\nabla[P(\varrho_t)]\big)$. Next, we establish the reverse inequality yielding the whole chain rule only if we know that

$$\int_0^T \mathrm{d}t \int_X |V_t|^\alpha \varrho_t \ \mathrm{d}\mathbf{x}, \quad \int_0^T \mathrm{d}t \int_X |\varrho_t^{-1}\nabla[P(\varrho_t)]|^{\alpha'} \varrho_t \ \mathrm{d}\mathbf{x} < +\infty \tag{1.0.4}$$

for some $\alpha \in (1,\infty)$, $\alpha' = \alpha/(\alpha-1)$. In that case, we can conclude that

$$\varrho_t V_t := \varrho_t \nabla_{\mathbf{p}} H\big(\mathbf{x}, -\varrho_t^{-1}\nabla[P(\varrho_t)]\big)$$

and

$$\frac{d}{\mathrm{d}t}\mathcal{U}(\varrho_t) = \int_X \langle V_t, \nabla[P(\varrho_t)]\rangle \, \mathrm{d}\mathbf{x}.$$

In light of the energy inequality in (3.5.12), a sufficient condition to have the inequality (1.0.4) is that $L(\mathbf{x},\mathbf{v}) \sim |\mathbf{v}|^\alpha$. This is what we later impose in this work.

Suppose now that $X$ may be unbounded. As pointed out in Remark 3.5.5, by a simple scaling argument we can solve Equation (1.0.1) for general nonnegative densities, not necessarily of unit mass. Lemma 4.1.1 shows that if we require the bound in (4.0.23) on the negative part of $U$, then $\int_X U(\varrho(\mathbf{x}))\,\mathrm{d}\mathbf{x}$ is well-defined for $\varrho \in \mathcal{P}_1^{ac}(X)$. We assume that the initial condition $\varrho_0 \in \mathcal{P}_1^{ac}(X)$ and $\int_X |U(\varrho_0(\mathbf{x}))|\,\mathrm{d}\mathbf{x}$ is finite, and we start our approximation argument by replacing $X$ by $X_m := X \cap B_m(0)$ and $\varrho_0$ by $\varrho_0^m := \varrho_0 \chi_{B_m(0)}$. Here, $B_m(0)$ is the open ball of radius $m$, centered at the origin. The previous argument provides us with a solution of Equation (1.0.1), starting at $\varrho_0^m$, for which we show that

$$\max_{t\in[0,T]} \left\{ \int_{X_m} |\mathbf{x}|\varrho_t^m \,\mathrm{d}\mathbf{x} + \int_{X_m} |U(\varrho_t^m)|\,\mathrm{d}\mathbf{x} \right\}$$

is bounded by a constant independent of $m$. Using the fact that for each $m$, $\varrho^m$ satisfies the energy inequality (1.0.3), we obtain that a subsequence of $\{\varrho^m\}$ converges to a solution of Equation (1.0.1) starting at $\varrho_0$. Moreover, as we will see, our approximation argument also allows to relax the regularity assumptions on the Hamiltonian $H$. This shows a remarkable feature of the existence scheme described before, as it allows us to construct solutions of a highly nonlinear PDE as in (1.0.1) by approximating at the same time the initial datum and the Hamiltonian (and the same strategy could also be applied to relax the assumptions on $U$, cf. Chapter 4). This completes the existence part.

In order to prove uniqueness of the solution in Equation (1.0.1) we make several additional assumptions on $P$ and $H$. First of all, we assume that $L(\mathbf{x}, \mathbf{v}) > L(\mathbf{x}, \mathbf{0})$ for all $\mathbf{x}, \mathbf{v} \in \mathbb{R}^d$ such that $\mathbf{v} \neq \mathbf{0}$ to ensure that the maximum principle (cf. Theorem 3.1.1) holds. Next, let $Q$ be the inverse of $P$ and set $u(t, \cdot) := P(\varrho_t)$. Then Equation (1.0.1) is equivalent to

$$\partial_t Q(u) = \operatorname{div} \mathbf{a}(\mathbf{x}, Q(u), \nabla u) \quad \text{in} \quad \mathcal{D}'((0,T) \times X), \tag{1.0.5}$$

that is a quasilinear elliptic-parabolic equation. Here $\mathbf{a}$ is given by Equation (5.2.2). The study in [18] addresses contraction properties of solutions of Equation (1.0.5) even when $\partial_t Q(u)$ is not a bounded measure but is merely a distribution, as in our case. Our vector field $\mathbf{a}$ does not necessarily satisfy the assumptions in [18]. (Indeed one can check that it violates drastically the strict monotonicity condition of [18], for large $Q(u)$.) For this reason, we only study uniqueness of solutions with bounded initial conditions even if, for this class of solution, $\mathbf{a}$ is still not strictly monotone in the sense of [2] or [18].

The strategy consists first in showing that there exists a Hamiltonian $\bar{H} \equiv \bar{H}(x, \varrho, \mathbf{z})$ (cf. Equation (5.2.3)) such that for each $\mathbf{x}$, $-\mathbf{a}(\mathbf{x}, \varrho, -\mathbf{z})$ is contained in the subdifferential of $\bar{H}(\mathbf{x}, \cdot, \cdot)$ at $(\varrho, \mathbf{z})$. Then, assuming $\bar{H}(\mathbf{x}, \cdot, \cdot)$ convex and $Q$ Lipschitz, we establish a contraction property for bounded solutions of Equation (1.0.1). As a by product we conclude uniqueness of bounded solutions.

The thesis is organized as follows: in Chapter 2 we start with some preliminaries and set up the general framework for our study. The proof of the existence of solutions is then split into two cases. Chapter 3 is concerned with the case where $X$ is bounded, and we prove existence of solutions of Equation (1.0.1) by applying the discrete algorithm described before. In chapter 4 we relax the assumption that $X$ is bounded: under the hypotheses that $\varrho_0 \in \mathcal{P}_1^{ac}(X)$ and $\int_X |U(\varrho_0)| \,\mathrm{d}\mathbf{x}$ is finite, we construct by approximation a solution of Equation (1.0.1) as described above. Chapter 5 is concerned with uniqueness and stability in $L^1$ of bounded solutions of Equation (1.0.1) when $Q$ is Lipschitz. To achieve that goal, we impose the stronger condition (5.2.5) on the Hamiltonian $H$. In Appendix, we prove some supplementary results.

# CHAPTER II

# PRELIMINARIES

## 2.1 Notation, Definitions and Main Assumptions

We fix a convex superlinear function $\theta : [0, +\infty) \to [0, +\infty)$ such that $\theta(0) = 0$. The example we have in mind is a function $\theta$ which behaves like $t^\alpha$ with $\alpha > 1$ (for more general behaviors, like $t(\ln t)^+$ or $e^t$, cf. Remark 3.5.6). We consider a function $L : \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}$ which we call *Lagrangian*. We assume that:

(L1) $L \in C^2(\mathbb{R}^d \times \mathbb{R}^d)$, and $L(\mathbf{x}, \mathbf{0}) = 0$ for all $\mathbf{x} \in \mathbb{R}^d$.

(L2) The matrix $\nabla_{\mathbf{vv}} L(\mathbf{x}, \mathbf{v})$ is strictly positive definite for all $\mathbf{x}, \mathbf{v} \in \mathbb{R}^d$.

(L3) There exist constants $A^*, A_*, C^* > 0$ such that

$$\theta(|\mathbf{v}|) - A_* \leq L(\mathbf{x}, \mathbf{v}) \leq C^* \theta(|\mathbf{v}|) + A^* \qquad \forall\, \mathbf{x}, \mathbf{v} \in \mathbb{R}^d.$$

Remark that the condition $L(\mathbf{x}, \mathbf{0}) = 0$ is not restrictive, as we can always replace $L$ by $L(\mathbf{x}, \mathbf{v}) - L(\mathbf{x}, \mathbf{0})$, and this would not affect the study of the problem we are going to consider. We also note that (L1), (L2) and (L3) ensure that $L$ is a so-called *Tonelli Lagrangian* (cf. for instance [9, Appendix B]). To prove a maximum principle for the solutions of (1.0.1), we will also need the assumption:

(L4) $L(\mathbf{x}, \mathbf{v}) \geq L(\mathbf{x}, \mathbf{0})$ for all $\mathbf{x}, \mathbf{v} \in \mathbb{R}^d$.

The *global Legendre transform* $\mathcal{L} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ of $L$ is defined by

$$\mathcal{L}(\mathbf{x}, \mathbf{v}) := (\mathbf{x}, \nabla_{\mathbf{v}} L(\mathbf{x}, \mathbf{v}))\,.$$

We denote by $\mathbf{\Phi}^L : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ the *Lagrangian flow* defined by

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\big[\nabla_{\mathbf{v}} L\big(\mathbf{\Phi}^L(t,\mathbf{x},\mathbf{v})\big)\big] = \nabla_{\mathbf{x}} L\big(\mathbf{\Phi}^L(t,\mathbf{x},\mathbf{v})\big), \\ \mathbf{\Phi}^L(0,\mathbf{x},\mathbf{v}) = (\mathbf{x},\mathbf{v}). \end{cases} \tag{2.1.1}$$

Furthermore, we denote by $\mathbf{\Phi}_1^L : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ the first component of the flow:

$$\mathbf{\Phi}_1^L := \pi_1 \circ \mathbf{\Phi}^L, \qquad \pi_1(\mathbf{x},\mathbf{v}) := \mathbf{x}.$$

The Legendre transform of $L$, called the *Hamiltonian* of $L$, is defined by

$$H(\mathbf{x},\mathbf{p}) := \sup_{\mathbf{v}\in\mathbb{R}^d} \big\{\langle \mathbf{v},\mathbf{p}\rangle - L(\mathbf{x},\mathbf{v})\big\}.$$

Moreover we define the Legendre transform of $\theta$ by

$$\theta^*(s) := \sup_{t\geq 0}\big\{st - \theta(t)\big\}, \qquad s \in \mathbb{R}.$$

It is well-known that $L$ satisfies (L1), (L2) and (L3) if and only if $H$ satisfies the following conditions:

(H1) $H \in C^2(\mathbb{R}^d \times \mathbb{R}^d)$, and $H(\mathbf{x},\mathbf{p}) \geq 0$ for all $\mathbf{x},\mathbf{p} \in \mathbb{R}^d$.

(H2) The matrix $\nabla_{\mathbf{pp}} H(\mathbf{x},\mathbf{p})$ is strictly positive definite for all $\mathbf{x},\mathbf{p} \in \mathbb{R}^d$.

(H3) $\theta^* : \mathbb{R} \to [0,+\infty)$ is convex, superlinear at $+\infty$, and we have

$$-A^* + C^*\theta^*\left(\frac{|\mathbf{p}|}{C^*}\right) \leq H(\mathbf{x},\mathbf{p}) \leq \theta^*(|\mathbf{p}|) + A_* \qquad \forall\, \mathbf{x},\mathbf{v} \in \mathbb{R}^d.$$

Moreover, (L4) is equivalent to:

(H4) $\nabla_{\mathbf{p}} H(\mathbf{x},\mathbf{0}) = 0$ for all $\mathbf{x} \in \mathbb{R}^d$.

We also introduce some weaker conditions on $L$, which combined with (L3) make it a *weak Tonelli Lagrangian*:

(L1$^w$) $L \in C^1(\mathbb{R}^d \times \mathbb{R}^d)$, and $L(\mathbf{x},\mathbf{0}) = 0$ for all $\mathbf{x} \in \mathbb{R}^d$.

(L2$^w$) For each $\mathbf{x} \in \mathbb{R}^d$, $L(\mathbf{x}, \cdot)$ is strictly convex.

Under (L1$^w$), (L2$^w$) and (L3), the global Legendre transform is an homeomorphism, and the Hamiltonian associated with $L$ satisfies (H3) and

(H1$^w$) $H \in C^1(\mathbb{R}^d \times \mathbb{R}^d)$, and $H(\mathbf{x}, \mathbf{p}) \geq 0$ for all $\mathbf{x}, \mathbf{p} \in \mathbb{R}^d$.

(H2$^w$) For each $\mathbf{x} \in \mathbb{R}^d$, $H(\mathbf{x}, \cdot)$ is strictly convex.

(cf. for instance [9, Appendix B].) In this work, we mainly work assuming (L1), (L2) and (L3), except in Chapter 4 where we relax the assumptions on $L$ ( accordingly on $H$) to (L1$^w$), (L2$^w$) and (L3).

Let $U : [0, +\infty) \to \mathbb{R}$ be a given function such that

$$U \in C^2((0, +\infty)) \cap C([0, +\infty)), \qquad U'' > 0, \tag{2.1.2}$$

and

$$U(0) = 0, \qquad \lim_{t \to \infty} \frac{U(t)}{t} = +\infty. \tag{2.1.3}$$

We set $U(t) = +\infty$ for $t \in (-\infty, 0)$, so that $U$ remains convex and lower-semicontinuous on the whole $\mathbb{R}$. We denote by $U^*$ the Legendre transform of $U$ :

$$U^*(s) := \sup_{t \in \mathbb{R}} \big\{ st - U(t) \big\} = \sup_{t \geq 0} \big\{ st - U(t) \big\}. \tag{2.1.4}$$

When $\varrho$ is a Borel probability density of $\mathbb{R}^d$ such $U^-(\varrho) \in L^1(\mathbb{R}^d)$, we define the *internal energy*

$$\mathcal{U}(\varrho) := \int_{\mathbb{R}^d} U(\varrho(\mathbf{x}))\, \mathrm{d}\mathbf{x}.$$

If $\varrho$ represents the *density* of a fluid, one interpretes $P(\varrho)$ as a *pressure*, where

$$P(t) := tU'(t) - U(t). \tag{2.1.5}$$

Note that $P'(t) = tU''(t)$, so that $P$ is increasing on $[0, +\infty)$.

If $\varrho$ is a probability density and $\alpha > 0$, we write

$$M_\alpha(\varrho) := \int_{\mathbb{R}^d} |\mathbf{x}|^\alpha \varrho(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

for its moment of order $\alpha$. If $X \subset \mathbb{R}^d$ is a Borel set, we denote by $\mathcal{P}^{ac}(X)$ the set of all Borel probability densities on $X$. If $\varrho \in \mathcal{P}^{ac}(X)$, we implicitly identify it with its extension defined to be 0 outside $X$. We denote by $\mathcal{P}(X)$ the set of Borel probability measures $\mu$ on $\mathbb{R}^d$ that are concentrated on $X$: $\mu(X) = 1$. Finally, we denote by $\mathcal{P}^{ac}_\alpha(X) \subset \mathcal{P}^{ac}(X)$ the set of probability densities $\varrho$ on $X$ such that $M_\alpha(\varrho)$ is finite. When $\alpha \geq 1$, this is a metric space when endowed with the Wasserstein distance $W_\alpha$ (cf. Equation (2.1.9) below).

Let $u : X \subset \mathbb{R}^d \to \mathbb{R} \cup \{\pm\infty\}$. The set of points $\mathbf{x}$ such that $u(\mathbf{x}) \in \mathbb{R}$ is called the domain of $u$ and denoted by $\mathrm{dom}u$. We denote by $\partial_- u(\mathbf{x})$ the subdifferential of $u$ at $\mathbf{x}$. Similarly, we denote by $\partial^+ u(\mathbf{x})$ the superdifferential of $u$ at $\mathbf{x}$. The set of points where $u$ is differentiable is called the domain of $\nabla u$ and is denoted by $\mathrm{dom}\nabla u$.

Let $u : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$. Its Legendre transform is $u^* : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ defined by

$$u^*(\mathbf{y}) = \sup_{\mathbf{x} \in X} \big\{ \langle \mathbf{x}, \mathbf{y} \rangle - u(\mathbf{x}) \big\}.$$

In case $u : X \subset \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, its Legendre transform is defined by identifying $u$ with its extension which takes the value $+\infty$ outside $X$.

For $f : (a, b) \to \mathbb{R}$, we set

$$\frac{d^+ f}{\mathrm{d}t}(t) := \limsup_{h \downarrow 0^+} \frac{f(t+h) - f(t)}{h}.$$

For $h > 0$, we define the *action* $\mathcal{A}_h(\sigma)$ of an absolutely continuous curve $\sigma : [0, h] \to \mathbb{R}^d$ by

$$\mathcal{A}_h(\sigma) := \int_0^h L(\sigma(\tau), \dot{\sigma}(\tau))\,\mathrm{d}\tau$$

and the *cost function*

$$c_h(\mathbf{x}, \mathbf{y}) := \inf_\sigma \Big\{ \mathcal{A}_h(\sigma) \;:\; \sigma \in W^{1,1}(0, h; \mathbb{R}^d),\; \sigma(0) = \mathbf{x},\; \sigma(h) = \mathbf{y} \Big\}. \tag{2.1.6}$$

**Definition 2.1.1** *A Borel map* $T : \mathbb{R}^d \to \mathbb{R}^d$ *pushes* $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ *forward to* $\mu_1 \in \mathcal{P}(\mathbb{R}^d)$ *if* $\mu_1(B) = \mu_0(T^{-1}(B))$ *holds for any Borel set* $B \subset \mathbb{R}^d$*. In integral form this is equivalent to*

$$\int_{\mathbb{R}^d} f(\mathbf{y})\, \mathrm{d}\mu_1(\mathbf{y}) = \int_{\mathbb{R}^d} f(T(\mathbf{x}))\, \mathrm{d}\mu_0(\mathbf{x})$$

*for all* $f \in L^1(\mathbb{R}^d, \mu_1)$*. In short, we write* $T_{\#}\mu_0 = \mu_1$*.*

**Definition 2.1.2** *For* $\mu_0, \mu_1 \in \mathcal{P}(\mathbb{R}^d)$*, a Borel probability measure* $\gamma$ *on* $\mathbb{R}^d \times \mathbb{R}^d$ *is said to have* $\mu_0$ *and* $\mu_1$ *as its marginals if for any Borel set* $B \subset \mathbb{R}^d$*,*

$$\gamma(B \times \mathbb{R}^d) = \mu_0(B), \qquad \gamma(\mathbb{R}^d \times B) = \mu_1(B).$$

*Equivalently,*

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} (f(\mathbf{x}) + g(\mathbf{y}))\, \mathrm{d}\gamma(\mathbf{x}, \mathbf{y}) = \int_{\mathbb{R}^d} f(\mathbf{x})\, \mathrm{d}\mu_0(\mathbf{x}) + \int_{\mathbb{R}^d} g(\mathbf{y})\, \mathrm{d}\mu_1(\mathbf{y})$$

*holds for* $f \in L^1(\mathbb{R}^d, \mu_0)$ *and* $g \in L^1(\mathbb{R}^d, \mu_1)$*.* $\Gamma(\mu_0, \mu_1)$ *denotes the set of joint probability measures on* $\mathbb{R}^d \times \mathbb{R}^d$ *which have* $\mu_0$ *and* $\mu_1$ *as marginals. If* $\mu_0$ *and* $\mu_1$ *are absolutely continuous with respect to the Lebesgue measure* $\mathcal{L}^d$ *on* $\mathbb{R}^d$*, we write* $\Gamma(\varrho_0, \varrho_1)$ *in place of* $\Gamma(\mu_0, \mu_1)$*, where* $\varrho_0$ *and* $\varrho_1$ *are the density functions of* $\mu_0$ *and* $\mu_1$ *respectively.*

Set

$$\mathcal{C}_h(\mu_0, \mu_1) := \inf_{\gamma} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} c_h(\mathbf{x}, \mathbf{y})\, \mathrm{d}\gamma(\mathbf{x}, \mathbf{y}) \;:\; \gamma \in \Gamma(\mu_0, \mu_1) \right\} \tag{2.1.7}$$

and

$$W_{\theta,h}(\mu_0, \mu_1) := \inf_{\gamma} \left\{ h \int_{\mathbb{R}^d \times \mathbb{R}^d} \theta\left( \frac{|\mathbf{y} - \mathbf{x}|}{h} \right) \mathrm{d}\gamma(\mathbf{x}, \mathbf{y}) \;:\; \gamma \in \Gamma(\mu_0, \mu_1) \right\}. \tag{2.1.8}$$

We also recall the definition of the $\alpha$-Wasserstein distance, $\alpha \geq 1$:

$$W_\alpha(\mu_0, \mu_1) := \inf_{\gamma} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |\mathbf{y} - \mathbf{x}|^\alpha\, \mathrm{d}\gamma(\mathbf{x}, \mathbf{y}) \;:\; \gamma \in \Gamma(\mu_0, \mu_1) \right\}^{1/\alpha}. \tag{2.1.9}$$

It is well-known (cf. for instance [3]) that $W_\alpha$ metrizes the weak$^*$ topology of measures on bounded sets.

**Definition 2.1.3 ($c$-transform)** *Let $X \subset \mathbb{R}^d$ and let $u, v : X \to \mathbb{R} \cup \{-\infty\}$. The* first $c$-transform *of $u$, $u^c : X \to \mathbb{R} \cup \{-\infty\}$, and the* second $c$-transform *of $v$, $v_c : X \to \mathbb{R} \cup \{-\infty\}$, are defined by*

$$u^c(\mathbf{y}) := \inf_{\mathbf{x} \in X} \{c(\mathbf{x}, \mathbf{y}) - u(\mathbf{x})\}, \qquad v_c(\mathbf{x}) := \inf_{\mathbf{y} \in X} \{c(\mathbf{x}, \mathbf{y}) - v(\mathbf{y})\}. \tag{2.1.10}$$

**Remark 2.1.4** *We directly obtain from the definitions that $u \leq v_c$ and $v \leq u^c$.*

**Definition 2.1.5 ($c$-convexity)** *We say that $u : X \to \mathbb{R} \cup \{-\infty\}$ is* first $c$-concave *if there exists $v : X \to \mathbb{R} \cup \{-\infty\}$ such that $u = v_c$. Similarly, $v : X \to \mathbb{R} \cup \{-\infty\}$ is* second $c$-concave *if there exists $u : X \to \mathbb{R} \cup \{-\infty\}$ such that $v = u^c$.*

For simplicity we will omit the words "first" and "second" when referring to $c$-transform and $c$-concavity. Let's recall some well-known results:

**Lemma 2.1.6** *The functions $u^c$ and $v_c$ satisfy the following properties:*

$$(a)\ \ (v_c)^c \geq v, \qquad (b)\ \ (u^c)_c \geq u, \qquad (c)\ \ ((v_c)^c)_c = v_c, \qquad (d)\ \ ((u^c)_c)^c = u^c.$$

Proof This can be found in [15, 21, 22]. For completeness we'll sketch the proofs. Part $(a)$ follows from the following observation:

$$\begin{aligned}
(v_c)^c(\mathbf{y}) &= \inf_{\mathbf{x} \in X} \{c(\mathbf{x}, \mathbf{y}) - v_c(\mathbf{x})\} \\
&= \inf_{\mathbf{x} \in X} \{c(\mathbf{x}, \mathbf{y}) - \inf_{\mathbf{z} \in \mathbf{X}} \{c(\mathbf{x}, \mathbf{z}) - v(\mathbf{z})\}\} \\
&\geq \inf_{\mathbf{x} \in X} \{c(\mathbf{x}, \mathbf{y}) - (c(\mathbf{x}, \mathbf{y}) - v(\mathbf{y}))\} \\
&= \inf_{\mathbf{x} \in X} \{v(\mathbf{y})\} \\
&= v(\mathbf{y})
\end{aligned}$$

which proves (a). The proof of part $(b)$ is likewise. Let's prove part (c). Substituting $u = v_c$ in part (b), we get $((v_c)^c)_c \geq v_c$. Therefore proving part (c) amounts to showing the reverse inequality: $v_c \geq ((v_c)^c)_c$. By $(a)$, $(v_c)^c \geq v$, we observe that

$$\begin{aligned} ((v_c)^c)_c(\mathbf{x}) &= \inf_{\mathbf{y}\in X}\{c(\mathbf{x},\mathbf{y}) - (v_c)^c(\mathbf{y})\} \\ &\leq \inf_{\mathbf{y}\in X}\{c(\mathbf{x},\mathbf{y}) - v(\mathbf{y})\} \\ &= v_c(\mathbf{x}), \end{aligned}$$

concluding the proof of part (c). Similarly, we can prove part $(d)$. □

## *2.2 Properties of Enthalpy and Pressure Functionals*

In this section, we assume that (2.1.2) and (2.1.3) hold. The following lemma is immediate but important.

**Lemma 2.2.1** *The following properties hold:*

*(i)* $U' : [0, +\infty) \to \mathbb{R}$ *is strictly increasing, and so invertible. Its inverse is of class* $C^1$ *and* $\lim_{t\to+\infty} U'(t) = +\infty$.

*(ii)* $U^* \in C^1(\mathbb{R})$ *is nonnegative, and* $(U^*)'(s) \geq 0$ *for all* $s \in \mathbb{R}$.

*(iii)* $\lim_{s\to+\infty} (U^*)'(s) = +\infty$.

*(iv)* $U^*$ *is superlinear, i.e.,* $\lim_{s\to+\infty} \frac{U^*(s)}{s} = +\infty$.

*(v)* $P : [0, +\infty) \to [0, +\infty)$ *is strictly increasing, bijective,* $\lim_{t\to+\infty} P(t) = +\infty$*, and its inverse* $Q : [0, +\infty) \to [0, +\infty)$ *satisfies* $\lim_{s\to+\infty} Q(s) = +\infty$.

Proof (i) Since $U$ is convex and $U(0) = 0$, we have $U'(t) \geq \frac{U(t)}{t}$. This and $U'' > 0$ easily imply the result.

(ii) $U^* \geq 0$ follows from $U(0) = 0$. The remaining part is a consequence of $(U^*)'(U'(t)) = t$ for $t > 0$, together with $U^*(s) = 0$ (and so $(U^*)'(s) = 0$) for $s \leq U'(0^+)$.

(iii) follows from (i) and the identity $(U^*)'(U'(t)) = t$ for $t > 0$. Also, this property can be obtained from (iv). Let $s > 0$. Since $U^*$ is convex and $C^1(\mathbb{R})$, we have

$U^*(0) \geq U^*(s) + (0-s)(U^*)'(s)$ which yields

$$(U^*)'(s) \geq \frac{U^*(s)}{s} - \frac{U^*(0)}{s}.$$

Thus, the required result follows from (iv).

(iv) Since $U^*$ is convex and nonnegative we have

$$U^*(s) \geq U^*\left(\frac{s}{2}\right) + \left(s - \frac{s}{2}\right)(U^*)'\left(\frac{s}{2}\right) \geq \frac{s}{2}(U^*)'\left(\frac{s}{2}\right).$$

Therefore, the result follows from (iii). Here is another proof of (iv) that directly follows from the definition of $U^*$: Let $a > 0$ be a real number and fix $s \in \mathbb{R}$. Then, taking $t = a$ we have

$$\frac{U^*(s)}{s} = \frac{1}{s}\sup_{t\in\mathbb{R}}\{ts - U(t)\} \geq a - \frac{U(a)}{s} \geq a - \frac{\max\{0, U(a)\}}{s}$$

which implies that

$$\lim_{s\to+\infty} \frac{U^*(s)}{s} \geq a$$

holds for any $a > 0$. Therefore (iv) follows.

(v) Observe that $P(t) = U^*(U'(t)) \geq 0$ by (ii). Since $U'$ is monotone nondecreasing then for $t < 1$ we have $P(t) \leq tU'(1) - U(t)$. We conclude that $\lim_{t\to0^+} P(t) = 0$. The remaining statements then follow. □

**Remark 2.2.2** *Let $X \subset \mathbb{R}^d$ be a bounded set, and let $\varrho \in \mathcal{P}^{ac}(X)$ be a probability density. Recall that we extend $\varrho$ outside $X$ by setting its value to be 0 there. If $R > 0$ is such that $X \subset B_R(0)$, we have $\int_{\mathbb{R}^d} \theta(|\mathbf{x}|)\varrho(\mathbf{x})\,\mathrm{d}\mathbf{x} \leq \theta(R)$. Moreover, since by convexity $U(t) \geq U(1) + U'(1)(t-1) \equiv at + b$ for $t \geq 0$, $\int_{\mathbb{R}^d} U^-(\varrho)\,\mathrm{d}\mathbf{x}$ is bounded on $\mathcal{P}^{ac}(X)$ by $|a| + |b|\mathcal{L}^d(X)$. Hence, $\int_{\mathbb{R}^d} U(\varrho)\,\mathrm{d}\mathbf{x}$ is always well-defined on $\mathcal{P}^{ac}(X)$, and is finite if and only if $U^+(\varrho) \in L^1(X)$.*

The following lemma is a standard result of the calculus of variations, cf. for instance [6] (for a more general result on unbounded domains, cf. Chapter 4):

**Lemma 2.2.3** *Let $X \subset \mathbb{R}^d$ and suppose $\{\varrho^n\}_{n\in\mathbb{N}} \subset \mathcal{P}^{ac}(X)$ converges weakly to $\varrho$ in $L^1(X)$. If either $X$ is bounded, or $X$ is unbounded and $U \geq 0$, then $\mathcal{U}$ is weakly lower semicontinuous, i.e.,*

$$\liminf_{n\to\infty} \mathcal{U}(\varrho^n) \geq \mathcal{U}(\varrho).$$

## *2.3 Properties of $H$ and the Cost Functions*

**Lemma 2.3.1** *The following properties hold:*

*(i)* $c_h(\mathbf{x},\mathbf{x}) \leq 0$ *for all* $h > 0$, $\mathbf{x} \in \mathbb{R}^d$.

*(ii) For all* $h > 0$, $\mathbf{x},\mathbf{y} \in \mathbb{R}^d$,

$$C^* h\,\theta\left(\frac{|\mathbf{x}-\mathbf{y}|}{h}\right) + A^* h \geq c_h(\mathbf{x},\mathbf{y}) \geq h\,\theta\left(\frac{|\mathbf{x}-\mathbf{y}|}{h}\right) - A_* h \geq -A_* h.$$

Proof (i) Set $\sigma(t) \equiv \mathbf{x}$ for $t \in [0,h]$ and recall that $L(\mathbf{x},\mathbf{0}) = 0$ to get

$$c_h(\mathbf{x},\mathbf{x}) \leq \mathcal{A}_h(\sigma) = 0.$$

(ii) The first inequality is obtained using (L3), and $c_h(\mathbf{x},\mathbf{y}) \leq \mathcal{A}_T(\sigma)$ with

$$\sigma(t) = \left(1 - \frac{t}{h}\right)\mathbf{x} + \frac{t}{h}\mathbf{y}$$

while the second one follows from Jensen's inequality. □

The following proposition is classical (cf. for instance [9, Appendix B]):

**Proposition 2.3.2** *Under the assumptions (L1), (L2) and (L3),* (2.1.6) *admits a minimizer $\sigma_{\mathbf{x},\mathbf{y}}$ for any $\mathbf{x},\mathbf{y} \in \mathbb{R}^d$. We have that $\sigma_{\mathbf{x},\mathbf{y}}$ is of class $C^2([0,h])$ and satisfies the Euler-Lagrange equation*

$$(\sigma_{\mathbf{x},\mathbf{y}}(\tau), \dot\sigma_{\mathbf{x},\mathbf{y}}(\tau)) = \mathbf{\Phi}^L(\tau, \mathbf{x}, \dot\sigma(0)) \qquad \forall \tau \in [0,h], \tag{2.3.1}$$

*where $\mathbf{\Phi}^L$ is the Lagrangian flow defined in Equation* (2.1.1)*. Moreover, for any $h, r > 0$, there exists a constant $k_h(r)$, depending on $h$ and $r$ only, such that $||\sigma_{\mathbf{x},\mathbf{y}}||_{C^2([0,h])} \leq k_h(r)$ if $|\mathbf{x}|, |\mathbf{y}| \leq r$.*

**Remark 2.3.3** *Let $\sigma$ be a minimizer of the problem* (2.1.6), *and set*

$$\mathbf{p}(\tau) := \nabla_{\mathbf{v}} L\big(\sigma(\tau), \dot{\sigma}(\tau)\big).$$

*(a) The Euler-Lagrange Equation* (2.3.1) *implies that $\sigma$ and $\mathbf{p}$ are of class $C^1$ and satisfy the ordinary differential equation*

$$\begin{cases} \dot{\sigma}(\tau) = \nabla_{\mathbf{p}} H(\sigma(\tau), \mathbf{p}(\tau)), \\ \dot{\mathbf{p}}(\tau) = -\nabla_{\mathbf{x}} H(\sigma(\tau), \mathbf{p}(\tau)). \end{cases} \tag{2.3.2}$$

*(b) The Hamiltonian is constant along the integral curve $(\sigma(\tau), \mathbf{p}(\tau))$, i.e.*

$$H(\sigma(\tau), \mathbf{p}(\tau)) = H(\sigma(0), \mathbf{p}(0))$$

*for $\tau \in [0, h]$.*

**Lemma 2.3.4** *Under the assumptions in Proposition 2.3.2, let $\sigma$ be a minimizer of (2.1.6), and define $\mathbf{p}_i := \nabla_{\mathbf{v}} L(\sigma(i), \dot{\sigma}(i))$ for $i = 0, h$. For $r, m > 0$ there exists a constant $\ell_h(r, m)$, depending only on $h, r, m$, such that if $\mathbf{x}, \mathbf{y} \in B_r(\mathbf{0})$ and $\mathbf{w} \in B_m(\mathbf{0})$,*

*(a) $c_h(\mathbf{x} + \mathbf{w}, \mathbf{y}) \leq c_h(\mathbf{x}, \mathbf{y}) - \langle \mathbf{p}_0, \mathbf{w} \rangle + \frac{1}{2} \ell_h(r, m) |\mathbf{w}|^2$;*

*(b) $c_h(\mathbf{x}, \mathbf{y} + \mathbf{w}) \leq c_h(\mathbf{x}, \mathbf{y}) + \langle \mathbf{p}_h, \mathbf{w} \rangle + \frac{1}{2} \ell_h(r, m) |\mathbf{w}|^2$.*

Proof In order to prove (a), we set

$$\sigma_{\mathbf{w}}(t) = \sigma(t) + \frac{h - t}{h} \mathbf{w}$$

so that

$$\sigma_{\mathbf{w}}(0) = \mathbf{x} + \mathbf{w}, \qquad \sigma_{\mathbf{w}}(h) = \mathbf{y}.$$

Set

$$\alpha(a, t) := L\Big(\sigma(t) + a\frac{h - t}{h} \mathbf{w}, \dot{\sigma}(t) - \frac{a}{h} \mathbf{w}\Big).$$

Employing $\alpha$ in the identity,

$$\alpha(1, t) = \alpha(0, t) + \partial_a \alpha(0, t) + \int_0^1 \left( \int_0^s \partial_a^2 \alpha(a, t) \mathrm{d}a \right) \mathrm{d}s \tag{2.3.3}$$

We obtain

$$
\begin{aligned}
c_h(\mathbf{x}+\mathbf{w},\mathbf{y}) &\leq \int_0^h L(\sigma_{\mathbf{w}},\dot{\sigma}_{\mathbf{w}})\,\mathrm{d}t \\
&= \int_0^h L(\sigma,\dot{\sigma})\,\mathrm{d}t + \int_0^h \Big(\langle \nabla_{\mathbf{x}} L(\sigma,\dot{\sigma}),\mathbf{w}\rangle \frac{h-t}{h} - \langle \nabla_{\mathbf{v}} L(\sigma,\dot{\sigma}),\mathbf{w}\rangle \frac{1}{h}\Big)\,\mathrm{d}t \\
&+ \int_0^h \mathrm{d}t \int_0^1 \mathrm{d}s \int_0^s \partial_a^2 \alpha(a,t)\,\mathrm{d}a, \qquad (2.3.4)
\end{aligned}
$$

One readily checks that

$$
|\partial_a^2 \alpha(a,t)| \leq \frac{1+(h-t)^2}{h^2} \ell_h(r,m)|\mathbf{w}|^2,
$$

where

$$
\ell_h(r,m) := \sup_{\mathbf{x},\mathbf{v},\mathbf{z}} \left\{ |\langle \nabla^2 L(\mathbf{x},\mathbf{v})\,\mathbf{z},\mathbf{z}\rangle| \ :\ \mathbf{x},\mathbf{v} \in B_R(\mathbf{0}) \subset \mathbb{R}^d, \mathbf{z} \in B_1(\mathbf{0}) \subset \mathbb{R}^{2d} \right\}
$$

and $R := k_h(r+m)$ and $k_h$ is as in Proposition 2.3.2. Thus,

$$
\int_0^h \mathrm{d}t \int_0^1 \mathrm{d}s \int_0^s \partial_a^2 \alpha(a,t)\,\mathrm{d}a \leq \frac{1}{2}\Big(\frac{1}{h}+\frac{h}{3}\Big)\ell(r,m)|\mathbf{w}|^2. \qquad (2.3.5)
$$

The Euler-Lagrange Equation (2.3.1) together with the Fundamental Theorem of Calculus gives that the second expression in Equation (2.3.4) is $-\langle \mathbf{p}_0,\mathbf{w}\rangle$. Indeed,

$$
\int_0^h \Big(\langle \nabla_{\mathbf{x}} L(\sigma,\dot{\sigma}),\mathbf{w}\rangle \frac{h-t}{h} - \langle \nabla_{\mathbf{v}} L(\sigma,\dot{\sigma}),\mathbf{w}\rangle \frac{1}{h}\Big)\,\mathrm{d}t = \int_0^h \frac{\mathrm{d}}{\mathrm{d}t}\Big[\langle \nabla_{\mathbf{v}} L(\sigma,\dot{\sigma}), \frac{h-t}{h}\mathbf{w}\rangle\Big]\,\mathrm{d}t
$$

$$
= \Big[\langle \nabla_{\mathbf{v}} L(\sigma(t),\dot{\sigma}(t)), \frac{h-t}{h}\mathbf{w}\rangle\Big]_{t=0}^{t=h} = \langle -\nabla_{\mathbf{v}} L(\sigma(0),\dot{\sigma}(0)),\mathbf{w}\rangle = -\langle \mathbf{p}_0,\mathbf{w}\rangle,
$$

which together with equations (2.3.4) and (2.3.5) yields

$$
c_h(\mathbf{x}+\mathbf{w},\mathbf{y}) \leq c_h(\mathbf{x},\mathbf{y}) - \langle \mathbf{p}_0,\mathbf{w}\rangle + \frac{1}{2}\Big(\frac{1}{h}+\frac{h}{3}\Big)\ell(r,m)|\mathbf{w}|^2. \qquad (2.3.6)
$$

This proves (a). The proof of (b) is analogous. □

**Remark 2.3.5** *This lemma says that* $-\mathbf{p}_0 \in \partial^+ c_h(\cdot,\mathbf{y})(\mathbf{x})$, *and for* $\mathbf{y} \in B_r(\mathbf{0})$ *the restriction of* $c(\cdot,\mathbf{y})$ *to* $B_r(\mathbf{0})$ *is* $\ell_h(r,m)$*-concave. Similarly,* $\mathbf{p}_h \in \partial^+ c_h(\mathbf{x},\cdot)(\mathbf{y})$, *and for* $\mathbf{x} \in B_r(\mathbf{0})$ *the restriction of* $c(\mathbf{x},\cdot)$ *to* $B_r(\mathbf{0})$ *is* $\ell_h(r,m)$*-concave.*

## *2.4 Total Works and Their Properties*

In this section, we assume that (2.1.2) and (2.1.3) hold.

**Remark 2.4.1** *By Remark 2.3.5 $c_h$ is continuous. In particular, there always exists a minimizer for* (2.1.7) *(cf. for example, Theorem 2.4. in [13]). It is trivial if $\mathcal{C}_h$ is identically $+\infty$ on $\Gamma(\varrho_0, \varrho_h)$. We denote the set of minimizers by $\Gamma_h(\varrho_0, \varrho_h)$. Similarly, there is a minimizer for* (2.1.8)*, and we denote the set of its minimizers by $\Gamma_h^\theta(\varrho_0, \varrho_h)$.*

**Lemma 2.4.2** *Suppose $0 < s < t < +\infty$. Then the cost function satisfies the following inequality:*

$$c_t(\mathbf{x}, \mathbf{z}) \leq c_s(\mathbf{x}, \mathbf{y}) + c_{t-s}(\mathbf{y}, \mathbf{z}).$$

Proof Let's suppose $\sigma_1$ and $\sigma_2$ are the minimizers for $c_s(\mathbf{x}, \mathbf{y})$ and $c_{t-s}(\mathbf{y}, \mathbf{z})$, respectively. That is to say,

$$c_s(\mathbf{x}, \mathbf{y}) = \int_0^s L(\sigma_1(\tau), \dot{\sigma}_1(\tau))\, \mathrm{d}\tau, \qquad \sigma_1(0) = \mathbf{x}, \quad \sigma_1(s) = \mathbf{y}$$

and

$$c_{t-s}(\mathbf{y}, \mathbf{z}) = \int_0^{t-s} L(\sigma_2(\tau), \dot{\sigma}_2(\tau))\, \mathrm{d}\tau, \qquad \sigma_2(0) = \mathbf{y}, \quad \sigma_2(t-s) = \mathbf{z}.$$

By a change of variable $\delta = \tau + s$ and then using substitution $\tilde{\sigma}_2(\delta) = \sigma_2(\delta - s)$ together with $\tilde{\sigma}_2(s) = \sigma_2(0) = \mathbf{y}$ and $\tilde{\sigma}_2(t) = \sigma_2(t-s) = \mathbf{z}$, we can rewrite $c_{t-s}$ in terms of $\tilde{\sigma}_2$ as

$$c_{t-s}(\mathbf{y}, \mathbf{z}) = \int_s^t L(\tilde{\sigma}_2(\delta), \dot{\tilde{\sigma}}_2(\delta))\, \mathrm{d}\delta, \qquad \tilde{\sigma}_2(s) = \mathbf{y}, \quad \tilde{\sigma}_2(t) = \mathbf{z}.$$

Thus, observe that

$$\begin{aligned}
c_s(\mathbf{x}, \mathbf{y}) + c_{t-s}(\mathbf{y}, \mathbf{z}) &= \int_0^s L(\sigma_1(\tau), \dot{\sigma}_1(\tau))\, \mathrm{d}\tau + \int_s^t L(\tilde{\sigma}_2(\tau), \dot{\tilde{\sigma}}_2(\tau))\, \mathrm{d}\tau \\
&\geq \inf_{\sigma \in \Lambda} \left\{ \int_0^s L(\sigma, \dot{\sigma})\, \mathrm{d}\tau + \int_s^t L(\sigma, \dot{\sigma})\, \mathrm{d}\tau \right\} \\
&\geq \inf_{\sigma} \left\{ \int_0^t L(\sigma, \dot{\sigma})\, \mathrm{d}\tau : \ \sigma(0) = \mathbf{x},\ \sigma(t) = \mathbf{z} \right\} \\
&= c_t(\mathbf{x}, \mathbf{z})
\end{aligned}$$

where $\Lambda = \{\sigma \in W^{1,1}[0,t] \ : \ \sigma(0) = \mathbf{x}, \quad \sigma(s) = \mathbf{y}, \quad \sigma(t) = \mathbf{z}\}$. □

Here is an immediate corollary:

**Corollary 2.4.3** *Let $\varrho_1, \varrho_2, \varrho_3 \in \mathcal{P}^{ac}(\mathbb{R}^d)$. Then,*

$$\mathcal{C}_h(\varrho_1, \varrho_3) \le \mathcal{C}_t(\varrho_1, \varrho_2) + \mathcal{C}_{h-t}(\varrho_2, \varrho_3). \tag{2.4.1}$$

*holds for all $t \in [0, h]$.*

Proof Let $T_t^{1,2}{}_{\#}\varrho_1 = \varrho_2$ and $T_{h-t\#}^{2,3}\varrho_2 = \varrho_3$ (cf. [9],[10]) such that

$$\mathcal{C}_t(\varrho_1, \varrho_2) = \int_{\mathbb{R}^d \times \mathbb{R}^d} c_t(\mathbf{x}, T_t^{1,2}(\mathbf{x}))\varrho_1(\mathbf{x})\,\mathrm{d}\mathbf{x},$$

$$\mathcal{C}_{h-t}(\varrho_2, \varrho_3) = \int_{\mathbb{R}^d \times \mathbb{R}^d} c_{h-t}(\mathbf{y}, T_{h-t}^{2,3}(\mathbf{y}))\varrho_2(\mathbf{y})\mathrm{d}\mathbf{y}.$$

Define $T_h^{1,3} = T_{h-t}^{2,3} \circ T_t^{1,2}$. Clearly, $T_h^{1,3}{}_{\#}\varrho_1 = \varrho_3$. By using Lemma 2.4.2

$$\begin{aligned}
\mathcal{C}_h(\varrho_1, \varrho_3) &\le \int_{\mathbb{R}^d \times \mathbb{R}^d} c_h(\mathbf{x}, T_h^{1,3}(\mathbf{x}))\varrho_1(\mathbf{x})\,\mathrm{d}\mathbf{x} \\
&\le \int_{\mathbb{R}^d \times \mathbb{R}^d} (c_t(\mathbf{x}, T_t^{1,2}(\mathbf{x})) + c_{h-t}(T_t^{1,2}(\mathbf{x}), T_h^{1,3}(\mathbf{x})))\varrho_1(\mathbf{x})\,\mathrm{d}\mathbf{x} \\
&= \int_{\mathbb{R}^d} (c_t(\mathbf{x}, T_t^{1,2}(\mathbf{x})))\varrho_1(\mathbf{x})\,\mathrm{d}\mathbf{x} + \int_{\mathbb{R}^d} c_{h-t}(\mathbf{y}, T_{h-t}^{2,3}(\mathbf{y})))\varrho_2(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&= \mathcal{C}_t(\varrho_1, \varrho_2) + \mathcal{C}_{h-t}(\varrho_2, \varrho_3)
\end{aligned}$$

which concludes (2.4.1). □

As a result, we are in position to state following very useful upshot:

**Lemma 2.4.4** *The following properties hold:*

*(i) For any $\mu \in \mathcal{P}(\mathbb{R}^d)$ we have $\mathcal{C}_h(\mu, \mu) \le 0$. In particular, for any $\mu, \bar{\mu} \in \mathcal{P}(\mathbb{R}^d)$, and $h < \bar{h}$*

$$\mathcal{C}_{\bar{h}}(\mu, \bar{\mu}) \le \mathcal{C}_h(\mu, \bar{\mu}).$$

*(ii) For any $h > 0$, $\mu, \bar{\mu} \in \mathcal{P}(\mathbb{R}^d)$,*

$$-A_* h \le -A_* h + W_{\theta,h}(\mu, \bar{\mu}) \le \mathcal{C}_h(\mu, \bar{\mu}) \le C^* W_{\theta,h}(\mu, \bar{\mu}) + A^* h.$$

*(iii) For any $K > 0$ there exists a constant $C(K) > 0$ such that*

$$W_1(\mu,\bar\mu) \le \frac{1}{K} W_{\theta,h}(\mu,\bar\mu) + \frac{C(K)}{K} h \qquad \forall\, h > 0,\ \mu,\bar\mu \in \mathcal{P}(\mathbb{R}^d). \tag{2.4.2}$$

Proof (i) The first part follows from $c_h(\mathbf{x},\mathbf{x}) \le 0$, while the second statement is a consequence of the first one and $\mathcal{C}_{\bar h}(\mu,\bar\mu) \le \mathcal{C}_h(\mu,\bar\mu) + \mathcal{C}_{\bar h - h}(\bar\mu,\bar\mu)$ (see Corollary 2.4.3.)
(ii) It follows directly from Lemma 2.3.1(iii).
(iii) Thanks to the superlinearity of $h$, for any $K > 0$ there exists a constant $C(K) > 0$ such that

$$\theta(s) \ge Ks - C(K) \qquad \forall\, s \ge 0. \tag{2.4.3}$$

Let now $\gamma \in \Gamma_h^\theta(\mu_0,\mu_1)$. Then

$$\begin{aligned}
W_1(\mu,\bar\mu) &\le \int_{\mathbb{R}^d\times\mathbb{R}^d} |\mathbf{x}-\mathbf{y}|\,\mathrm{d}\gamma(\mathbf{x},\mathbf{y})\\
&\le \frac{h}{K}\int_{\mathbb{R}^d\times\mathbb{R}^d} \left(K\frac{|\mathbf{x}-\mathbf{y}|}{h} - C(K)\right)\mathrm{d}\gamma(\mathbf{x},\mathbf{y}) + \frac{C(K)}{K}h\\
&\le \frac{1}{K}\int_{\mathbb{R}^d\times\mathbb{R}^d} \theta\left(\frac{|\mathbf{x}-\mathbf{y}|}{h}\right)\mathrm{d}\gamma(\mathbf{x},\mathbf{y}) + \frac{C(K)}{K}h\\
&= \frac{1}{K}W_{\theta,h}(\mu,\bar\mu) + \frac{C(K)}{K}h
\end{aligned}$$

which is exactly (2.4.2). □

**Remark 2.4.5** *The map $\varrho \mapsto \mathcal{C}_h[\varrho_0,\varrho]$ is convex on $\mathcal{P}^{ac}(\mathbb{R}^d)$.*

Proof Let $\varrho_1,\varrho_2 \in \mathcal{P}^{ac}(\mathbb{R}^d)$. By Remark 2.4.1, there exist $\gamma_1 \in \Gamma_h(\varrho_0,\varrho_1)$ and $\gamma_2 \in \Gamma_h(\varrho_0,\varrho_2)$ such that

$$\mathcal{C}_h[\varrho_0,\varrho_1] = \int_{\mathbb{R}^d\times\mathbb{R}^d} c_h(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma_1(x,y), \qquad \mathcal{C}_h[\varrho_0,\varrho_2] = \int_{\mathbb{R}^d\times\mathbb{R}^d} c_h(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma_2(x,y).$$

Set $\varrho_\lambda = (1-\lambda)\varrho_1 + \lambda\varrho_2$ for $\lambda \in [0,1]$. Define $\gamma_\lambda = (1-\lambda)\gamma_1 + \lambda\gamma_2$. Observe that for $g \in L^1(\mathbb{R}^d, \varrho_0\mathcal{L}^d) \cap L^1(\mathbb{R}^d, \varrho_1\mathcal{L}^d) \cap L^1(\mathbb{R}^d, \varrho_2\mathcal{L}^d)$,

$$\begin{aligned}
\int_{\mathbb{R}^d\times\mathbb{R}^d} g(\mathbf{x})\mathrm{d}\gamma_\lambda(\mathbf{x},\mathbf{y}) &= (1-\lambda)\int_{\mathbb{R}^d\times\mathbb{R}^d} g(\mathbf{x})\,\mathrm{d}\gamma_1(\mathbf{x},\mathbf{y}) + \lambda\int_{\mathbb{R}^d\times\mathbb{R}^d} g(\mathbf{x})\,\mathrm{d}\gamma_2(\mathbf{x},\mathbf{y})\\
&= (1-\lambda)\int_{\mathbb{R}^d} g(\mathbf{x})\varrho_0(\mathbf{x})\,\mathrm{d}\mathbf{x} + \lambda\int_{\mathbb{R}^d} g(\mathbf{x})\varrho_0(\mathbf{x})\,\mathrm{d}\mathbf{x}\\
&= \int_{\mathbb{R}^d} g(\mathbf{x})\varrho_0(\mathbf{x})\,\mathrm{d}\mathbf{x}
\end{aligned}$$

$$
\begin{aligned}
\int_{\mathbb{R}^d\times\mathbb{R}^d} g(\mathbf{y})\mathrm{d}\gamma_\lambda(\mathbf{x},\mathbf{y}) &= (1-\lambda)\int_{\mathbb{R}^d\times\mathbb{R}^d} g(\mathbf{y})\,\mathrm{d}\gamma_1(\mathbf{x},\mathbf{y}) + \lambda\int_{\mathbb{R}^d\times\mathbb{R}^d} g(\mathbf{y})\,\mathrm{d}\gamma_2(\mathbf{x},\mathbf{y}) \\
&= (1-\lambda)\int_{\mathbb{R}^d} g(\mathbf{y})\varrho_1(\mathbf{y})\,\mathrm{d}\mathbf{y} + \lambda\int_{\mathbb{R}^d} g(\mathbf{y})\varrho_2(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&= \int_{\mathbb{R}^d} g(\mathbf{y})\varrho_\lambda(\mathbf{y})\,\mathrm{d}\mathbf{y}
\end{aligned}
$$

So $\gamma_\lambda \in \Gamma(\varrho_\lambda, \varrho_0)$. Therefore, since $\gamma_1 \in \Gamma_h(\varrho_0, \varrho_1)$ and $\gamma_2 \in \Gamma_h(\varrho_0, \varrho_2)$ we get

$$
\begin{aligned}
\mathcal{C}_h[\varrho_\lambda, \varrho_0] &= \inf_{\gamma\in\Gamma(\varrho_0,\varrho_\lambda)} \left\{ \int_{\mathbb{R}^d\times\mathbb{R}^d} c_h(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma(\mathbf{x},\mathbf{y}) \right\} \\
&\leq \int_{\mathbb{R}^d\times\mathbb{R}^d} c(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma_\lambda(\mathbf{x},\mathbf{y}) \\
&= (1-\lambda)\int_{\mathbb{R}^d\times\mathbb{R}^d} c_h(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma_1(\mathbf{x},\mathbf{y}) + \lambda\int_{\mathbb{R}^d\times\mathbb{R}^d} c_h(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma_2(\mathbf{x},\mathbf{y}) \\
&= (1-\lambda)\mathcal{C}_h[\varrho_0,\varrho_1] + \lambda\mathcal{C}_h[\varrho_0,\varrho_2]
\end{aligned}
$$

This finishes the proof. □

**Lemma 2.4.6** *Let $h > 0$. Suppose that $\{\varrho^n\}_{n\in\mathbb{N}}$ converges weakly to $\varrho$ in $L^1(\mathbb{R}^d)$ and that $\{M_1(\varrho^n)\}_{n\in\mathbb{N}}$ is bounded. Then $M_1(\varrho)$ is finite, and we have*

$$
\liminf_{n\to\infty} \mathcal{C}_h(\bar{\varrho}, \varrho^n) \geq \mathcal{C}_h(\bar{\varrho}, \varrho) \qquad \forall\, \bar{\varrho} \in \mathcal{P}_1^{ac}(X).
$$

Proof The fact that $M_1(\varrho)$ is finite follows from the weak lower-semicontinuity in $L^1(\mathbb{R}^d)$ of $M_1$. Let now $\gamma^n \in \Gamma_h(\bar{\varrho}, \varrho^n)$. Since $\{M_1(\varrho^n)\}_{n\in\mathbb{N}}$ is bounded

$$
\sup_{n\in\mathbb{N}} \int_{\mathbb{R}^d} \big(|\mathbf{x}| + |\mathbf{y}|\big)\,\mathrm{d}\gamma^n(\mathbf{x},\mathbf{y}) < +\infty. \tag{2.4.4}
$$

As $|\mathbf{x}|+|\mathbf{y}|$ is coercive, Equation (2.4.4) implies that $\{\gamma^n\}_{n\in\mathbb{N}}$ admits a cluster point $\gamma$ for the topology of narrow convergence. Furthermore, it is easy to see that $\gamma \in \Gamma(\bar{\varrho}, \varrho)$ and so, since $c_h$ is continuous and bounded below, we get

$$
\begin{aligned}
\liminf_{n\to\infty} \mathcal{C}_h(\bar{\varrho}, \varrho^n) &= \liminf_{n\to\infty} \int_{\mathbb{R}^d\times\mathbb{R}^d} c_h(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma^n(\mathbf{x},\mathbf{y}) \\
&\geq \int_{\mathbb{R}^d\times\mathbb{R}^d} c_h(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma(\mathbf{x},\mathbf{y}) \\
&\geq \mathcal{C}_h(\bar{\varrho}, \varrho).
\end{aligned}
$$

which completes the proof of lemma. (See Lemma 4.3 in [22] for its generalization.)□

# CHAPTER III

# EXISTENCE OF SOLUTIONS IN A BOUNDED DOMAIN

Throughout this chapter, we assume that (2.1.2) and (2.1.3) hold. We recall that $L$ satisfies (L1), (L2) and (L3). We also assume that $X \subset \mathbb{R}^d$ is an open bounded set whose boundary $\partial X$ is of zero Lebesgue measure, and we denote by $\overline{X}$ its closure. The goal is to prove existence of distributional solutions to Equation (1.0.1) by using an approximation by discretization in time. More precisely, in Section 3.1 we construct approximate solutions at discrete times $\{h, 2h, 3h, \ldots\}$ by an implicit Euler scheme, which involves the minimization of a functional. Then, in Section 3.2 we explicitly characterize the minimizer, introducing a dual problem. We then study the properties of an augmented action functional which allows us to prove a priori bounds on De Giorgi's variational and geodesic interpolations (cf. Section 3.4). Finally, using these bounds, we can take the limit as $h \to 0$, and prove existence of distributional solutions to Equation (1.0.1) when $\theta$ behaves like $t^\alpha$, $\alpha > 1$ for large values of $t > 0$.

## *3.1 The Discrete Variational Problem*

We fix a time step $h > 0$, and for simplicity of notation we set $c = c_h$. We fix $\varrho_0 \in \mathcal{P}^{ac}(X)$, and we consider the variational problem

$$\inf_{\varrho \in \mathcal{P}^{ac}(X)} \{\mathcal{C}_h(\varrho_0, \varrho) + \mathcal{U}(\varrho)\} := \inf_{\varrho \in \mathcal{P}^{ac}(X)} \{\Phi[h, \varrho_0, \varrho]\}. \tag{3.1.1}$$

**Theorem 3.1.1** *There exists a unique minimizer $\varrho_h$ of Problem* (3.1.1)*. Suppose in addition that (L4) holds. If $M \in (0, \infty)$ and $\varrho_0(\mathbf{x}) \leq M$* a.e.*, then $\varrho_h(\mathbf{x}) \leq M$* a.e.. *In other words, we have a maximum principle.*

Proof First note that the technique used here is analogous to the ones in [1, 19]. Existence of a minimizer $\varrho_h$ follows by classical methods in the calculus of variation, thanks to the lower-semicontinuity of the functional

$$\varrho \mapsto \Phi[h, \varrho_0, \varrho] = \mathcal{C}_h(\varrho_0, \varrho) + \mathcal{U}(\varrho)$$

in the weak topology of measures and to the superlinearity of $U$ (which implies that any limit point of a minimizing sequence still belongs to $\mathcal{P}^{ac}(X)$).

To prove uniqueness, let $\varrho_1$ and $\varrho_2$ be two minimizers, and take $\gamma_1 \in \Gamma_h(\varrho_0, \varrho_1)$, $\gamma_2 \in \Gamma_h(\varrho_0, \varrho_2)$ (cf. Remark 2.4.1). Then $\dfrac{\gamma_1 + \gamma_2}{2} \in \Gamma\left(\varrho_0, \frac{\varrho_1+\varrho_2}{2}\right)$, so that

$$\mathcal{C}_h\Big(\varrho_0, \frac{\varrho_1 + \varrho_2}{2}\Big) \leq \int_{X\times X} c_h(\mathbf{x}, \mathbf{y}) \,\mathrm{d}\Big(\frac{\gamma_1 + \gamma_2}{2}\Big) = \frac{\mathcal{C}_h(\varrho_0, \varrho_1) + \mathcal{C}_h(\varrho_0, \varrho_2)}{2}.$$

Moreover, by strict convexity of $U$,

$$\mathcal{U}\Big(\frac{\varrho_1 + \varrho_2}{2}\Big) \leq \frac{\mathcal{U}(\varrho_1) + \mathcal{U}(\varrho_2)}{2},$$

with equality if and only if $\varrho_1 = \varrho_2$. This implies uniqueness.

Thanks to (L1) and (L4) one easily gets that $c_h(\mathbf{x}, \mathbf{x}) < c_h(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in X$, $\mathbf{x} \neq \mathbf{y}$. Let's show the maximum principle: For the sake of contradiction assume that $\Omega := \{\mathbf{x} \in X \ : \ \varrho_h(\mathbf{x}) > M\}$ is of positive Lebesgue measure. Let $\gamma \in \Gamma(\varrho_0, \varrho_h)$. Set $\Omega^c := X \setminus \Omega$. Then, we obtain

$$\gamma(\Omega^c \times \Omega) > 0. \tag{3.1.2}$$

Suppose on the contrary that $\gamma(\Omega^c \times \Omega) = 0$. In this case, we have

$$\begin{aligned} M\mathcal{L}^d(\Omega) &< \int_\Omega \varrho_h(\mathbf{y}) \,\mathrm{d}\mathbf{y} = \gamma(X \times \Omega) \\ &= \gamma(\Omega \times \Omega) \leq \gamma(\Omega \times X) \\ &= \int_\Omega \varrho_0(\mathbf{x}) \,\mathrm{d}\mathbf{x} \leq M\mathcal{L}^d(\Omega) \end{aligned}$$

which is a contradiction, concluding (3.1.2). Now, consider the restricted probability measure

$$\gamma_\Omega := \frac{1}{\gamma(\Omega^c \times \Omega)} \gamma\big|_{\Omega^c \times \Omega}$$

defined by

$$\gamma_\Omega(B) = \frac{\gamma\,(\,B \cap (\Omega^c \times \Omega)\,)}{\gamma(\Omega^c \times \Omega)}$$

for Borel sets $B \subset X \times X$. Note that

$$\gamma(\Omega^c \times \Omega) \int_{X\times X} \zeta(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma_\Omega(\mathbf{x},\mathbf{y}) = \int_{\Omega^c \times \Omega} \zeta(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma(\mathbf{x},\mathbf{y})$$

for $\zeta \in C(X \times X)$. Let $\omega_0$ and $\omega_1$ be the marginals of $\gamma_\Omega$, more precisely,

$$\int_{X\times X} (\,\varphi(\mathbf{x}) + \psi(\mathbf{y})\,)\ \mathrm{d}\gamma_\Omega(\mathbf{x},\mathbf{y}) = \int_{\Omega^c} \varphi(\mathbf{x})\,\mathrm{d}\omega_0(\mathbf{x}) + \int_\Omega \psi(\mathbf{y})\,\mathrm{d}\omega_1(\mathbf{y}),$$

for $\varphi$, $\psi \in C(X)$. It is easy to check that $\gamma_\Omega \ll \gamma \in \Gamma(\varrho_0, \varrho_h)$. Therefore $\omega_0 \ll \varrho_0(\mathbf{x})\mathcal{L}^d$ and $\omega_1 \ll \varrho_h(\mathbf{y})\mathcal{L}^d$. That is to say, $\omega_0$ and $\omega_1$ are absolutely continuous with respect to Lebesgue measure $\mathcal{L}^d$. Let $w_0$ and $w_1$ be their density functions respectively. Observe that

$$0 \le w_0 \le \frac{1}{\gamma(\Omega^c \times \Omega)}\varrho_0 \le \frac{M}{\gamma(\Omega^c \times \Omega)} \ \text{ a.}e. \quad \text{and} \quad 0 \le w_1 \le \frac{1}{\gamma(\Omega^c \times \Omega)}\varrho_h \ \text{ a.}e.$$

Also, we see that

$$\begin{aligned}
\omega_0(\Omega) &= \gamma_\Omega(((\Omega \times X) \cap (\Omega^c \times \Omega)) \\
&= \gamma_\Omega((\Omega \cap \Omega^c) \times (X \cap \Omega)) \\
&= 0 \\
&= \gamma_\Omega((X \cap \Omega^c) \times (\Omega \cap \Omega^c)) \\
&= \gamma_\Omega(((X \times \Omega^c) \cap (\Omega^c \times \Omega)) \\
&= \omega_1(\Omega^c).
\end{aligned}$$

In other words, we observe that $\mathrm{spt}(\omega_0) \subseteq \Omega^c$, and $\mathrm{spt}(\omega_1) \subseteq \Omega$. Next, define

$$\varrho_\alpha := \varrho_h - \alpha(w_1 - w_0)$$

for $\alpha \in (0, \gamma(\Omega^c \times \Omega))$. Note that $\varrho_\alpha \in \mathcal{P}^{ac}(X)$. Indeed,

$$\varrho_{\alpha|_\Omega} = \varrho_{h|_\Omega} - \alpha w_1 > M - \alpha w_1 > 0 \qquad \text{and} \qquad \varrho_{\alpha|_{\Omega^c}} = \varrho_{h|_{\Omega^c}} + \alpha w_0 \ge \alpha w_0 \ge 0.$$

and

$$\int_X \varrho_\alpha(\mathbf{y})\,\mathrm{d}\mathbf{y} = 1$$

Accordingly, define a measure $\gamma_\alpha$ on $X \times X$ by

$$\begin{aligned}\int_{X\times X} \zeta(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma_\alpha(\mathbf{x},\mathbf{y}) &:= \int_{X\times X} \zeta(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma(\mathbf{x},\mathbf{y}) \\ &\quad + \alpha \int_{\Omega^c\times\Omega} \left(\zeta(\mathbf{x},\mathbf{x}) - \zeta(\mathbf{x},\mathbf{y})\right)\,\mathrm{d}\gamma(\mathbf{x},\mathbf{y}),\end{aligned}$$

for all $\zeta \in C(X \times X)$. Notice that $\gamma_\alpha \in \Gamma(\varrho_0, \varrho_\alpha)$. Indeed, for $\eta \in C(X)$

$$\int_{X\times X} \eta(\mathbf{x})\,\mathrm{d}\gamma_\alpha(\mathbf{x},\mathbf{y}) = \int_X \eta(\mathbf{x})\varrho_0(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

$$\begin{aligned}\int_{X\times X} \eta(\mathbf{y})\,\mathrm{d}\gamma_\alpha(\mathbf{x},\mathbf{y}) &= \int_X \eta(\mathbf{y})\varrho_h(\mathbf{y})\,\mathrm{d}\mathbf{y} \\ &\quad +\alpha\left(\int_{\Omega^c} \eta(\mathbf{x})w_0(\mathbf{x})\,\mathrm{d}\mathbf{x} - \int_{\Omega} \eta(\mathbf{y})w_1(\mathbf{y})\,\mathrm{d}\mathbf{y}\right) \\ &= \int_X \eta(\mathbf{y})\varrho_\alpha(\mathbf{y})\,\mathrm{d}\mathbf{y}\end{aligned}$$

Next, we will show that there exists $\alpha \in (0, \gamma(\Omega^c \times \Omega)) \subset (0, 1)$ such that

$$\Phi_h[\varrho_0, \varrho_\alpha] < \Phi_h[\varrho_0, \varrho_h] \tag{3.1.3}$$

To this end, remember that $c_h(\mathbf{x},\mathbf{y}) - c_h(\mathbf{x},\mathbf{x}) > 0$ on $\Omega^c \times \Omega$, and so we arrive at

$$\begin{aligned}\mathcal{C}_h(\varrho_0,\varrho_\alpha) - \mathcal{C}_h(\varrho_0,\varrho_h) &\leq \int_{X\times X} c_h(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma_\alpha(\mathbf{x},\mathbf{y}) - \int_{X\times X} c_h(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma(\mathbf{x},\mathbf{y}) \\ &= \alpha \int_{\Omega^c\times\Omega} \left(c_h(\mathbf{x},\mathbf{x}) - c_h(\mathbf{x},\mathbf{y})\right)\,\mathrm{d}\gamma(\mathbf{x},\mathbf{y}) \\ &< 0\end{aligned} \tag{3.1.4}$$

On the other hand, since $U'$ is strictly increasing because of the assumption $U'' > 0$ (cf. Equation (2.1.2)), we have

$$U(t) - U(s) \leq U'(t)(t-s), \quad s, t \in [0, +\infty)$$

and so, we get

$$\begin{aligned}
\mathcal{U}(\varrho_\alpha) - \mathcal{U}(\varrho_1) &= \int_X [U(\varrho_\alpha(\mathbf{x})) - U(\varrho_h(\mathbf{x}))]\, \mathrm{d}\mathbf{x} \\
&= \int_{\Omega^c} [\, U\,(\varrho_h + \alpha w_0) - U(\varrho_h)\,] + \int_\Omega [\, U\,(\varrho_h - \alpha w_1) - U(\varrho_h)\,] \\
&\leq \alpha \int_{\Omega^c} U'(\varrho_h(\mathbf{x}) + \alpha w_0(\mathbf{x})) w_0(\mathbf{x})\, \mathrm{d}\mathbf{x} \\
&\quad -\alpha \int_\Omega U'(\varrho_h(\mathbf{y}) - \alpha w_1(\mathbf{y})) w_1(\mathbf{y})\, \mathrm{d}\mathbf{y} \\
&\leq \alpha \left[ \int_{\Omega^c} U'(M + \alpha w_0(\mathbf{x})) w_0(\mathbf{x})\, \mathrm{d}\mathbf{x} - \int_\Omega U'(M - \alpha w_1(\mathbf{y})) w_1(\mathbf{y})\, \mathrm{d}\mathbf{y} \right] \\
&= \alpha \int_{X \times X} \left( U'(M + \alpha w_0(\mathbf{x})) - U'(M - \alpha w_1(\mathbf{y})) \right) \mathrm{d}\gamma_\Omega(\mathbf{x}, \mathbf{y})\,.
\end{aligned}$$

Since $U \in C^2\left((0,\infty)\right)$, by the Mean Value Theorem,

$$\Big( U'(M + \alpha w_0(\mathbf{x})) - U'(M - \alpha w_1(\mathbf{y})) \Big) \in 0(\alpha)$$

and so we have

$$\mathcal{U}(\varrho_\alpha) - \mathcal{U}(\varrho_1) = 0(\alpha^2). \tag{3.1.5}$$

Choosing $\alpha \in (0, \gamma(\Omega^c \times \Omega))$ sufficiently small, and using (3.1.4) and (3.1.5) we obtain (3.1.3). This proves (3.1.3), which therefore shows that if $\mathcal{L}^d(\Omega) > 0$, then $\varrho_h$ is not a minimizer, contradicting the assumption. □

**Remark 3.1.2** *It is immediate that* $\|\rho_h\|_\infty \leq \|\varrho_0\|_\infty$.

## *3.2 Characterization of Minimizers Via a Dual Problem*

The aim of this section is to completely characterize the minimizer $\varrho_*$ provided by Theorem 3.1.1. We are going to identify a problem, dual to Problem (3.1.1), and use it to achieve that goal.

We define $\mathcal{E} \equiv \mathcal{E}_c$ to be the set of pairs $(u, v) \in C(\overline{X}) \times C(\overline{X})$ such that

$$u(\mathbf{x}) + v(\mathbf{y}) \leq c(\mathbf{x}, \mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in \overline{X}$, and we write $u \oplus v \leq c$. We consider the functional

$$J(u, v) := \int_X u(\mathbf{x})\varrho_0(\mathbf{x})\,\mathrm{d}\mathbf{x} - \int_X U^*(-v(\mathbf{y}))\,\mathrm{d}\mathbf{y}.$$

To lighten the notation, we have omitted to display the $\varrho_0$ dependence in $J$.

**Lemma 3.2.1** *Let $u \in C_b(X)$. Moreover:*

*(i) If $u = v_c$ for some $v \in C(\overline{X})$, then:*

*(a) There exists a constant $A = A(c, X)$, independent of $u$, such that $u$ is $A$-Lipschitz and $A$-semiconcave.*

*(b) If $\bar{\mathbf{x}} \in X$ is a point of differentiability of $u$, $\bar{\mathbf{y}} \in \overline{X}$, and $u(\bar{\mathbf{x}}) + v(\bar{\mathbf{y}}) = c(\bar{\mathbf{x}}, \bar{\mathbf{y}})$, then $\bar{\mathbf{x}}$ is a point of differentiability of $c(\cdot, \bar{\mathbf{y}})$ and $\nabla u(\bar{\mathbf{x}}) = \nabla_{\mathbf{x}} c(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ $= -\nabla_{\mathbf{v}} L(\sigma(0), \dot{\sigma}(0))$. Furthermore $\bar{\mathbf{y}} = \mathbf{\Phi}_1^L\left(h, \bar{\mathbf{x}}, \nabla_{\mathbf{p}} H\left(\bar{\mathbf{x}}, -\nabla u(\bar{\mathbf{x}})\right)\right)$, and $\bar{\mathbf{y}}$ is uniquely determined.*

*(ii) If $v = u^c$ for some $u \in C(\overline{X})$, then:*

*(a) There exists a constant $A = A(c, X)$, independent of $v$, such that $v$ is $A$-Lipschitz and $A$-semiconcave.*

*(b) If $\bar{\mathbf{x}} \in \overline{X}$, $\bar{\mathbf{y}} \in X$ is a point of differentiability of $v$, and $u(\bar{\mathbf{x}}) + v(\bar{\mathbf{y}}) = c(\bar{\mathbf{x}}, \bar{\mathbf{y}})$, then $\bar{\mathbf{y}}$ is a point of differentiability of $c(\bar{\mathbf{x}}, \cdot)$ and $\nabla v(\bar{\mathbf{y}}) = \nabla_{\mathbf{y}} c(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ $= \nabla_{\mathbf{v}} L(\sigma(h), \dot{\sigma}(h))$. Furthermore, $\bar{\mathbf{x}} = \mathbf{\Phi}_1^L\left(-h, \bar{\mathbf{y}}, \nabla_{\mathbf{p}} H\left(\bar{\mathbf{y}}, \nabla v(\bar{\mathbf{y}})\right)\right)$, and $\bar{\mathbf{y}}$ is uniquely determined.*

*In particular, if $K \subset \mathbb{R}$ is bounded, the set $\{v_c : v \in C(\overline{X}),\ v_c(X) \cap K \neq \emptyset\}$ is compact in $C(\overline{X})$, and weak* compact in $W^{1,\infty}(X)$.*

Proof Although the assertions made in the lemma are now part of the folklore of the Monge-Kantorovich theory, we sketch the main steps of the proof. Regarding (i)-(a), we observe that by Remark 2.3.5 the functions $c(\cdot, \mathbf{y})$ are uniformly semiconcave for

$\mathbf{y} \in X$, so that $u$ is semiconcave as the infimum of uniformly semiconcave functions (cf. for instance [9, Appendix A]). In particular $u$ is Lipschitz, with a Lipschitz constant bounded by $\|\nabla_{\mathbf{x}} c\|_{L^\infty(\overline{X}\times\overline{X})}$ (see below for reasoning).

To prove (i)-(b), we note that $\partial_- u(\bar{\mathbf{x}}) \subset \partial_- c(\cdot, \bar{\mathbf{y}})(\bar{\mathbf{x}})$. Since by Remark 2.3.5 $\partial^+ c(\cdot, \bar{\mathbf{y}})(\bar{\mathbf{x}})$ is nonempty, we conclude that $c(\cdot, \bar{\mathbf{y}})$ is differentiable at $\bar{\mathbf{x}}$ if $u$ is. Hence,

$$\nabla u(\bar{\mathbf{x}}) = \nabla_{\mathbf{x}} c(\bar{\mathbf{x}}, \bar{\mathbf{y}}) = -\nabla_{\mathbf{v}} L(\sigma(0), \dot{\sigma}(0))$$

where $\sigma : [0, h] \to X$ is (the unique curve minimizing the action) such that

$$c(\bar{\mathbf{x}}, \bar{\mathbf{y}}) = \int_0^h L(\sigma(t), \dot{\sigma}(t))\,\mathrm{d}t$$

(cf. [9, Section 4 and Appendix B]). This together with Equation (2.3.1) implies

$$\bar{\mathbf{y}} = \mathbf{\Phi}_1^L\big(h, \mathbf{x}, \nabla_{\mathbf{p}} H\big(\mathbf{x}, -\nabla u(\bar{\mathbf{x}})\big)\big). \tag{3.2.1}$$

The proof of (ii) is analogous.

Let's show that the set $C_v := \{v_c : v \in C(\overline{X}),\ v_c(X) \cap K \neq \emptyset\}$ is compact in $C(\overline{X})$. Take a sequence $v_c^n \in C_v$. Let $\mathbf{x}_1, \mathbf{x}_2 \in \overline{X}$ and

$$\begin{aligned} v_c^n(\mathbf{x}_1) &= \min_{\mathbf{y}\in\overline{X}} \{c(\mathbf{x}_1, \mathbf{y}) - v^n(\mathbf{y})\} \\ &:= c(\mathbf{x}_1, \mathbf{y}_n) - v^n(\mathbf{y}_n) \end{aligned}$$

for $\mathbf{y}_n \in \overline{X}$. Then

$$\begin{aligned} v_c^n(\mathbf{x}_2) - v_c^n(\mathbf{x}_1) &\leq (c(\mathbf{x}_2, \mathbf{y}_n) - v^n(\mathbf{y}_n)) - (c(\mathbf{x}_1, \mathbf{y}_n) - v^n(\mathbf{y}_n)) \\ &= c(\mathbf{x}_2, \mathbf{y}_n) - c(\mathbf{x}_1, \mathbf{y}_n) \\ &= \langle \nabla_{\mathbf{x}} c(\bar{\mathbf{x}}, \mathbf{y}_n), \mathbf{x}_1 - \mathbf{x}_2 \rangle \\ &\leq \|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{y})\|_{L^\infty(\overline{X}\times\overline{X})} |\mathbf{x}_2 - \mathbf{x}_1| \end{aligned}$$

where $\bar{\mathbf{x}}$ is on the line segment joining $\mathbf{x}_1$ and $\mathbf{x}_2$. Similarly we prove

$$v_c^n(\mathbf{x}_2) - v_c^n(\mathbf{x}_1) \geq -\|\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{y})\|_{L^\infty(\overline{X}\times\overline{X})} |\mathbf{x}_2 - \mathbf{x}_1|.$$

This concludes that $\{v_c^n\}$ is equi-Lipshitz, therefore, equicontinuous on $\overline{X}$. Furthermore, for any $\mathbf{x} \in \overline{X}$ there exists $\bar{\mathbf{y}}_n \in \overline{X}$

$$v_c^n(\mathbf{x}) = \min_{\mathbf{y}\in\overline{X}} \{c(\mathbf{x},\mathbf{y}) - v^n(\mathbf{y})\} = c(\mathbf{x},\bar{\mathbf{y}}_n) - v^n(\bar{\mathbf{y}}_n) \tag{3.2.2}$$

On the other hand, there exists, $\mathbf{x}_0 \in \overline{X}$ such that $v_c^n(\mathbf{x}_0) \cap K \neq \emptyset$, which implies $|v_c^n(\mathbf{x}_0)| \leq \sup|K|$. Notice that

$$v_c^n(\mathbf{x}_0) = \min_{\mathbf{y}\in\overline{X}} \{c(\mathbf{x}_0,\mathbf{y}) - v^n(\mathbf{y})\} \leq c(\mathbf{x}_0,\bar{\mathbf{y}}_n) - v^n(\bar{\mathbf{y}}_n)$$

and hence

$$v^n(\bar{\mathbf{y}}_n) \leq c(\mathbf{x}_0,\bar{\mathbf{y}}_n) - v_c^n(\mathbf{x}_0)$$

and so

$$|v^n(\bar{\mathbf{y}}_n)| \leq \max_{\mathbf{y}\in\overline{X}} |v^n(\mathbf{y})| \leq \max_{(\mathbf{x},\mathbf{y})\in\overline{X}\times\overline{X}} |c(\mathbf{x},\mathbf{y})| + \sup|K|. \tag{3.2.3}$$

Using (3.2.2) and (3.2.3) we conclude that

$$\begin{aligned} |v_c^n(\mathbf{x})| &= |c(\mathbf{x},\bar{\mathbf{y}}_n) - v^n(\bar{\mathbf{y}}_n)| \\ &\leq |c(\mathbf{x},\bar{\mathbf{y}}_n)| + |v^n(\bar{\mathbf{y}}_n)| \\ &\leq 2 \max_{(\mathbf{x},\mathbf{y})\in\overline{X}\times\overline{X}} |c(\mathbf{x},\mathbf{y})| + \sup|K|, \end{aligned}$$

implying that $\{v_c^n\}$ is equi-bounded in $C(\overline{X})$. By the Arzelà-Ascoli Theorem, we conclude that the set $C_v$ is a compact subset of $C(\overline{X})$. □

**Remark 3.2.2** *By Lemma 3.2.1, if* $u = v_c$ *for some* $v \in C_b(\overline{X})$*, we can uniquely define* $\mathcal{L}^d$*-a.e. a map* $T : \mathrm{dom}\nabla u \to \overline{X}$ *such that* $u(\mathbf{x}) + v(T\mathbf{x}) = c(\mathbf{x}, T\mathbf{x})$*. This map is continuous on* $\mathrm{dom}\nabla u$*, and since* $\nabla u$ *can be extended to a Borel map on* $X$ *we conclude that* $T$ *can be extended to a Borel map on* $X$*, too. Moreover we have* $\nabla u(\mathbf{x}) = \nabla_{\mathbf{x}} c(\mathbf{x}, T\mathbf{x})$ $\mathcal{L}^d$*-a.e., and* $T$ *is the unique optimal map pushing a density* $\varrho \in \mathcal{P}^{ac}(X)$ *forward to* $\bar{\mu} := T_{\#}(\varrho\mathcal{L}^d) \in \mathcal{P}(X)$ *(cf. for instance [15, 21, 22]).*

**Lemma 3.2.3** *If $(u,v) \in \mathcal{E}$ and $\varrho \in \mathcal{P}^{ac}(X)$, then $J(u,v) \le \mathcal{C}_h(\varrho_0,\varrho) + \mathcal{U}(\varrho)$.*

Proof Let $\gamma \in \Gamma(\varrho_0,\varrho)$. Since $U(\varrho(\mathbf{y})) + U^*(-v(\mathbf{y})) \ge -\varrho(\mathbf{y})v(\mathbf{y})$ and $(u,v) \in \mathcal{E}$, we have by integration

$$\begin{aligned}\int_X \big(U(\varrho(\mathbf{y})) \quad &+ \quad U^*(-v(\mathbf{y}))\big)\,\mathrm{d}\mathbf{y} \ge -\int_X \varrho(\mathbf{y})v(\mathbf{y})\,\mathrm{d}\mathbf{y} \\ &\ge \quad -\int_{X\times X} c(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma(\mathbf{x},\mathbf{y}) + \int_X \varrho_0(\mathbf{x})u(\mathbf{x})\,\mathrm{d}\mathbf{x}. \qquad (3.2.4)\end{aligned}$$

Rearranging the expressions in Equation (3.2.4), and optimizing over $\Gamma(\varrho_0,\varrho)$, we obtain the result. □

**Lemma 3.2.4** *There exists $(u_*,v_*) \in \mathcal{E}$ maximizing $J(u,v)$ over $\mathcal{E}$ and satisfying $u_*^c = v_*$ and $(v_*)_c = u_*$. Furthermore:*

*(i) $u_*$ and $v_*$ are Lipschitz with a Lipschitz constant bounded by $\|\nabla c\|_{L^\infty(X\times X)}$.*

*(ii) We have that $\varrho_{v_*} := (U^*)'(-v_*)$ is a probability density on $X$, and the optimal map $T$ associated with $u_*$ (cf. Remark 3.2.2) pushes $\varrho_0\mathcal{L}^d$ forward to $\varrho_{v_*}\mathcal{L}^d$.*

Proof Note that if $u_*^c = v_*$ and $(v_*)_c = u_*$, then (i) is a direct consequence of Lemma 3.2.1. Before proving the first statement of the lemma, let us show that it implies (ii). Let $\varphi \in C(\overline{X})$ and set

$$v_\varepsilon := v_* + \varepsilon\varphi, \quad u_\varepsilon := (v_\varepsilon)_c.$$

Remark 3.2.2 says that for $\mathcal{L}^d$-a.e. $\mathbf{x} \in X$, the equation $u_*(\mathbf{x}) + v_*(\mathbf{y}) = c(\mathbf{x},\mathbf{y})$ admits a unique solution $\mathbf{y} = T\mathbf{x}$. As in [12] (cf. also [14]) we have

$$\|u_\varepsilon - u_*\|_\infty \le \varepsilon\|\varphi\|_\infty \qquad \text{and} \qquad \lim_{\varepsilon\to 0}\frac{u_\varepsilon(\mathbf{x}) - u_*(\mathbf{x})}{\varepsilon} = -\varphi(T\mathbf{x}) \qquad (3.2.5)$$

for $\mathcal{L}^d$-a.e. $\mathbf{x} \in X$. Indeed, by the definition of infimum, for any $\delta > 0$ there exists $\mathbf{z}_\delta$

such that

$$\begin{aligned}
u_\varepsilon(\mathbf{x}) &= (v_\varepsilon)_c(\mathbf{x}) \\
&= \inf_{\mathbf{z}}\{c(\mathbf{x},\mathbf{z}) - v_\varepsilon(\mathbf{z})\} \\
&\geq c(\mathbf{x},\mathbf{z}_\delta) - v_\varepsilon(\mathbf{z}_\delta) - \delta \\
&\geq c(\mathbf{x},\mathbf{z}_\delta) - (v_*(\mathbf{z}_\delta) + \varepsilon\varphi(\mathbf{z}_\delta)) - \delta \\
&\geq u_*(\mathbf{x}) - \varepsilon\varphi(\mathbf{z}_\delta) - \delta
\end{aligned}$$

which in turn yields

$$\begin{aligned}
u_\varepsilon(\mathbf{x}) - u_*(\mathbf{x}) &\geq -\varepsilon\varphi(\mathbf{z}_\delta) - \delta \\
u_\varepsilon(\mathbf{x}) - u_*(\mathbf{x}) + \delta &\geq -\varepsilon\varphi(\mathbf{z}_\delta) \\
u_\varepsilon(\mathbf{x}) - u_*(\mathbf{x}) + \delta &\geq -\varepsilon\|\varphi\|_\infty \qquad \forall\delta > 0
\end{aligned}$$

because $v_*(\mathbf{z}_\delta) + u_*(\mathbf{x}) \leq c(\mathbf{x},\mathbf{z}_\delta)$. Hence,

$$u_\varepsilon(\mathbf{x}) - u_*(\mathbf{x}) \geq -\varepsilon\|\varphi\|_\infty \tag{3.2.6}$$

Also, using the definition of infimum and $u \oplus v \leq c$ and the fact that $c(\mathbf{x}, T\mathbf{x}) = u_*(\mathbf{x}) + v_*(T\mathbf{x})$, we obtain

$$\begin{aligned}
u_\varepsilon(\mathbf{x}) &= (v_\varepsilon)_c(\mathbf{x}) \\
&= \inf_{\mathbf{z}}\{c(\mathbf{x},\mathbf{z}) - v_\varepsilon(\mathbf{z})\} \\
&\leq c(\mathbf{x},T\mathbf{x}) - v_\varepsilon(T\mathbf{x}) \\
&= c(\mathbf{x},T\mathbf{x}) - v_*(T\mathbf{x}) - \varepsilon\varphi(T\mathbf{x}) \\
&= u_*(T\mathbf{x}) - \varepsilon\varphi(T\mathbf{x}) \qquad (3.2.7)
\end{aligned}$$

and so, we end up with

$$u_\varepsilon(\mathbf{x}) - u_*(\mathbf{x}) \leq -\varepsilon\varphi(T\mathbf{x}) \leq \varepsilon\|\varphi\|_\infty \tag{3.2.8}$$

Combining Equations (3.2.6) and (3.2.8) we obtain the first inequality in (3.2.5). Now let's show the limit on the right in (3.2.5). It follows from (3.2.7) (or (3.2.7)) that

$$u_\varepsilon(\mathbf{x}) - u_*(\mathbf{x}) \leq -\varepsilon\varphi(T\mathbf{x}) \tag{3.2.9}$$

Moreover, for $\varepsilon > 0$ there exists $\mathbf{z}_\varepsilon$ such that

$$\begin{aligned} u_\varepsilon(\mathbf{x}) &= (v_\varepsilon)_c(\mathbf{x}) \\ &= \inf_{\mathbf{z}}\{c(\mathbf{x},\mathbf{z}) - v_\varepsilon(\mathbf{z})\} \\ &\geq c(\mathbf{x},\mathbf{z}_\varepsilon) - v_\varepsilon(\mathbf{z}_\varepsilon) - \varepsilon^2 \\ &\geq c(\mathbf{x},\mathbf{z}_\varepsilon) - (v_*(\mathbf{z}_\varepsilon) + \varepsilon\varphi(\mathbf{z}_\varepsilon)) - \varepsilon^2 \end{aligned} \tag{3.2.10}$$

Passing to a subsequence we can say that $\mathbf{z}_\varepsilon \to \mathbf{z}$ as $\varepsilon \to 0$ for some $\mathbf{z}$ in $\overline{X}$. Since $\|u_\varepsilon - u_*\|_\infty \leq \varepsilon\|\varphi\|_\infty$, (i.e., $u_\varepsilon \to u_*$ uniformly as $\varepsilon \to 0$) passing to limit in (3.2.10) yields

$$u_*(\mathbf{x}) \geq c(\mathbf{x},\mathbf{z}) - v_*(\mathbf{z})$$

which means that $\mathbf{z} = T\mathbf{x}$. Writing (3.2.10) explicitly we have

$$u_\varepsilon(\mathbf{x}) \geq u_*(\mathbf{x}) - \varepsilon\varphi(\mathbf{z}_\varepsilon) - \varepsilon^2 \tag{3.2.11}$$

Combining (3.2.11) with (3.2.9) gives

$$-\varphi(\mathbf{z}_\varepsilon) - \varepsilon \leq \frac{u_\varepsilon(\mathbf{x}) - u_*(\mathbf{x})}{\varepsilon} \leq -\varphi(T\mathbf{x}),$$

which produces the limit in (3.2.5) as $\varepsilon \to 0$.

Therefore, by the Lebesgue Dominated Convergence Theorem we get

$$\lim_{\varepsilon\to 0}\int_X \frac{u_\varepsilon(\mathbf{x}) - u_*(\mathbf{x})}{\varepsilon}\varrho_0(\mathbf{x})\,\mathrm{d}\mathbf{x} = -\int_X \varphi(T\mathbf{x})\varrho_0(\mathbf{x})\,\mathrm{d}\mathbf{x}. \tag{3.2.12}$$

Since $(u_*, v_*)$ maximizes $J$ over $\mathcal{E}$, by Equation (3.2.12) we obtain

$$0 = \lim_{\varepsilon\to 0}\frac{J(u_\varepsilon, v_\varepsilon) - J(u_*, v_*)}{\varepsilon} = -\int_X \varphi(T\mathbf{x})\varrho_0(\mathbf{x})\,\mathrm{d}\mathbf{x} + \int_X (U^*)'(-v_*(\mathbf{x}))\varphi(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Therefore, we are left with

$$\int_X \varphi(T\mathbf{x})\varrho_0(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_X (U^*)'(-v_*(\mathbf{x}))\varphi(\mathbf{x})\,\mathrm{d}\mathbf{x}. \tag{3.2.13}$$

Choosing $\varphi \equiv 1$ in Equation (3.2.13) and recalling that $(U^*)' \geq 0$ (cf. Lemma 2.2.1(ii)), we discover that $\varrho_{v_*} := (U^*)'(-v_*)$ is a probability density on $X$. Moreover Equation (3.2.13) means that $T$ pushes $\varrho_0\mathcal{L}^d$ forward to $\varrho_{v_*}\mathcal{L}^d$. This proves (ii).

We finally proceed with the proof of the first statement. Observe that the functional $J$ is linear, and so it is continuous on $\mathcal{E}$, which is a closed subset of $C(\overline{X})\times C(\overline{X})$. Thus it suffices to show the existence of a compact set $\mathcal{E}' \subset \mathcal{E}$ such that

$$\mathcal{E}' \subset \{(u,v):\ u^c = v, v_c = u\} \qquad \text{and} \qquad \sup_{\mathcal{E}} J = \sup_{\mathcal{E}'} J.$$

If $(u,v) \in \mathcal{E}$ then, by Remark 2.1.4, $u \leq v_c$, and so $J(u,v) \leq J(v_c, v)$. But, as pointed out in Lemma 2.1.6, $v \leq (v_c)^c$, and since by Lemma 2.2.1(ii) $U^* \in C^1(\mathbb{R})$ is monotone nondecreasing, we have

$$J(u,v) \leq J(v_c, v) \leq J(v_c, (v_c)^c).$$

Set $\bar{u} = v_c$ and $\bar{v} = (v_c)^c$. By Lemma 3.2.1, $\bar{u} = \bar{v}_c$ and $\bar{v} = \bar{u}^c$. As $U^* \in C^1(\mathbb{R})$ and $(U^*)' \geq 0$, the functional

$$\lambda \mapsto e(\lambda) := \int_X U^*(-\bar{v}(\mathbf{x}) + \lambda)\,\mathrm{d}\mathbf{x}$$

is differentiable, and

$$e'(\lambda) = \int_X (U^*)'(-\bar{v}(\mathbf{x}) + \lambda)\,\mathrm{d}\mathbf{x} \geq 0.$$

Since by Lemma 2.2.1(iv) $U^*$ grows superlinearly at infinity, so does $e(\lambda)$. Hence

$$\lim_{\lambda\to+\infty} J(\bar{u}+\lambda, \bar{v}-\lambda) = \lim_{\lambda\to+\infty} \int_X \bar{u}\varrho_0 \mathrm{d}\mathbf{x} + \lambda - e(\lambda) = -\infty. \tag{3.2.14}$$

Moreover, as $U^* \geq 0$ (cf. Lemma 2.2.1(ii)),

$$\lim_{\lambda\to-\infty} J(\bar{u}+\lambda, \bar{v}-\lambda) \leq \lim_{\lambda\to-\infty} \int_X \bar{u}\varrho_0 \mathrm{d}\mathbf{x} + \lambda = -\infty. \tag{3.2.15}$$

Since $\lambda \to J(\bar{u}+\lambda, \bar{v}-\lambda)$ is differentiable, by Equations (3.2.14) and (3.2.15) $J(\bar{u}+\lambda, \bar{v}-\lambda)$ achieves its maximum at a certain $\bar{\lambda}$ which satisfies $e'(\bar{\lambda}) = 1$. Therefore we obtain the following conclusions:

$$(\tilde{u}, \tilde{v}) := (\bar{u}+\bar{\lambda}, \bar{v}-\bar{\lambda}) \in \mathcal{E}, \quad J(\bar{u}, \bar{v}) \leq J(\tilde{u}, \tilde{v}), \quad \text{and} \quad \int_X (U^*)'(-\tilde{v})\,\mathrm{d}\mathbf{x} = 1.$$

This last equality and the fact that $(U^*)'(-\tilde{v})$ is continuous on the compact set $\overline{X}$ ensure the existence of a point $\mathbf{x}_0 \in \overline{X}$ such that $-\tilde{v}(\mathbf{x}_0) = U'\big(1/\mathcal{L}^d(X)\big)$. In light of Lemma 3.2.1 and the above reasoning, we have established that the set

$$\mathcal{E}' := \big\{(u,v) \in \mathcal{E} \ : \ u^c = v, \ v_c = u, \ v(\mathbf{x}_0) = -U'\left(1/\mathcal{L}^d(X)\right) \text{ for some } \mathbf{x}_0 \in \overline{X}\big\}$$

satisfies the required conditions. Indeed, take a sequence $(u_n, v_n) \in \mathcal{E}'$. Set

$$M_1 = \sup_{(\mathbf{x},\mathbf{y}) \in \overline{X}\times\overline{X}} |\nabla_{\mathbf{y}} c(\mathbf{x},\mathbf{y})|, \qquad M_2 = \max_{(\mathbf{x},\mathbf{y}) \in \overline{X}\times\overline{X}} |c(\mathbf{x},\mathbf{y})|.$$

Let $\overline{X} \subset B_R$ for some $R > 0$. Since $v_n = u_n^c$, we know that $\mathrm{Lip}(v_n) \leq M_1$. Thus, we have

$$\begin{aligned} |v_n(\mathbf{y})| &\leq |v_n(\mathbf{x}_0)| + |v_n(\mathbf{y}) - v_n(\mathbf{x}_0)| \\ &\leq |v_n(\mathbf{x}_0)| + M_1|\mathbf{y} - \mathbf{x}_0| \\ &\leq |U'(1/\mathcal{L}^d(X))(\mathbf{x}_0)| + M_1 R. \end{aligned}$$

Thus, $v_n$ is uniformly bounded. Therefore, using $u_n = (v_n)_c$, we obtain

$$u_n(\mathbf{x}) = \inf_{\mathbf{y} \in \overline{X}} \{c(\mathbf{x},\mathbf{y}) - v_n(\mathbf{y})\} := c(\mathbf{x},\mathbf{y}_n) - v_n(\mathbf{y}_n).$$

$$\begin{aligned} |u_n(\mathbf{x})| &\leq |c(\mathbf{x},\mathbf{y}_n)| + |v_n(\mathbf{y}_n)| \\ &\leq M_2 + \left|U'\left(1/\mathcal{L}^d(X)\right)\right| + M_1 R, \end{aligned}$$

implying that $u_n$ is also uniformly bounded. On the other hand, $u_n$ and $v_n$ are equi-Lipshitz, and therefore equicontinuous. Employing the Arzelà-Ascoli theorem, we conclude that $\mathcal{E}'$ is compact in $C(\overline{X}) \times C(\overline{X})$ with respect to the uniform norm. $\square$

Set

$$\Phi(\varrho) := \Phi(h, \varrho_0, \varrho) = \mathcal{C}_h(\varrho_0, \varrho) + \mathcal{U}(\varrho).$$

**Lemma 3.2.5** *Let* $\varrho_*$ *be the unique minimizer of* $\Phi$ *provided by Theorem 3.1.1, and let* $(u_*, v_*)$ *be a maximizer of* $J$ *obtained in Lemma 3.2.4. Then* $\varrho_* = (U^*)'(-v_*)$*, and*

$$\max_{\mathcal{E}} J = J(u_*, v_*) = \Phi(\varrho_*) = \min_{\varrho \in \mathcal{P}^{ac}(X)} \Phi(\varrho).$$

Proof Let $T$ be as in Lemma 3.2.4(ii), and define $\varrho_{v_*} := (U^*)'(-v_*)$. Note that since $T$ pushes $\varrho_0 \mathcal{L}^d$ forward to $\varrho_{v_*} \mathcal{L}^d$, we have

$$(\mathrm{id} \times T)_{\#}(\varrho_0 \mathcal{L}^d) \in \Gamma(\varrho_0, \varrho_{v_*}).$$

Therefore, as $c(\mathbf{x}, T\mathbf{x}) = u_*(\mathbf{x}) + v_*(T\mathbf{x})$ for $\varrho_0 \mathcal{L}^d$-a.e. $\mathbf{x} \in X$,

$$\begin{aligned} \mathcal{C}_h(\varrho_0, \varrho_{v_*}) &\leq \int_X c(\mathbf{x}, Tx) \varrho_0(\mathbf{x})\, \mathrm{d}\mathbf{x} \\ &= \int_X \big(u_*(\mathbf{x}) + v_*(T\mathbf{x})\big) \varrho_0(\mathbf{x})\, \mathrm{d}\mathbf{x} \\ &= \int_X u_*(\mathbf{x}) \varrho_0(\mathbf{x})\, \mathrm{d}\mathbf{x} + \int_X v_*(\mathbf{x}) \varrho_{v_*}(\mathbf{x})\, \mathrm{d}\mathbf{x}. \end{aligned} \tag{3.2.16}$$

Since

$$\mathcal{C}_h(\varrho_0, \varrho_{v_*}) \geq \int_X u_*(\mathbf{x}) \varrho_0(\mathbf{x})\, \mathrm{d}\mathbf{x} + \int_X v_*(\mathbf{x}) \varrho_{v_*}(\mathbf{x})\, \mathrm{d}\mathbf{x}$$

trivially holds (as $u \oplus v \leq c$), Inequality (3.2.16) is in fact an equality, and so

$$\mathcal{C}_h(\varrho_0, \varrho_{v_*}) + \mathcal{U}(\varrho_{v_*}) = \int_X u_*(\mathbf{x}) \varrho_0(\mathbf{x})\, \mathrm{d}\mathbf{x} + \int_X \big(v_*(\mathbf{x}) \varrho_{v_*}(\mathbf{x}) + U(\varrho_{v_*})\big)\, \mathrm{d}\mathbf{x}. \tag{3.2.17}$$

Combining the equality $-v_* \varrho_{v_*} = U(\varrho_{v_*}) + U^*(-v_*)$ (which follows from $\varrho_{v_*} = (U^*)'(-v_*)$) with Equation (3.2.17) we get

$$\mathcal{C}_h(\varrho_0, \varrho_*) + \mathcal{U}(\varrho_*) = J(u_*, v_*),$$

which together with Lemma 3.2.3 gives that $\varrho_{v_*}$ minimizes $\Phi$ over $\mathcal{P}^{ac}(X)$ and

$$\sup_{\mathcal{E}} J = \Phi(\varrho_{v_*}).$$

Since the minimizer of $\Phi$ over $\mathcal{P}^{ac}(X)$ is unique (cf. Theorem 3.1.1), this concludes the proof. □

**Remark 3.2.6** *By Lemma 3.2.1, we can uniquely define a map* $S$ *on* $\mathrm{dom}\nabla v_*$ *by*

$$u_*(S\mathbf{y}) + v_*(\mathbf{y}) = c(S\mathbf{y}, \mathbf{y})$$

*and we have*

$$\nabla v_*(\mathbf{y}) = \nabla_{\mathbf{y}} c(S\mathbf{y}, \mathbf{y}).$$

*This map is the inverse of* $T$ *up to a set of zero measure, it pushes* $\varrho_* \mathcal{L}^d$ *forward to* $\varrho_0 \mathcal{L}^d$*, and*

$$S\mathbf{y} = \mathbf{\Phi}_1^L\big(-h, \mathbf{y}, \nabla_{\mathbf{p}} H\big(\mathbf{y}, \nabla v_*(\mathbf{y})\big)\big).$$

*Moreover, by Lemma 3.2.5,* $U'(\varrho_*) = -v_*$ *is Lipschitz, and*

$$-\nabla v_*(\mathbf{y}) = \nabla[U'(\varrho_*)](\mathbf{y}).$$

*In particular,*

$$S\mathbf{y} = \mathbf{\Phi}_1^L\big(-h, \mathbf{y}, \nabla_{\mathbf{p}} H\big(\mathbf{y}, -\nabla[U'(\varrho_*)](\mathbf{y})\big)\big).$$

We observe that the duality method allows us to deduce the Euler-Lagrange Equation associated with the functional $\Phi$ by passing many technical problems due to regularity issues. Moreover it gives the equality

$$\nabla_{\mathbf{y}} c(S\mathbf{y}, \mathbf{y}) = -\nabla[U'(\varrho_*)](\mathbf{y}) \quad \mathcal{L}^d\text{-a.e. in } X$$

(and not only $\varrho_* \mathcal{L}^d$-a.e.).

## *3.3 Augmented Actions*

To ease the notation we introduce the functional

$$\Phi(\tau, \varrho_0, \varrho) := \mathcal{C}_\tau(\varrho_0, \varrho) + \mathcal{U}(\varrho) \qquad \varrho_0, \varrho \in \mathcal{P}^{ac}(X),$$

and we define

$$\phi_\tau(\varrho_0) := \inf_{\varrho \in \mathcal{P}^{ac}(X)} \Phi(\tau, \varrho_0, \varrho).$$

The goal of this section is to study the properties of $\Phi$ and $\phi_\tau$, in the same spirit as in [3, Chapter 3].

In the sequel, we fix $\varrho_0 \in \mathcal{P}^{ac}(X)$. Lemma 3.2.5 provides existence of a unique minimizer of $\Phi(\tau, \varrho_0, \varrho)$ over $\mathcal{P}^{ac}(X)$, which we call $\varrho_\tau$.

**Lemma 3.3.1** *The function $\tau \mapsto \phi_\tau(\varrho_0)$ is nonincreasing, and satisfies*

$$\frac{\phi_{\tau_1}(\varrho_0) - \phi_{\tau_0}(\varrho_0)}{\tau_1 - \tau_0} \leq \frac{\mathcal{C}_{\tau_1}(\varrho_0, \varrho_{\tau_0}) - \mathcal{C}_{\tau_0}(\varrho_0, \varrho_{\tau_0})}{\tau_1 - \tau_0} \qquad \forall\, 0 \leq \tau_0 \leq \tau_1. \tag{3.3.1}$$

*In particular* $\dfrac{d^+\phi_\tau(\varrho_0)}{d\tau}(\tau) \leq 0$ *for all* $\tau \geq 0$, $\dfrac{d^+\phi_\tau(\varrho_0)}{d\tau} \in L^1_{loc}([0, +\infty))$, *and*

$$\phi_{\tau_1}(\varrho_0) - \phi_{\tau_0}(\varrho_0) \leq \int_{\tau_0}^{\tau_1} \frac{d^+\phi_\tau(\varrho_0)}{d\tau}(\tau)\, d\tau \qquad \forall\, 0 \leq \tau_0 \leq \tau_1. \tag{3.3.2}$$

Proof It is an immediate consequence of the definition of $\phi_\tau$ and $\varrho_\tau$ that, for all $\tau_0, \tau_1 > 0$, we have

$$\mathcal{C}_{\tau_1}(\varrho_0, \varrho_{\tau_0}) - \mathcal{C}_{\tau_0}(\varrho_0, \varrho_{\tau_0}) \geq \phi_{\tau_1}(\varrho_0) - \phi_{\tau_0}(\varrho_0).$$

This gives Equation (3.3.1), which together with Lemma 2.4.4(i) implies that $\tau \mapsto \phi_\tau(\varrho_0)$ is nonincreasing. The last part of the lemma follows from the general fact that, if $f : [a, b] \to \mathbb{R}$ is a nonincreasing function, then

$$\frac{\mathrm{d}^+ f}{\mathrm{d}t} \leq 0, \qquad \frac{\mathrm{d}^+ f}{\mathrm{d}t} \in L^1(a, b), \quad \text{and} \quad \left\| \frac{\mathrm{d}^+ f}{\mathrm{d}t} \right\|_{L^1(a,b)} \leq f(a) - f(b)$$

(cf. Lemma A.1.6). □

For $h > 0$, we denote by $T_h$ the optimal map that pushes $\varrho_0 \mathcal{L}^d$ forward to $\varrho_h \mathcal{L}^d$ as provided by the last paragraph. We have

$$T_h \mathbf{x} = \mathbf{\Phi}_1^L\big(h, \mathbf{x}, \nabla_{\mathbf{p}} H\big(\mathbf{x}, -\nabla u_h(\bar{\mathbf{x}})\big)\big),$$

with $(u_h, v_h)$ a maximizer of

$$(u, v) \to \int_X \varrho_0(\mathbf{x}) u(\mathbf{x})\, \mathrm{d}\mathbf{x} - \int_X U^*(-v(\mathbf{y}))\, \mathrm{d}\mathbf{y}.$$

We recall that $(U^*)'(-v_h) = \varrho_h$ (cf. Lemma 3.2.5). Moreover, if we define the interpolation map between $\varrho_0$ and $\varrho_h$ by using

$$T_h^s \mathbf{x} := \mathbf{\Phi}_1^L\big(s, \mathbf{x}, \nabla_{\mathbf{p}} H\big(\mathbf{x}, -\nabla u_h(\bar{\mathbf{x}})\big)\big), \qquad s \in [0, h], \tag{3.3.3}$$

we have

$$c_h(\mathbf{x}, T_h\mathbf{x}) = \int_0^h L\big(\sigma_0^{\mathbf{x}}(s), \dot{\sigma}_0^{\mathbf{x}}(s)\big)\,\mathrm{d}s \quad \text{with} \quad \sigma_0^{\mathbf{x}}(s) := T_h^s\mathbf{x}. \tag{3.3.4}$$

Finally, since $v_h = -U'(\varrho_h)$, denoting by $S_h$ the inverse of $T_h$ we also have

$$\nabla_{\mathbf{y}} c_h(S_h\mathbf{y}, \mathbf{y}) = -\nabla[U'(\varrho_h)](\mathbf{y}) = \nabla_{\mathbf{v}} L\big(\sigma_0^{S_h\mathbf{y}}(h), \dot{\sigma}_0^{S_h\mathbf{y}}(h)\big) \quad \text{for } \mathcal{L}^d\text{-a.e. } \mathbf{y} \in X. \tag{3.3.5}$$

**Lemma 3.3.2** *We have*

$$\int_X H\big(\mathbf{y}, -\nabla[U'(\varrho_h)](\mathbf{y})\big)\varrho_h(\mathbf{y})\,\mathrm{d}\mathbf{y} \leq -\frac{d^+\phi_t(\varrho_0)}{\mathrm{d}t}(h).$$

Proof For $s \in [0, h+\varepsilon]$ we define

$$\sigma_\varepsilon^{\mathbf{x}}(s) := \Phi_1^L\left(\frac{sh}{h+\varepsilon}, \mathbf{x}, \nabla_{\mathbf{p}} H\big(\mathbf{x}, -\nabla u_h(\mathbf{x})\big)\right).$$

Since $\sigma_\varepsilon^{\mathbf{x}}(0) = \mathbf{x}$ and $\sigma_\varepsilon(h+\varepsilon) = T_h\mathbf{x}$, by the definition of $\mathcal{C}_{h+\varepsilon}$ we get

$$\begin{aligned} \mathcal{C}_{h+\varepsilon}(\varrho_0, \varrho_h) &\leq \int_X \varrho_0(\mathbf{x}) \int_0^{h+\varepsilon} L\big(\sigma_\varepsilon^{\mathbf{x}}, \dot{\sigma}_\varepsilon^{\mathbf{x}}\big)\,\mathrm{d}s\,\mathrm{d}\mathbf{x} \\ &= \frac{h+\varepsilon}{h} \int_X \varrho_0(\mathbf{x}) \int_0^h L\Big(\sigma_0^{\mathbf{x}}, \frac{h}{h+\varepsilon}\dot{\sigma}_0^{\mathbf{x}}\Big)\,\mathrm{d}s\,\mathrm{d}\mathbf{x}. \end{aligned} \tag{3.3.6}$$

Moreover, since $L(\mathbf{x}, \cdot)$ is convex,

$$L\big(\sigma_0^{\mathbf{x}}, \dot{\sigma}_0^{\mathbf{x}}\big) \geq L\Big(\sigma_0^{\mathbf{x}}, \frac{h}{h+\varepsilon}\dot{\sigma}_0^{\mathbf{x}}\Big) + \frac{\varepsilon}{h+\varepsilon}\Big\langle \nabla_{\mathbf{v}} L\big(\sigma_0^{\mathbf{x}}, \frac{h}{h+\varepsilon}\dot{\sigma}_0^{\mathbf{x}}\big), \dot{\sigma}_0^{\mathbf{x}}\Big\rangle. \tag{3.3.7}$$

Therefore, recalling that

$$\Big\langle \nabla_{\mathbf{v}} L\Big(\sigma_0^{\mathbf{x}}, \frac{h}{h+\varepsilon}\dot{\sigma}_0^{\mathbf{x}}\Big), \dot{\sigma}_0^{\mathbf{x}}\Big\rangle = L\Big(\sigma_0^{\mathbf{x}}, \frac{h}{h+\varepsilon}\dot{\sigma}_0^{\mathbf{x}}\Big) + H\Big(\sigma_0^{\mathbf{x}}, \nabla_{\mathbf{v}} L\Big(\sigma_0^{\mathbf{x}}, \frac{h}{h+\varepsilon}\dot{\sigma}_0^{\mathbf{x}}\Big)\Big),$$

combining Equations (3.3.6) and (3.3.7) we obtain

$$\begin{aligned}\mathcal{C}_{h+\varepsilon}(\varrho_0,\varrho_h) &\leq \left(1-\frac{\varepsilon^2}{h(h+\varepsilon)}\right)\mathcal{C}_h(\varrho_0,\varrho_h)\\ &- \frac{\varepsilon}{h+\varepsilon}\int_X \varrho_0(\mathbf{x})\int_0^h H\Big(\sigma_0^{\mathbf{x}},\nabla_{\mathbf{v}}L\big(\sigma_0^{\mathbf{x}},\frac{h}{h+\varepsilon}\dot\sigma_0^{\mathbf{x}}\big)\Big)\,\mathrm{d}s\,\mathrm{d}\mathbf{x}.\end{aligned}$$

Hence, as $H \geq 0$ by (H1), by Fatou's Lemma we get

$$\limsup_{\varepsilon\to 0^+}\frac{\mathcal{C}_{h+\varepsilon}(\varrho_0,\varrho_h)-\mathcal{C}_h(\varrho_0,\varrho_h)}{\varepsilon} \leq -\frac{1}{h}\int_X \varrho_0(\mathbf{x})\int_0^h H\Big(\sigma_0^{\mathbf{x}},\nabla_{\mathbf{v}}L\big(\sigma_0^{\mathbf{x}},\dot\sigma_0^{\mathbf{x}}\big)\Big)\,\mathrm{d}s\,\mathrm{d}\mathbf{x}. \tag{3.3.8}$$

Using the conservation of the Hamiltonian $H$ recalled in Remark 2.3.3(ii) and Equations (3.3.1) and (3.3.8) we obtain

$$\frac{d^+\phi_t(\varrho_0)}{\mathrm{d}t}(h) \leq -\int_X H\Big(\sigma_0^{\mathbf{x}}(h),\nabla_{\mathbf{v}}L\big(\sigma_0^{\mathbf{x}}(h),\dot\sigma_0^{\mathbf{x}}(h)\big)\Big)\varrho_0(\mathbf{x})\,\mathrm{d}\mathbf{x},$$

and recalling that $\mathbf{y} = \sigma_0^{\mathbf{x}}(h) = T_h\mathbf{x}$, where $T_h$ pushes $\varrho_0\mathcal{L}^d$ forward to $\varrho_h\mathcal{L}^d$, the desired result follows from Equation (3.3.5). □

**Remark 3.3.3** *Note that*

$$\phi_\tau(\varrho_0) \leq \Phi(\tau,\varrho_0,\varrho_0) \leq \mathcal{U}(\varrho_0)$$

*(since $\mathcal{C}_\tau(\varrho_0,\varrho_0) \leq 0$, cf. Lemma 2.3.1). Therefore setting $\phi_0(\varrho_0) = \mathcal{U}(\varrho_0)$ ensures that $\tau \mapsto \phi_\tau$ remains monotone nonincreasing on $[0,\infty)$, and we have*

$$\mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_h) = \phi_0(\varrho_0) - \phi_h(\varrho_h) + \mathcal{C}_h(\varrho_0,\varrho_h).$$

## *3.4 De Giorgi's Variational and Geodesic Interpolations*

We fix $\varrho_0 \in \mathcal{P}^{ac}(X)$ and a time step $h > 0$, and set $\varrho_0^h = \varrho_0$. We consider $\varrho_h = \varrho_h^h \in P^{ac}(X)$ the (unique) minimizer of $\Phi(\tau,\varrho_0,\cdot)$ provided by Theorem 3.1.1, and we interpolate between $\varrho_0^h$ and $\varrho_h^h$ along paths minimizing the action $\mathcal{A}_h$: By [9, Theorem 5.1] there exists a the unique solution $\bar\varrho_s^h \in \mathcal{P}^{ac}(X)$ of

$$\mathcal{C}_s\big(\varrho_0^h,\bar\varrho_s^h\big) + \mathcal{C}_{h-s}\big(\bar\varrho_s^h,\varrho_h^h\big) = \mathcal{C}_h\big(\varrho_0^h,\varrho_h^h\big),$$

which is also given by

$$\bar{\varrho}_s^h \mathcal{L}^d := (T_h^s)_\# \varrho_0 \mathcal{L}^d, \qquad 0 \leq s \leq h.$$

Moreover [9, Theorem 5.1] ensures that $T_h^s$ is invertible $\bar{\varrho}_s^h$-a.e., so that in particular there exists a unique vector field $V_s^h$ defined on the measure theoretical support of $\bar{\varrho}_s^h$ such that

$$V_s^h(T_h^s) = \partial_s T_h^s \qquad \bar{\varrho}_s\text{-a.e.}$$

Recall that by Lemma 3.2.4(i)

$$\|\nabla u_h\|_{L^\infty(X)} \leq \|\nabla c_h\|_{L^\infty(X \times X)}.$$

Exploiting Equation (3.3.3) and the fact that $\partial_s \mathbf{\Phi}^L$ maps bounded subsets of $\mathbb{R}^d \times \mathbb{R}^d$ onto bounded subsets $\mathbb{R}^d \times \mathbb{R}^d$, we obtain that

$$\sup_{0 \leq s \leq h} \|\partial_s T_h^s\|_{L^\infty(\bar{\varrho}_s)} < +\infty.$$

Therefore,

$$\sup_{0 \leq s \leq h} \|V_s^h\|_{L^\infty(\bar{\varrho}_s)} < +\infty.$$

Finally a direct computation gives that

$$\partial_s \bar{\varrho}_s^h + \operatorname{div}(\bar{\varrho}_s^h V_s^h) = 0 \tag{3.4.1}$$

in the sense of distribution on $(0, h) \times \mathbb{R}^d$. Observe that $\bar{\varrho}_0^h = \varrho_0$ and $\bar{\varrho}_h^h = \varrho_h^h$.

**Remark 3.4.1** *Note that although the range of $T_h$ is contained in $X$, that of $T_h^s$ may fail to be in that set.*

We set

$$\varrho_s := \operatorname{argmin}\Big\{\mathcal{C}_s(\varrho_0, \varrho) + \mathcal{U}(\varrho) : \ \varrho \in \mathcal{P}^{ac}(X)\Big\}, \qquad 0 \leq s \leq h.$$

In a metric space $(\mathcal{S}, \operatorname{dist})$ with $s\mathcal{C}_s = \operatorname{dist}^2$, the interpolation $s \mapsto \varrho_s$ is due to De Giorgi [8] (cf. also [3, 7]).

**Theorem 3.4.2** *We have the energy inequality*

$$\mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_h) \geq \int_0^h \mathrm{d}s \int_{\mathbb{R}^d} L(\mathbf{y}, V_s)\bar{\varrho}_s \,\mathrm{d}\mathbf{y} + \int_0^h \mathrm{d}s \int_X H\left(\mathbf{x}, -\nabla\left[U'(\varrho_s)\right]\right) \varrho_s \,\mathrm{d}\mathbf{x}.$$

Proof We combine Lemmas 3.3.1 and 3.3.2 with Remark 3.3.3 to conclude that

$$\begin{aligned}
\int_0^h \mathrm{d}t \int_X H\left(\mathbf{x}, -\nabla\left[U'(\varrho_s)\right]\right) \varrho_s \,\mathrm{d}\mathbf{x} \;&\leq\; \phi_0(\varrho_0) - \phi_h(\varrho_h) \\
&=\; \mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_h) - \mathcal{C}_h(\varrho_0, \varrho_h) \\
&=\; \mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_h) - \int_0^h \mathrm{d}t \int_X L(T_h^s(\mathbf{x}), \partial_s T_h^s(\mathbf{x}))\varrho_0(\mathbf{x}) \,\mathrm{d}\mathbf{x} \\
&=\; \mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_h) - \int_0^h \mathrm{d}t \int_{\mathbb{R}^d} L(\mathbf{y}, V_s(\mathbf{y}))\bar{\varrho}_s(\mathbf{y}) \,\mathrm{d}\mathbf{y}.
\end{aligned}$$

(We remark that the last integral has to be taken on the whole $\mathbb{R}^d$, as we do not know in general that the measures $\bar{\varrho}_s$ are concentrated on $X$, cf. Remark 3.4.1). □

We now iterate the argument above: Lemma 3.2.5 ensures existence of a sequence $\{\varrho_{kh}^h\}_{k=0}^\infty \subset \mathcal{P}^{ac}(X)$ such that

$$\varrho_{(k+1)h}^h := \operatorname{argmin}\Big\{\mathcal{C}_h(\varrho_{kh}^h, \varrho) + \mathcal{U}(\varrho) \;:\; \varrho \in \mathcal{P}^{ac}(X)\Big\}.$$

As above, we define

$$\varrho_{kh+s}^h := \operatorname{argmin}\Big\{\mathcal{C}_s(\varrho_{kh}^h, \varrho) + \mathcal{U}(\varrho) \;:\; \varrho \in \mathcal{P}_1^{ac}\Big\}, \qquad 0 \leq s \leq h.$$

By the same argument applied to $(\varrho_{kh}, \varrho_{(k+1)h})$ in place of $(\varrho_0, \varrho_h)$, we obtain a unique map $T_{kh} : X \to X$ such that

$$(\mathrm{id} \times T_{kh})_\#(\varrho_{kh}\mathcal{L}^d) \in \Gamma_h(\varrho_{kh}, \varrho_{(k+1)h}).$$

Moreover, for $s \in (0, h)$ we define $\bar{\varrho}_{kh+s}^h$ to be the interpolation along paths minimizing the action $\mathcal{A}_h$, that is $\bar{\varrho}_{kh+s}$ is the unique solution of

$$\mathcal{C}_s\big(\varrho_{kh}^h, \bar{\varrho}_{kh+s}^h\big) + \mathcal{C}_{h-s}\big(\bar{\varrho}_{kh+s}^h, \varrho_{(k+1)h}^h\big) = \mathcal{C}_h\big(\varrho_{kh}^h, \varrho_{(k+1)h}^h\big).$$

We denote by $(u^h_{kh+s}, v^h_{kh+s})$ the solution to the dual problem provided by Lemma 3.2.4, and we define $V^h_{kh+s}$ by

$$V^h_{kh+s}(T^s_{kh}) = \partial_s T^s_{kh},$$

where $T^s_{kh}$ is the interpolation map:

$$(T^s_{kh})_\# \varrho^h_{kh} = \varrho^h_{kh+s}.$$

**Corollary 3.4.3** *For $h > 0$, for any $j \leq k \in \mathbb{N}$, we have*

$$\mathcal{U}(\varrho^h_{jh}) - \mathcal{U}(\varrho^h_{kh}) \geq \int_{jh}^{kh} \mathrm{d}s \int_{\mathbb{R}^d} L(\mathbf{y}, V^h_s)\bar{\varrho}^h_s \,\mathrm{d}\mathbf{y} + \int_{jh}^{kh} \mathrm{d}s \int_X H\big(\mathbf{x}, -\nabla[U'(\varrho^h_s)]\big)\varrho^h_s \,\mathrm{d}\mathbf{x}.$$

Proof The proof is a direct consequence of Theorem 3.4.2. □

## 3.5 Stability Property and Existence of Solutions

We fix $T > 0$ and want to prove existence of solutions to Equation (1.0.1) on $[0, T]$. Recall that by Lemma 2.4.4(i) $\mathcal{C}_s(\varrho, \varrho) \leq 0$ for any $s \geq 0$, $\varrho \in \mathcal{P}^{ac}_1$. This, together with the definition of $\varrho^h_{kh+s}$, yields

$$\mathcal{C}_h\big(\varrho^h_{kh}, \varrho^h_{kh+s}\big) + \mathcal{U}\big(\varrho^h_{kh+s}\big) \leq \mathcal{U}\big(\varrho^h_{kh}\big), \qquad 0 \leq s \leq h.$$

By adding the above inequality over $k \in \mathbb{N}$, invoking Remark 2.2.2 we get

$$\sum_{k=0}^{\infty} \mathcal{C}_h\big(\varrho^h_{kh}, \varrho^h_{(k+1)h}\big) \leq \mathcal{U}\big(\varrho^h_0\big) - \liminf_{n \to +\infty} \mathcal{U}\big(\varrho^h_{nh}\big) \leq \mathcal{U}\big(\varrho^h_0\big) + |a| + |b|\mathcal{L}^d(X). \tag{3.5.1}$$

We also recall that $v^h_t : X \to \mathbb{R}$ is a Lipschitz function (cf. Lemma 3.2.4(i)) which satisfies $v^h_t = -U'(\varrho^h_t)$, so that setting

$$\beta^h_t := U^*(-v^h_t) = P(\varrho^h_t)$$

we have

$$\varrho^h_t \nabla[U'(\varrho^h_t)] = -(U^*)'(-v^h_t)\nabla v^h_t = \nabla[U^*(-v^h_t)] = \nabla[P(\varrho^h_t)] = \nabla\beta^h_t \quad \mathcal{L}^d\text{-a.e.} \tag{3.5.2}$$

We start with the following:

**Lemma 3.5.1** *We have*

$$\mathcal{U}(\varrho_t^h) \le \mathcal{U}(\varrho_0) + A_* t. \tag{3.5.3}$$

*Moreover, for any $K > 0$ there exists a constant $C(K) > 0$ such that, for any $h \in (0,1]$,*

$$W_1(\bar\varrho_t^h, \varrho_t^h) \le \frac{C_0}{K} + 2\frac{A_* + C(K)}{K} h \qquad \forall\, t \in [0,T], \tag{3.5.4}$$

$$W_1(\bar\varrho_t^h, \varrho_{kh}^h) \le \frac{C_0}{K} + \frac{A_* + C(K)}{K} h \qquad \forall\, t \in [kh,(k+1)h],\ k \in \mathbb{N}, \tag{3.5.5}$$

$$W_1\big(\varrho_t^h, \varrho_s^h\big) \le \frac{C_0}{K} + 2\frac{A_* + C(K)}{K}[(t-s)+h] \qquad \forall\, 0 \le s \le t, \tag{3.5.6}$$

$$W_1\big(\bar\varrho_t^h, \bar\varrho_s^h\big) \le \frac{C_0}{K} + \frac{A_* + C(K)}{K}[(t-s)+h] \qquad \forall\, 0 \le s \le t. \tag{3.5.7}$$

*Here $C_0$ is a positive constant independent of $t$, $K$, and $h \in (0,1]$.*

Proof Let $t \in [kh,(k+1)h]$ for some $k \in \mathbb{N}$. Then by Lemma 2.4.4(ii) we have

$$\mathcal{U}(\varrho_t^h) - A_*(t-kh) \le \mathcal{U}(\varrho_t^h) + \mathcal{C}_{t-kh}(\varrho_{kh}^h, \varrho_t^h) \le \mathcal{U}(\varrho_{kh}^h). \tag{3.5.8}$$

In particular

$$\mathcal{U}(\varrho_{(k+1)h}^h) \le \mathcal{U}(\varrho_{kh}^h) + A_* h \tag{3.5.9}$$

for all $k \in \mathbb{N}$. Using (3.5.9) repeatedly, we obtain that

$$\begin{aligned}
\mathcal{U}(\varrho_t^h) &\le \mathcal{U}(\varrho_{kh}^h) + A_*(t-kh) \\
&\le \mathcal{U}(\varrho_{(k-1)h}^h) + A_*[h + (t-kh)] \\
&\ \vdots \\
&\le \mathcal{U}(\varrho_0) + A_*[kh + (t-kh)] \\
&= \mathcal{U}(\varrho_0) + A_* t.
\end{aligned}$$

This proves Equation (3.5.3). Now, since $\mathcal{C}_h \le \mathcal{C}_{t-kh}$ (cf. Lemma 2.4.4(i)), we have

$$\mathcal{C}_h(\varrho_{kh}^h, \varrho_t^h) \le \mathcal{U}(\varrho_{kh}^h) - \mathcal{U}(\varrho_t^h) \qquad \forall\, t \in [kh,(k+1)h].$$

which combined with Equation (3.5.8) and Remark 2.2.2 gives

$$\mathcal{C}_h(\varrho^h_{kh}, \varrho^h_t) \le \mathcal{U}(\varrho_0) + A_* h + |a| + |b|\mathcal{L}^d(X) \qquad \forall\, t \in [kh, (k+1)h]. \tag{3.5.10}$$

Moreover, as $\varrho^h_{kh} = \bar\varrho^h_{kh}$ for any $k \in \mathbb{N}$, using again Lemma 2.4.4(ii) we get

$$\begin{aligned}
\mathcal{C}_h(\varrho^h_{kh}, \varrho^h_{(k+1)h}) &= \mathcal{C}_{t-kh}(\varrho^h_{kh}, \bar\varrho^h_t) + \mathcal{C}_{(k+1)h-t}(\bar\varrho^h_t, \varrho^h_{(k+1)h}) \\
&\ge \mathcal{C}_{t-kh}(\varrho^h_{kh}, \bar\varrho^h_t) - A_* h \\
&\ge \mathcal{C}_h(\varrho^h_{kh}, \bar\varrho^h_t) - A_* h.
\end{aligned}$$

Thanks to Lemma 2.4.4(ii)-(iii), for any $K > 0$ there exists a constant $C(K) > 0$ such that

$$W_1(\varrho^h_{kh}, \varrho^h_t) \le \frac{1}{K}\mathcal{C}_h(\varrho^h_{kh}, \varrho^h_t) + \frac{A_* + C(K)}{K} h,$$
$$W_1(\varrho^h_{kh}, \bar\varrho^h_t) \le \frac{1}{K}\mathcal{C}_h(\varrho^h_{kh}, \bar\varrho^h_t) + \frac{A_* + C(K)}{K} h,$$

which combined with the above estimates and the triangle inequality proves Equations (3.5.4) and (3.5.5).

Finally, to prove Equations (3.5.6) and (3.5.7), we observe that Equation (3.5.1) combined with Lemma 2.4.4(iii) gives

$$\begin{aligned}
W_1\big(\varrho^h_{Nh}, \varrho^h_{Mh}\big) &\le \sum_{j=M}^{N-1} W_1\big(\varrho^h_{(j+1)h}, \varrho^h_{jh}\big) \\
&\le \frac{1}{K}\sum_{j=M}^{N-1} \mathcal{C}_h(\varrho^h_{(j+1)h}, \varrho^h_{jh}) + \frac{A_* + C(K)}{K} h(N-M) \\
&\le \frac{1}{K}\big[\mathcal{U}\big(\varrho^h_0\big) + |a| + |b|\mathcal{L}^d(X)\big] + \frac{A_* + C(K)}{K} h(N-M).
\end{aligned}$$

Combining this estimate with Equations (3.5.4) and (3.5.5), we obtain the desired result. □

We can now prove the compactness of our discrete solutions.

**Proposition 3.5.2** *There exists a sequence $h_n \to 0$, a density $\varrho \in \mathcal{P}^{ac}([0,T] \times X)$, and a Borel function $V : [0,T] \times X \to \mathbb{R}^d$ such that:*

*(i) The measure valued curves* $t \mapsto \varrho_t^{h_n} \in \mathcal{P}^{ac}(X)$ *and* $t \mapsto \bar{\varrho}_t^{h_n} \in \mathcal{P}^{ac}(\mathbb{R}^d)$ *converge uniformly (locally in time) to the curve* $t \mapsto \varrho_t := \varrho(t, \cdot)$, *and this curve is uniformly continuous with respect to the narrow topology. Moreover* $w^*$-$\lim_{t\to 0^+} \varrho_t = \varrho_0$.

*(ii) The vector-valued measures* $\bar{\varrho}_t^n(\mathbf{x})V_t^{h_n}(\mathbf{x})\mathrm{d}\mathbf{x}\,\mathrm{d}t$ *converge narrowly to the vector-valued measure* $\varrho_t(\mathbf{x})V_t(\mathbf{x})\mathrm{d}\mathbf{x}\,\mathrm{d}t$, *where* $V_t := V(t, \cdot)$.

*(iii)* $\partial_t \varrho_t + \mathrm{div}(\varrho_t V_t) = 0$ *holds on* $(0,T) \times X$ *in the sense of distribution.*

Proof By Equations (3.5.6) and (3.5.7), as $K > 0$ is arbitrary, it is easy to see that the curves $t \mapsto \varrho_t^h$ and $t \mapsto \bar{\varrho}_t^h$ are equicontinuous with respect to the 1-Wasserstein distance. Since bounded sets with respect to $W_1$ are precompact with respect to the narrow topology on $\mathbb{R}^d$ (cf. for instance [3, Chapter 7]), by the Arzelà-Ascoli Theorem we can find a sequence $h_n \to 0$ such that $t \mapsto \varrho_t^{h_n} \in \mathcal{P}^{ac}(X)$ and $t \mapsto \bar{\varrho}_t^{h_n} \in \mathcal{P}^{ac}(\mathbb{R}^d)$ converge uniformly (locally in time) to a narrow-continuous curve $t \mapsto \mu_t \in \mathcal{P}(X)$ (which is the same for both $\varrho_t^{h_n}$ and $\bar{\varrho}_t^{h_n}$ thanks to Equation (3.5.4)). Moreover $t \mapsto \mu_t$ is supported in $X$ as so is $\varrho_t^{h_n}$, and the initial condition $w^*$-$\lim_{t\to 0^+} \bar{\varrho}_t^{h_n} = \varrho_0$ holds in the limit.

Concerning the vector-valued measure $V_t^h \varrho_t^h$, recalling that $H \geq 0$, using Corollary 3.4.3 and Remark 2.2.2 we have

$$\int_0^T \mathrm{d}t \int_{\mathbb{R}^d} L(\mathbf{x}, V_t^h)\bar{\varrho}_t^h \,\mathrm{d}\mathbf{x} \leq \mathcal{U}(\varrho_0) + |a| + |b|\mathcal{L}^d(X).$$

By (L3) this gives

$$\int_0^T \mathrm{d}t \int_{\mathbb{R}^d} \theta(|V_t^h|)\bar{\varrho}_t^h \,\mathrm{d}\mathbf{x} \leq \mathcal{U}(\varrho_0) + |a| + |b|\mathcal{L}^d(X) + A_* T =: C_1.$$

The above inequality, together with the superlinearity of $\theta$ and the uniform convergence of $\bar{\varrho}^h$ to $\mu_t$, implies easily that the vector-valued measure $V_t^h \bar{\varrho}^h$ have a limit point $\lambda$ that is concentrated on $[0,T] \times X$. Moreover, the superlinearity and the convexity of

$\theta$ insure that $\lambda \ll \mu$, and there exists a $\mu$-measurable vector field $V : [0,T] \times \overline{X} \to \mathbb{R}^d$ such that $\lambda = V\mu$, and

$$\int_0^T \mathrm{d}t \int_X \theta(|V_t|)\, d\mu_t \le C_1.$$

To conclude the proof of (i) and (ii), we have to show that $\mu \ll \mathcal{L}^{d+1}$. We observe that thanks to Equation (3.5.3)

$$\int_0^T \mathrm{d}t \int_X U(\varrho_t^{h_n})\,\mathrm{d}\mathbf{x} = \int_0^T \mathcal{U}(\varrho_t^{h_n})\,\mathrm{d}t \le T\mathcal{U}(\varrho_0) + A_* \frac{T^2}{2},$$

so that by the superlinearity of $U$ any limit point of $\varrho^{h_n}$ is absolutely continuous. Hence $\mu = \varrho\mathcal{L}^d$, and (i) and (ii) are proved.

Finally, from Equation (3.4.1) we deduce that

$$\partial_t \bar{\varrho}_t^{h_n} + \mathrm{div}(\bar{\varrho}_t^{h_n} V_t^{h_n}) = 0 \qquad \text{on } (0,T) \times X \tag{3.5.11}$$

in the sense of distributions, so that (iii) follows taking the limit as $n \to +\infty$. □

**Remark 3.5.3** *In the proof of the lemma above we have seen that each curve* $t \mapsto \bar{\varrho}_t^{h_n}, t \mapsto \varrho_t^{h_n} \in \mathcal{P}^{ac}(X)$ *admits a representative which is uniformly continuous on* $[0,T]$ *with respect to the weak* topology. We will always implicitly refer to such a representative, so that in particular* $\bar{\varrho}_t^{h_n}$ *is well-defined for every* $t \in [0,T]$*. Moreover we conclude that both* $\bar{\varrho}_t^{h_n}$ *and* $\varrho_t^{h_n}$ *converge weakly* to* $\varrho_t$ *for every fixed* $t \in [0,T]$*.*

We are now ready to prove the following existence result. To simplify the notation, given two nonnegative functions $f$ and $g$, we write $f \gtrsim g$ if there exist two nonnegative constants $c_0, c_1$ such that $c_0 f + c_1 \ge g$. If both $\gtrsim$ and $\lesssim$ hold, we write $f \sim g$.

**Theorem 3.5.4** *Let* $X \subset \mathbb{R}^d$ *be an open bounded set whose boundary is of zero Lebesgue measure, and assume that* $H$ *satisfies (H1), (H2) and (H3). Let* $\varrho_t$ *and* $V_t$ *be as in Proposition 3.5.2. Then we have*

$$P(\varrho_t) \in L^1(0,T; W^{1,1}(X)),$$

$\nabla[P(\varrho_t)]$ *is absolutely continuous with respect to* $\varrho_t$, *and*

$$\mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_T) \geq \int_0^T \mathrm{d}t \int_X \Big[L\big(\mathbf{x}, V_t\big) + H\big(\mathbf{x}, -\varrho_t^{-1}\nabla[P(\varrho_t)]\big)\Big]\varrho_t \,\mathrm{d}\mathbf{x}. \tag{3.5.12}$$

*Furthermore, if* $\theta(t) \sim t^\alpha$ *for some* $\alpha > 1$ *and* $U$ *satisfies the doubling condition*

$$U(t+s) \leq C(U(t) + U(s) + 1) \qquad \forall\, t, s \geq 0, \tag{3.5.13}$$

*then* $\varrho_t \in AC_\alpha(0, T; \mathcal{P}_\alpha^{ac}(X))$,

$$\mathcal{U}(\varrho_{T_1}) - \mathcal{U}(\varrho_{T_2}) = -\int_{T_1}^{T_2} \mathrm{d}t \int_X \big\langle \nabla[P(\varrho_t)], V_t\big\rangle \,\mathrm{d}\mathbf{x}, \tag{3.5.14}$$

*for* $T_1, T_2 \in [0, T]$. *In particular*

$$V_t(\mathbf{x}) = \nabla_{\mathbf{p}} H\big(\mathbf{x}, -\varrho_t^{-1}\nabla[P(\varrho_t)]\big) \qquad \varrho_t\text{-}a.e.,$$

*and* $\varrho_t$ *is a distributional solution of Equation* (1.0.1) *starting from* $\varrho_0$.

*Suppose in addition that (H4) holds. Then we have the maximum principle: if* $\varrho_0 \leq M$ *for some* $M \geq 0$, *then* $\varrho_t \leq M$ *for every* $t \in (0, T)$.

Proof The maximum principle is a direct consequence of Theorem 3.1.1. We first remark that the last part of the statement is a simple consequence of Equations (3.5.12) and (3.5.14) combined with Proposition 3.5.2(i)-(iii). So it suffices to prove to prove Equations (3.5.12) and (3.5.14).

We first prove Inequality (3.5.12). Corollary 3.4.3 implies that, if $T \in [kh_n, (k+1)h_n]$ for some $k \in \mathbb{N}$, since $L \geq -A_*$ and $H \geq 0$ we have

$$\mathcal{U}(\varrho_0^{h_n}) - \mathcal{U}(\varrho_{(k+1)h_n}^{h_n}) \geq \int_0^T \mathrm{d}t \int_{\mathbb{R}^d} \Big[L(\mathbf{x}, V_t^{h_n})\bar{\varrho}_t^{h_n} + H\big(\mathbf{x}, -\nabla[U'(\varrho_t^{h_n})]\big)\varrho_t^{h_n}\Big] \,\mathrm{d}\mathbf{x} - A_* h_n.$$

We now consider two continuous functions $w, \bar{w} : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ with compact support. Then

$$\begin{aligned}\int_0^T \mathrm{d}t \int_{\mathbb{R}^d} \Big[L(\mathbf{x}, V_t^{h_n})\bar{\varrho}_t^{h_n} &+ H\big(\mathbf{x}, -\nabla[U'(\varrho_t^{h_n})]\big)\varrho_t^{h_n}\Big] \,\mathrm{d}\mathbf{x} \\ &\geq \int_0^T \mathrm{d}t \int_{\mathbb{R}^d} \Big[\langle V_t^{h_n}, \bar{w}(t, \mathbf{x})\rangle \bar{\varrho}_t^{h_n} - H(\mathbf{x}, \bar{w}(t, \mathbf{x}))\bar{\varrho}_t^{h_n}\Big] \,\mathrm{d}\mathbf{x} \\ &+ \int_0^T \mathrm{d}t \int_X \Big[\langle -\nabla[U'(\varrho_t^{h_n})], w(t, \mathbf{x})\rangle \varrho_t^{h_n} - L(\mathbf{x}, w(t, \mathbf{x}))\varrho_t^{h_n}\Big] \,\mathrm{d}\mathbf{x}.\end{aligned}$$

By Proposition 3.5.2(i)-(ii) we immediately get

$$\lim_{n\to+\infty}\int_0^T \mathrm{d}t\int_{\mathbb{R}^d}\Big[\langle V_t^{h_n},\bar w(t,\mathbf{x})\rangle\bar\varrho_t^{h_n}-H(\mathbf{x},\bar w(t,\mathbf{x}))\bar\varrho_t^{h_n}\Big]\,\mathrm{d}\mathbf{x}$$

$$=\int_0^T \mathrm{d}t\int_{\mathbb{R}^d}\Big[\langle V_t,\bar w(t,\mathbf{x})\rangle\varrho_t-H(\mathbf{x},\bar w(t,\mathbf{x}))\varrho_t\Big]\,\mathrm{d}\mathbf{x},$$

so that taking the supremum among all continuous functions $\bar w:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ with compact support we obtain

$$\liminf_{n\to+\infty}\int_0^T \mathrm{d}t\int_{\mathbb{R}^d}L(\mathbf{x},V_t^{h_n})\bar\varrho_t^{h_n}\,\mathrm{d}\mathbf{x}\ge\int_0^T \mathrm{d}t\int_{\mathbb{R}^d}L(\mathbf{x},V_t)\varrho_t\,\mathrm{d}\mathbf{x}.$$

As for the other term, we observe that, by Remark 2.2.2, as $L\ge -A_*$ we have that

$$\int_0^T \mathrm{d}t\int_{\mathbb{R}^d}H\big(\mathbf{x},-\nabla[U'(\varrho_t^{h_n})]\big)\varrho_t^{h_n}\,\mathrm{d}\mathbf{x}$$

is uniformly bounded. In particular, since (by (H3))

$$H(\mathbf{x},\mathbf{p})\ge|\mathbf{p}|-C_1$$

for some constant $C_1$ and Equation (3.5.2), we arrive at the conclusion that

$$\int_0^T \mathrm{d}t\int_{\mathbb{R}^d}|\nabla[P(\varrho_t^{h_n})]|\,\mathrm{d}\mathbf{x}=\int_0^T \mathrm{d}t\int_{\mathbb{R}^d}|\nabla[U'(\varrho_t^{h_n})]|\varrho_t^{h_n}\,\mathrm{d}\mathbf{x}$$

is uniformly bounded. This implies that, up to a subsequence, the vector-valued measures $\nabla[P(\varrho_t^{h_n})]\,\mathrm{d}\mathbf{x}\,\mathrm{d}t$ converges weakly to a measure $\nu$ of finite total mass. Therefore we obtain

$$\begin{aligned}
+\infty &> \liminf_{n\to\infty}\int_0^T \mathrm{d}t\int_{\mathbb{R}^d}H\big(\mathbf{x},-\nabla[U'(\varrho_t^{h_n})]\big)\varrho_t^{h_n}\,\mathrm{d}\mathbf{x}\\
&\ge \lim_{n\to+\infty}\int_0^T \mathrm{d}t\int_X\Big[\langle-\nabla[U'(\varrho_t^{h_n})],w(t,\mathbf{x})\rangle\varrho_t^{h_n}-L(\mathbf{x},w(t,\mathbf{x}))\varrho_t^{h_n}\Big]\,\mathrm{d}\mathbf{x}\\
&= \int_0^T \mathrm{d}t\int_X -\langle w(t,\mathbf{x}),\nu(\mathrm{d}t,\mathrm{d}\mathbf{x})\rangle-\int_0^T \mathrm{d}t\int_X L(\mathbf{x},w(t,\mathbf{x}))\varrho_t\,\mathrm{d}\mathbf{x}.
\end{aligned}$$

Since $w$ is arbitrary, we easily get that the measure $\nu(\mathrm{d}t,\mathrm{d}\mathbf{x})$ is absolutely continuous with respect to $\varrho_t\,\mathrm{d}\mathbf{x}\,\mathrm{d}t$, so that $\nu(\mathrm{d}t,\mathrm{d}\mathbf{x})=e_t(\mathbf{x})\varrho_t(\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}t$ for some Borel function $e:[0,T]\times X\to\mathbb{R}^d$. We now observe that by Fatou's Lemma we also have

$$\int_0^T\left(\liminf_{n\to+\infty}\int_{\mathbb{R}^d}|\nabla[P(\varrho_t^{h_n})]|\,\mathrm{d}\mathbf{x}\right)\mathrm{d}t<+\infty,$$

which gives

$$\liminf_{n\to+\infty}\int_{\mathbb{R}^d}|\nabla[P(\varrho_t^{h_n})]|\,\mathrm{d}\mathbf{x}<+\infty \qquad \text{for } t\in[0,T]\setminus\mathcal{N}, \tag{3.5.15}$$

with $\mathcal{L}^1(\mathcal{N})=0$. This fact easily implies that, for any $t\in[0,T]\setminus\mathcal{N}$, there exists a subsequence $\varrho_t^{h_{n_k(t)}}$ such that

$$\liminf_{n\to+\infty}\int_{\mathbb{R}^d}|\nabla[P(\varrho_t^{h_n})]|\,\mathrm{d}\mathbf{x}=\lim_{k\to+\infty}\int_{\mathbb{R}^d}|\nabla[P(\varrho_t^{h_{n_k(t)}})]|\,\mathrm{d}\mathbf{x},$$

and $P(\varrho_t^{h_{n_k(t)}})$ converges weakly in $BV(X)$ and $\mathcal{L}^d$-a.e. to a function $\beta_t$. As a consequence $D\beta_t=e_t\varrho_t\mathcal{L}^d$, so that $\beta_t\in W^{1,1}(X)$. Since $Q$ is continuous, we deduce that $\varrho_t^{h_{n_k(t)}}=Q\big(P(\varrho_t^{h_{n_k(t)}})\big)$ converges $\mathcal{L}^d$-a.e. to $Q(\beta_t)$. Recalling that $\varrho_t^{h_{n_k(t)}}$ also converges weakly to $\varrho_t$, we obtain $Q(\beta_t)=\varrho_t$, that is $\beta_t=P(\varrho_t)$. Moreover, from the equality $\nabla\beta_t=e_t\varrho$, we get $\nabla[P(\varrho_t)]=e_t\varrho_t$. We have proved that $P(\varrho_t)\in L^1(0,T;W^{1,1}(X))$ and $\nabla[P(\varrho_t)]$ is absolutely continuous with respect to $\varrho_t$. Finally $\varrho_{(k+1)h_n}^{h_n}$ converges weakly* to $\varrho_T$, and the term $\mathcal{U}(\varrho_T^{h_n})$ is lower-semicontinuous under weak* convergence, and this concludes the proof of Equation (3.5.12).

We now prove Equation (3.5.14). Let us recall that by assumption $\theta\sim t^\alpha$, which implies that $L(\mathbf{x},\mathbf{v})\gtrsim|\mathbf{v}|^\alpha$ and $H(\mathbf{x},\mathbf{p})\gtrsim|\mathbf{p}|^{\alpha'}$, where $\alpha'=\alpha/(\alpha-1)$. Let us observe that, thanks to Equation (3.5.12), we have

$$+\infty>\int_0^T\mathrm{d}t\int_X L(\mathbf{x},V_t)\varrho_t\,\mathrm{d}\mathbf{x}\gtrsim\int_0^T\|V_t\|^\alpha_{L^\alpha(\varrho_t)}\,\mathrm{d}t \tag{3.5.16}$$

and

$$+\infty>\int_0^T\mathrm{d}t\int_X H(\mathbf{x},-e_t)\varrho_t\,\mathrm{d}\mathbf{x}\gtrsim\int_0^T\|e_t\|^{\alpha'}_{L^{\alpha'}(\varrho_t)}\,\mathrm{d}t. \tag{3.5.17}$$

Since

$$\partial_t\varrho_t+\operatorname{div}\big(\varrho_tV_t\big)=0,$$

Equation (3.5.16) implies that the curve $t\to\varrho_t$ is absolutely continuous with values in the $\alpha$-Wasserstein space $P_\alpha(X)$, and we denote by $\bar{V}$ its velocity of minimal norm (cf. [3, Chapter 8]). Moreover, thanks to Equation (3.5.17), $e_t\in L^{\alpha'}(\varrho_t)$ for a.e. $t\in(0,T)$.

Denoting by $|\varrho'|$ the metric derivative of the curve $t \to \varrho_t$ (with respect to the $\alpha$-Wasserstein distance, cf. Equation (2.1.9)), by Equation (3.5.16) and [3, Theorem 8.3.1] we have that

$$|\varrho'|(t) \leq \|\bar{V}_t\|_{L^\alpha(\varrho_t)} \leq \|V_t\|_{L^\alpha(\varrho_t)} < +\infty. \tag{3.5.18}$$

Since $e_t\varrho_t = \nabla P(\varrho_t)$ with $P(\varrho_t) \in W^{1,1}(X)$ for a.e. $t$, we can apply [3, Theorem 10.4.6] to conclude that, for $\mathcal{L}^1$-a.e. $t$, $\mathcal{U}$ has a finite slope at $\varrho\mathcal{L}^d$, $|\partial\mathcal{U}|(\varrho_t) = \|e_t\|_{L^{\alpha'}(\varrho_t)}$ and $e_t = \partial^o\mathcal{U}(\varrho_t)$. The last statement means that $e_t$ is the element of minimal norm of the convex set $\partial\mathcal{U}(\varrho_t)$ and so, it belongs to the closure of $\{\nabla\varphi : \varphi \in C_c^\infty(X)\}$ in $L^{\alpha'}(\varrho_t)$. Let $\Lambda \subset (0,T)$ be the set of $t$ such that

(a) $\partial\mathcal{U}(\varrho_t) \neq \emptyset$;

(b) $\mathcal{U}$ is approximately differentiable at $t$;

(c) (8.4.6) of [3] holds.

We use equations (3.5.17), (3.5.18), and the fact that $|\partial\mathcal{U}|(\varrho_t) = \|e_t\|_{L^{\alpha'}(\varrho_t)}$ for $\mathcal{L}^1$-a.e. $t \in (0,T)$, to conclude that

$$\int_0^T |\partial\mathcal{U}|(\varrho_t)|\varrho'|(t)\mathrm{d}t \leq \frac{1}{\alpha'}\int_0^T \|e_t\|^{\alpha'}_{L^{\alpha'}(\varrho_t)}\mathrm{d}t + \frac{1}{\alpha}\int_0^T \|V_t\|^{\alpha}_{L^{\alpha}(\varrho_t)}\mathrm{d}t < +\infty. \tag{3.5.19}$$

By [3, Proposition 9.3.9] $\mathcal{U}$ is convex along $\alpha$-Wasserstein geodesics, and so exploiting Equation (3.5.19) and invoking [3, Proposition 10.3.18] we conclude that $\mathcal{L}^1\big((0,T) \setminus \Lambda\big) = 0$ and $t \mapsto \mathcal{U}(\varrho_t)$ is absolutely continuous. Thus its pointwise, distributional, and approximate derivatives coincide almost everywhere, and by [3, Proposition 10.3.18] and the fact that $e_t \in \partial\mathcal{U}(\varrho_t)$ we get

$$\frac{d}{\mathrm{d}t}\mathcal{U}(\varrho_t) = \int_X \langle e_t, \bar{V}_t\rangle \varrho_t \mathrm{d}\mathbf{x}. \tag{3.5.20}$$

Because $V$ and $\bar{V}$ are velocities for $\varrho$ we have

$$\int_X \langle \nabla\phi, V_t - \bar{V}_t\rangle \varrho_t \mathrm{d}\mathbf{x} = 0$$

for all $\phi \in C_c^\infty(X)$ for $\mathcal{L}^1$-a.e. $t \in (0,T)$, and since $e_t$ belongs to the closure of $\{\nabla\varphi : \varphi \in C_c^\infty(X)\}$ in $L^{\alpha'}(\varrho_t)$, by a density argument we conclude that

$$\int_X \langle e_t, V_t - \bar{V}_t\rangle \varrho_t \mathrm{d}\mathbf{x} = 0$$

for $\mathcal{L}^1$-a.e. $t \in (0,T)$. This, together with Equation (3.5.20), finally yields

$$\mathcal{U}(\varrho_{T_1}) - \mathcal{U}(\varrho_{T_2}) = -\int_{T_1}^{T_2} \mathrm{d}t \int_X \langle e_t, \bar{V}_t\rangle \varrho_t \mathrm{d}\mathbf{x} = -\int_{T_1}^{T_2} \mathrm{d}t \int_X \langle e_t, V_t\rangle \varrho_t \mathrm{d}\mathbf{x}, \tag{3.5.21}$$

as desired. □

**Remark 3.5.5** *If $\varrho_0$ is a general nonnegative integrable function on $X$ which does not necessarily have a unit mass, we can still prove existence of solutions to Equation* (1.0.1)*. Indeed, defining $c := \int_X \varrho_0 \,\mathrm{d}\mathbf{x}$, we consider $\varrho_t^c \in \mathcal{P}^{ac}(X)$ a solution of Equation* (1.0.1) *for the Hamiltonian*

$$H^c(\mathbf{x}, \mathbf{p}) := cH\left(\mathbf{x}, \frac{\mathbf{p}}{c}\right)$$

*and the internal energy $U^c(t) := U(ct)$, starting from $\varrho_0^c := \varrho_0/c$. Then $\varrho_t := c\varrho_t^c$ solves Equation* (1.0.1)*. Moreover, using this scaling argument also at a discrete level, we can also construct discrete solutions starting from $\varrho_0$.*

**Remark 3.5.6** *We believe that the above existence result could be extend to more general functions $\theta$ by introducing some Orlicz-type spaces as follows: for $\theta : [0,+\infty) \to [0,+\infty)$ convex, superlinear, and such that $\theta(0) = 0$, we define the Orlicz-Wasserstein distance*

$$\mathcal{W}_\theta(\mu_0, \mu_1) := \inf\left\{\lambda > 0 \,:\, \inf_{\gamma \in \Gamma(\mu_0,\mu_1)} \int_{X\times X} \theta\left(\frac{|\mathbf{x}-\mathbf{y}|}{\lambda}\right) \mathrm{d}\gamma(\mathbf{x}, \mathbf{y}) \le 1\right\}.$$

*We also define the Orlicz-type norm*

$$\|f\|_{\theta,\mu} := \inf\left\{\lambda > 0 \,:\, \int_X \theta\left(\frac{|f|}{\lambda}\right) \mathrm{d}\mu(\mathbf{x}) \le 1\right\}.$$

*It is not difficult to prove that the following dynamical formulation of the Orlicz-Wasserstein distance holds:*

$$\mathcal{W}_\theta(\mu_0, \mu_1) = \inf \left\{ \int_0^1 \|V_t\|_{\theta,\mu_t} \, \mathrm{d}t \, : \, \partial_t \mu_t + \mathrm{div}(\mu_t V_t) = 0 \right\}. \tag{3.5.22}$$

*Now, in order to prove the identity in* (3.5.14) *of the previous theorem in the case where* $\theta$ *does not necessarily behave as a power function, one should extend the results of [3] to a more general setting. We believe such an extension to be reachable although not straightforward.*

# CHAPTER IV

# EXISTENCE OF SOLUTIONS IN UNBOUNDED DOMAINS FOR WEAK TONELLI LAGRANGIANS

The aim of this part is to extend the existence result proved in the previous section to unbounded domains $X$, using an approximation argument where we construct our solutions in $X \cap B_m(0)$ for smoothed Lagrangians $L_m$, and then we let $m \to +\infty$. In order to be able to pass to the limit in the estimates and find a solution, we require that there exist $c > 0$ and $a \in (d/(d+1), 1)$ such that

$$U^-(t) = \max\{-U(t), 0\} \leq ct^a \qquad \forall\, t \geq 0. \tag{4.0.23}$$

The above assumption, together with (2.1.2) and (2.1.3), is satisfied by positive multiples of the following functions: $t \ln t$, or $t^\alpha$ with $\alpha > 1$. Under this additional assumption, we now prove some lemmas and propositions which easily allow us to construct our solution as a limit of solutions in bounded domains (cf. Section 4.3).

By Assumption (4.0.23), we can prove that if $M_1(\varrho)$ is finite, then $\int_{\mathbb{R}^d} U^-(\varrho)\, \mathrm{d}\mathbf{x}$ is finite, and so $\int_{\mathbb{R}^d} U(\varrho)\, \mathrm{d}\mathbf{x}$ is well-defined.

## 4.1 Lower-semicontinuity of $\mathcal{U}$

**Lemma 4.1.1** *There exists $C \equiv C(d, a)$ such that*

$$\mathcal{U}^-(\varrho) \leq C(M_1(\varrho)^a + 1).$$

*Consequently $\mathcal{U}(\varrho)$ is well defined whenever $M_1(\varrho)$ is finite. Furthermore $C$ can be chosen so that*

$$\int_{B_R(0)^c} U^-(\varrho)\, \mathrm{d}\mathbf{x} \leq C M_1(\varrho)^a R^{d(1-a)-a} \qquad \forall\, R \geq 0.$$

Proof We use Assumption (4.0.23) to obtain

$$\begin{aligned}
\int_{B_R(0)^c} U^-(\varrho)\,\mathrm{d}\mathbf{x} &\leq c\int_{B_R(0)^c} \varrho^a\,\mathrm{d}\mathbf{x} \qquad (4.1.1)\\
&= c\int_{B_R(0)^c} (|\mathbf{x}|\varrho)^a \frac{1}{|\mathbf{x}|^a}\,\mathrm{d}\mathbf{x}\\
&\leq c\Big(\int_{B_R(0)^c} |\mathbf{x}|\varrho\,\mathrm{d}\mathbf{x}\Big)^a\Big(\int_{B_R(0)^c} |\mathbf{x}|^{-a/(1-a)}\,\mathrm{d}\mathbf{x}\Big)^{1-a}\\
&\leq cM_1(\varrho)^a\Big(\int_R^\infty r^{\left(d-1-\frac{a}{1-a}\right)}\,\mathrm{d}r\Big)^{1-a}\\
&=: c_1(d,a)M_1(\varrho)^a R^{d(1-a)-a}. \qquad (4.1.2)
\end{aligned}$$

This proves the second statement of the lemma. Observing that

$$\begin{aligned}
\int_{B_R(0)} U^-(\varrho)\,\mathrm{d}\mathbf{x} &\leq c\int_{B_R(0)} \varrho^a\,\mathrm{d}\mathbf{x}\\
&\leq c\int_{B_R(0)} (1+\varrho)\,\mathrm{d}\mathbf{x}\\
&\leq c\big(\mathcal{L}^d(B_R(0))+1\big)\\
&=: \tilde{c}R^d + c. \qquad (4.1.3)
\end{aligned}$$

We use Equations (4.1.2) and (4.1.3) to conclude the proof. □

We now prove a lower-semicontinuity result.

**Proposition 4.1.2** *Suppose that $\{\varrho_n\}_{n\in\mathbb{N}} \subset \mathcal{P}_1^{ac}(\mathbb{R}^d)$ converges weakly in $L^1(\mathbb{R}^d)$ to $\varrho$, and that*

$$\sup_{n\in\mathbb{N}} M_1(\varrho_n) < +\infty.$$

*Then $\varrho \in \mathcal{P}_1^{ac}(\mathbb{R}^d)$, and*

$$\liminf_{n\to\infty} \mathcal{U}(\varrho_n) \geq \mathcal{U}(\varrho).$$

Proof The fact that $\varrho \in \mathcal{P}_1^{ac}(\mathbb{R}^d)$ follows from the lower-semicontinuity with respect to the weak $L^1$-topology of the first moment.

We now suppose without loss of generality that

$$\liminf_{n\to\infty} \mathcal{U}(\varrho_n) < +\infty.$$

Fix $\varepsilon > 0$. We have to prove that

$$\liminf_{n\to\infty} \mathcal{U}(\varrho_n) \geq \mathcal{U}(\varrho) - \varepsilon.$$

By Lemma 4.1.1 we can find $R > 0$ such that

$$\sup_{n\in\mathbb{N}} \int_{B_R(0)^c} U^-(\varrho_n)\,\mathrm{d}\mathbf{x} \leq \varepsilon. \tag{4.1.4}$$

By Lemma 2.2.3 and the fact that $U$ and $U^+ \geq 0$ are convex we get

$$\liminf_{n\to\infty} \int_{B_R(0)} U(\varrho_n)\,\mathrm{d}\mathbf{x} \geq \int_{B_R(0)} U(\varrho)\,\mathrm{d}\mathbf{x} \tag{4.1.5}$$

$$\liminf_{n\to\infty} \int_{B_R(0)^c} U^+(\varrho_n)\,\mathrm{d}\mathbf{x} \geq \int_{B_R(0)^c} U^+(\varrho)\,\mathrm{d}\mathbf{x} \tag{4.1.6}$$

Combining Equations (4.1.4), (4.1.5) and (4.1.6) we obtain

$$\liminf_{n\to\infty} \mathcal{U}(\varrho_n) \geq \mathcal{U}(\varrho) - \varepsilon.$$

This concludes the proof. □

We remark that, using the results above, it is easy to see that Lemma 3.2.3 holds assuming $\varrho \in \mathcal{P}_1^{ac}(\mathbb{R}^d)$.

## *4.2 Properties of Moments in the Unbounded Case*

Fix $T > 0$, and for any $h > 0$ we suppose that a sequence $\{\varrho_k^h\}_{0\leq k\leq T/h} \subset \mathcal{P}_1^{ac}(\mathbb{R}^d)$ is given such that

$$\mathcal{C}_h(\varrho_k^h, \varrho_{k+1}^h) + \mathcal{U}(\varrho_{k+1}^h) \leq \mathcal{U}(\varrho_k^h). \tag{4.2.1}$$

Assume that

$$m^*(1) := \sup_h \Big\{ M_1(\varrho_0^h) + \int_{\mathbb{R}^d} |U(\varrho_0^h)|\,\mathrm{d}\mathbf{x} \Big\} < +\infty. \tag{4.2.2}$$

For instance, if $\varrho_0^h = \varrho_0$ for all $h > 0$, Equation (4.2.2) holds if $M_1(\varrho_0)$ and $\int_{\mathbb{R}^d} |U(\varrho_0)|\,\mathrm{d}\mathbf{x}$ are both finite.

By Equations (2.4.1) and (4.2.1)

$$\mathcal{C}_{lh}(\varrho_0^h, \varrho_l^h) \leq \sum_{k=0}^{l-1} \mathcal{C}_h(\varrho_k^h, \varrho_{k+1}^h) \leq \mathcal{U}(\varrho_0^h) + \mathcal{U}^-(\varrho_l^h) - \mathcal{U}^+(\varrho_l^h), \tag{4.2.3}$$

which together with Lemma 2.4.4(ii), implies

$$-A_* hl + W_{\theta,lh}(\varrho_0^h, \varrho_l^h) + \mathcal{U}^+(\varrho_l^h) \le \mathcal{U}(\varrho_0^h) + \mathcal{U}^-(\varrho_l^h). \tag{4.2.4}$$

**Lemma 4.2.1** *If $\varrho, \bar\varrho \in \mathcal{P}_1^{ac}(\mathbb{R}^d)$, then*

$$M_1(\bar\varrho) \le [A_* + C(1)]h + \mathcal{C}_h(\varrho, \bar\varrho) + M_1(\varrho) \qquad \forall\, h > 0,$$

*where $C(1)$ is the constant provided by Lemma 2.4.4(iii).*

Proof For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ we have

$$|\mathbf{y}| \le |\mathbf{y} - \mathbf{x}| + |\mathbf{x}|$$

so that integrating the above inequality with respect to $\gamma \in \Gamma(\varrho, \bar\varrho)$ we obtain

$$M_1(\bar\varrho) \le \int_{\mathbb{R}^d \times \mathbb{R}^d} |\mathbf{y} - \mathbf{x}| \,\mathrm{d}\gamma(\mathbf{x}, \mathbf{y}) + M_1(\varrho), \tag{4.2.5}$$

and since $\gamma \in \Gamma(\varrho, \bar\varrho)$ is arbitrary we conclude that

$$M_1(\bar\varrho) \le W_1(\varrho, \bar\varrho) + M_1(\varrho).$$

This together with Lemma 2.4.4(ii)-(iii) gives the desired estimate. □

The following proposition shows that $M_1(\varrho_k^h)$ is uniformly bounded for $kh \le T$, provided that it is bounded for $k = 0$.

**Proposition 4.2.2** *There exists a constant $\bar C$, depending on $m^*(1)$ and $T$ only, such that the following holds:*

$$M_1(\varrho_k^h) + \int_{\mathbb{R}^d} |U(\varrho_k^h)| \,\mathrm{d}\mathbf{x} \le \bar C \qquad \forall\, k, h, \ \textit{with } kh \le T.$$

Proof We recall that by assumption $\varrho_k^h \in \mathcal{P}_1^{ac}$ for all $k, h$, so that $M_1(\varrho_k^h) < +\infty$.

To ease the notation, we drop the superscript $h$. Suppose $kh \le T$. By Lemma 4.2.1 and by Equation (4.2.3)

$$\begin{aligned} M_1(\varrho_k) &\le \mathcal{C}_{kh}(\varrho_0, \varrho_k) + [A_* + C(1)]hk + M_1(\varrho_0) \\ &\le \mathcal{U}(\varrho_0) - \mathcal{U}(\varrho_k) + [A_* + C(1)]hk + M_1(\varrho_0). \end{aligned} \tag{4.2.6}$$

Let $C$ be the constant provided by Lemma 4.1.1. We use that lemma and Equation (4.2.6) to obtain

$$M_1(\varrho_k) + \mathcal{U}^+(\varrho_k) \le \mathcal{U}^+(\varrho_0) + C\big(1 + M_1^a(\varrho_k)\big) + [A_* + C(1)]hk + M_1(\varrho_0). \quad (4.2.7)$$

Define for $t \ge 0$

$$f(t) := \sup_{s\ge 0}\{s \,:\, s - C(s^a + 1) \le t\}.$$

Observe that $f(t) \ge t$, and $f$ is nondecreasing. Thus, recalling that $M_1(\varrho_k) < +\infty$, by Equation (4.2.7) we get

$$M_1(\varrho_k) \le f\Big(\mathcal{U}^+(\varrho_0) + [A_* + C(1)]T + M_1(\varrho_0)\Big) := f_0 \quad (4.2.8)$$

and

$$\mathcal{U}^+(\varrho_k) \le \mathcal{U}^+(\varrho_0) + C\Big(1 + M_1^a(\varrho_k)\Big) + [A_* + C(1)]T + M_1(\varrho_0).$$

By Lemma 4.1.1 and Equation (4.2.8)

$$\mathcal{U}^-(\varrho_k) \le \tilde{C}\big(f_0^a + 1\big) \qquad \text{for } kh \le T,$$

where $\tilde{C}$ depends on $C$, $T$, $m^*(1)$, $A_*$ and $C(1)$ only. This concludes the proof. □

**Remark 4.2.3** *It is easy to check that the estimates proved in this section depend on $L$ only through the function $\theta$ and the constants $A^*, A_*, C^*$ appearing in (L3). Hence such estimates are uniform if $\{L_m\}_{m\in\mathbb{N}}$ is a sequence of Lagrangians satisfying (L1), (L2) and (L3) with the same function $\theta$ and the same constants $A^*, A_*, C^*$.*

## 4.3 Existence of Solutions

In this paragraph we briefly sketch how to prove existence of solutions in the case when $X$ is not necessarily bounded and $L$ satisfies (L1$^w$), (L2$^w$) and (L3), leaving the details to the interested reader. We remark that our approximation argument could also be used to relax some of the assumptions on $U$.

Let $X \subset \mathbb{R}^d$ be an open set whose boundary has zero Lebesgue measure. We fix $\varrho_0 \in \mathcal{P}^{ac}(X)$, and we assume that $M_1(\varrho_0)$ and $\int_X |U(\varrho_0)|\,\mathrm{d}\mathbf{x}$ are both finite. Assuming that $L$ satisfies (L1$^w$), (L2$^w$) and (L3), we consider a sequence of Lagrangians $\{L_m\}_{m\in\mathbb{N}}$ converging to $L$ in $C^1(\mathbb{R}^d \times \mathbb{R}^d)$ and which satisfy (L1), (L2) and (L3) with the same function $\theta$ as for $L$ and constants $A^* + 1, A_* + 1, C^* + 1$ (we slightly increase the constants of $L$ to ensure that one can construct such a sequence). We denote by $H_m$ the Hamiltonians associated with $L_m$. Consider now the increasing sequence of bounded sets $X_m$ defined by $X_m := X \cap B_m(0)$, and observe that, for each $m \in \mathbb{N}$, the set $X_m$ is open and its boundary has zero Lebesgue measure (since $\partial X_m \subset \partial X \cup \partial B_m(0)$). We now apply the variational scheme in $X_m$ starting from $\varrho_0^m := \varrho_0 \chi_{B_m(0)}$ (cf. Remark 3.5.5) with Lagrangian $L_m$. In this way we construct approximate discrete solutions $\{\varrho_{kh}^{h,m}\}$ on $X_m$ which satisfy the discrete energy inequality

$$\mathcal{U}(\rho_0^m) - \mathcal{U}(\rho_{(k+1)h}^{h,m}) \geq \int_0^T \int_{\mathbb{R}^d} \Big[ L_m(\mathbf{x}, V_t^{h,m}) \bar{\varrho}_t^{h,m} + H_m\big(\mathbf{x}, -\nabla[U'(\varrho_t^{h,m})]\big) \varrho_t^{h,m} \Big] \,\mathrm{d}\mathbf{x}\,\mathrm{d}t - A_* h.$$

Moreover, thanks to Proposition 4.2.2 (cf. Remark 4.2.3), we obtain that the measures $\{\varrho_{kh}^{h,m}\}$ have uniformly bounded first moments for all $k, h, m$, with $kh \leq T$. This fact, together with Lemma 4.1.1, implies that also $\mathcal{U}^-(\varrho_{kh}^{h,m})$ is uniformly bounded. Therefore, taking the limit as $h \to 0$ (cf. Section 3.5), we obtain that a family of curves $t \mapsto \varrho_t^m$ satisfying the energy bound in (3.5.12) and such that

$$\sup_{m\in\mathbb{N},\, t\in[0,T]} \left\{ M_1(\varrho_t^m) + \int_{\mathbb{R}^d} |U(\varrho_t^m)|\,\mathrm{d}\mathbf{x} \right\} < +\infty.$$

(Indeed $\mathcal{U}^-(\varrho_t^m)$ are uniformly bounded, and $t \mapsto \mathcal{U}(\varrho_t^m)$ is bounded too, cf. Equation (3.5.3).) Moreover

$$\partial_t \varrho_t^m + \mathrm{div}(\rho_t^m V_t^m) = 0 \qquad \text{on } (0,T) \times \mathbb{R}^d,$$

with

$$\sup_{m\in\mathbb{N}} \int_0^T \int_{\mathbb{R}^d} \theta(|V_t^m|) \varrho_t^m \,\mathrm{d}\mathbf{x}\,\mathrm{d}t < +\infty$$

(by Equation (3.5.12)), which implies a uniform continuity in time of the curves $t \mapsto \varrho_t^m$ on $[0,T]$. Using these bounds, it is not difficult to take the limit as $m \to +\infty$ (cf. the arguments in Section 3.5), and find a uniform continuous curve $t \mapsto \varrho_t$ which satisfies

$$\partial_t \varrho_t + \operatorname{div}(\varrho_t V_t) = 0 \qquad \text{on } (0,T) \times \mathbb{R}^d$$

in the sense of distributions, and

$$\mathcal{U}(\rho_0) - \mathcal{U}(\rho_T) \geq \int_0^T \int_{\mathbb{R}^d} \Big[ L(\mathbf{x}, V_t)\varrho_t + H\big(\mathbf{x}, -\nabla[U'(\varrho_t)]\big)\varrho_t \Big] \, \mathrm{d}\mathbf{x} \, \mathrm{d}t$$

(here we used $\mathcal{U}(\rho_0^m) \to \mathcal{U}(\rho_0)$ and Proposition 4.1.2). Once this estimate is established, the proof of Equation (3.5.14) is the same as in the bounded case. Hence we state the following result.

**Theorem 4.3.1** *Let $X \subset \mathbb{R}^d$ be an open set whose boundary is of zero Lebesgue measure, and assume that $H$ satisfies (H1$^w$), (H2$^w$) and (H3). Let $\varrho_0 \in \mathcal{P}_1^{ac}(X)$ be such that $\int_X |U(\varrho_0)| \, \mathrm{d}\mathbf{x} < +\infty$, and assume the $U$ satisfies Equations* (2.1.2), (2.1.3), (4.0.23)*. Then there exists a narrowly continuous curve $t \mapsto \varrho_t \in \mathcal{P}_1^{ac}(X)$ on $[0,T]$, starting from $\varrho_0$, such that $M_1(\varrho_t)$ is bounded on $[0,T]$ (so that in particular $\mathcal{U}^-(\rho_t)$ is bounded),*

$$\partial_t \varrho_t + \operatorname{div}(\varrho_t V_t) = 0 \qquad \text{on } (0,T) \times \mathbb{R}^d$$

*in the sense of distributions, and Equation (3.5.12) holds. Also, we have*

$$\nabla[P(\varrho)] \in L^1(0,T; L^1(X))$$

*and $\nabla[P(\varrho)]$ is absolutely continuous with respect to $\varrho$.*

*Furthermore, if $\theta(t) \sim t^\alpha$ for some $\alpha > 1$ and $U$ satisfies the doubling condition (3.5.13) then $\varrho$ is a solution of Equation* (1.0.1) *starting from $\varrho_0$, satisfying Equation* (1.0.2)*. In other words, Equation (3.5.14) holds and*

$$V_t(\mathbf{x}) = \nabla_{\mathbf{p}} H\big(\mathbf{x}, -\varrho_t^{-1}(\mathbf{x}) \nabla[P(\varrho_t(\mathbf{x}))]\big) \qquad \varrho\text{-a.e.}.$$

*If* $\varrho_0 \in \mathcal{P}_\alpha(X)$ *then* $\varrho \in AC_\alpha(0,T;\mathcal{P}_\alpha(X))$.

*In addition, suppose that (H4) holds. If* $\varrho_0 \leq M$ *for some* $M \geq 0$, *then* $\varrho_t \leq M$ *for all* $t \in [0,T]$ *(maximum principle).*

**Remark 4.3.2** *When* $\theta(t) \sim t^\alpha$ *with* $\alpha > 1$, *it is not difficult to see that if* $\int_X |\mathbf{x}|^\alpha \varrho_0 \,\mathrm{d}\mathbf{x}$ *is finite so is* $\int_X |\mathbf{x}|^\alpha \varrho_t \,\mathrm{d}\mathbf{x}$ *(here* $\varrho_t$ *is any limit curve constructed using the minimizing movement scheme). Hence one can generalize Lemma 4.1.1 proving that the* $\alpha$*-moment of* $\varrho$ *controls* $\mathcal{U}^-(\rho)$ *assuming only that Condition* (4.0.23) *holds for some* $a \in (\frac{d}{d+\alpha}, 1)$, *and the theorem above still holds under this weaker assumption on* $U$.

# CHAPTER V

# UNIQUENESS OF DISTRIBUTIONAL SOLUTIONS

## *5.1 Assumptions*

Throughout this chapter, we assume that $H$ satisfies (H1), (H2$^w$), (H3) and (H4). We assume that $U$ satisfies (2.1.2), (2.1.3). $X \subset \mathbb{R}^d$ is an open whose boundary $\partial X$ is of zero Lebesgue measure, and we denote by $\overline{X}$ its closure. We suppose that either $X$ is bounded or $X$ is unbounded but Condition (4.0.23) holds. We suppose that $\theta(t) \sim t^\alpha$ for some $\alpha > 1$ and $U$ satisfies the doubling condition (3.5.13). Our goal is to prove uniqueness of distributional solutions of Equation (1.0.1) when the initial condition $\varrho_0$ is bounded. The ellipticity conditions we impose seem to be different from what is usually imposed in the literature. Our proof of uniqueness of solutions follows the same line as in [18], except that most of our assumptions are not always comparable with the ones there. In the sequel, we use the following notation:

$$\Omega := (0,T) \times X, \qquad \tilde{\Omega} := (0,T) \times \Omega.$$

## *5.2 A new Hamiltonian*

We consider the density function $\varrho_t$ of Equation (1.0.1) provided by Theorem 4.3.1, which satisfies the property that $\nabla[P(\varrho_t)] \in L^1(\Omega)$ and is absolutely continuous with respect to $\varrho$. If we set

$$u(t,\cdot) := P(\varrho_t(\cdot))$$

we have

$$\partial_t Q(u) = \text{div }\mathbf{a}(\mathbf{x}, Q(u), \nabla u) \quad \text{in} \quad \mathcal{D}'(\Omega), \tag{5.2.1}$$

where

$$\mathbf{a}(\mathbf{x},s,\mathbf{z}) := \begin{cases} -\nabla_{\mathbf{z}}\bar{H}(\mathbf{x},s,-\mathbf{z}) & \text{if } s>0 \\ 0 & \text{if } s=0, \mathbf{z}=\mathbf{0}, \end{cases} \tag{5.2.2}$$

and $\bar{H} : \mathbb{R}^d \times [0,+\infty) \times \mathbb{R}^d \to [0,+\infty)$ is defined by

$$\bar{H}(\mathbf{x},s,\mathbf{z}) := \begin{cases} s^2 H\left(\mathbf{x}, \frac{\mathbf{z}}{s}\right) & \text{if } s>0 \\ 0 & \text{if } s=0, \mathbf{z}=\mathbf{0} \\ +\infty & \text{if } s=0, \mathbf{z}\neq\mathbf{0} \end{cases} \tag{5.2.3}$$

Here, $\mathbf{0} := (0,\dots,0)$.

For each $\mathbf{x} \in \mathbb{R}^d$, $\bar{H}(\mathbf{x},\cdot,\cdot)$ is of class $C^2((0,+\infty)\times\mathbb{R}^d)$, and the gradient of $\bar{H}(\mathbf{x},\cdot,\cdot)$ at $(s,\mathbf{z})$ is easily computed as follows

$$\nabla\bar{H}(\mathbf{x},s,\mathbf{z}) = \begin{pmatrix} 2sH\left(\mathbf{x},\frac{\mathbf{z}}{s}\right) - s\left\langle \nabla_{\mathbf{p}}H\left(\mathbf{x},\frac{\mathbf{z}}{s}\right), \frac{\mathbf{z}}{s}\right\rangle \\ s\nabla_{\mathbf{p}}H\left(\mathbf{x},\frac{\mathbf{z}}{s}\right) \end{pmatrix}$$

for $s>0$ and $\mathbf{z}\in\mathbb{R}^d$. Also, observe that

$$\nabla^2\bar{H}(\mathbf{x},\cdot,\cdot) = \begin{pmatrix} 2H - 2\left\langle \nabla_{\mathbf{p}}H, \frac{\mathbf{z}}{s}\right\rangle + \left\langle \nabla_{\mathbf{pp}}H\cdot\frac{\mathbf{z}}{s}, \frac{\mathbf{z}}{s}\right\rangle & \nabla_{\mathbf{p}}H - \nabla_{\mathbf{pp}}H\cdot\frac{\mathbf{z}}{s} \\ \nabla_{\mathbf{p}}H - \nabla_{\mathbf{pp}}H\cdot\frac{\mathbf{z}}{s} & \nabla_{\mathbf{pp}}H \end{pmatrix}. \tag{5.2.4}$$

Here $H, \nabla_{\mathbf{p}}H, \nabla_{\mathbf{pp}}H$ are all evaluated at $\left(\mathbf{x},\frac{\mathbf{z}}{s}\right)$. Since $H(\mathbf{x},\cdot)$ is convex we have that

$$\langle \nabla^2\bar{H}(\mathbf{x},\cdot,\cdot)(0,\lambda),(0,\lambda)\rangle = \langle \nabla_{pp}H\cdot\lambda,\lambda\rangle \geq 0.$$

for $\lambda\in\mathbb{R}^d$. Hence, the matrix in Equation (5.2.4) is nonnegative definite if and only if for every $\lambda\in\mathbb{R}^d$

$$\begin{aligned} 0 \leq \langle \nabla^2\bar{H}(\mathbf{x},\cdot,\cdot)(1,\lambda),(1,\lambda)\rangle &= 2H - \left\langle \nabla_{\mathbf{p}}H, \frac{\mathbf{z}}{s}\right\rangle + \left\langle \nabla_{\mathbf{pp}}H\cdot\frac{\mathbf{z}}{s}, \frac{\mathbf{z}}{s}\right\rangle \\ &+ 2\langle \nabla_{\mathbf{p}}H,\lambda\rangle - 2\left\langle \nabla_{\mathbf{pp}}H\cdot\frac{\mathbf{z}}{s},\lambda\right\rangle + \langle \nabla_{\mathbf{pp}}H\cdot\lambda,\lambda\rangle \\ &= 2H - 2\left\langle \nabla_{\mathbf{p}}H, \lambda-\frac{\mathbf{z}}{s}\right\rangle + \left\langle \nabla_{\mathbf{pp}}H\cdot\left(\lambda-\frac{\mathbf{z}}{s}\right), \lambda-\frac{\mathbf{z}}{s}\right\rangle. \end{aligned}$$

Hence $\bar{H}(\mathbf{x},\cdot,\cdot)$ is convex on $(0,\infty)\times\mathbb{R}^d$ if and only if

$$2H - 2\langle \nabla_{\mathbf{p}}H,\mathbf{w}\rangle + \langle \nabla_{\mathbf{pp}}H\cdot\mathbf{w},\mathbf{w}\rangle \geq 0 \qquad \forall\, \mathbf{w}\in\mathbb{R}^d. \tag{5.2.5}$$

This is what we assume in the sequel.

**Remark 5.2.1** *$H(\mathbf{x},\mathbf{p}) = |\mathbf{p}|^r$ satisfies Condition (5.2.5) if and only if $r \geq 2$. If $A(\mathbf{x})$ is a symmetric nonnegative definite matrix then $H(\mathbf{x},\mathbf{p}) = \langle A(\mathbf{x})\mathbf{p},\mathbf{p}\rangle$ satisfies Condition (5.2.5). Moreover, by linearity, if $H_1$ and $H_2$ satisfy Condition (5.2.5) so does $H_1 + H_2$.*

**Remark 5.2.2** *Suppose (5.2.5) holds.*

*(a) Since $\bar{H} \geq 0$ we have that $(0,\mathbf{0})$ belongs to the subdifferential of $\bar{H}(\mathbf{x},\cdot,\cdot)$ at $(0,\mathbf{0})$. In other words, $-\mathbf{a}(\mathbf{x},0,\mathbf{0})$ belongs to the subdifferential of $\bar{H}(\mathbf{x},\cdot,\cdot)$ at $(0,\mathbf{0})$.*

*(b) The convexity of $\bar{H}(\mathbf{x},\cdot,\cdot)$ is equivalent to*

$$\langle \mathbf{a}(\mathbf{x},s_1,\mathbf{z}_1) - \mathbf{a}(\mathbf{x},s_2,\mathbf{z}_2), \mathbf{z}_1 - \mathbf{z}_2\rangle \geq -(s_1 - s_2)\Big\{ 2\Big(s_1 H\big(\mathbf{x}, -\frac{\mathbf{z}_1}{s_1}\big) - s_2 H\big(\mathbf{x}, -\frac{\mathbf{z}_2}{s_2}\big)\Big) + \big\langle \nabla_{\mathbf{p}} H\big(\mathbf{x}, -\frac{\mathbf{z}_1}{s_1}\big), \mathbf{z}_1\big\rangle - \big\langle \nabla_{\mathbf{p}} H\big(\mathbf{x}, -\frac{\mathbf{z}_2}{s_2}\big), \mathbf{z}_2\big\rangle \Big\}$$

## 5.3 Additional Properties Satisfied by Bounded Solutions

We assume that (5.2.5) holds. Let $\varrho_t \in AC_1(0,T;\mathcal{P}_1^{ac}(X))$ be a solution of Equation (1.0.1) satisfying Equation (1.0.2) such that $t \to \int_{\mathbb{R}^d} U(\varrho_t)d\mathbf{x}$ is absolutely continuous, monotone nonincreasing, and $\nabla[P(\varrho_t)] \in L^1(\Omega)$ and is absolutely continuous with respect to $\varrho_t$. Observe that $\varrho_t$ satisfies Equation (3.5.12) and the inequality there becomes an equality. Suppose there exists $M > 0$ such that $\varrho_t \leq M$. Because $\theta(t) \sim t^\alpha$, (H3) implies that, for $\bar{c} > 0$ sufficiently small,

$$\bar{c}\Big(\big|\varrho_t^{-1}\nabla[P(\varrho)]\big|^{\alpha'} - 1\Big) \leq H(\mathbf{x}, -\varrho_t^{-1}\nabla[P(\varrho_t)])$$

and so, multiplying by both sides of the inequality by $\varrho_t$ we have

$$\bar{c}\Big(M^{1-\alpha'}\big|\nabla[P(\varrho_t)]\big|^{\alpha'} - \varrho_t\Big) \leq \varrho_t H(\mathbf{x}, -\varrho_t^{-1}\nabla[P(\varrho_t)]) \tag{5.3.1}$$

Taking $\bar{c} > 0$ small enough, (L3) ensures

$$\bar{c}(\varrho_t|V_t|^\alpha - \varrho_t) \leq \varrho_t L(\mathbf{x}, V_t), \qquad \bar{c}|\varrho_t V_t|^\alpha \leq M^{\alpha-1}\varrho_t\Big(\bar{c} + L(\mathbf{x}, V_t)\Big). \tag{5.3.2}$$

We use the fact that equality holds in Equation (3.5.12), exploit Equations (5.3.1) and (5.3.2) to obtain existence of a constant $C_M$, which depends only on $M$ and $\theta$, such that

$$\int_0^T \mathrm{d}t \int_X \Big|\nabla[P(\varrho_t)]\Big|^{\alpha'} \mathrm{d}\mathbf{x}, \quad \int_0^T \mathrm{d}t \int_X \varrho|V_t|^\alpha \varrho_t \,\mathrm{d}\mathbf{x}, \quad \int_0^T\int_X |\varrho_t V_t|^\alpha \varrho_t \,\mathrm{d}\mathbf{x} \leq C_M \tag{5.3.3}$$

where $V_t := \nabla_{\mathbf{p}} H\big(\mathbf{x}, -\varrho_t^{-1}\nabla[P(\varrho_t)]\big)$. Also, choosing $C_M$ large enough and using (L3), (H3) and Equation (5.3.3), we have

$$\int_0^T \mathrm{d}t \int_X \varrho_t\big|H(\mathbf{x}, -\varrho_t^{-1}\nabla[P(\varrho_t)])\big| \,\mathrm{d}\mathbf{x}, \quad \int_0^T \mathrm{d}t \int_X \varrho_t\big|L(\mathbf{x}, V_t)\big| \,\mathrm{d}\mathbf{x} \leq C_M. \tag{5.3.4}$$

**Remark 5.3.1** *Since $\varrho_t \in AC_1(0,T;\mathcal{P}_1^{ac}(X))$, $U$ is strictly convex, and*

$$t \to \int_{\mathbb{R}^d} U(\varrho_t)\mathrm{d}\mathbf{x}$$

*is absolutely continuous, we conclude that $\varrho \in C([0,T]; L^1(X))$.*

Observe that, by Equation (5.3.3), $u(t,\cdot) = P(\varrho_t(\cdot))$ satisfies $\nabla u \in L^{\alpha'}(\Omega)$, while the last inequality in (5.3.3) reads $\mathbf{a}(\cdot, Q(u), \nabla u) \in L^\alpha(\Omega)$. Since $\varrho_t$ satisfies Equation (1.0.1), by an approximation argument and Remark 5.3.1 we have

$$\int_\Omega Q(u)\partial_t \mathcal{E} = \int_\Omega \langle \mathbf{a}(\mathbf{x}, Q(u), \nabla u), \nabla \mathcal{E}\rangle \tag{5.3.5}$$

for any $\mathcal{E} \in W^{1,\alpha'}(\Omega)$ such that $\mathcal{E}(t,\cdot) \equiv 0$ for $t$ near 0 and $T$.

As in [18], for $\eta \in C^2(\mathbb{R})$ convex monotone nondecreasing such that $\eta'$ and $\eta''$ are bounded, we define $q_\eta, \eta_* : \mathbb{R}^2 \to \mathbb{R}$ by

$$q_\eta(z, z^o) = \int_{z^o}^z \eta'(s - z^o)Q'(s)ds, \quad z, z^o \in \mathbb{R}$$

$$\eta_*(w, z^o) = \sup_{z\in\mathbb{R}}\{\eta'(z - z^o)(w - Q(z)) + q_\eta(z, z^o)\}, \quad w, z^o \in \mathbb{R}.$$

**Lemma 5.3.2** *Suppose $v^o \in W^{1,\alpha'}(X) \cap L^\infty(X)$ and $\gamma \in C_c^\infty((0,T)\times\mathbb{R}^d)$ is nonnegative. Then*

$$\int_\Omega -q_\eta(u,v^o)\partial_t\gamma + \langle \mathbf{a}(\mathbf{x},Q(u),\nabla u), \nabla[\eta'(u-v^o)\gamma]\rangle \leq 0. \tag{5.3.6}$$

Proof The proof is identical to that of [18, Lemma 1]. □

## 5.4 Uniqueness of Bounded Solutions

In this section, for $i = 1, 2$, we consider $\varrho_t^i \in AC_1(0,T;\mathcal{P}_1^{ac}(X))$ solutions of Equation (1.0.1) satisfying (1.0.2) and such that $t \to \int_{\mathbb{R}^d} U(\varrho_t^i)\mathrm{d}\mathbf{x}$ is absolutely continuous monotone nonincreasing. We require that $\nabla[P(\varrho_t^i)] \in L^1(\Omega)$ be absolutely continuous with respect to $\varrho_t^i$. We further assume existence of an $M > 0$ such that $\varrho_t^i \leq M$. The goal of the section is to show that

$$t \to \int_X |\varrho_t^1 - \varrho_t^2|\mathrm{d}\mathbf{x} \quad \text{is monotone nondecreasing.}$$

Once such an estimate is proved, it extends immediately to solutions whose initial data belong to $L^1$ and have bounded first moment, and which are constructed by approximation (cf. Chapter 4) as a limit of solutions with bounded initial data.

We define $u_1, u_2$ on $\tilde{\Omega}$ by

$$u_1(t_1,t_2,\mathbf{x}) := P(\varrho^1(t_1,\mathbf{x})), \qquad u_2(t_1,t_2,\mathbf{x}) := P(\varrho^2(t_2,\mathbf{x})).$$

If $\alpha \in \mathbb{R}$ we set $\alpha^+ = \max\{0,\alpha\}$ and $\alpha^- = \max\{0,-\alpha\}$.

To achieve the main goal of this section, we first prove a lemma whose proof is more or less a repetition of the arguments presented on [18, pages 31-33]. Since $\mathbf{a}$ does not satisfy the assumptions imposed in that paper, we show that the arguments there go through for completeness.

**Lemma 5.4.1** *If $\min_{[0,M]} P' > 0$ and $\tilde{\gamma} \in C_c^\infty((0,T)^2)$ is nonnegative, then*

$$-\int_{\tilde{\Omega}} \big(Q(u_1) - Q(u_2)\big)^+ (\partial_{t_1}\tilde{\gamma} + \partial_{t_2}\tilde{\gamma}) \leq 0. \tag{5.4.1}$$

Proof Let $f_n \in C_c^\infty(\mathbb{R}^d)$ be such that $0 \le f_n \le 1$, $f_n(\mathbf{x}) = 1$ for $|\mathbf{x}| \le n$, $f_n(\mathbf{x}) = 0$ for $|\mathbf{x}| \ge n+2$, and $|\nabla f_n| \le 1$. Let $\eta \in C^2(\mathbb{R})$ be a convex nonnegative function such that $\eta(z) = 0$ for $z \le 0$, $\eta(z) = z - 1/2$ for $z \ge 1$. Set

$$\eta_\delta^+(z) = \delta\eta\left(\frac{z}{\delta}\right), \qquad \eta_\delta^-(z) = \delta\eta\left(-\frac{z}{\delta}\right), \qquad q_\delta^\pm = q_{\eta_\delta^\pm}$$

so that

$$(\eta_\delta^-)'(z) = -(\eta_\delta^+)'(-z). \tag{5.4.2}$$

We fix $t_2$ and apply Lemma 5.3.2 to

$$v^o = u_2(\cdot, t_2, \cdot) \equiv u_2(t_2, \cdot), \quad \eta = \eta_\delta^+, \quad \gamma = \tilde{\gamma}(\cdot, t_2) f_n.$$

Then we integrate the subsequent inequality with respect to $t_2$ over $(0,T)$ to obtain

$$\int_{\tilde{\Omega}} -q_\delta^+(u_1, u_2)\partial_{t_1}(\tilde{\gamma} f_n) + \Big\langle \mathbf{a}(\mathbf{x}, Q(u_1), \nabla u_1), \nabla[(\eta_\delta^+)'(u_1 - u_2)\tilde{\gamma} f_n]\Big\rangle \le 0. \tag{5.4.3}$$

Similarly,

$$\int_{\tilde{\Omega}} -q_\delta^-(u_2, u_1)\partial_{t_2}(\tilde{\gamma} f_n) + \Big\langle \mathbf{a}(\mathbf{x}, Q(u_2), \nabla u_2), \nabla[(\eta_\delta^-)'(u_2 - u_1)\tilde{\gamma} f_n]\Big\rangle \le 0. \tag{5.4.4}$$

We exploit Equations (5.4.2), (5.4.3) and (5.4.4) to obtain

$$\int_{\tilde{\Omega}} \tilde{\gamma}\Big\langle \mathbf{a}(\mathbf{x}, Q(u_1), \nabla u_1) - \mathbf{a}(\mathbf{x}, Q(u_2), \nabla u_2), (\nabla u_1 - \nabla u_2)(\eta_\delta^+)''(u_1 - u_2) f_n +$$
$$(\eta_\delta^+)'(u_1 - u_2)\nabla f_n\Big\rangle \le \int_{\tilde{\Omega}} \Big(q_\delta^+(u_1, u_2)\partial_{t_1}\tilde{\gamma} + q_\delta^-(u_2, u_1)\partial_{t_2}\tilde{\gamma}\Big) f_n$$

This, together with Remark 5.2.2 (b), yields

$$\int_{\tilde{\Omega}} f_n(\eta_\delta^+)''(u_1 - u_2)(Q(u_2) - Q(u_1))(E_1 - E_2)\tilde{\gamma} + \int_{\tilde{\Omega}} R_n^1$$
$$\le \int_{\tilde{\Omega}} (q_\delta^+(u_1, u_2)\partial_{t_1}\tilde{\gamma} + q_\delta^-(u_2, u_1)\partial_{t_2}\tilde{\gamma}) f_n. \tag{5.4.5}$$

Here,

$$R_n^1 := \tilde{\gamma}\Big\langle \mathbf{a}(\mathbf{x}, Q(u_1), \nabla u_1) - \mathbf{a}(\mathbf{x}, Q(u_2), \nabla u_2), (\eta_\delta^+)'(u_1 - u_2)\nabla f_n\Big\rangle.$$

$$E_i(t_1,t_2,\mathbf{x}) := \varrho^i(t_i,\mathbf{x})\Big(2H(\mathbf{x},-e^i(t_i,\mathbf{x})) + \Big\langle \nabla_p H\big(\mathbf{x},-e^i(t_i,\mathbf{x})\big), e^i(t_i,\mathbf{x})\Big\rangle\Big),$$

with $\varrho^i(t_i,\mathbf{x})e^i(t_i,\mathbf{x}) := \nabla[P(\varrho^i)](t_i,\mathbf{x})$ for $i = 1,2$. We observe that, by exploiting Equation (5.3.4), one obtains $E_1, E_2 \in L^1(\tilde{\Omega})$. The second inequality in (5.3.3) gives that

$$V_t^1 := \nabla_{\mathbf{p}} H\big(\mathbf{x}, -(\varrho_t^1)^{-1}\nabla[P(\varrho_t^1)]\big) \in L^\alpha(\varrho_t^1) \subset L^1(\varrho_t^1)$$

and so, $\mathbf{a}(\mathbf{x}, Q(u_1), \nabla u_1) \in L^1(\tilde{\Omega})$. Similarly, $\mathbf{a}(\mathbf{x}, Q(u_2), \nabla u_2) \in L^1(\tilde{\Omega})$. Hence

$$|R_n^1| \leq A^1 |\nabla f_n| \leq A^1$$

where $A^1 \in L^1(\tilde{\Omega})$. Since $|\nabla f_n| \to 0$ as $n \to +\infty$, we use the Lebesgue Dominated Convergence Theorem to conclude that $\int_{\tilde{\Omega}} T_n^1 \to 0$ as $n \to +\infty$. Since $u_1$ and $u_2$ are bounded, we may apply the Lebesgue Dominated Convergence Theorem to the first expression on the left hand side of Equation (5.4.5) and the one on its right hand side, to conclude that

$$-\int_{\tilde{\Omega}} (\eta_\delta^+)''(u_1-u_2)|Q(u_1)-Q(u_2)||E_1-E_2|\tilde{\gamma} \leq \int_{\tilde{\Omega}} \Big(q_\delta^+(u_1,u_2)\partial_{t_1}\tilde{\gamma} + q_\delta^-(u_2,u_1)\partial_{t_2}\tilde{\gamma}\Big). \tag{5.4.6}$$

Recall that $u_1$ and $u_2$ have their ranges contained in the compact set $[0, P(M)]$. Since $\min_{[0,M]} P' > 0$ we conclude that $Q$ is Lipschitz on $[0, P(M)]$. Let $\bar{C}_M$ be its Lipschitz constant there. By Equation (5.4.6)

$$-\bar{C}_M \int_{\tilde{\Omega}} (\eta_\delta^+)''(u_1-u_2)|u_1-u_2||E_1-E_2|\tilde{\gamma} \leq \int_{\tilde{\Omega}} \Big(q_\delta^+(u_1,u_2)\partial_{t_1}\tilde{\gamma} + q_\delta^-(u_2,u_1)\partial_{t_2}\tilde{\gamma}\Big). \tag{5.4.7}$$

Recall that

$$|q_\delta^\pm(z,z^o)| \leq (Q(z)-Q(z^o))^\pm, \quad |(\eta_\delta^+)'(z)| \leq z^+, \quad |z(\eta_\delta^+)''(z)| \leq \sup_{a\in\mathbb{R}} |a\eta''(a)| \tag{5.4.8}$$

and that, as $\delta \to 0^+$,

$$q_\delta^\pm(z,z^o) \to (Q(z)-Q(z^o))^\pm, \quad (\eta_\delta^+)'(z) \to z^+, \quad z(\eta_\delta^+)''(z) \to 0. \tag{5.4.9}$$

We can now conclude the proof of the lemma by combining Equations (5.4.7), (5.4.8) and (5.4.9). □

**Theorem 5.4.2** *Suppose $H$ satisfies (H1), (H2$^w$), (H3) and (H4). Suppose $U$ satisfies* (2.1.2), (2.1.3) *and the doubling condition in* (3.5.13). *Assume* $\min_{[0,M]} P' > 0$ *for any $M > 0$, $X \subset \mathbb{R}^d$ is an open set whose boundary $\partial X$ is of zero Lebesgue measure and $\theta(t) \sim t^\alpha$ with $\alpha > 1$. Suppose for $i = 1, 2$ that $\varrho_t^i \in AC_\alpha(0, T; \mathcal{P}_\alpha^{ac}(X))$ are solutions of Equation (1.0.1) satisfying (1.0.2). Assume further $t \to \int_{\mathbb{R}^d} U(\varrho_t^i)\mathrm{d}\mathbf{x}$ is absolutely continuous, monotone nonincreasing, and $\nabla[P(\varrho_t^i)] \in L^1(\Omega)$ and is absolutely continuous with respect to $\varrho_t^i$. If $\varrho_0^1, \varrho_0^2$ are bounded, then $t \to \int_X |\varrho_t^1 - \varrho_t^2|\mathrm{d}\mathbf{x}$ is monotone nondecreasing.*

Proof As shown in [18] this theorem is a direct consequence of Equation (5.4.1). □

# APPENDIX A

# SOME SUPPLEMENTARY RESULTS

## *A.1 Functions of Bounded Variations*

Firstly, let's recall some well-known definitions and results.

**Definition A.1.1** *Let $\Omega \subseteq \mathbb{R}^d$. and $f \in L^1(\Omega)$. We say that $f$ is a function of bounded variation in $\Omega$ if the distributional derivative of $f$ is representable by a finite Radon measure in $\Omega$, i.e., if*

$$\int_\Omega f \frac{\partial \varphi}{\partial x_i} \, \mathrm{d}\mathbf{x} = -\int_\Omega \varphi D_i f \qquad \forall \varphi \in C_c^1(\Omega), \ \ i = 1, 2,$$

*for some $\mathbb{R}^d$-valued measure $Df$ in $\Omega$. The vector space of all functions of bounded variation in $\Omega$ is denoted by $BV(\Omega)$.*

**Definition A.1.2 (variation)** *Let $f \in L^1(\Omega)$. The variation of $f$ in $\Omega$ is defined by*

$$V(f,\Omega) := \sup\left\{ \int_\Omega f(\mathbf{x}) \operatorname{div} \varphi(\mathbf{x}) \, \mathrm{d}\mathbf{x} \ : \ \ \varphi \in \mathrm{C}_\mathrm{c}^1(\Omega), \|\varphi\|_{\mathrm{L}^\infty(\Omega)} \le 1 \right\}.$$

**Proposition A.1.3** *Let $f \in L^1(\Omega)$. Then $f \in BV(\Omega)$ if and only if $V(f,\Omega) < \infty$. In addition, $V(f,\Omega) = |Df|(\Omega)$ for any $f \in BV(\Omega)$.*

**Remark A.1.4** *Note that if $f \in W^{1,1}(\Omega)$, then $Df = \nabla f \, \mathrm{d}\mathbf{x}$, and*

$$|Df|(\Omega) = \int_\Omega |\nabla f|(\mathbf{x}) \, \mathrm{d}\mathbf{x}.$$

*Moreover, $BV(\Omega)$ equipped with the norm $\|f\|_{BV(\Omega)} := \|f\|_{L^1(\Omega)} + V(f,\Omega)$ becomes a Banach space and $W^{1,1}(\Omega) \subsetneq BV(\Omega)$.*

**Remark A.1.5** *If $f \in \mathrm{BV}(a,b)$, $I = (a,b)$, then*

$$V(f,I) = \int_I |f'|(t) \, \mathrm{d}t \ \ = \ \ \sup_\varphi \left\{ \int_I f(t)\varphi'(t) \, \mathrm{d}t \ : \ \varphi \in C_c^1(I) \ \ \text{and} \ \ |\varphi| \le 1 \right\}$$

**Lemma A.1.6** *Suppose* $f : [a,b] \to \mathbb{R}$ *is monotone nonincreasing and set*

$$g(t) := \limsup_{h \to 0^+} \frac{f(t+h) - f(t)}{h} \leq 0, \qquad a < t < b.$$

*Then* $g \in L^1(a,b)$ *and*

$$f(b) - f(a) \leq \int_a^b g(t)\,\mathrm{d}t. \tag{A.1.1}$$

Proof For each $\delta > 0$, there are points $a_\delta \in (a, a+\delta)$, $b_\delta \in (b-\delta, b)$ in the points of continuity of $f$. Let $\rho_\epsilon$ be the standard mollifier such that

$$\rho_\epsilon(t) = \frac{1}{\epsilon}\rho\left(\frac{t}{\epsilon}\right), \qquad \int_{\mathbb{R}} \rho(t)\,\mathrm{d}t = 1, \qquad \mathrm{spt}\rho = [-1,1].$$

Notice that $f_\epsilon = \rho_\epsilon * f$ is well-defined in $(a+\delta, b-\delta)$ for $0 < \epsilon < \delta$. Note that

$$\lim_{\epsilon \to 0} f_\epsilon(t) = f(t)$$

wherever $t$ is a point of continuity of $f$. Since $f$ is monotone nonincreasing, $f_\epsilon$ is also monotone nonincreasing, as directly follows from the definition of convolution. Moreover, we have

$$\liminf_{\epsilon \to 0} \int_I |f_\epsilon'|(t)\,\mathrm{d}t \geq \int_I |f'|(t)\,\mathrm{d}t. \tag{A.1.2}$$

Indeed, for a number $a > 0$ there exists $\varphi \in C_c^1(I)$ with $|\varphi| \leq 1$ such that

$$\begin{aligned}
\int_I |f'|\,(t)\,\mathrm{d}t &\leq \int_I f(t)\varphi'(t)\,\mathrm{d}t + a \\
&= \liminf_{\epsilon \to 0} \int_I f_\epsilon(t)\varphi'(t)\,\mathrm{d}t + a \\
&= \liminf_{\epsilon \to 0} \int_I |f_\epsilon'|(t)\,\mathrm{d}t + a
\end{aligned}$$

holds for all $a > 0$. This proves (A.1.2).

1. Observe that $f'_\epsilon \leq 0$ means $f'_\epsilon = -|f'_\epsilon|$ and so

$$
\begin{aligned}
f(b) - f(a) &\leq f(b_\delta) - f(a_\delta) \\
&= \lim_{\epsilon \to 0} f_\epsilon(b_\delta) - f_\epsilon(a_\delta) \\
&= \lim_{\epsilon \to 0} \int_{a_\delta}^{b_\delta} f'_\epsilon(t)\,\mathrm{d}t \\
&= \limsup_{\epsilon \to 0} -\int_{a_\delta}^{b_\delta} |f'_\epsilon|(t)\,\mathrm{d}t \\
&= -\liminf_{\epsilon \to 0} \int_{a_\delta}^{b_\delta} |f'_\epsilon|(t)\,\mathrm{d}t \\
&\leq -\int_{a_\delta}^{b_\delta} |f'|(t)\,\mathrm{d}t. \qquad \text{(A.1.3)}
\end{aligned}
$$

2. Set

$$
\varphi(t) = \begin{cases} \dfrac{t - a_\delta}{h} & \text{if } a_\delta \leq t \leq a_\delta + h \\ 1 & \text{if } a_\delta + h \leq t \leq b_\delta - h \\ \dfrac{b_\delta - t}{h} & \text{if } b_\delta - h \leq t \leq b_\delta \end{cases} .
$$

Note that $0 \leq \varphi \leq 1$ and $\varphi(a_\delta) = \varphi(b_\delta) = 0$. Then,

$$
\begin{aligned}
\int_{a_\delta}^{b_\delta - h} \frac{f(t+h) - f(t)}{h}\,\mathrm{d}t &= \frac{1}{h}\left\{\int_{a_\delta}^{b_\delta - h} f(t+h)\,\mathrm{d}t - \int_{a_\delta}^{b_\delta - h} f(t)\,\mathrm{d}t\right\} \\
&= \frac{1}{h}\left\{\int_{a_\delta + h}^{b_\delta} f(t)\,\mathrm{d}t - \int_{a_\delta}^{b_\delta - h} f(t)\,\mathrm{d}t\right\}.
\end{aligned}
$$

After rearranging the bounds of integrals, we end up with the following inequality:

$$
\begin{aligned}
\pm \int_{a_\delta}^{b_\delta - h} \frac{f(t+h) - f(t)}{h}\,\mathrm{d}t &= \pm\frac{1}{h}\left\{\int_{b_\delta - h}^{b_\delta} f(t)\,\mathrm{d}t - \int_{a_\delta}^{a_\delta + h} f(t)\,\mathrm{d}t\right\} \\
&= \mp \int_{a_\delta}^{b_\delta} \varphi'(t) f(t)\,\mathrm{d}t \\
&\leq \int_{a_\delta}^{b_\delta} |f'|(t)\,\mathrm{d}t,
\end{aligned}
$$

implying that

$$
-\int_{a_\delta}^{b_\delta} |f'|(t)\,\mathrm{d}t \leq \int_{a_\delta}^{b_\delta - h} \frac{f(t+h) - f(t)}{h}\,\mathrm{d}t. \qquad \text{(A.1.4)}
$$

Note that $\varphi$ defined above is continuous but not $C^1$. Replacing it with $\varphi_\epsilon = \varphi * \rho_\epsilon$ using the Dominated Convergence Theorem yields the inequality in (A.1.4).

3. Therefore, using (A.1.3) and (A.1.4), we observe that

$$\begin{aligned}
f(b)-f(a) &\leq \limsup_{h\to 0^+}\int_{a_\delta}^{b_\delta}\frac{f(t+h)-f(t)}{h}\chi_{(a_\delta,b_\delta-h)}\,\mathrm{d}t\\
&= -\liminf_{h\to 0^+}\int_{a_\delta}^{b_\delta}\frac{f(t)-f(t+h)}{h}\chi_{(a_\delta,b_\delta-h)}\,\mathrm{d}t\\
&\leq -\int_{a_\delta}^{b_\delta}\liminf_{h\to 0^+}\frac{f(t)-f(t+h)}{h}\chi_{(a_\delta,b_\delta-h)}\,\mathrm{d}t\\
&= \int_{a_\delta}^{b_\delta}\limsup_{h\to 0^+}\frac{f(t+h)-f(t)}{h}\chi_{(a_\delta,b_\delta-h)}\,\mathrm{d}t\\
&= \int_{a_\delta}^{b_\delta} g(t)\,\mathrm{d}t
\end{aligned}$$

where we used Fatou's Lemma to get the second inequality. Thus,

$$\int_{a_\delta}^{b_\delta} -g(t)\,\mathrm{d}t \leq f(a)-f(b).$$

and since $g \leq 0$, we obtain that

$$\int_{a_\delta}^{b_\delta} |g(t)|\,\mathrm{d}t \leq f(a)-f(b)$$

which holds for any $\delta > 0$. Hence $g \in L^1(a,b)$ with $\|g\|_{L^1(a,b)} \leq f(a)-f(b)$. Since $g \leq 0$, the inequality in (A.1.1) follows immediately. □

## A.2 Minimum Principle

Following the proof of the maximum principle (cf. Theorem 3.1.1) and Lemma 1.1.2 in [20], we have the following result:

**Theorem A.2.1** *[Minimum principle]*

*Let $X \subset \mathbb{R}^d$ be an open, bounded set and $\varrho_0 \in \mathcal{P}^{ac}(X)$ be such that $\varrho_0 \geq K$* a.e. *for some $K > 0$. Suppose that $\varrho_h$ is the minimizer of*

$$\Phi[h,\varrho_0,\varrho] := \mathcal{C}_h(\varrho_0,\varrho) + \mathcal{U}(\varrho)$$

*over $\mathcal{P}^{ac}(X)$. Then $\varrho_h \geq K$* a.e..

Proof For the sake of contradiction assume that $B := \{\mathbf{x} \in X \,:\, \varrho_h(\mathbf{x}) < K\}$ is of positive Lebesgue measure. The clincher here is to construct a measure that minimizes our functional on $X \times X$, which is constructed in a way that the new measure remains positive on a subset of $X \times X$ that is formed by two disjoint subsets of $X$. To this end, set $B^c := X \setminus B$ and let $\gamma \in \Gamma(\varrho_0, \varrho_h)$.

**Claim 1.**

$$\gamma(B \times B^c) > 0. \tag{A.2.1}$$

**Proof of Claim 1.** Suppose that $\gamma(B \times B^c) = 0$. Then, notice that

$$\begin{aligned} K\mathcal{L}^d(B) &> \int_B \varrho_h(\mathbf{y})\,\mathrm{d}\mathbf{y} = \gamma(X \times B) \\ &> \gamma(B \times B) = \gamma(B \times X) \\ &= \int_B \varrho_0(\mathbf{x})\,\mathrm{d}\mathbf{x} \geq K\mathcal{L}^d(B) \end{aligned}$$

which is a contradiction, concluding Claim 1.

Now, consider the restricted probability measure $\gamma_B := \frac{1}{\gamma(B\times B^c)}\gamma\big|_{B\times B^c}$ defined by

$$\gamma_B(A) = \frac{\gamma\,(\,A \cap (B \times B^c)\,)}{\gamma(B \times B^c)}$$

for Borel sets $A \subset X \times X$. Note that

$$\gamma(B \times B^c) \int_{X\times X} \eta(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma_B(\mathbf{x},\mathbf{y}) = \int_{B\times B^c} \eta(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma(\mathbf{x},\mathbf{y})$$

for $\eta \in C(X \times X)$. Let $\nu_0$ and $\nu_1$ be the marginals of $\gamma_B$, more precisely,

$$\int_{X\times X} (\,\varphi(\mathbf{x}) + \psi(\mathbf{y})\,)\,\mathrm{d}\gamma_B(\mathbf{x},\mathbf{y}) = \int_B \varphi(\mathbf{x})\,\mathrm{d}\nu_0(\mathbf{x}) + \int_{B^c} \psi(\mathbf{y})\,\mathrm{d}\nu_1(\mathbf{y}),$$

for $\varphi$, $\psi \in C(X)$. It is easy to check that $\gamma_B \ll \gamma \in \Gamma(\varrho_0, \varrho_h)$. Therefore $\nu_0 \ll \varrho_0(\mathbf{x})\mathcal{L}^d$ and $\nu_1 \ll \varrho_h(\mathbf{y})\mathcal{L}^d$. That is to say, $\nu_0$ and $\nu_1$ are absolutely continuous with respect to Lebesgue measure $\mathcal{L}^d$. Also, we see that

$$
\begin{aligned}
\nu_0(B^c) &= \gamma_B(((B^c \times X) \cap (B \times B^c)) \\
&= \gamma_B((B^c \cap B) \times (X \cap B^c)) \\
&= 0 \\
&= \gamma_B((X \cap B) \times (B \cap B^c)) \\
&= \gamma_B(((X \times B) \cap (B \times B^c)) \\
&= \nu_1(B).
\end{aligned}
$$

In other words, we observe that $\mathrm{spt}(\nu_0) \subseteq B$, and $\mathrm{spt}(\nu_1) \subseteq B^c$. Let $v_0$ and $v_1$ be their density functions respectively. Observe that

$$
0 \leq v_0(\mathbf{x}) \leq \frac{1}{\gamma(B \times B^c)} \varrho_0(\mathbf{x}) \ \text{ a.}e. \qquad \text{and} \qquad 0 \leq v_1(\mathbf{y}) \leq \frac{1}{\gamma(B \times B^c)} \varrho_h(\mathbf{y}) \ \text{ a.}e.
$$

Let $\beta \in (0, \gamma(B \times B^c))$. Then, define

$$
\varrho_\beta(\mathbf{y}) := \varrho_h(\mathbf{y}) - \beta(v_1(\mathbf{y}) - v_0(\mathbf{y})), \qquad \mathbf{y} \in X.
$$

Note that $\varrho_\beta \in \mathcal{P}^{ac}(X)$. Indeed,

$$
\varrho_{\beta_{|B^c}} = \varrho_{h_{|B^c}} - \beta v_1 \geq 0 \qquad \text{and} \qquad \varrho_{\beta_{|B}} = \varrho_{h_{|B}} + \beta v_0 \geq 0,
$$

and by $\nu_0(B^c) = \nu_1(B)$ we have

$$
\int_X \varrho_\beta(\mathbf{y}) \, \mathrm{d}\mathbf{y} = \int_X \varrho_h(\mathbf{y}) \, \mathrm{d}\mathbf{y} = 1.
$$

Accordingly, define a measure $\gamma_\beta$ on $X \times X$ by

$$
\begin{aligned}
\int_{X \times X} \eta(\mathbf{x}, \mathbf{y}) \, \mathrm{d}\gamma_\beta(\mathbf{x}, \mathbf{y}) \ &:= \ \int_{X \times X} \eta(\mathbf{x}, \mathbf{y}) \, \mathrm{d}\gamma(\mathbf{x}, \mathbf{y}) \\
&\quad + \beta \int_{B \times B^c} (\eta(\mathbf{x}, \mathbf{x}) - \eta(\mathbf{x}, \mathbf{y})) \, \mathrm{d}\gamma(\mathbf{x}, \mathbf{y}),
\end{aligned}
$$

for all $\eta \in C(X \times X)$. Notice that $\gamma_\beta \in \Gamma(\varrho_0, \varrho_\beta)$. Indeed, for $\psi \in C(X)$

$$
\int_{X \times X} \psi(\mathbf{x}) \, \mathrm{d}\gamma_\beta(\mathbf{x}, \mathbf{y}) = \int_X \psi(\mathbf{x}) \varrho_0(\mathbf{x}) \, \mathrm{d}\mathbf{x}
$$

$$\begin{aligned}\int_{X\times X}\psi(\mathbf{y})\,\mathrm{d}\gamma_\beta(\mathbf{x},\mathbf{y}) &= \int_X \psi(\mathbf{y})\varrho_h(\mathbf{y})\,\mathrm{d}\mathbf{y} \\ &\quad +\beta\left(\int_B \psi(\mathbf{x})w_0(\mathbf{x})\,\mathrm{d}\mathbf{x} - \int_{B^c}\psi(\mathbf{y})w_1(\mathbf{y})\,\mathrm{d}\mathbf{y}\right) \\ &= \int_X \psi(\mathbf{y})\varrho_\beta(\mathbf{y})\,\mathrm{d}\mathbf{y}\end{aligned}$$

**Claim 2.** There exists $\beta \in (0, \gamma(B\times B^c))$ such that

$$\Phi[h,\varrho_0,\varrho_\beta] < \Phi[h,\varrho_0,\varrho_h] \tag{A.2.2}$$

**Proof of Claim 2:** Remember that $c_h(\mathbf{x},\mathbf{x}) - c_h(\mathbf{x},\mathbf{y}) < 0$ on $B\times B^c$. Then by (A.2.2) we get

$$\begin{aligned}\mathcal{C}_h(\varrho_0,\varrho_\beta) - \mathcal{C}_h(\varrho_0,\varrho_h) &\leq \int_{X\times X} c_h(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma_\beta(\mathbf{x},\mathbf{y}) - \int_{X\times X} c_h(\mathbf{x},\mathbf{y})\,\mathrm{d}\gamma(\mathbf{x},\mathbf{y}) \\ &= \beta\int_{B\times B^c}\left(c_h(\mathbf{x},\mathbf{x}) - c_h(\mathbf{x},\mathbf{y})\right)\mathrm{d}\gamma(\mathbf{x},\mathbf{y}) \\ &< 0 \end{aligned}\tag{A.2.3}$$

The convexity of $U$ implies $U'$ is increasing and $U(t) - U(s) \leq U'(t)(t-s)$ and so, we get

$$\begin{aligned}\mathcal{U}(\varrho_\beta) - \mathcal{U}(\varrho_h) &= \int_X \left[U(\varrho_\beta(\mathbf{x})) - U(\varrho_h(\mathbf{x}))\right]\mathrm{d}\mathbf{x} \\ &= \int_B \left[U(\varrho_h + \beta v_0) - U(\varrho_h)\right] + \int_{B^c}\left[U(\varrho_h - \beta v_1) - U(\varrho_h)\right] \\ &\leq \beta\int_B U'(\varrho_h(\mathbf{x}) + \beta v_0(\mathbf{x}))v_0(\mathbf{x})\,\mathrm{d}\mathbf{x} \\ &\quad -\beta\int_{B^c} U'(\varrho_h(\mathbf{y}) - \beta v_1(\mathbf{y}))v_1(\mathbf{y})\,\mathrm{d}\mathbf{y} \\ &\leq \beta\left[\int_B U'(K+\beta v_0(\mathbf{x}))v_0(\mathbf{x})\,\mathrm{d}\mathbf{x} - \int_{B^c} U'(K-\beta v_1(\mathbf{y}))v_1(\mathbf{y})\,\mathrm{d}\mathbf{y}\right] \\ &= \beta\int_{X\times X}\left(U'(K+\beta v_0(\mathbf{x})) - U'(K-\beta v_1(\mathbf{y}))\right)\mathrm{d}\gamma_A(\mathbf{x},\mathbf{y})\,.\end{aligned}$$

Since $U \in C^2((0,\infty))$, invoking the Mean Value Theorem, we obtain

$$\Big(U'(K+\beta v_0(\mathbf{x})) - U'(K-\beta v_1(\mathbf{y})\Big) \in o(\beta)$$

and so we have

$$\mathcal{U}(\varrho_\beta) - \mathcal{U}(\varrho_h) = 0(\beta^2). \tag{A.2.4}$$

Choosing $\beta \in (0, \gamma(B \times B^c))$ sufficiently small, and using (A.2.3) and (A.2.4) we obtain (A.2.2). This proves Claim 2.

As a result of Claim 1 and Claim 2, if $\mathcal{L}^d(B) > 0$, then $\varrho_h$ is not a minimizer, which contradicts the assumption. □

## A.3 Weak Compactness in $W^{1,1}(X)$

Although the following compactness result is standard, we still give its proof for the sake of completeness.

**Theorem A.3.1** *Suppose that $X \subset \mathbb{R}^d$ is open, bounded, connected and $\partial X$ is smooth. Let $\{v_n\}_{n=1}^\infty$ be a sequence satisfying*

$$\sup_n \int_X \theta(|\nabla v_n(\mathbf{x})|)\, \mathrm{d}\mathbf{x} := C_1 < +\infty \tag{A.3.1}$$

*and*

$$\sup_n \int_X \ell(|v_n(\mathbf{x})|)\, \mathrm{d}\mathbf{x} := C_2 < +\infty \tag{A.3.2}$$

*where $\theta, \ell \in C[0, +\infty)$, $\theta \geq 0$, $\theta(0) = 0$, $\theta$ is convex, $\lim_{t\to\infty} \frac{\theta(t)}{t} = +\infty$ and $\ell \geq 0$, $\lim_{t\to\infty} \ell(t) = +\infty$. Then $\{v_n\}_{n=1}^\infty$ is weakly compact in $W^{1,1}(X)$.*

Proof Let us first show that

$$\sup_n \|v_n\|_{L^1(X)} < +\infty. \tag{A.3.3}$$

To see this set $\lambda_n = \|v_n\|_{L^1(X)}$ and suppose on the contrary that $1 \leq \lambda_n \to +\infty$ as $n \to \infty$. Define a function

$$u_n(\mathbf{x}) = \frac{v_n(\mathbf{x})}{\lambda_n}.$$

Then, clearly

$$\|u_n\|_{L^1(X)} = 1.$$

Since $\theta$ is superlinear, there exists $A_0 > 0$ such that

$$\theta(t) \geq t - A_0 \tag{A.3.4}$$

Using the inequalities in (A.3.1) and (A.3.4), we conclude that

$$\|\nabla v_n\|_{L^1(X)} = \int_X |\nabla v_n(\mathbf{x})| \, \mathrm{d}\mathbf{x} \leq A_0 \mathcal{L}^d(X) + C_1 < +\infty \tag{A.3.5}$$

which implies that

$$\|\nabla u_n\|_{L^1(X)} = \frac{1}{\lambda_n} \|\nabla v_n\|_{L^1(X)} \leq \frac{A_0 \mathcal{L}^d(X) + C_1}{\lambda_n}$$

and $\nabla u_n \to 0$ in $L^1(X)$. On the other hand, since $u_n \in W^{1,1}(X)$ is bounded and $W^{1,1}(X) \hookrightarrow L^1(X)$, passing to a subsequence if necessary, there exists a function $u$ such that $u_n \to u$ in $L^1(X)$. Note that,

$$\|u\|_{L^1(X)} \leq \|u - u_n\|_{L^1(X)} + \|u_n\|_{L^1(X)}$$

implies that

$$\|u\|_{L^1(X)} \leq 1. \tag{A.3.6}$$

Also,

$$1 = \|u_n\|_{L^1(X)} \leq \|u\|_{L^1(X)} + \|u_n - u\|_{L^1(X)}$$

gives

$$\|u\|_{L^1(X)} \geq 1. \tag{A.3.7}$$

Then using (A.3.6) and (A.3.7), we conclude that

$$\|u\|_{L^1(X)} = 1. \tag{A.3.8}$$

Since $\nabla u_n \to 0$ and $u_n \to u$ in $L^1(X)$ , by uniqueness of the weak derivatives, we observe that for all $\phi \in C_c^\infty(X)$

$$0 = \lim_{n \to \infty} \int_X \frac{\partial u_n}{\partial x_i} \phi = - \lim_{n \to \infty} \int_X u_n \frac{\partial \phi}{\partial x_i} = - \int_X u \frac{\partial \phi}{\partial x_i},$$

where the limit follows from the fact that $u_n \to u$ in $L^1(X)$ implies $u_n \rightharpoonup u$ in $L^1(X)$. Hence, we conclude that $\partial_{x_i} u(\mathbf{x}) = 0$ for all $1 \le i \le d$ i.e., the weak gradient $\nabla u = 0$ on $X$. Therefore,

$$u = \text{constant} = C \ \text{ on } \ X. \tag{A.3.9}$$

Namely,

$$u_n \to C \ \text{ in } \ L^1(X).$$

We can extract a subsequence $\{u_{k_n}\}_{n=1}^{\infty}$ such that

$$u_{k_n}(\mathbf{x}) \to C \text{ a.e.}$$

By (A.3.8), we get $C\mathcal{L}^d(X) = 1$ and hence $C \neq 0$. However,

$$\begin{aligned} \sup_n \int_X \ell(|v_{k_n}(\mathbf{x})|)\,\mathrm{d}\mathbf{x} &\geq \liminf_{n\to\infty} \int_X \ell(|v_{k_n}(\mathbf{x})|)\,\mathrm{d}\mathbf{x} \\ &= \liminf_{n\to\infty} \int_X \ell(|\lambda_{k_n}||u_{k_n}(\mathbf{x})|)\,\mathrm{d}\mathbf{x} \\ &\overset{\text{Fatou's Lemma}}{\geq} \int_X \liminf_{n\to\infty} \ell(|\lambda_{k_n}||u_{k_n}(\mathbf{x})|)\,\mathrm{d}\mathbf{x} = +\infty \end{aligned}$$

because $C \neq 0$ and $\lambda_{k_n} u_{k_n}(\mathbf{x}) \to \infty$ as $n \to \infty$. This contradicts (A.3.2). As a result of (A.3.3) and (A.3.5), we deduce that

$$\sup_n \|v_n\|_{W^{1,1}(X)} < +\infty \tag{A.3.10}$$

By the Dunford-Pettis Theorem together with (A.3.1), we conclude that there exists a subsequence of $\{\nabla v_n\}$, still denoted by $\{\nabla v_n\}$, converging weakly in $L^1(X)$. On the other hand, the fact that the embedding of $W^{1,1}(X)$ into $L^1(X)$ is compact and (A.3.10) give us that $v_n$ is precompact in $L^1(X)$. Therefore, there exists a subsequence of $v_n$ , still denoted by $\{v_n\}$, and $v$ such that $v_n \to v$ in $L^1(X)$. Since the strong convergence implies the weak convergence, $v_n \rightharpoonup v$ in $L^1(X)$. By (A.3.1) and uniqueness of the weak derivative and the weak limit, passing to another subsequence, we obtain that $\nabla v_{k_n} \rightharpoonup \nabla v$ in $L^1(X)$. That is to say, $v_{k_n} \rightharpoonup v$ in $W^{1,1}(X)$. Using (A.3.10), we also obtain that $v \in W^{1,1}(X)$. Indeed, this follows from the fact that

$$\|\nabla v\|_{L^1(X)} \leq \liminf_{n\to\infty} \|\nabla v_{k_n}\|_{L^1(X)} \leq \sup_n \|\nabla v_{k_n}\|_{L^1(X)} := C_3 < +\infty$$

and

$$\|v\|_{L^1(X)} \leq \liminf_{n\to\infty} \|v_{k_n}\|_{L^1(X)} \leq \sup_n \|v_{k_n}\|_{L^1(X)} := C_4 < +\infty.$$

Since $v_n \to v$ in $L^1(X)$, there exists a subsequence $\{v_{m_n}\}$ such that $v_{m_n}(\mathbf{x}) \to v(\mathbf{x})$ a.e.. Therefore, Fatou's Lemma gives

$$\int_X \ell(|v(\mathbf{x})|)\, dx \leq \liminf_{n\to\infty} \int_X \ell(|v_{m_n}(\mathbf{x})|)\, \mathrm{d}\mathbf{x} \leq C_2 < +\infty. \tag{A.3.11}$$

Moreover, since $\theta$ is convex and $\theta(0) = 0$, the map $t \mapsto \theta(t)$ is increasing. Indeed, for $0 < s < t$ we use the convexity and positivity of $\theta$ to get

$$\theta(s) = \theta\left(\frac{s}{t}t + \left(1 - \frac{s}{t}\right)0\right) \leq \frac{s}{t}\theta(t) \leq \theta(t).$$

Also observe that for $\lambda \in [0,1]$ we obtain

$$\theta(|(1-\lambda)\mathbf{x} + \lambda\mathbf{y}|) \leq \theta((1-\lambda)|\mathbf{x}| + \lambda|\mathbf{y}|) \leq (1-\lambda)\theta(|\mathbf{x}|) + \lambda\theta(|\mathbf{y}|),$$

meaning that $\mathbf{x} \mapsto \theta(|\mathbf{x}|)$ is convex. This way we arrive at the convexity of the map $v \mapsto \theta(|\nabla v|)$ and consequently the functional

$$v \mapsto I[v] := \int_X \theta(|\nabla v(\mathbf{x})|)\, \mathrm{d}\mathbf{x}$$

is convex too. Therefore, $I[\cdot]$ is weakly lower semicontinuous in $W^{1,1}(X)$, (cf. [6] Theorem 3.4., given originally by E. De Giorgi) and hence $v_{k_n} \rightharpoonup v$ in $W^{1,1}(X)$ yields

$$\int_X \theta(|\nabla v(\mathbf{x})|)\, \mathrm{d}\mathbf{x} \leq \liminf_{n\to\infty} \int_X \theta(|\nabla v_{k_n}(\mathbf{x})|)\, \mathrm{d}\mathbf{x} \leq C_1 < +\infty. \tag{A.3.12}$$

By (A.3.11) and (A.3.12), we conclude that the weak limit $v \in W^{1,1}(X)$ satisfies Conditions (A.3.1) and (A.3.2), completing the proof. □

# REFERENCES


[1] Agueh, M., "Existence of solutions to degenerate parabolic equations via the monge-kantorovich theory," *Adv. Differential Equations*, vol. 10, no. 3, pp. 309–360, 2005.

[2] Alt, H.-W. and Luckhaus, S., "Quasilinear elliptic–parabolic differential equations," *Math. Z.*, vol. 183, pp. 311–341, 1983.

[3] Ambrosio L., Gigli N., S. G., *Gradient flows in metric spaces and the Wasserstein spaces of probability measures.* ETH Zurich: Birkhäuser, 2005. Lectures in Mathematics.

[4] Attouch H., Buttazzo G., M. G., *Variational Analysis in Sobolev and BV Spaces:Applications to PDEs and Optimization.* MPS-SIAM Series on Optimization, 2005.

[5] Bernard, P. and Buffoni, B., "Optimal mass transportation and mather theory," *J. Eur. Math. Soc.*, vol. 9, no. 1, pp. 85–121, 2007.

[6] Dacorogna, B., *Direct methods in the calculus of variations.* Berlin: Springer-Verlag, 1989.

[7] De Giorgi E., Mariono A., T. M., "Problems of evolutions in metric spaces and maximal decreasing curve," *Atti. Accad. Naz. Lincei Rend. Cl. Sci. Mat. Natur. (8)*, vol. 68, pp. 180–187, 1980.

[8] E., D. G., "New problems on minimizing movements," *In Boundary Value Problems for PDE and Applications, C. Baiocchi and J.L. Lions, eds.,Masson*, pp. 81–98, 1993.

[9] Fathi, A. and Figalli, A., "Optimal transportation on non-compact manifolds," *Israel J. Math.*, To appear.

[10] Figalli, A., "Existence, uniqueness, and regularity of optimaltransport maps.," *SIAM J. Math. Anal.*, vol. 39, no. 1, pp. 126–137, 2007.

[11] Figalli A., Gangbo W., Y. T., "A variational method for class of parabolic pdes," *Preprint*, 2009.

[12] Gangbo, W., "Quelques problèmes d'analyse convexe," *Rapport d'habilitation à diriger des recherches*, Jan. 1995.

[13] Gangbo, W., "An introduction to the mass transportation theory and its applications," *http://www.math.gatech.edu/∼gangbo/publications/notecmu.pdf*, Lecture Notes, 2004, last visited: June 5th, 2009.

[14] Gangbo, W. and McCann, R., "Optimal maps in monge's mass transport problem," *C. R. Acad. Sci. Paris Série I Math.*, vol. 321, pp. 1653–1658, 1995.

[15] Gangbo, W. and McCann, R., "The geometry of optimal transportation," *Acta Math.*, vol. 177, pp. 113–161, 1996.

[16] Jordan R., Kinderlehrer D., O. F., "The variational formulation of the fokker–planck equation," *SIAM J. Math Anal.*, vol. 29, pp. 1–17, 1998.

[17] Lisini, S., "Nonlinear diffusion equations with variable coefficients as gradient flows in wasserstein spaces," *ESAIM Control Optim. Calc. Var.*, To appear.

[18] Otto, F., "$L^1$-contraction and uniqueness for quasilinear elliptic–parabolic equations," *J. Differential Equations*, vol. 131, pp. 20–38, 1996.

[19] Otto, F., "Dynamics of labyrinthine pattern formation in magnetic fluids: A mean-field theory," *Arch.Rational Mech. Anal.*, vol. 141, pp. 63–103, 1998.

[20] Otto, F., "Doubly degenerate diffusion equations as steepest descent," *http://www-mathphys.iam.uni-bonn.de/∼otto/publications.html*, Preprint 1996, Last visited June 05, 2009.

[21] Villani, C., *Topics in optimal transportation.* American Mathematical Society, 2003. Graduate Studies in Mathematics **58**.

[22] Villani, C., "Optimal transport, old and new, notes for the 2005 saint-flour summer school," *Grundlehren der mathematischen Wissenschaften*, To appear.