Displacement Convexity for First-Order Mean-Field Games

Thesis by
Tommaso Seneci

In Partial Fulfillment of the Requirements

For the Degree of

Masters of Science

King Abdullah University of Science and Technology
Thuwal, Kingdom of Saudi Arabia

April, 2018
The thesis of Tommaso Seneci is approved by the examination committee.

Committee Chairperson: Professor Diogo Gomes
Committee Members: Professor Hernando Ombao, Professor Peter Markowich
ABSTRACT

Displacement convexity for first-order mean-field games
Tommaso Seneci

In this thesis, we consider the planning problem for first-order mean-field games (MFG). These games degenerate into optimal transport when there is no coupling between players. Our aim is to extend the concept of displacement convexity from optimal transport to MFGs. This extension gives new estimates for solutions of MFGs.

First, we introduce the Monge-Kantorovich problem and examine related results on rearrangement maps. Next, we present the concept of displacement convexity. Then, we derive first-order MFGs, which are given by a system of a Hamilton-Jacobi equation coupled with a transport equation.

Finally, we identify a large class of functions, that depend on solutions of MFGs, which are convex in time. Among these, we find several norms. This convexity gives bounds for the density of solutions of the planning problem.
I would like to start this acknowledgement by thanking my advisor, professor Diogo Gomes, for guiding me throughout my stay here at KAUST. He helped me to understand the direction of my studies and the direction of my future, in a moment when I could not firmly decide my path. Moreover, he pushed me to accomplish my thesis; without his perseverance, I would have never finished my work. Thanks also go to my previous advisor, professor Matteo Parsani. He helped during the first weeks at KAUST, by providing useful suggestions and being someone I could rely on.

Next, I would like to thank my friends here at KAUST. My colleague Mohammed, for the laughs at the workstation. My housemate Francisco, for the funny moments we had at home. My mathematical colleagues, for sharing charming thoughts about maths, even on things we did not know. To all my other friends, thanks for the time spent together and for cheering me up. A special thank goes to Alfiya, for being always there for me and bringing an immense joy in my life.

At last, but not least, I would like to thank my family. Even if they were distant from me, they have always been present in my life during the past two years. Thanks go to my mother Silvia and my father Massimo, for taking care of every one of my needs. To my brother Pietro and my sister Cecilia, for surprising me every time I were coming home. To my uncle Roberto and my aunt Titti, for cheering me up and treating me like a baby, even if I am 23 years old. And finally, special thanks go to my grandparents, Mariella, Armando and Marisa, for celebrating every small achievement in my life.
# TABLE OF CONTENTS

Examination Committee Page 2  
Copyright 3  
Abstract 4  
Acknowledgements 5  
1 Introduction 7  
2 Optimal Transport 13  
2.0.1 Monge-Kantorovich Problem 13  
2.0.2 The Existence of an Optimal Map 15  
2.1 The Dual Problem 17  
3 Displacement Convexity 19  
3.1 A Variational Problem 19  
3.2 Displacement Convex Functions 21  
3.3 The Time-Dependent Optimal Transport Problem 22  
3.4 The Benamou-Brenier Formulation 23  
4 Mean-field Games 28  
4.1 Hamilton-Jacobi Equation 28  
4.2 Transport Equation 30  
4.3 First-Order Mean-Field Games 30  
5 Displacement Convexity for Mean-Field Games 34  
5.1 Convex Functionals for First-Order Mean-Field Games 34  
5.1.1 $L^p$ Estimates 41  
5.1.2 Convexity of $L^q$ Norms 45  
5.1.3 Entropy Estimate 47  
References 49
Chapter 1

Introduction

Optimal transport studies the cost of moving mass between different locations. This mass distribution is composed of infinite many particles. Different trajectories transport particles into a target distribution. The starting and ending points of each particle determine an infinitesimal cost. We seek to identify the optimal trajectory that minimizes the total cost. Deterministic optimal transport was introduced by Monge in 1781 in [1]. Unfortunately, Monge’s problem is hard to study. In 1942, Kantorovich relaxed Monge’s problem as an infinite dimensional linear programming problem [2].

Mean-field games (MFG) are systems that study the interaction between identical rational agents. Each agent chooses its path to minimize a value function. In classical MFG, agents choose their trajectories given an initial configuration and a final cost. In the planning problem, the initial and final configurations of the agents are prescribed, and the final cost is unknown. These systems were introduced independently in 2006/2007 by Lasry and Lions in [3, 4] and, approximately at the same time, by Huang, Malhame and Caines in [5].

We define displacement convexity as a property for functions defined on measure spaces. It is an alternative concept of convexity which provides new tools for the study of minimization problems. Displacement convexity was introduced by Robert J. McCann in 1997 in [6], in the context of a non-convex variational problem. Displacement convexity turned his problem into a convex minimization problem.
In this thesis, we focus on the planning problem for first-order MFG. These games are given by an Hamilton-Jacobi equation coupled with a transport equation

\[
\begin{aligned}
-u_t + H(x, Du) &= g(m) \\
m_t - \text{div}(mD_pH(x, Du)) &= 0 \quad \forall (x, t) \in \mathbb{T}^d \times (0, T) \\
\int_{\mathbb{T}^d} m(x,t)dx &= 0, \quad m(\cdot, t) \geq 0 \quad \forall t \in (0, T) \\
m(\cdot,0) &= m^0(\cdot), \quad m(\cdot, T) = m^T(\cdot).
\end{aligned}
\]

(1.1)

In classical first-order MFG, (1.1) is endowed with initial \( m(\cdot, 0) = m^0(\cdot) \) and terminal \( u(\cdot, T) = u^T(\cdot) \).

A classical solution of (1.1) is a periodic \((u(x,t), m(x,t)) \in C^\infty(\mathbb{T}^d \times [0,T])\), where \( \mathbb{T}^d \) is the \( n \)-dimensional torus. The function \( m \) represents the statistical distribution of the agents in the space, whereas \( u \) represents their value function. \( H(x, Du) \) accounts for the spatial cost of the agents, their cost and their preferred direction of motion. \( g(m) \) determines the interactions between agents. The probability measures \( m^0 \) and \( m^T \) are, respectively, the initial and final configuration of the mass.

There are few results, in the literature, on the planning problem for MFG. Consider the case where \( H(p) = a|p|^2, a \in \mathbb{R} \) and \( g(m) \) is non-decreasing. Existence and uniqueness of smooth solutions for second-order MFG were studied by Lions in [7]. Second-order MFGs are systems of Hamilton-Jacobi-Bellman equation and Fokker-Plank equation

\[
\begin{aligned}
-u_t - \eta \Delta u + H(x, Du) &= g(m) \\
m_t - \eta \Delta m - \text{div}(mD_pH(x, Du)) &= 0 \quad \forall (x, t) \in \mathbb{T}^d \times (0, T) \\
\int_{\mathbb{T}^d} m(x,t)dx &= 0, \quad m(\cdot, t) \geq 0 \quad \forall t \in (0, T) \\
m(\cdot,0) &= m^0(\cdot), \quad m(\cdot, T) = m^T(\cdot),
\end{aligned}
\]

(1.2)
for $\eta > 0$. Concerning more general Hamiltonians, the existence of weak solutions for second order MFG was examined by Porretta in [8, 9]. Numerical schemes were developed by Achdou, Camilli, and Capuzzo-Dolcetta in [10].

About classical second-order MFG, existence and uniqueness of solutions were studied by Gomes and Pimentel in [11], and by Gomes, Pimentel and Sanchez Morgado in [12, 13].

Chapter 2 is devoted to the study of the optimal transport problem. At each pair $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$, $x$ being the initial state and $y$ the final state, we associate a lower semi-continuous cost $c(x, y)$. The optimal transport problem consists in minimizing the cost functional

$$
\pi \mapsto \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y) d\pi(x, y), \quad \pi \in \Pi[\mu, \nu].
$$

(1.3)

The probability measures $\mu$ and $\nu$ give the initial and final configurations of the mass. $\Pi[\mu, \nu]$ is the set of probabilities measures with marginals $\mu$ and $\nu$. Each $\gamma \in \Pi[\mu, \nu]$ is interpreted as a generalized path connecting $\mu$ to $\nu$.

In the time-dependent optimal transport problem, the mass evolves, in time, from an initial configuration $\mu$ to a final $\nu$. Particles follow piecewise $C^1$ trajectories $\{\{T_x(t)\}_{0 \leq t \leq 1}\}_{x \in \mathbb{R}^d}$. To each trajectory, we associate an infinitesimal differential cost $C(\{T_x(t)\}_{0 \leq t \leq 1}) = \int_0^1 c(\dot{T}_x(t)) dt$. Again, the total cost is given by adding together the infinitesimal costs of all particles. In Chapter 3, we study the minimization of the differential cost functional

$$
\{\{T_x(t)\}_{0 \leq t \leq 1}\}_{x \in \mathbb{R}^d} \mapsto \int_{\mathbb{R}^d} \int_0^1 c(\dot{T}_x(t)) dt d\mu(x), \quad T_x(0) = x, \quad \nu = (T_{(\cdot)}(1))_{#\mu}.
$$

(1.4)
The following optimality conditions characterize minimizers

\[
\begin{aligned}
    u_t + \frac{|Du|^2}{2} &= \bar{H} \\
    (\rho^t)_t + \text{div}(\rho^t Du) &= 0,
\end{aligned}
\]  

(1.5)

where \( Du = \dot{T}_x(t) \) and \( \rho^t = (T^t(\cdot))(\#\mu) \).

The optimal transport problem has been widely studied and applied to different fields. A general introduction to the topic was developed in [14, 15]. Applications range from physical sciences, such as transport networks [16, 17, 18, 19] and classical mechanics [20, 21, 22], to economics [23, 24]. Optimal transport also applies to other mathematical fields. The link with Riemannian geometry was established in [25, 26]. In [27, 28, 29], optimal control was used to derive new Sobolev-type inequalities.

The primary goal of this thesis is to identify functions \( U : \mathbb{R}^+ \to \mathbb{R} \) such that

\[
U : [0, T] \to \mathbb{R}, t \mapsto \int U(m(x, t))dx \quad \text{is convex in time,} \quad (1.6)
\]

where \( m(x, t) \) is the density of solutions of the MFG (1.1). We find a class of convex functions for which this property holds. Among these, we find functionals of the form \( U(t) = \int m(x, t)^pdx \). As a corollary, we get new estimates for the density of solutions of MFG.

**Proposition 1.0.1.** Let \( u, m \in C^\infty(\mathbb{T}^d \times [0, T]) \) be periodic solutions of (5.12). Suppose that \( g : \mathbb{R}^+ \to \mathbb{R}, H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R} \) are smooth, with \( g \) non-decreasing, \( H(x, \cdot) \) convex and

\[
D^2_{xp}H(x, p)D_pH(x, p) = D^2_{pp}H(x, p)D_xH(x, p) \quad \forall \ (x, p) \in \mathbb{T}^d \times \mathbb{R}^d.
\]  

(1.7)
Then, for all $1 \leq p \leq \infty$

$$\|m(\cdot, t)\|_{L^p(\mathbb{T}^d)} \leq \left(1 - \frac{t}{T}\right)\|m^0(\cdot)\|_{L^p(\mathbb{T}^d)} + \frac{t}{T}\|m^T(\cdot)\|_{L^p(\mathbb{T}^d)}, \quad \forall \ t \in [0, T]. \quad (1.8)$$

Then, we improve this result by showing that inequality (1.8) is strict, provided that $H(x, \cdot)$ is uniformly convex, $g'(m) \geq Cm^\alpha$, for some $C > 0, \alpha \geq 0$, and $\|m^0\|_{L^p}, \|m^T\|_{L^p} > 1$.

Our proof of convexity is based on the following idea. In [6] the author showed that

$$U : \mathcal{P}_{ac}(\mathbb{R}^d) \to \mathbb{R}, \ m \mapsto \int U(m(x))dx \quad (1.9)$$

is displacement convex, given that $x \mapsto x^dU(x^{-d})$ is convex and non-increasing. Based on this result, in [14] another theorem states that if $m$ is a solution of (1.5) and $P(x) = U'(x)x - U(x) \geq 0$, with $P$ growing fast enough, then $U(m(\cdot, t))$ is convex in time. In the case of first-order MFG, we have a coupling of (1.5) through a function of the mass $g = g(m)$. This suggested that, if $g(m)$ behaves nicely, the same result holds for $m$ density of solutions of first-order MFG.

We remark that convexity of $U$ for mean-field games is obtained through differentiation. Nevertheless, $U$ needs not be $C^2$, but only $C^1$. In the case of $L^p$ norms, this means that the proposition applies for all $1 \leq p < \infty$. Convexity for $p = \infty$ is obtained by passing to the limit as $p \to \infty$.

We end this introduction with an outline of this thesis. The existence of minimizers for Kantorovich’s problem is discussed in Chapter 2, where a characterization of minimizers is also derived. In Chapter 3, we connect displacement convexity with the time-dependent optimal transport problem. Next, we show how these two concepts lead to a system similar to first-order MFG. In Chapter 4, we provide a heuristic derivation of first-order MFG. Finally, in Chapter 5, we prove our main result about
displacement convexity of norms and other functionals.
Chapter 2

Optimal Transport

2.0.1 Monge-Kantorovich Problem

Optimal transport studies the cost of moving mass between different locations. This problem was considered for the first time by Monge in the 18-th century. Subsequently, the problem was reformulated by Kantorovich in a more general way.

We consider two measure spaces \( X, Y \) endowed with probabilities \( \mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y) \). For each \((x,y) \in X \times Y\), we associate a positive lower semi-continuous cost function \( c(x,y) \). Suppose that every particle \( x \in X \) is transported to a unique particle \( y = T(x) \in Y \). Moreover, assume that \( \nu \) is the pushforward of \( \mu \) through \( T \), i.e.

\[
\forall \, B \subset Y \text{ measurable}, \quad \nu(B) = T_\# \mu(B) \equiv \mu(T^{-1}(B)). \tag{2.1}
\]

The Monge’s problem is formulated as the minimization of the total cost over all measurable maps \( T \) moving \( X \) to \( Y \)

\[
\inf_{T \text{ such that } T_\# \mu = \nu} \int_X c(x, T(x))d\mu(x). \tag{2.2}
\]

The existence of a \( T \) such that \( \nu = T_\# \mu \) is not trivial. Moreover, the total cost is a nonlinear function in \( T \). These two issues make the Monge’s problem hard to study.

We now relax this setting by allowing "splitting" of mass. Particles move from \( X \) to \( Y \) according to a joint probability measure \( \pi \in \Pi[\mu, \nu] = \{ \pi \in P(X \times Y) : \gamma \text{ has marginals } \mu \text{ and } \nu \} \). Every \( \pi \in \Pi[\mu, \nu] \) is characterized by the following equa-
\[ \int_{X \times Y} \psi(x) + \phi(y) d\pi(x, y) = \int_X \psi(x) d\mu(x) + \int_Y \phi(y) d\nu(y) \quad \forall \psi \in L^1(X), \phi \in L^1(Y). \] (2.3)

To each joint probability \( \pi \) we associate a total cost

\[ I[\pi] = \int_{X \times Y} c(x, y) d\pi(x, y). \] (2.4)

The Kantorovich problem consists in minimizing of the total cost over all \( \pi \in \Pi[\mu, \nu] \), i.e.

\[ \inf_{\pi \in \Pi[\mu, \nu]} I[\pi] = \inf_{\pi \in \Pi[\mu, \nu]} \int_{X \times Y} c(x, y) \pi(x, y). \] (2.5)

Note that \( \Pi[\mu, \nu] \) is always non-empty, for \( \mu \otimes \nu \in \Pi[\mu, \nu] \).

The change of variable formula suggests that Monge’s problem can be written as a differential equation. Recall that, for all \( \phi \in L^1(Y) \), \( T : X \to Y \) measurable

\[ \int_Y \phi(y) dT_\#\mu = \int_X \phi(T(x)) d\mu. \] (2.6)

If we plug the above equation into (2.3) and simplify the expression, we obtain

\[ \int_Y \phi(y) d\nu(y) = \int_X \phi(T(x)) d\mu(x). \] (2.7)

Now suppose that \( T \) is a diffeomorphism \( T : \mathbb{R}^d \to \mathbb{R}^d \) and \( \mu \ll \mathcal{L}^d \), \( \nu \ll \mathcal{L}^d \), with given densities, respectively, \( f \) and \( g \). We can perform a standard change of variable

\[ \int_{\mathbb{R}^d} \phi(T(x)) f(x) dx = \int_{\mathbb{R}^d} \phi(x) g(x) dx = \int_{\mathbb{R}^d} \phi(T(x)) g(T(x)) |\det(DT(x))| dx \] (2.8)
and recover a partial differential equation for $T$

$$g(T(x))\det |(DT(x))| = f(x). \quad (2.9)$$

We treat the theory of optimal transport only in the case $X = Y = \mathbb{R}^d$. Most of the results still hold if $X$ and $Y$ are replaced by more general spaces, such as Polish spaces\footnote{Separable completely metrizable topological space. See \cite{30} for more information on measure theory in metric spaces.}. We also consider measures that are absolutely continuous with respect to the Lebesgue measure. This assumption can often be replaced by taking measures that do not give mass to small sets\footnote{Set of Hausdorff dimension at most $n - 1$.} We refer to \cite{14} for a treatment on optimal transport that takes into account these generalizations.

### 2.0.2 The Existence of an Optimal Map

Monge’s minimization problem requires the existence of a transport map $T$ such that $\nu = T_# \mu$. On the other side, the minimization of Kantorovich’s problem is performed over a non-empty set. For this reason, we study first Kantorovich’s problem. It turns out that, if $\mu, \nu \in \mathcal{P}_{ac}(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d)$, a minimizing measure for Kantorovich’s problem exists and can be written as $\gamma = (Id \times T)_# \mu$, for some $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$. In particular, this implies $\nu = T_# \mu$, which means that $T$ is a minimizer for Monge’s problem. So, we talk about the Monge-Kantorovich problem in general.

We present two results concerning the existence of a minimizer for (2.5), with $c(x, y) = |x - y|^2$. Initial and final measures are taken to be $\mu, \nu \in \mathcal{P}_{ac}(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d)$. We refer to \cite{14}, section 2, Theorem 2.12, for their proofs.

**Theorem 2.0.1 (Knott-Smith).** A measure $\pi \in \Pi[\mu, \nu]$ is optimal for (2.5) if and only if it is concentrated on the graph of the subdifferential of convex lower semi-
continuous function \( \phi \), i.e.

\[
supp(\pi) \subset Graph(\partial \phi).
\] (2.10)

**Theorem 2.0.2** (Brenier). Let \( \pi \in \Pi[\mu, \nu] \) be optimal for (2.5), and \( \phi \) be a convex lower semi-continuous function such that

\[
supp(\pi) \subset Graph(\partial \phi).
\] (2.11)

Then \( \phi \) is differentiable \( d\mu \)-almost everywhere. Moreover, the optimal measure \( \pi \) is unique and can be represented as

\[
\pi = (Id \times D\phi)_{\#}\mu.
\] (2.12)

The existence of a transport map \( D\phi \) such that

\[\nu = D\phi_{\#}\mu\] (2.13)

is an immediate corollary of Brenier’s Theorem. In the literature, \( \phi \) is called the Brenier’s map.

Given probability measures \( \mu \) and \( \nu \), the existence of a ”rearrangement” map \( T \) such that \( \nu = T_{\#}\mu \) is a more general problem than the Monge-Kantorovich problem. For this reason, the transport map could exist even if \( \mu \) and \( \nu \) do not have a finite second moment. Brenier, followed by a refinement carried out by McCann \[31\], proved the following.

**Theorem 2.0.3** (McCann-Brenier). Given \( \mu \in \mathcal{P}_{ac}(\mathbb{R}^d) \) and \( \nu \in \mathcal{P}(\mathbb{R}^d) \), there exists
a unique gradient of a convex map $D\phi$ such that

$$\nu = D\phi\#\mu. \quad (2.14)$$

The uniqueness of $\phi$ must be understood in the following way: if a gradient of a convex function $D\psi$ realizes the pushforward, then $D\phi = D\psi \mu$-almost everywhere. In principle, there could exist another map $T$ such that $\nu = T\#\mu$. For example, fix a measurable map $T : \mathbb{R}^d \to \mathbb{R}^d$ which is not a gradient of a convex function. Set $\nu = T\#\mu$. Then, Brenier’s theorem still applies to $\mu$ and $\nu$, and yields the existence of $D\phi$, $\phi$ convex, such that $T\#\mu = \nu = D\phi\#\mu$.

2.1 The Dual Problem

As we already observed, Kantorovich’s problem is a linear minimization problem. We briefly introduce its dual formulation, which explains why the Brenier’s map is the gradient of a convex function.

**Theorem 2.1.1.** Let $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$ and $c : \mathbb{R}^d \to \mathbb{R}^+ \cup \{\infty\}$ lower semi-continuous. Define the following set

$$\Phi_c = \{ (\phi, \psi) \in L^1(d\mu) \times L^1(d\nu) : \phi(x) + \psi(y) \leq c(x, y) \quad d\mu \otimes d\nu\text{-almost everywhere} \} \quad (2.15)$$

and the dual cost

$$J(\phi, \psi) = \int_X \phi(x) d\mu(x) + \int_Y \psi(y) d\nu(y). \quad (2.16)$$

Then

$$\inf_{\pi \in \Pi[\mu, \nu]} I[\pi] = \sup_{(\phi, \psi) \in \Phi_c} J(\phi, \psi). \quad (2.17)$$

If the cost function is $c(x, y) = |x - y|^2$ and the marginals $\mu, \nu$ are in $\mathcal{P}_2(\mathbb{R}^d) \cap$
\(P_{ac}(\mathbb{R}^d)\), we can rewrite

\[
\int_{X \times Y} \frac{|x - y|^2}{2} d\pi(x, y) = \int_{X \times Y} -x \cdot y \, d\pi(x, y) + \frac{M_2(\mu)}{2} + \frac{M_2(\nu)}{2},
\]

(2.18)

where \(M_2(\mu)\) stands for the second moment of \(\mu\). An infimum on the left-hand side corresponds to a supremum on the right-hand side. We recover an auxiliary dual formulation:

\[
\inf_{\pi \in \Pi[\mu, \nu]} I[\pi] = -\inf_{(\phi, \psi) \in \Phi_{x \cdot y}} J(\phi, \psi) + \frac{M_2(\mu)}{2} + \frac{M_2(\nu)}{2}.
\]

(2.19)

If a pair \((\phi, \psi)\) belongs to \(\Phi_{x \cdot y}\), we can construct a new pair in the following way

\[
x \cdot y - \phi(y) \leq \psi(x) \iff \sup_{y} (x \cdot y - \phi(y)) = \phi^*(x) \leq \psi(x).
\]

(2.20)

The same transformation on \(\phi\) gives a pair \((\phi^{**}, \phi^*)\) such that

\[
J(\phi, \psi) \geq J(\phi, \phi^*) \geq J(\phi^{**}, \phi^*).
\]

(2.21)

This trick is known as "double convexification". So, if minimizers for the dual problem exist, we can assume they are a pair of convex conjugate functions \((\phi, \phi^*)\). The map that pushes \(\mu\) towards \(\nu\) is exactly \(D\phi\). Nevertheless, the maximum for the (non-auxiliary) dual problem is achieved by \(\left(\frac{|x|^2}{2} - \phi(x), \frac{|y|^2}{2} - \phi^*(y)\right)\).
Chapter 3

Displacement Convexity

3.1 A Variational Problem

The concept of Displacement Convexity was introduced by Robert McCann in [6], to study a non-convex variational problem. His problem considered an "interacting gas model in which the force of attraction increases with distance (between particles)."

The density of the gas is a probability density $\rho \in P_{ac}(\mathbb{R}^d)$. Each particle is subject to two different forces. One is given by an interaction potential $W(x - y)$ which increases with the distance between particles. The other is the internal energy, denoted by $U$. The total potential is $\mathcal{W}(\rho) = \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} W(x - y) d\rho(x) d\rho(y)$, and the total internal energy is $\mathcal{U}(\rho) = \int_{\mathbb{R}^d} U(\rho(x)) dx$. The density of the steady-state gas minimizes the total energy

$$E(\rho) = \mathcal{U}(\rho) + \mathcal{W}(\rho).$$

(3.1)

In the problem analysed by McCann, $W$ is symmetric, strictly convex, and increasing; for our purpose, $W$ needs only to be convex. The assumptions on $U$ will be derived later.

Given the variational nature of the problem, we are interested in studying the convexity of the total energy $E$. If $U$ is convex, then $\mathcal{U}$ is convex as a function of the mass. However, even if $W$ is convex, we can not conclude the same about $\mathcal{W}$. To see this, take a convex combination of Dirac’s deltas $\delta_x$ and $\delta_y$, $x \neq y$.

Thus, we look for a new way of interpolating two probabilities densities $\mu, \nu \in P(\mathbb{R}^d)$ that reveals a hidden convexity in $\mathcal{U}$ and $\mathcal{W}$. The change of variable formula,
suggests looking for interpolants through pushforward measures. For a given family \( \{ T(t) \}_{t \in (0, 1)} \), \( T(t) : \mathbb{R}^d \to \mathbb{R}^d \), we define \( \rho^t = T(t) \# \rho \). Then we compute

\[
W(\rho^t) = \int \int W(x - y) dT(t) \# \rho(y) dT(t) \# \rho(x) = \int \int W(T_x(t) - T_y(t)) d\rho(y) d\rho(x).
\]

If \( T(t) \) were to linearly interpolate two given densities \( \mu \) and \( \nu \), \( W \) would become convex. Brenier’s theorem guarantees that there exists a unique gradient of a convex function \( T = D\phi : \mathbb{R}^d \to \mathbb{R}^d \) such that \( \nu = T \# \mu \). Hence, we define \( \rho^t \) to be a linear ”pushforward” interpolant of \( \mu \) and \( \nu \) in the following way.

**Definition 3.1.1.** Let \( \mu \in \mathcal{P}_{ac}(\mathbb{R}^d) \), \( \nu \in \mathcal{P}(\mathbb{R}^d) \) and \( T : \mathbb{R}^d \to \mathbb{R}^d \) the unique gradient of a convex function such that \( \nu = T \# \mu \). The displacement interpolant between \( \mu \) and \( \nu \) is

\[
\rho^t = ((1 - t)x + tD\phi(x)) \# \mu.
\]

For simplicity, we set \( T(t) = (1 - t)(\cdot) + tT(\cdot) \). Furthermore, we write \([\mu, \nu]^t\) instead of \( \rho^t \) whenever we want to emphasize that \( \rho^t \) is interpolating between \( \mu \) and \( \nu \).

We can use the preceding definition to generalize the notion of convexity for sets and functions.

**Definition 3.1.2.** A set \( Q \subset \mathcal{P}_{ac}(\mathbb{R}^d) \) is displacement convex if, for all \( \mu, \nu \in Q \), the displacement interpolant \( \rho^t \) stays in \( Q \).

**Definition 3.1.3.** Let \( Q \subset \mathcal{P}_{ac}(\mathbb{R}^d) \) be a displacement convex set. A function \( F : Q \to \mathbb{R} \) is displacement convex if, for all \( \mu, \nu \in Q \),

\[
t \mapsto F(\rho^t) \quad \text{is convex.}
\]

In light of these definitions, we can now proceed to revisit (3.2) and obtain a
convex estimate

\[\mathcal{W}(\rho^t) = \iint W((1-t)(x-y) + t(T(x) - T(y)))d\mu(y)d\mu(x)\]
\[\leq (1-t) \iint W(x-y)d\mu(y)d\mu(x) + t \iint W(T(x) - T(y))d\mu(y)d\mu(x)\]
\[= (1-t)\mathcal{W}(\mu) + t \iint W(x-y)d\mathbb{T}_#\rho(y)d\mathbb{T}_#\mu = (1-t)\mathcal{W}(\mu) + t\mathcal{W}(\nu).\]

3.2 Displacement Convex Functions

The energy functional studied by McCann takes into account internal energy \(U(\rho) = \int U(\rho(x))dx\) and interaction potential \(\mathcal{W}(\rho) = \iint W(x - y)d\rho(x)d\rho(y)\). Internal energy is convex in the usual sense, but the interaction potential is not. By taking a displacement interpolant \(\rho^t\), we turned \(\mathcal{W}(\rho^t)\) into a convex function. We now focus on the internal energy \(U\).

It is important to remark that, if either \(\mu \in \mathcal{P}_{ac}(\mathbb{R}^d)\) or \(\nu \in \mathcal{P}_{ac}(\mathbb{R}^d)\), then \([\mu, \nu]^t\) stays absolutely continuous. This is so because the convexity of the pushforward does not allow concentration of mass on single points. For this reason, the internal energy \(U(\rho^t)\) makes sense for all times \(t \in [0, 1]\).

The sole convexity of \(U\) is not enough for proving displacement convexity of \(U\). Intuitively, we expect the internal energy to decrease with the scattering of the gas. For this reason, we require \(U\) to have the following property:

\[x \mapsto x^dU(x^{-d}), x \in \mathbb{R}^+ \text{ is convex, non-increasing and } U(0) = 0.\]  

(3.5)

The scaling factor \(d\) is justified by the fact that the expansion of the gas takes place in a \(d\)-dimensional space. Under this assumption, we obtain the following result

**Theorem 3.2.1** (McCann). *Consider two probability measures \(\mu, \nu \in \mathcal{P}_{ac}(\mathbb{R}^d)\) and*
their displacement interpolant \( \rho^t = [\mu, \nu]^t \in \mathcal{P}_{ac}(\mathbb{R}^d) \). If (3.5) holds, then

\[
t \mapsto \mathcal{U}(\rho^t) = \int U(\rho^t(x)) \, dx \quad \text{is convex.} \tag{3.6}
\]

The convexity of \( \phi \), which gradient pushes \( \mu \) to \( \nu \), plays an essential role in the proof. This enforces its use in the definition of displacement interpolation.

We remark that strict convexity of \( W \) implies strict displacement convexity of \( W \) only when \( \mu \) is not a translate of \( \nu \). About \( \mathcal{U} \), the strict displacement convexity fails whenever \( D^2 \phi = I \) in the Aleksandrov sense \(^1\).

At last, we mention the following simple case of a displacement convex functional. Let \( V : \mathbb{R} \to \mathbb{R}^+ \) be convex, and consider the functional

\[
\mathcal{V}(\rho) = \int_{\mathbb{R}^d} V(x) \, d\rho(x). \tag{3.7}
\]

Change of variable (2.7) and the convexity of \( V \) imply displacement convexity of \( \mathcal{V} \).

### 3.3 The Time-Dependent Optimal Transport Problem

Consider two probability densities \( \mu, \nu \in P_{ac}(\mathbb{R}^d) \). Suppose that each particle moves from \( \mu \) to \( \nu \) according to a piecewise \( C^1 \) trajectory \( T_x(t) : [0, 1] \to \mathbb{R}^d \). At time \( t = 0 \), particles stay in place, so that \( T_x(0) = x \). At time \( t = 1 \), they reach \( \nu \), so \( \nu = T_x(1) \# \mu \). The time-dependent optimal transport problem consists of minimizing a displacement cost \( C = C(T_x(\cdot)) \) over all trajectories \( \{T_x(\cdot)\}_{x \in \mathbb{R}^d} \), i.e.

\[
\inf \left\{ \int_X C(\{T_x(t)\}_{t \in (0,1)}) \, d\mu(x) : T_x(0) = x, \ T_x(1) \# \mu = \nu \right\}. \tag{3.8}
\]

If

\[
c(x, y) = \inf \{ C(\{T_x(t)\}_{t \in (0,1)}) \mid T_x(0) = x, \ T_x(1) = y \} \tag{3.9}
\]

\(^1\)Equivalently, \( D^2 \phi = I \) in the sense of distributions.
this problem predicts the same cost as the time-independent Monge problem.

An interesting case is represented by differential cost functions

\[
C(\{T_x(t)\}_{t \in (0,1)}) = \int_0^1 c(\dot{T}_x(t))dt,
\]

(3.10)

where \(c\) is a convex function. Thanks to Jensen’s inequality we find

\[
\int_0^1 c(\dot{T}_x(t))dx \geq c \left( \int_0^1 \dot{T}_x(t)dx \right) = c(y - x).
\]

(3.11)

Since straight lines are also admissible trajectories, they are minimizers.

In particular, in the case \(c(x) = |x|^2\), minimizing straight lines are exactly displacement interpolants.

**Proposition 3.3.1.** Let \(\mu, \nu \in \mathcal{P}_{ac}(\mathbb{R}^d)\) and \(D\phi\) such that \(\nu = D\phi \# \mu\), \(\phi\) convex. The differential cost

\[
\inf \left\{ \int_{\mathbb{R}^d} \int_0^1 |\dot{T}_x(t)|^2 dtd\mu(x) : T_x(0) = x, \ \nu = T_x(1) \# \mu \right\}
\]

(3.12)

is minimized by the displacement interpolant

\[
\rho_t = ((1 - t)x + tD\phi(x)) \# \mu.
\]

(3.13)

The proof is just a rewriting of the considerations above.

### 3.4 The Benamou-Brenier Formulation

Here we describe a continuous time formulation of optimal transport, that is due to Benamou and Brenier [32].

For a given map \(T(x, t) = T_x(t)\) such that \(T_x(0) = x\), we consider the pushforward
measure
\[ \rho^t = T(x)(t) \# \mu. \] (3.14)

We aim to get a differential equation solved by \( \rho^t \). To simplify the computations, we consider a Lipschitz velocity field \( v_x(t) = v(x, t) \) and a trajectory \( T_x(t) \) solving the initial value problem
\[
\begin{cases}
\dot{T}_x(t) = v(T_x(t), t) = v_{T_x(t)}(t) \\
T_x(0) = x.
\end{cases}
\] (3.15)

**Theorem 3.4.1.** Let \( \{T(x)(t)\}_{0 \leq t < T} \) be a locally Lipschitz family of diffeomorphisms, which trajectories have velocity \( v(x, t) \). Given \( \mu \in \mathcal{P}(\mathbb{R}^d) \), the pushforward \( \rho^t = T(x)(t) \# \mu \) is the unique solution of the transport equation
\[
\frac{\partial \rho^t}{\partial t} + \text{div}(\rho^t v) = 0
\] (3.16)
in \( C([0, T); \mathcal{P}(\mathbb{R}^d)); \mathcal{P}(\mathbb{R}^d) \) endowed with the weak topology.

Equation \( (3.16) \) is known as the "mass conservation equation".

As an intermediate step, we recover a PDE solved by constant speed trajectories.

**Theorem 3.4.2.** Consider a continuous, a.e. differentiable function \( v(x)(0) : \mathbb{R}^d \to \mathbb{R}^d \) and the time-displacement interpolant \( T_x(t) = x - tv_x(0) \). Suppose also that \( \{T(x)(t)\}_{0 \leq t < T} \) are diffeomorphisms. Then, the velocity field \( v(x)(t) = T(x)(t)^{-1} \circ \frac{\partial T(x)(t)}{\partial t} \) solves
\[
\frac{\partial v}{\partial t} + v \cdot Dv = 0.
\] (3.17)

Suppose that \( v \) is the gradient of a function \( u \). If \( v \) solves \( (3.17) \), then \( u \) solves an Hamilton-Jacobi equation
\[
0 = \frac{\partial v}{\partial t} + v \cdot Dv = Du_t + Du \cdot D^2u = D \left( u_t + \frac{|Du|^2}{2} \right) \iff u_t + \frac{|Du|^2}{2} = \bar{H}, \ \bar{H} \in \mathbb{R}.
\] (3.18)
If we combine (3.18) with (3.16), we get the following PDE system

\[
\begin{cases}
\frac{\partial u}{\partial t} + |Du|^2 = \bar{H} \\
\frac{\partial \rho}{\partial t} + \text{div}(\rho^t Du) = 0
\end{cases}
\]  

(3.19)

The system is uncoupled. One can solve the first equation for \(u\) and then solve the second for a given datum \(Du\). First-order mean-field games are recovered by coupling the two equations by adding a function of the mass, \(g = g(m)\), to the Hamilton-Jacobi equation.

Let \(v(x, t) = v_x(t)\) be a Lipshitz velocity field and consider \(\{T(t)\}_0^1\) such that \(T_x(t)\) solves (3.15).

The pushforward measure \(\rho^t = (T(t))_{\#}\mu\) is the unique solution of the PDE (3.4.1)

\[
\begin{cases}
\frac{\partial \rho^t}{\partial t} + \text{div}(\rho^t v_x(t)) = 0 \\
\rho(0, \cdot) = \mu(\cdot).
\end{cases}
\]  

(3.20)

Moreover, if \(T = D\phi\) is the Brenier’s map between \(\mu \in \mathcal{P}_{ac}(\mathbb{R}^d)\) and \(\nu \in \mathcal{P}_{ac}(\mathbb{R}^d)\) and \(v_x(t) = T(x) - x\), then \(\rho^1 = (T(1))_{\#}\mu = \nu\). This means that, among the solutions of (3.19), we can always find \((\rho, v)\) such that \(\rho^t = [\mu, \nu]^t\).

The idea of Benamou and Brenier [32] was to consider the kinetic energy of the particles

\[E(t) = \int_{\mathbb{R}^d} \rho^t(x) \frac{|v_x(t)|^2}{2} dx,\]  

(3.21)

and minimize the total energy over time

\[A[\rho, v] = \int_0^1 E(t) dt = \int_0^1 \int_{\mathbb{R}^d} \rho^t(x) \frac{|v_x(t)|^2}{2} dx dt.\]  

(3.22)
The following theorem was proved for the first time by Benamou-Brenier in [32]. A relaxed version can be found in [14].

**Theorem 3.4.3.** Let $\mu, \nu \in P_{ac}(\mathbb{R}^d)$ be compactly supported. Set $V$ to be the family of smooth functions $(\rho, v)$ solving (3.20) with $\rho^0 = \mu$ and $\rho^1 = \nu$. Then, for a quadratic cost function $c(x, y) = \frac{|x-y|^2}{2}$,

\[
\inf_{\pi \in \Pi[\mu, \nu]} I[\pi] = \inf_{(\rho, v) \in V} A[\rho, v]. \tag{3.23}
\]

**Proof.** From (2.0.2), we know that Kantorovich’s problem admits a minimizer of the form $\pi = (Id \times D\phi)\#\mu$. This shows that $\inf A \geq \inf I$, since

\[
A[\rho, v] = \int_{\mathbb{R}^d} \int_0^1 \rho^t(x)|v_x(t)|^2 dt dx = \int_{\mathbb{R}^d} \int_0^1 \rho^0(x)|v(T_x(t), t)|^2 dt dx \tag{3.24}
\]

\[
= \int_{\mathbb{R}^d} \int_0^1 |\dot{T}_x(t)|^2 dt d\mu \geq \int_{\mathbb{R}^d} |T_x(1) - T_x(0)|^2 d\mu \tag{3.25}
\]

\[
= \int_{\mathbb{R}^d} |T_x(1) - x|^2 d\mu \geq \int_{\mathbb{R}^d} |D\phi(x) - x|^2 d\mu = \inf I[\pi], \tag{3.26}
\]

We recover the opposite inequality by noticing that the vector field $T_x(t) = (1 - t)x + tD\phi(x)$ has optimal velocity for $I[\pi]$

\[
v_x(t) = \dot{T}_x(t) = \frac{\partial}{\partial t}((1-t)x + tD\phi(x)) = D\phi(x) - x. \tag{3.27}
\]

The new formulation has a remarkable property: if we use as variables $(\rho, \rho v) \equiv (\rho, m)$, the minimization problem becomes a convex minimization problem

\[
\inf_{(\rho, v) \in V} A[\rho, v] = \inf_{(\rho, m)} \int_0^1 \int_{\mathbb{R}^d} \frac{|m(x, t)|^2}{2\rho^t(x)} dx dt. \tag{3.28}
\]

This allows to extend (3.4.3) to more general $\rho^0, \rho^1$ and $v$. Moreover, it facilitates
numerical simulations, as shown in the paper by Benamou-Brenier [32].
Chapter 4

Mean-field Games

Mean-Field games describe the interaction among identical rational agents in competition. In these systems, each agent seeks to optimize a given cost function, taking into account statistical information about other agents. These systems comprise two partial differential equations. One encodes an optimal control problem that agents seek to optimize (Hamilton-Jacobi equation); the other accounts for the movement of the agents (mass transport equation). This second equation describes the evolution of a mass distribution driven by a flow.

Before, we presented first-order MFG as a variation of the optimality equations of the Monge-Kantorovich problem. Now, we show how to derive first-order MFG from a heuristic point of view. The following presentation is based on [33].

4.1 Hamilton-Jacobi Equation

The Hamilton-Jacobi equation comes from an optimal control problem. Consider a very large population of agents. The state of an agent at time $t \in [0, T]$ is determined by the space variable $x = x(t) \in \mathbb{R}^d$.

Suppose that agents change their state according to a velocity field

$$\dot{x}(t) = v(t), \quad v \in L^\infty([t, T], \mathbb{R}^d), 0 < t < T. \quad (4.1)$$

At each time $t \in (0, T)$, each agent pays an infinitesimal cost that depends both on the position and velocity of particles. This cost is encoded by a Lagrangian
\( L(x, v) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \). At time \( t = T \), a continuous and bounded from below \( u^T : \mathbb{R}^d \to \mathbb{R} \) determines a terminal cost. The total cost of moving an agent along a trajectory \( \{x(t)\}_{0 \leq t \leq T} \) is given by adding together the infinitesimal costs and the terminal cost

\[
J(x, v, t) = \int_t^T L(x(s), v(s))ds + u^T(x(T)). \tag{4.2}
\]

Agents seek to minimize the total energy along every admissible trajectory. Thus, we define the value function to be the optimal cost

\[
u(x, t) = \inf_{v \in L^{\infty}([t, T], \mathbb{R}^d)} J(v, x, t). \tag{4.3}\]

Let \( H(x, p) \) be the Hamiltonian associated to \( L(x, v) \), i.e. the Legendre’s transform of \( L(x, v) \) w.r.t. the velocity variable

\[H(x, p) = \sup_{v \in \mathbb{R}^d} (p \cdot v - L(x, v)). \tag{4.4}\]

If \( u \) is \( C^1 \), then it solves the Hamilton-Jacobi equation

\[
\begin{cases}
-u_t(x, t) + H(x, D_x u(x, t)) = 0 & (x, t) \in \mathbb{R}^d \times (0, T) \\
u(x, T) = u^T(x) & x \in \mathbb{R}^d.
\end{cases} \tag{4.5}
\]

If \( L(x, v) \) is uniformly Lipschitz in the second variable, its Hamiltonian \( H(x, p) \) is differentiable. Furthermore, we recover the following relation for the optimal control

\[
\tilde{v}(t) = -D_p H(\tilde{x}(t), D_x u(\tilde{x}(t), t)). \tag{4.6}
\]

So, \(-D_p H(\tilde{x}(t), D_x u(\tilde{x}(t), t))\) gives the optimal direction for the agents located at \( \tilde{x}(t) \) at the time \( t \).
4.2 Transport Equation

Let $b : \mathbb{T}^d \times [0, T] \to \mathbb{R}$ be a Lipschitz vector field and consider the Cauchy’s problem

$$
\begin{aligned}
&\dot{x}(t) = b(x(t), t) \quad t > 0 \\
x(0) = x.
\end{aligned}
$$

(4.7)

Suppose that the agents move with velocity $b$. The flow $\Phi^t(x)$ determines the trajectories of the agents. Assume furthermore that agents do not disappear nor increase in number. This means that the total mass is conserved. Hence, at time $t \in (0, T]$, the location of each agent is determined by the pushforward of $m^0$ through the flow $\Phi^t(x)$, namely

$$
m(t, x) \equiv m^t(x) = \Phi^t(x)\#m^0(x).
$$

(4.8)

By Theorem (3.4.1), we recover that the dynamic of the particles solves uniquely the transport equation

$$
\begin{aligned}
&m_t(x, t) + \text{div}(b(x, t)m(x, t)) = 0 \quad (x, t) \in \mathbb{T}^d \times [0, T] \\
m(x, 0) = m^0(x) \quad x \in \mathbb{T}^d.
\end{aligned}
$$

(4.9)

4.3 First-Order Mean-Field Games

We now couple the Hamilton-Jacobi equation and the transport equation to model a system where agents move according to an optimal velocity.

From the theory of Hamilton-Jacobi equation, we know that if $\bar{v}$ maximizes $p \cdot v - L(x, v)$, that is, $H(x, p) = \sup_v (p \cdot v - L(x, v)) = p \cdot \bar{v} - L(x, \bar{v})$, then

$$
\bar{v} = -D_p H(x, p).
$$

(4.10)

Hence, we consider trajectories that move with velocity $b = \bar{v}$. We obtain the following
uncoupled system
\[
\begin{cases}
-u_t + H(x, D_x u) = 0 \\
m_t - \text{div}(m D_p H(x, D_x u)) = 0.
\end{cases}
\tag{4.11}
\]
Consider now a smooth function \( g : \mathbb{R}^+ \to \mathbb{R} \) that "encodes the interactions between each agent and the mean-field". We modify the Hamilton-Jacobi equation, adding \( g \) to the r.h.s of (4.11). Thus, agent’s motion now depends on the distribution of the other agents, and we obtain the first-order MFG
\[
\begin{cases}
-u_t + H(x, D_x u) = g(m) \\
m_t - \text{div}(m D_p H(x, D_x u)) = 0.
\end{cases}
\tag{4.12}
\]

The uniqueness of solutions of first-order MFG was proved by Lasry and Lions in [34].

**Theorem 4.3.1.** Consider a first-order MFG
\[
\begin{cases}
-u_t + H(x, Du) = g(m) \\
m_t - \text{div}(m D_p H(x, Du)) = 0 \quad \forall (x,t) \in \mathbb{T}^d \times [0, T].
\end{cases}
\tag{4.13}
\]
Assume that \( g(m) \) is strictly increasing and \( H(x, \cdot) \) is \( C^1 \) strictly convex for all \( x \in \mathbb{T}^d \).

The following holds:

- if (4.13) is endowed with initial condition \( m(x,0) = m^0(x) \) and final condition \( u(x,T) = u^T(x) \) (classical problem), there exists at most one classical solution \((u(x,t), m(x,t))\);

- if (4.13) is endowed with initial \( m(x,0) = m^0(x) \) and final \( m(x,T) = m^T(x) \)
planning problem), there is at most one classical solution \( m(x, t) \), while \( u(x, t) \) is unique up to an additive constant.

**Proof.** Let \((m_1, u_1)\) and \((m_2, u_2)\) be two solutions of (4.13) and define \( \bar{m} = m_1 - m_2 \) and \( \bar{u} = u_1 - u_2 \). The couple \((\bar{m}, \bar{u})\) solves

\[
\begin{align*}
-\bar{u}_t + H(x, Du_1) - H(x, Du_2) &= g(m_1) - g(m_2) \\
\bar{m}_t - \text{div}(m_1 D_p H(x, Du_1) - m_2 D_p H(x, Du_2)) &= 0.
\end{align*}
\]

(4.14)

Now, we multiply the first equation by \( \bar{m} \) and the second by \( \bar{u} \), subtract them, and integrate by parts over \( \mathbb{T}^d \), to conclude that

\[
\int_{\mathbb{T}^d} \bar{m}_t \bar{u} + \bar{u}_t \bar{m} = \int_{\mathbb{T}^d} -(m_1 D_p H(x, Du_1) - m_2 D_p H(x, Du_2)) D\bar{u} \\
+ \bar{m}(H(x, Du_1) - H(x, Du_2)) - \bar{m}(g(m_1) - g(m_2)).
\]

The left-hand side equals 0 because \( \bar{m}_0 = \bar{m}_T = 0 \) in the planning problem, and \( \bar{m}_0 = \bar{u}_T = 0 \) in the classical problem. Furthermore, since \( H(x, \cdot) \) is strictly convex,

\[
H(x, Du_1) > H(x, Du_2) + D_p H(x, Du_2) D\bar{u}, \quad Du_1 \neq Du_2; \quad (4.15)
\]

mutatis mutandis for \( H(x, Du_2) \). Then, we estimate

\[
\bar{m}(H(x, Du_1) - H(x, Du_2)) = m_1 H(x, Du_1) - m_2 H(x, Du_1) - m_1 H(x, Du_2) + m_2 H(x, Du_2) \\
\leq (m_1 D_p H(x, Du_1) - m_2 D_p H(x, Du_2)) D\bar{u},
\]

with strict inequality unless \( D\bar{u} = Du_1 - Du_2 = 0 \). This means that

\[
\int_{\mathbb{T}^d} \bar{m}(g(m_1) - g(m_2)) \leq 0; \quad (4.16)
\]
that is, $m_1 = m_2$, which in turn implies $D u_1 = D u_2$. 
Chapter 5

Displacement Convexity for Mean-Field Games

5.1 Convex Functionals for First-Order Mean-Field Games

Before stating the main theorem of this section, we recall few preliminary results.

Theorem 3.2.1 states that if $x \mapsto x^d U(x^{-d})$, $x \in \mathbb{R}^+$ is convex and non-increasing and $U(0) = 0$, then $U : \rho \in P_{ac}(\mathbb{R}^d) \mapsto \int U(\rho) dx$ is displacement convex.

Besides, Theorem 3.4.1 characterizes displacement interpolants as unique solutions of mass transport equations. First-order MFGs are recovered by varying the value function that transports the mass in the transport equation.

These three facts suggest that densities of solutions of first-order MFG behave similarly to displacement interpolants. Here we prove that if $m(\cdot, t) \in P_{ac}(\mathbb{R}^d)$ solves a first-order MFG, then $t \to U(m(\cdot, t))$ is convex.

However, the assumptions on $U$ are not suitable for lengthy computations. In [14], the author derives nicer equivalent conditions to 5.1.

Suppose that $U$ is the density of internal energy of a system that depends only on the pressure $P$. $U$ satisfies the thermodynamic equation

$$P(x) = x U'(x) - U(x).$$  \hspace{1cm} (5.2)
Formally, \( U \) can be written in an integral form

\[
U(x) = x \int_0^x \frac{P(s)}{s^2} ds.
\] (5.3)

The convexity condition on \( U \) translates into a (first order) differential condition on the pressure \( P \)

\[
xP'(x) \geq \left(1 - \frac{1}{d}\right) P(x) \iff \frac{P(x)}{x^{1-1/d}} \text{ non-decreasing.} \quad (5.4)
\]

Consequently, given \( U \in C^1(\mathbb{R}^+) \), if there is \( P \) satisfying

\[
\begin{align*}
P(x) &= U'(x)x - U(x), \\
P &\in C^1(\mathbb{R}^+), \quad P(x) \geq 0, \\
\frac{P(x)}{x^{1-1/d}} &\text{ non-decreasing},
\end{align*}
\] (5.5)

then \( U \) is displacement convex.

Notice that \( P \) non-negative implies that \( P' \) is also non-negative because

\[
xP'(x) \geq \left(1 - \frac{1}{d}\right) P(x) \geq 0.
\] (5.6)

Moreover, if we differentiate \( P \) we obtain

\[
P'(x) = U''(x)x + U'(x) - U'(x) = U''(x)x \geq 0 \iff U''(x) \geq 0, \ x \geq 0, \quad (5.7)
\]
i.e. \( U \) is convex.

\( U \) needs to be continuously differentiable because we need to differentiate it once and perform integration by parts. However, \( U \) needs not be \( C^2(\mathbb{R}^+) \), so functions of the form \( U(x) = x^p, 1 < p < 2 \) satisfy (5.5).

We need the following lemma.
Lemma 5.1.1. Let $A \in \mathbb{R}^{d \times d}$, then $\text{tr}(A^2) \geq \frac{1}{d} \text{tr}(A)^2$

Proof. Decompose $A$ into its Jordan’s form, i.e. $A = SJS^{-1}$. If $\{\lambda_i\}_{i=1}^d$ are eigenvalues of $A$, then

$$\text{tr}(A) = \text{tr}(SJS^{-1}) = \text{tr}(JS^{-1}S) = \sum_{i=1}^d \lambda_i. \quad (5.8)$$

Similarly, for $A^2$,

$$\text{tr}(A^2) = \text{tr}(SJ^2S^{-1}) = \text{tr}(J^2S^{-1}S) = \sum_{i=1}^d \lambda_i^2. \quad (5.9)$$

Cauchy-Schwartz inequality yields the final result

$$\text{tr}(A)^2 = \left(\sum_{i=1}^d \lambda_i\right)^2 \leq \sum_{i,j=1}^d \lambda_i^2 + \frac{\lambda_j^2}{2} = d \sum_{i=1}^d \lambda_i^2 = d \text{tr}(A^2). \quad (5.10)$$

We can now prove the main theorem.

Theorem 5.1.2. Consider smooth functions $g : \mathbb{R}^+ \to \mathbb{R}$, $H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$, with $g$ non-decreasing, $H(x,\cdot)$ convex, and

$$D^2_{xp}H(x,p)D_pH(x,p) = D^2_{pp}H(x,p)D_xH(x,p) \quad \forall (x,p) \in \mathbb{T}^d \times \mathbb{R}^d. \quad (5.11)$$

Let $m, u \in C^\infty(\mathbb{T}^d \times [0,T])$ be periodic solutions of the first-order mean-field game

$$\begin{cases}
-u_t + H(x,Du) = g(m) \\
m_t - \text{div}(mD_pH(x,Du)) = 0 \\
m(x,t) \geq 0.
\end{cases} \quad (5.12)$$
If $U : \mathbb{R}^+ \to \mathbb{R}$ is such that (5.5) holds, then

$$t \mapsto \int_{T^d} U(m(x,t))dx \text{ is convex.} \quad (5.13)$$

Proof. We compute one derivative of $t \mapsto \int U(m(x,t))$

$$\frac{d}{dt} \int U(m) = \int U'(m)m_t = \int U'(m) \text{div}(mD_pH) \quad (5.14)$$

$$= \int U'(m)m \text{div}(D_pH) + U'(m)DmD_pH \quad (5.15)$$

$$= \int U'(m)m \text{div}(D_pH) + D(U(m))D_pH \quad (5.16)$$

$$= \int U'(m)m \text{div}(D_pH) - U(m) \text{div}(D_pH) = \int P(m) \text{div}(D_pH). \quad (5.17)$$

The second derivative is

$$\frac{d^2}{dt^2} \int U(m) = \int P'(m)m_t \text{div}(D_pH) + P(m) \text{div}(D_pH_t) \quad (5.18)$$

$$= \int P'(m) \text{div}(mD_pH) \text{div}(D_pH) + P(m) \text{div}(D^2_{pp}H Du_t) \quad (5.19)$$

$$= \left( A \right) + \left( B \right) + \left( C \right) + \left( D \right) \quad (5.20)$$

We aim to generalize Lemma 5.43 in [14]; thus, we expect to get an inequality of the form

$$\frac{\partial^2}{\partial t^2} \int U(m) \geq \int \left( mP'(m) - P(m) + \frac{P(m)}{d} \right) \text{div}(D_pH)^2 dx + \text{ non-negative terms} \quad (5.22)$$

$$= \left( A' \right) + \left( B' \right) + \left( C' \right) + \frac{P(m)}{d} \text{div}(D_pH)^2 dx \quad (5.23)$$

$$+ \text{ non-negative terms.} \quad (5.24)$$
\( A = A' \) already appeared. We integrate by parts \( B \) to get

\[
\int P'(m) Dm D_p H \text{ div}(D_p H) = \int D(P(m)) D_p H \text{ div}(D_p H) = - \int P(m) \text{ div}(D_p H \text{ div}(D_p H))
\]

(5.25)

\[
= \int (-P(m) \text{ div}(D_p H)^2) - P(m) D_p H D(\text{div}(D_p H)) .
\]

(5.26)

The assumption on the partial derivatives of \( H \) is equivalent to the following (in Einstein notation)

\[
D^2_{pp} H D(H) = H_{p_k, p_j} (H) x_j = H_{p_k, p_j} (H x_j + H_{p_k} u_{x_k, x_j}) = H_{p_k, p_j} u_{x_j, x_k} H_{p_k} + H_{p_k, p_j} H x_j
\]

(5.27)

\[
= (H_{p_k}) x_j H_{p_k} x_j + H_{p_k, p_j} H x_j
\]

(5.28)

\[
= D(D_p H) D_p H - D^2_{xp} H D_p H + D^2_{pp} H D x H = D(D_p H) D_p H.
\]

(5.29)

Thus, \( C \) is equal to \( P(m) \text{ div}(D(D_p H) D_p H) \). To proceed, we need the following equality

\[
\text{div}(D(D_p H) D_p H) = (H_{p_k}) x_j H_{p_j} x_j = (H_{p_k}) x_j, x_j H_{p_j} + (H_{p_k}) x_j (H_{p_j}) x_j
\]

(5.30)

\[
= D(\text{div}(D_p H)) D_p H + \text{tr}((D(D_p H))^2).
\]

(5.31)
Then, we can expand $C$

$$
\int P(m) \text{div}(D_{pp}^2 HD(H)) = \int P(m) \text{div}(D(D_p H) D_p H) \\
= \int \underbrace{P(m)D(\text{div}(D_p H))}_{Q} D_p H + P(m) \text{tr}((D(D_p H))^2)
$$

(5.32)

and cancel $Q$ from (5.25). Finally, $D$ is

$$
\int (-P(m) \text{div}(D_{pp}^2 HD(g(m)))) = \int P'(m)g'(m) Dm D_{pp}^2 HDm.
$$

(5.33)

Considering the preceding identities, we get

$$
\frac{d^2}{dt^2} \int U(m) = \int \underbrace{P'(m)m \text{div}(D_p H)^2}_{A'} + \underbrace{(-P(m) \text{div}(D_p H)^2)}_{B'} \\
+ P(m) \text{tr}((D(D_p H))^2) + P'(m)g'(m) Dm D_{pp}^2 HDm.
$$

(5.34)

Because $P$ is non-negative, we can apply lemma (5.1.1) to $(D(D_p H))^2 \in \mathbb{R}^{d \times d}$ to obtain

$$
\int P(m) \text{tr}((D(D_p H))^2) \geq \int \frac{1}{d} P(m) \text{tr}(D(D_p H))^2 = \int \underbrace{\frac{1}{d} P(m) \text{div}(D_p H)^2}_{C'}.
$$

(5.35)

Since $g(x)$, $P(x)$ and $\frac{P(x)}{x^{\frac{1}{1-q}}}$ are non-decreasing and $H(x, \cdot)$ is convex, $D$ is non-negative, so we achieve the desired result

$$
\frac{d^2}{dt^2} \int U(m) \geq \int \left( P'(m)m - P(m) + \frac{1}{d} P(m) \right) \text{div}(D_p H)^2 \\
+ P'(m)g'(m) Dm D_{pp}^2 HDm \geq 0
$$

(5.36)
The assumptions on the derivatives of \( H \) apply to various Hamiltonians, as the next two example illustrate.

**Example 5.1.3.** Consider \( A, B \in \mathbb{R}^{d \times d} \) with \( AB^T \) symmetric, \( c \in \mathbb{R}^d \) and let \( H(x, p) = \frac{|Ax+Bp+c|^2}{2} \). We compute the partial derivatives

\[
D_x H = A^T(Ax+Bp+c), \quad D_p H = B^T(Ax+Bp+c), \quad D^2_{xp} H = B^T A, \quad \text{and} \quad D^2_{pp} H = B^T B. \tag{5.40}
\]

\( H(x, p) \) satisfies (5.5) because

\[
D^2_{xp}HD_pH = B^T AB^T(Ax+Bp+c) = B^T BA^T(Ax+Bp+c) = D^2_{pp}HD_xH, \tag{5.41}
\]

Moreover, \( H(x, \cdot) \) is always convex, since

\[
x^TD^2_{pp}Hx = x^TB^TBx = |Bx|^2 \geq 0. \tag{5.42}
\]

**Example 5.1.4.** Let \( \phi : \mathbb{R} \to \mathbb{R} \) be convex. Take any \( v \in \mathbb{R}^d \), and define

\[
H(x, p) = \phi(v \cdot (x + p)). \tag{5.43}
\]

Its partial derivatives are

\[
D_x H = D_p H = \phi' v, \quad D^2_{xp} H = D^2_{pp} H = \phi'' v \otimes v. \tag{5.44}
\]

This implies

\[
D^2_{xp}HD_pH = \phi' \phi'' v \|v\|^2 = D^2_{pp}HD_xH. \tag{5.45}
\]

Moreover, \( H(x, \cdot) \) is convex because

\[
x \cdot D^2_{pp}Hx = \phi'' x \cdot (v \otimes v x) = \phi'' |x \cdot v|^2 \geq 0. \tag{5.46}
\]
5.1.1 \(L^p\) Estimates

In the previous section, we identified conditions on \(U\) such that \(\int U(m(x,t))dx\) is convex in time. We now show that \(U(\cdot) = (\cdot)^q, 1 \leq q < \infty\) satisfy such conditions. This gives us new estimates on the density of solutions of first order MFG.

Proposition 5.1.5. Let \(u, m \in C^\infty(\mathbb{T}^d \times [0,T])\) be periodic solutions of (5.12). Suppose that \(g : \mathbb{R}^+ \to \mathbb{R}, H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}\) are smooth, with \(g\) non-decreasing, \(H(x,\cdot)\) convex and

\[
D^2_{xp}H(x,p)D_pH(x,p) = D^2_{pp}H(x,p)D_xH(x,p) \quad \forall (x,p) \in \mathbb{T}^d \times \mathbb{R}^d. \tag{5.47}
\]

Then, for all \(1 \leq q < \infty\)

\[
\|m(\cdot,t)\|_{L^q(\mathbb{T}^d)} \leq \left(1 - \frac{t}{T}\right) \|m^0(\cdot)\|_{L^q(\mathbb{T}^d)} + \frac{t}{T} \|m^T(\cdot)\|_{L^q(\mathbb{T}^d)}, \quad \forall t \in [0,T]. \tag{5.48}
\]

Moreover

\[
\|m(\cdot,t)\|_{L^\infty(\mathbb{T}^d)} \leq \max\{\|m^0(\cdot)\|_{L^\infty(\mathbb{T}^d)}, \|m^T(\cdot)\|_{L^\infty(\mathbb{T}^d)}\}, \quad \forall t \in [0,1]. \tag{5.49}
\]

Proof. First, we prove the inequality for \(q\) finite; then we pass to the limit to recover the bound for \(p = \infty\).

- \(q = 1\). This case is trivial since

\[
\frac{d}{dt} \int m = \int m_t = \int \text{div}(mD_pH) = 0. \tag{5.50}
\]

- \(1 < q < \infty\). According to theorem (5.1.2), we only need to show that the \(L^q\) norms meet conditions (5.5)

\[
U'(x)x - U(x) = qx^q - x^q = (q-1)x^q = P(x) \in C^1(\mathbb{R}^+),\text{ is non-negative}. \tag{5.51}
\]
Moreover, \( P'(x) = (q - 1)x^{q - 1 + \frac{1}{q}} \) is increasing because \( q - 1 + \frac{1}{q} > 0 \).

- \( q = \infty \). Take the convex estimate for \( q \) finite and elevate it to the power \( \frac{1}{q} \):

\[
\|m(\cdot, t)\|_{L^{q}(\mathbb{T}^d)} \leq \left( \left( 1 - \frac{t}{T} \right) \|m^0(\cdot)\|_{L^{q}(\mathbb{T}^d)}^{q} + \frac{t}{T} \|m^T(\cdot)\|_{L^{q}(\mathbb{T}^d)}^{q} \right)^{\frac{1}{q}}
\]

(5.52)

\[
\leq \left( 2 \max \left\{ \left( 1 - \frac{t}{T} \right) \|m^0(\cdot)\|_{L^{q}(\mathbb{T}^d)}^{q}, \frac{t}{T} \|m^T(\cdot)\|_{L^{q}(\mathbb{T}^d)}^{q} \right\} \right)^{\frac{1}{q}}
\]

(5.53)

\[
\leq 2^{\frac{1}{q}} \max \left\{ \|m^0(\cdot)\|_{L^{q}(\mathbb{T}^d)}, \|m^T(\cdot)\|_{L^{q}(\mathbb{T}^d)} \right\}
\]

(5.54)

Since the set \( \mathbb{T}^d \) is bounded, \( \|f\|_{L^{q}(\mathbb{T}^d)} \xrightarrow{p \to \infty} \|f\|_{L^{\infty}(\mathbb{T}^d)} \). In particular, we can pass to the limit on both sides:

\[
\|m(\cdot, t)\|_{L^{\infty}(\mathbb{T}^d)} \leq \max \left\{ \|m^0(\cdot)\|_{L^{\infty}(\mathbb{T}^d)}, \|m^T(\cdot)\|_{L^{\infty}(\mathbb{T}^d)} \right\}, \quad \forall \ t \in [0, T].
\]

(5.55)

The estimate can be improved even further in the case \( \int U(m) = \|m\|_{L^{q}(\mathbb{T}^d)}^q, 1 < p < \infty \). From (5.51) we obtain \( P(m) = (q - 1)m^q, P'(m) = q(q - 1)m^{q-1} \). Combine with (5.38) to get

\[
\frac{d^2}{dt^2} \int m^q \geq C \int m^{q-1}g'(m)DmD_{pp}^2HDm.
\]

(5.56)

Assume that \( g'(m) \geq Cm^\alpha \) for some \( C > 0, \alpha \geq 0 \) and \( D_{pp}^2H \) is uniformly positive definite. The above estimate becomes

\[
\frac{d^2}{dt^2} \int m(x,t)^q \geq C \int m(x,t)^{q-1+\alpha}|Dm(x,t)|^2 \geq C \int |D(m^{\frac{q+1+\alpha}{2}})|^2.
\]

(5.57)

The integral on the right is 0 if and only if \( m(\cdot, t) = 1 \). We arrive at the next lemma.

**Lemma 5.1.6.** Consider smooth functions \( g: \mathbb{R}^+ \to \mathbb{R}, H: \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}, \) with
\[ g'(x) \geq C x^\alpha, \ C > 0, \alpha \geq 0, \ H(x, \cdot) \text{ uniformly convex and} \]

\[ D_{xp}^2 H(x, p) D_p H(x, p) = D_{pp}^2 H(x, p) D_x H(x, p) \quad \forall \ (x, p) \in \mathbb{T}^d \times \mathbb{R}^d. \quad (5.58) \]

Let \( m, u \in C^\infty(\mathbb{T}^d \times [0,T]) \) be periodic solutions of the first-order MFG

\[
\begin{cases}
-u_t + H(x, Du) = g(m) \\
m_t - \text{div}(m D_p H(x, Du)) = 0 \\
m(x, t) \geq 0, \int_{\mathbb{T}^d} m(x, t) dx = 1.
\end{cases}
\quad (5.59)
\]

Then, for \( 1 < q < \infty \),

\[
\mathcal{U} : t \mapsto \|m(\cdot, t)\|_{L^q(\mathbb{T}^d)}^q \quad (5.60)
\]

is strictly convex at \( \bar{t} \) if and only if there is \( \delta > 0 \) such that \( D_x m(\cdot, t) \neq 0 \) for all \( t \in (\bar{t} - \delta, \bar{t} + \delta) \).

**Proof.** Convexity is due to theorem \( (5.1.2) \). We notice that if \( m(\cdot, t) \) is constant for all \( t \in (\bar{t} - \delta, \bar{t} + \delta) \), then

\[
\int_{\mathbb{T}^d} m(x, t) dx = 1 = m(x, t). \quad (5.61)
\]

Thus, \( m \) is also constant in time over such interval.

\[
\Rightarrow \text{ Suppose that } \mathcal{U}(t) \text{ is strictly convex over } (\bar{t} - \delta, \bar{t} + \delta). \text{ If } D_x m(x, t) = 0 \text{ for all } t \in (\bar{t} - \delta, \bar{t} + \delta), \text{ then } \frac{d}{dt} \|m(\cdot, t)\|_{L^q(\mathbb{T}^d)}^q = 0, \text{ which contradicts the hypothesis. So } D_x m(x, t) \neq 0.
\]

\[
\Leftarrow \text{ On the other side, assume that } D_x m(x, t) \neq 0 \text{ over a small interval } (\bar{t} - \delta, \bar{t} + \delta). \text{ If } \mathcal{U}(t) \text{ were to be equal to } 0, \text{ then equation } (5.57) \text{ would imply}
\]

\[
0 = C \int |D(m^{\frac{q+1+\alpha}{2}})|^2 \quad \iff \quad D_x m(x, t) = 0 \ \forall \ t \in (\bar{t} - \delta, \bar{t} + \delta). \quad (5.62)
\]
This contradiction shows that $\mathcal{U}(t)$ is strictly convex.

Equality (5.57) gives more information on the behaviour of $m(\cdot,t)$ around the extrema. Suppose that the initial density $m^0$ is "scattered", i.e. its gradient is "very big" in norm. According to (5.57), the second derivative of the norm is also "big". We deduce that the norm decays and approaches the constant function 1. On the other side, if $m^0$ is not scattered, then is already close to the constant function 1. The convex bound shows that $m(x,t)$ stays close to 1 as time increases. The same reasoning applies solutions $m(x,t)$ around $t = T^-$. In conclusion, solutions $m(x,t)$ "tend" to reach the constant function 1 for intermediate times in $[0, T]$.

**Remark 5.1.7.** Lemma (5.1.6) allows us to refine estimate (5.48). Suppose that initial and final data have norms $\|m^0\|_{L^q(T^d)}, \|m^T\|_{L^q(T^d)} > 1$. This is equivalent to $D_xm^0, D_xm^T \neq 0$ due to Jensen’s inequality. Thus, in neighborhoods of $0^+$ and $T^-$,

$$\frac{d^2}{dt^2} \int_{T^d} m(x,t)^q dx > 0.$$  \hspace{1cm} (5.63)

This implies that the convex estimate is strict:

$$\|m(\cdot,t)\|_{L^q(T^d)}^q < \left(1 - \frac{t}{T}\right) \|m^0(\cdot)\|_{L^q(T^d)}^q + \frac{t}{T} \|m^T(\cdot)\|_{L^q(T^d)}^q, \quad \forall \ t \in (0, T).$$  \hspace{1cm} (5.64)

We conclude that, for $g$ and $H$ as in (5.1.6), estimate (5.48) is always strict, unless $m^0 = 1$ or $m^T = 1$. This reasoning cannot be extended to the case $q = \infty$. 


5.1.2 Convexity of $L^q$ Norms

Before, we proved convexity of $t \to \|m(\cdot,t)\|^q_{L^q}$, for $m$ density of solutions of (5.12).

We now extend this result and show that $t \to \|m(\cdot,t)\|_{L^q}$ is convex.

Suppose that $f : \mathbb{R} \to \mathbb{R}^+$ is smooth and convex. We want to find conditions for which $f^{\frac{1}{q}}$, $q > 1$ is also convex. Let $p$ be the conjugate exponent of $q$, i.e. $p = \frac{q}{q-1}$.

Clearly $f^{\frac{1}{q}}$ is smooth, so that we can differentiate twice

$$
(f^{\frac{1}{q}})'' = \frac{1}{q} \left( f^{\frac{1}{q}-1} f' \right)' = \frac{1}{q} \left( \frac{f'}{f^{\frac{1}{q}}} \right)' = \frac{1}{q} \left( f'' f^{\frac{1}{q}} - \frac{1}{p} f^{\frac{1}{q}-1} (f')^2 \right) = \frac{1}{pq} f^{\frac{1}{q}} (q f'' f - (f')^2).
$$

(5.65)

Then, $f^{\frac{1}{q}}$ is convex if and only if $pf'' f \geq (f')^2$. We now show that such condition holds if $f(t) = \|m(\cdot,t)\|_{L^q}$. Thanks to this result, we can prove convexity for the $L^\infty$ norm.

**Proposition 5.1.8.** Let $u, m \in C^\infty(T^d \times [0,T])$ be periodic solutions of (5.12). Suppose that $g : \mathbb{R}^+ \to \mathbb{R}$, $H : T^d \times \mathbb{R}^d \to \mathbb{R}$ are smooth, with $g$ non-decreasing, $H(x, \cdot)$ convex and

$$
D^2_{xp} H(x,p) D_x H(x,p) = D^2_{pp} H(x,p) D_x H(x,p) \quad \forall (x,p) \in T^d \times \mathbb{R}^d.
$$

(5.66)

Then, for all $1 \leq q \leq \infty$

$$
\|m(\cdot,t)\|_{L^q(T^d)} \leq \left( 1 - \frac{t}{T} \right) \|m^0(\cdot)\|_{L^q(T^d)} + \frac{t}{T} \|m^T(\cdot)\|_{L^q(T^d)}, \quad \forall t \in [0,T].
$$

(5.67)

**Proof.** The case $q = 1$ is trivial. So let $1 < q < \infty$. Given (5.65), we only need to show that

$$
p \|m(\cdot,t)\|_{L^q(T^d)} \frac{d^2}{dt^2} \|m(\cdot,t)\|_{L^q(T^d)} \geq \left( \frac{d}{dt} \|m(\cdot,t)\|_{L^q(T^d)} \right)^2
$$

(5.68)
From (5.38) we get a lower bound for the second derivative

\[
\frac{d^2}{dt^2} \|m(\cdot, t)\|_{L^q} \geq (q - 1) \left( q - 1 + \frac{1}{d} \right) \int m^q \, \text{div}(D_pH)^2 + q(q - 1) \int m^{q-1} g'(m) Dm D_p^2 H Dm
\]

(5.69)

\[
\geq (q - 1) \left( q - 1 + \frac{1}{d} \right) \int m^q \, \text{div}(D_pH)^2 \geq (q - 1)^2 \int m^q \, \text{div}(D_pH)^2.
\]

(5.70)

Next, we compute the first derivative

\[
\frac{d}{dt} \|m(\cdot, t)\|_{L^q} = \int P(m) \, \text{div}(D_pH) = (q - 1) \int m^q \, \text{div}(D_pH).
\]

(5.71)

Using Hölder’s inequality, we arrive at the desired conclusion

\[
\left( \frac{d}{dt} \|m(\cdot, t)\|_{L^q(T)} \right)^2 = (q - 1)^2 \left( \int m^{\frac{2}{q}} m^\frac{q}{2} \, \text{div}(D_pH) \right)^2 \leq (q - 1)^2 \int m^q \, \int m^q \, \text{div}(D_pH)^2
\]

(5.72)

\[
\leq \|m(\cdot, t)\|_{L^q} \frac{d^2}{dt^2} \|m(\cdot, t)\|_{L^q} \leq p \|m(\cdot, t)\|_{L^q} \frac{d^2}{dt^2} \|m(\cdot, t)\|_{L^q}.
\]

(5.73)

This shows that \( \| \cdot \|_{L^q} \) are convex for \( 1 \leq q < \infty \). Passing to the limit as \( q \to \infty \) in

\[
\|m(\cdot, t)\|_{L^q(T)} \leq \left( 1 - \frac{t}{T} \right) \|m^0(\cdot)\|_{L^q(T)} + \frac{t}{T} \|m^T(\cdot)\|_{L^q(T)}, \quad \forall \ t \in [0, T],
\]

(5.74)

we get

\[
\|m(\cdot, t)\|_{L^\infty(T)} \leq \left( 1 - \frac{t}{T} \right) \|m^0(\cdot)\|_{L^\infty(T)} + \frac{t}{T} \|m^T(\cdot)\|_{L^\infty(T)}, \quad \forall \ t \in [0, T],
\]

(5.75)

which is convexity for the \( L^\infty \) norm.

\[\square\]

**Remark 5.1.9.** Note that (5.1.6) also applies to this case. To bound the second
derivative we kept aside the positive term $\int g'(m) Dm D^2_{pp} H Dm$. If we keep it, we can pass through the proof of the lemma and get the same result. We conclude that if $g'(x) \geq C x^\alpha$, $C > 0$, $\alpha \geq 0$, $H(x, \cdot)$ uniformly convex and $\|m^0\|_{L^q}, \|m^T\|_{L^q} > 1$, then

$$\|m(\cdot, t)\|_{L^q(T^d)} < \left(1 - \frac{t}{T}\right) \|m^0(\cdot)\|_{L^q(T^d)} + \frac{t}{T} \|m^T(\cdot)\|_{L^q(T^d)}, \quad \forall \ t \in (0, T), \quad (5.76)$$

for all $1 < q < \infty$.

### 5.1.3 Entropy Estimate

Here we prove convexity of the functional $t \mapsto \int m(x, t) \ln(m(t, x)) \, dx$. Notice that, if $\rho \in P_{ac}(\mathbb{R}^n)$ is a probability measure, Jensen’s inequality yields

$$\int_{\mathbb{R}^d} \rho(x) \ln(\rho(x)) \, dx \geq \left(\int_{\mathbb{R}^d} \rho(x) \, dx\right) \ln\left(\int_{\mathbb{R}^d} \rho(x) \, dx\right) = 1 \cdot 0 = 0. \quad (5.77)$$

This shows that convex estimates are not meaningless inequalities of the form $-\infty \leq -\infty$.

If we were to formally find $P$, we would get $P(x) = x$. Unfortunately, $U$ is not smooth at $x = 0$. To avoid this problem, we proceed with an approximation argument.

**Proposition 5.1.10.** Let $u, m \in C^\infty(T^d \times [0, T])$ be periodic solutions of (5.12). Suppose that $g : \mathbb{R}^+ \to \mathbb{R}$, $H : T^d \times \mathbb{R}^d \to \mathbb{R}$ are smooth, $g$ non-decreasing, $H(x, \cdot)$ convex and

$$D^2_{xp} H(x, p) D_p H(x, p) = D^2_{pp} H(x, p) D_x H(x, p) \quad \forall \ (x, p) \in T^d \times \mathbb{R}^d. \quad (5.78)$$

The function

$$t \mapsto \int_{T^d} m(x, t) \ln(m(x, t)) \, dx \quad \text{is convex.} \quad (5.79)$$
Proof. Consider the family of function $U^a = x \ln(x + a)$. We find $P^a$

$$P^a(x) = (U^a)'(x)x - U^a(x) = \ln(x+a)x + \frac{x^2}{x+a} - x \ln(x+a) = \frac{x^2}{x+a} \in C^1(\mathbb{R}^+), \quad P^a(x) \geq 0.$$  

(5.80)

$P^a$ satisfies the growing condition because

$$\frac{P(x)}{x^{1-\frac{1}{a}}} = \frac{x^{1+\frac{1}{a}}}{x+a} \text{ is non-decreasing} \iff \frac{x+a}{x^{1+\frac{1}{a}}} = \frac{1}{x^{\frac{1}{a}}} + \frac{a}{x^{1+\frac{1}{a}}} \text{ is non-increasing.}$$  

(5.81)

From (5.1.2), we get convexity for $t \mapsto \int U^a(m(x,t))dx$, for all $a \geq 0$, i.e.

$$\int_{\mathbb{T}^d} m(x,t) \ln(m(x,t)+a)dx \leq \left(1 - \frac{t}{T}\right) \int_{\mathbb{T}^d} m^0(x) \ln(m^0(x)+a)dx + \frac{t}{T} \int_{\mathbb{T}^d} m^T(x) \ln(m^T(x)+a)dx$$  

(5.82)

We want to pass to the limit as $a \to 0$ in the inequality. We immediately get a lower bound

$$x \ln(x+a) \geq x \ln(x) \geq -e^{-1} \geq -1.$$  

(5.83)

Moreover, if $a \leq 1$, then $m(x,t) \ln(m(x,t)+a) \leq m(x,t)^2 + am(x,t) \leq 2\|m(\cdot,t)\|_{L^\infty}^2$. Therefore

$$|m(x,t) \ln(m(x,t)+a)| \leq \max\{1, 2\|m(\cdot,t)\|_{L^\infty}^2\} = 2\|m(\cdot,t)\|_{L^\infty}^2.$$  

(5.84)

We can apply Dominated convergence theorem and pass to the limit as $a \to \infty$, obtaining the desired convexity.
REFERENCES


