Mass transport and geometric inequalities

by

Felipe García Ramos Aguilar

B. Sc. Mathematics, Universidad Nacional Autónoma de México, 2008

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

THE FACULTY OF GRADUATE STUDIES
(Mathematics)

The University Of British Columbia
(Vancouver)

October 2010

Abstract

In this thesis we will review some recent results of Optimal Mass Transportation emphasizing on the role of displacement interpolation and displacement convexity. We will show some of its recent applications, specially the ones by Bernard, and Agueh-Ghoussoub-Kang.

Table of Contents

Acknowledgments

I would like to thank Prof. Young-Heon Kim, for carefully reading this thesis and making very useful observations and explanations; Ramón Zárate and Craig Cowan, for the Functional Analysis, PDE and Mass Transport conversations that helped the completion of this thesis; Ignacio Rozada, who showed me several things, including how to compile this thesis; and Lee Yupitun for all her time and help at the Math Department.

This thesis was done under the supervision of Prof. Nassif Ghoussoub, and I was economically supported by a CONACyT fellowship.

Al pollo que llegó, dejó su legado
y se fué.

...y al Pollo también.

Chapter 1

Introduction

Mathematics is the art of giving the same name to different things.
— J.H. Poincaré

The most basic problem of modern mass transportation is the Monge-Kantorovich problem for quadratic cost function, that is, given the measures μ_0 and μ_1 of $\mathbb{R}^n$ find a measure γ_0 of $\mathbb{R}^n \times \mathbb{R}^n$, that satisfies

$$\inf_{\gamma \in \Gamma(\mu_0,\mu_1)} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x-y|^2 \, d\gamma(x,y) = \int_{\mathbb{R}^n \times \mathbb{R}^n} |x-y|^2 \, d\gamma_0(x,y),$$

where $\Gamma(\mu_0,\mu_1)$ is the set of measures in $\mathbb{R}^n \times \mathbb{R}^n$ with marginals μ_0 and μ_1. This actually defines a distance $d_w(\mu_0,\mu_1)$ in the probability space $\mathscr{P}(\mathbb{R}^n)$, which we shall denote by Wasserstein distance. In Chapter 1 we will review the basic results of Optimal Mass Transportation. A very good text book for this subject is the one written by Villani [13].

Brenier and Benamou [8] studied a different point of view of the transport problem, one involving time, which is to find a one parameter family of pairs (μ_t, V_t) , that minimize

$$\inf \int_0^1 \int |V_t|^2 \, d\mu_t dt,$$

where μ_t is a measure and V_t a vector field that depends on t and satisfy the continuity equation, or also known as the conservation of mass equation

$$\mu_t + \operatorname{div}(V\mu) = 0.$$

In Chapter 2 we will show the proof of Brenier's theorem, that under some circumstances both problems have the same infimum or optimal cost. In this chapter we will also review the concepts of displacement convexity and displacement interpolation, which are related to the time dependent point of view.

The main idea of chapters 3 and 4 is to explain some of the results given by Bernard [1] and Agueh, Ghoussoub and Kang [10] emphasizing on the role of time dependent mass transportation.

In Chapter 3 we will explain some basic results in displacement interpolation and displacement convexity. Then we will relax Brenier's problem by considering instead of classical flows V_t, *generalized flows* $\eta_{t,x}(v)$, which are probability measures that indicate the probability of having velocity $v \in TM$ at the point x in time t. To generalize the continuity equation we will define a *Transport measure*, as a *Young measure* η that, for a set of test functions g, satisfies

$$\int_{I\times TM} [\partial_t g + \partial_x g \cdot v]\, d\eta_{t,x}(v) d\mu(t,x) = \int_{I\times TM} [\partial_t g + \partial_x g \cdot v]\, d\eta(v,t,x) = 0. \quad (1.1)$$

For certain functionals L we will show existence of minimizers of $\int L d\eta$, first in the case where η is a generalized curve and, under some conditions, we will see that the infimum coincides with the one taken over the classical curves, hence showing a known result previously proved by Tonelli.

Using a simple optimal transport argument and the geometric-arithmetic inequality one can prove the isoperimetric inequality and Sobolev-Nirenberg inequalities. With the use of the Monge-Ampere equation it is also possible to prove log-Sobolev inequalities, HWI, Brunn-Minkowski, and many others (see See [3],[2],[11]). An interesting result is that actually many of this inequalities belong or can be obtained from a more general inequality. Using displacement convexity we will first show a general Sobolev inequality, that involves a positive measurable function ρ that represents density, force F, pressure P_F, the internal energy functional H, the Young's function c^*, and a constant K_c,

$$H^{F+nP_F}(\rho) \leq \int_\Omega c^*(-\nabla F'(\rho))\rho dx + K_c. \quad (1.2)$$

Afterwards we will show that if we choose different forms of the force, we can derive interesting inequalities from this one, like the Log-Sobolev, and the Gagliardo Nirenberg inequality.

Also in this chapter we will prove the general Agueh-Ghoussoub-Kang inequality shown in [10], that involves a Young's function c and its dual c^*, energy functionals $H_V^{F,W}$, relative entropy production $\mathscr{I}_{c^*}(\rho_0 \mid \rho_V)$, Wasserstein distance $d_w^2(\rho_{0,}\rho_1)$ and barycentre $b(\rho_0)$, namely

$$H_{V+c}^{F,W}(\rho_0 \mid \rho_1) + \frac{\lambda+\nu}{2} d_w^2(\rho_{0,}\rho_1) - \frac{\nu}{2} |b(\rho_0) - b(\rho_1)|^2 \\ \leq H_{c+\nabla V\cdot x}^{-nP_F, 2x\cdot\nabla W}(\rho_0) + \mathscr{I}_{c^*}(\rho_0 \mid \rho_V).$$

We will show this inequality generalizes (1.2) and other inequalities including the HWI inequality.

Chapter 2

Kantorovich problem

The original transport problem was proposed by Monge around the 1780´s, the question was how to move given pile of soil into an excavation with the least amount of work. Kantorovich relaxed the problem in terms of probability measures. In this chapter we explain some basic results in this direction, where a basic reference is [13].

Whenever T is map from a measure space (X,μ) to arbitrary space Y, we can equip Y with the pushforward measure $T_{\#}\mu$, where $T_{\#}\mu(B) = \mu(T^{-1}(B))$, for every set $B \subset Y$.

We will denote the space of probability measures of X, as $\mathscr{P}(X)$, and $\mathscr{P}_{AC}(X) \subset \mathscr{P}(X)$ the space of absolutely continuous probability measures. Let $M^{\pm} \subset \mathbb{R}^n$ be two compact sets , and $\mu^{\pm} \in \mathscr{P}(M^{\pm})$. We denote the projection functions as $\pi_+ : M^+ \times M^- \mapsto M^+$, where $\pi_+(x,y) = x$, and $\pi_- : M^+ \times M^- \mapsto M^-$, where $\pi_-(x,y) = y$. So we define

$$\Gamma(\mu^+,\mu^-) = \left\{\gamma \in \mathscr{P}(M^+ \times M^-) \mid (\pi_+)_{\#}\gamma = \mu^+ \text{ and } (\pi_-)_{\#}\gamma = \mu^-\right\}.$$

Let $c : M^+ \times M^- \mapsto \mathbb{R}$, be a continuous cost function. The Kantorovich problem is to minimize the total cost defined as

$$\mathscr{C}(\gamma) = \int_{M^+ \times M^-} c(x,y)d\gamma, \text{ where } \gamma \in \Gamma(\mu^+,\mu^-) .$$

As we shall see the existence of minimizers is not hard, as we have that $\mathscr{C}$ is linear in γ and we can use the Banach-Alaouglu theorem.

Theorem 2.1 *There exists a minimizer for the Kantorovich problem.*

Proof. Since the space of probability measures is contained in the unit ball of the dual space, which is weak* compact. Since $\Gamma(\mu^+, \mu^-)$ is closed, it is weak-* compact. Since $\mathscr{C}(\gamma)$ is continuous in the weak-* topology, it attains a minimum in the compact set. ■

We will say γ is optimal if

$$\gamma \in \Gamma_{op}(\mu^+, \mu^-) := \left\{ \gamma \in \Gamma(\mu^+, \mu^-) \mid \mathscr{C}(\gamma) = \inf_{\gamma \in \Gamma(\mu^+, \mu^-)} \int_{M^+ \times M^-} c(x,y) d\gamma \right\}.$$

2.0.1 Brenier's theorem

First we state two definitions and their relationship.

Definition 2.2 *Given a set $\Gamma \subset X \times Y$, and a cost function $c(x,y)$, we say that Γ is c-cyclically monotone if for any finite set of pairs $\{(x_i, y_i) \mid 1 \leq i \leq N\} \subset \Gamma$. we have that*

$$\sum_i^N c(x_{i,} y_i) \leq \sum_i^N c(x_i, y_{i+1}). \text{ (with N+1=1)}$$

Definition 2.3 *A function $\phi : X \mapsto \mathbb{R} \cup \{+\infty\}$ is said to be c-convex if it is not identical to $\{+\infty\}$, and there exists a function $g : Y \mapsto \mathbb{R} \cup \{+\infty\}$, such that*

$$\phi(x) = \sup_{y \in Y} (g(y) - c(x,y)) \quad \forall x \in X.$$

It´s c-transform is the function

$$\phi^c(y) := \inf_{x \in X} (\phi(x) + c(x,y)),$$

and it's c-subdifferential is the c-cyclically monotone set

$$\partial^c \phi := \{(x,y) \in X \times Y \mid \phi^c(y) - \phi(x) = c(x,y)\}.$$

Remark 2.4 *The fact that a set* Γ *is c-cyclically monotone if and only if* $\Gamma = \partial^c \phi$ *for a c-convex function* ϕ *is called Rockafellers's theorem, a proof can be found in [13].*

Remark 2.5 *If we take* $X = Y = \mathbb{R}^n$ *and* $c(x,y) = -x \cdot y$*, the the c-transform is the usual Legendre transform or dual, and c-convexity is just convexity.*

Lemma 2.6 *The support of an optimal mapping* γ *is c-cyclically monotone.*

Proof. If supp γ is not c-cyclically monotone then we have that, this means there exists a set of pairs $\{(x_i, y_i) \mid 1 \leq i \leq N\} \subset \Gamma$ such that

$$\sum_i^N c(x_i, y_i) - \sum_i^N c(x_i, y_{i+1}) > 0.$$

Even more, since c is continuous we can find a set of open neighbourhood $U_i \times V_i$, of (x_i, y_i) such that

$$\sum_i^N c(u_i, v_i) - \sum_i^N c(u_i, v_{i+1}) > 0 \quad \forall\ (u_i, v_i) \in U_i \times V_i.$$

We define the measures $\gamma_i(E) = \frac{\gamma(E \cap (U_i \times V_i))}{\gamma(U_i \times V_i)}$; let $\eta = \Pi \gamma_i$ a measure of $(X \times Y)^N$ and $H_i = (f_i, g_i\ )$ the projections such that $\gamma_i = H_{i\#}\eta$.

We define

$$\widetilde{\gamma} = \gamma + \frac{\lambda}{n} \sum_i^N (f_{i+1} \times g_i)_\# \eta - (f_i \times g_i)_\# \eta,$$

where $\lambda = \inf \gamma(U_i \times V_i) > 0$. We check that $(\pi_+)_\# \widetilde{\gamma} = \mu^+ + \frac{\lambda}{n} \sum_i^N (f_{i+1})_\# \eta - (f_i)_\# \eta = \mu^+$, similarly $(\pi_-)_\# \widetilde{\gamma} = \mu^-$, and $\widetilde{\gamma}(X \times Y) = 1 + \frac{\lambda}{n} \sum_i^N \frac{\gamma((U_{i+1} \times V_i))}{\gamma(U_{i+1} \times V_i)} - \frac{\gamma((U_i \times V_i))}{\gamma(U_i \times V_i)} = 1$.

Finally we compute

$$\mathscr{C}(\gamma) - \mathscr{C}(\widetilde{\gamma}) = \int_{M^+ \times M^-} \sum c(f_{i+1}, g_i) - c(f_i, g_i) d\eta > 0$$

■

Now we proceed to prove a theorem by Brenier, that applies to quadratic cost functions, i.e. $c(x,y) = \frac{1}{2} |x - y|^2$.

Theorem 2.7 *Let $M^{\pm} \subset \mathbb{R}^n$ be open and bounded sets, and $\mu^{\pm} \in \mathscr{P}_{AC}(M^{\pm})$. Hence there exists an optimal mapping $T_{\#}\mu^{+} = \mu^{-}$, and φ convex function such that*

$$T = \nabla\varphi$$

almost everywhere.

Proof. By the previous results, we know there exists an optimal measure $\gamma_{op} \in \Gamma_{op}(\mu^{+}, \mu^{-})$, let S be a maximal c-cyclical monotone set, hence $\mathrm{supp}\gamma_{op} \subset S$. By Rockafellar theorem, there exists a c-convex function ϕ, such that $S = \partial^c \phi(x)$, by definition of c-convexity we have that

$$\begin{aligned}\phi(x) &= \sup_{y \in M^-} \{-c(x,y) - \phi^c(y)\} = \sup_{y \in M^-} \left\{-\frac{1}{2}|x-y|^2 - \phi^c(y)\right\} \\ &= \sup_{y \in M^-} \left\{-\frac{1}{2}|x|^2 - \frac{1}{2}|y|^2 + xy - \phi^c(y)\right\}.\end{aligned}$$

Now we can define a function φ,

$$\varphi(x) := \phi(x) + \frac{1}{2}|x|^2 = \sup_{y \in M^-} \left\{-\frac{1}{2}|y|^2 + xy - \phi^c(y)\right\}.$$

This function is convex and bounded and moreover, $\partial\varphi(x) = \partial^c\phi$. This means φ is Lipschitz and hence differentiable almost everywhere by Rademacher's theorem So almost everywhere $\partial\varphi(x) = \nabla\varphi(x)$. ■

Monge-Ampere equation

Let $M^{\pm} \subset \mathbb{R}^n$ be open and bounded sets, and $\mu^{\pm} \in \mathscr{P}_{AC}(M^{\pm})$, that is $\mu^{+} = f dx$ and $\mu^{-} = g dy$, from the last theorem there exists an optimal mapping $\mathrm{T}_{\#}\mu^{+} = \mu^{-}$, and φ convex function such that $T = \nabla\varphi$, in other words we have that

$$\int \psi(y)g(y)dy = \int \psi(\nabla\varphi(x))f(x)dx \text{ for all test functions } \psi.$$

Alexandrov's theorem says that $D^2\varphi$ exists almost everywhere if φ is convex, so if we do a change of variables $y = \nabla\ \varphi(x)$ we get that

$$\int \psi(y)g(y)dy = \int \psi(\nabla\varphi(x))g(\nabla\varphi(x))\det D^2\varphi dx.$$

From this we get the *Monge-Ampere's* equation

$$g(\nabla\varphi)\det D^2\varphi = f(x).$$

2.1 Kantorovich-Rubinstein space

Now that we know about the existence of solutions of the Kantorovich problem, we can define a very useful metric. Let $\mathscr{P}_n(X)$ be the space of Borel probability measures on $\mu \in X$ with finite nth moment, which means

$$\int |d(x_0,x)|^n\, d\mu(x) < \infty$$

for some $x_0 \in X$, and hence for any $x_0 \in X$.

Definition 2.8 *We can define a distance in the space $\mathscr{P}_n(X)$ is defined as follows,*

$$d_n(\mu,\eta) = \left|\min_{\lambda} \int_{X\times X} |d(x,y)|^n\, d\lambda(x,y)\right|^{1/n}.$$

Where λ runs along de probability measures on $X \times X$, whose marginals are μ and η. Particularly for $n = 2$, we shall call it the Wasserstein distance and we will denote it as d_w.

We can characterize the topological space $\mathscr{P}_1(X)$, let´s consider $C_1(X)$, the space of functions $f : X \mapsto \mathbb{R}$, such that

$$\sup_{x\in X} \frac{|f(x)|}{1+d(x_0,x)} < \infty$$

for one (and hence for all) $x_0 \in X$. Then we have that $d_1(\mu_n,\mu) \mapsto 0$ if and only if

$$\int f d\mu_n \mapsto \int f d\mu$$

for every $f \in C_1(X)$.

We also define a weaker topology , called the narrow topology

Definition 2.9 *We say μ_n converges narrowly to μ if*

$$\int f d\mu_n \mapsto \int f d\mu$$

for each bounded continuous function f.

Chapter 3

Time dependent mass transportation

3.1 Displacement interpolation

So far we only minded the starting and ending point of the mass transportation problem, without giving any information of what could happen in the middle. This view point is related to fluid dynamics and has been studied principally by Brenier [8]

Definition 3.1 *Let $\mu^+, \mu^- \in \mathscr{P}_2(\mathbb{R}^n)$ and $\gamma_0 \in \Gamma(\mu^+, \mu^-)$ be the solution to the Kantorovich problem with quadratic cost . For every $s \in [0,1]$ we define $\pi_s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, as*

$$\pi_s(x,y) := (1-s)x + sy,$$

and we will call $\mu_s = (\pi_s)_\# \gamma_0$ the displacement interpolation between μ^+ and μ^-.

As we shall see the solutions of the time dependent minimization problems can be represents as displacement interpolation of two measures.

We now prove a result that shows that the displacement interpolation of two measures is a constant speed geodesic.

Theorem 3.2 *Let μ_s be the displacement interpolation between μ_0 and μ_1 then*

$\forall s,t \in [0,1]$, we have that

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

Proof. First we take

$$\pi_{st} = (\pi_s \times \pi_t)_{\#}\gamma_0 = ((1-s)x+sy),(1-t)x+ty)_{\#}\gamma_0 \in \Gamma(\mu_s,\mu_t).$$

So that

$$\begin{aligned} d_w^2(\mu_t,\mu_s) &\le \int |x-y|^2\, d\pi_{st} = \int |(1-s)x+sy-((1-t)x+ty)|^2\, d\gamma_0 \\ &= (t-s)^2 \int |x-y|^2\, d\gamma_0 = (t-s)^2 d_w^2(\mu_0,\mu_1). \end{aligned}$$

To get the equality use the triangle inequality

$$\begin{aligned} d_w(\mu_0,\mu_1) &\le d_w(\mu_0,\mu_s) + d_w(\mu_s,\mu_t) + d_w(\mu_t,\mu_1) \\ &\le s d_w(\mu_0,\mu_1) + d_w(\mu_s,\mu_t) + (1-t) d_w(\mu_0,\mu_1). \end{aligned}$$

So we conclude that

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

■

3.2 Displacement convexity

In this chapter we explain an important concept called displacement convexity, originally due to R. McCann, which inspired a lot of development in Optimal Transportation theory.

Definition 3.3 *We will say $H : dom(H) \subset \mathscr{P}_2 \to \mathbb{R}$ is displacement convex if*

$$H(\rho_s) \le (1-s)H(\rho_0) + sH(\rho_1),$$

for all ρ_s displacement interpolation of ρ_0 and $\rho_1 \in dom(H)$.

Lemma 3.4 *Suppose h:(0,∞)* $\to \mathbb{R}\cup\{\infty\}$ *is convex and non increasing, and g:*$[0,1]$ *→(0,∞) is concave. Then h∘g will be convex.*

Proof. Let $s, t_0, t_1 \in [0,1]$, then

$$\begin{aligned} h\circ g((1-s)t_0+st_1) &\leq h((1-s)g(t_0)+sg(t_1)) \\ &\leq (1-s)h\circ g(t_0)+sh\circ g(t_1). \end{aligned}$$

■

Definition 3.5 *Let* $F:[0,\infty)\to\mathbb{R}^n$ *differentiable, then we can define the associated* Internal Energy Functional *as*

$$H^F(\rho) := \int_\Omega F(\rho(x))dx\,.$$

Proposition 3.6 *Let* H^F *be the internal energy functional. If we suppose* $F:[0,\infty)\to\mathbb{R}^n$ *is differentiable with* $F(0)=0$, *and* $x\mapsto x^nF(\frac{r}{x^n})$ *is convex and non increasing for all* $r>0$, *then* H^F *is displacement convex.*

Proof. Let ρ_0 and $\rho_1 \in \mathscr{P}_{2,ac}$, , $\nabla\psi$ be the optimal mapping, and the displacement interpolation of ρ_0 and ρ_1.

Since $Supp(\rho_s) = Supp(\rho_0)((1-s)I+s\nabla\psi)$ we have that

$$\begin{aligned} H^F(\rho_s) &= \int_\Omega F(\rho_s(x))dx = \int_{\Omega\cap Supp(\rho_s)} F(\rho_s(x))dx \\ &= \int_{\Omega\cap Supp(\rho_0)} F(\rho_s(1-s)x+s\nabla\psi(x))\det((1-s)I+s\nabla\psi)dx. \end{aligned}$$

Using the Monge-Ampere formula, and defining $\lambda(s)=\det((1-s)I+s\nabla\psi)^{\frac{1}{n}}$ we can conclude that

$$\begin{aligned} H^F(\rho_s) &= \int_{\Omega\cap Supp(\rho_0)} F(\frac{\rho_0(x)}{\det((1-s)I+s\nabla\psi)})\det((1-s)I+s\nabla\psi)dx \\ &= \int_{\Omega\cap Supp(\rho_0)} F(\frac{\rho_0(x)}{\lambda^n})\lambda^n dx \end{aligned}$$

Using the fact that $\lambda(s)$ is concave, the remark and the lemma, we see that $s \mapsto F(\frac{\rho_0(x)}{\lambda^n})\lambda^n$ is convex, this means

$$H^F(\rho_s) \leq (1-s)H^F(\rho_0) + sH^F(\rho_1).$$

■

3.3 Benamou-Brenier formula

In physics if we have a density μ_t and a vector field V,and we assume the mass is conserved, then the density must satisfy the continuity equation. Inspired on this we have the following definition.

Definition 3.7 *We will call (μ_t, V_t), an admissible pair if*
$\cdot t \to \mu_t$ *is weak* continuous*
$\cdot t \to \int |x| \, d\mu_t$ *is continuous*
$\cdot \int \|V(t,x)\|^2 d\mu_t dt < \infty$.
$\cdot \partial_t \mu + \nabla \cdot (\mu V) = 0$ *in a weak sense.*

Theorem 3.8 *Let X be a complete smooth manifold, let μ_0 be a probability measure on X. If v is an integrable field, that is, there exists a locally Lipschitz family of diffeomorphisms $(T_t)_{0 \leq t \leq \overline{T}}$, such that*

$$\frac{dT_t}{dt}(x) = V_t(T(x)),$$

then (μ_t, V_t) is an admissible pairing, where $\mu_t = T_{t\#}\mu$ is the unique solution to the continuity equation .

Proof. Let φ be a test function and $t \in (0, \bar{t})$, by definition of push-forward we have

$$\int \varphi d\mu_t = \int (\varphi \circ T_t) d\mu$$

so for $h > 0$ we can write

$$\frac{1}{h}(\int \varphi d\mu_{t+h} - \int \varphi d\mu_t) = \int \frac{\varphi \circ T_{t+x}(x) - \varphi \circ T_t}{h} d\mu.$$

Since T_t^{-1} is continuous, then $\varphi \circ T_t$ is Lipschitz and compactly supported uniformly for $t \in [0,\bar{t}]$, so the right hand of the equation is uniformly bounded for $t \in [0,\bar{t}-h]$ and for almost all t,x converges point-wise to

$$\frac{\partial}{\partial t}(\varphi \circ T_t) = (\nabla\varphi \circ T_t) \cdot \frac{\partial}{\partial t} T_t = (\nabla\varphi \circ T_t) \cdot (v_t \circ T_t).$$

By Lebesgue´s dominated convergence theorem we deduce that for almost all t we have

$$\frac{d}{dt}\int \varphi d\mu_t = \int (\nabla\varphi \circ T_t) \cdot (v_t \circ T_t) = \int \nabla\varphi \cdot v_t d\mu_t.$$

To prove uniqueness we will prove that if μ_t satisfies the continuity equation then for any $T \in [0,\bar{t}]$, if $\mu_0 = 0$ then $\mu_T = 0$. We first assume we can find a Lipschitz compactly supported function $\varphi(t,x)$ that satisfies

$$\frac{\partial\varphi}{\partial t} + v_t \cdot \nabla\varphi = 0$$

$$\varphi\mid_{t=T} = \varphi_T.$$

Where $\varphi_T \in \mathscr{D}(X)$, the space of distribution, so we can compute for almost all t

$$\begin{aligned}\frac{d}{dt}\int \varphi_t d\mu_t &= \int \frac{\partial\varphi_t}{\partial t} d\mu_t + \int \varphi_t d(\frac{\partial\mu_t}{\partial t}) \\ &= -\int v_t \cdot \nabla\varphi + \int \varphi_t d(\nabla \cdot v_t \mu_t) = 0.\end{aligned}$$

Since $\mu_0 = 0$, then

$$\int \varphi_T d\mu_T = 0 \implies \mu_T = 0.$$

Finally we can check that $\varphi_t = \varphi_T \circ T_T \circ T_t^{-1}$ Lipschitz with compact support, and is a solution of

$$\frac{d}{dt}\varphi_t(T_t x) = \frac{\partial\varphi}{\partial t} + v \cdot \nabla\varphi = 0.$$

■

We will need the following lemma to prove the Benamou-Brenier theorem.

Lemma 3.9 *Let σ be a measure in $\mathbb{R}^n$, $f \in L^2(\sigma)$, and T a map such that $T_\#(f\sigma) = hT_\#(\sigma)$. Then*

$$\|h\|_{L^2(T_\#\sigma)} \leq \|f\|_{L^2(\sigma)}$$

Proof. Let $g \in L^2(T_\#\sigma)$, computing

$$\langle T_\#(f\sigma), g\rangle = \langle g \circ T, f\sigma\rangle \leq \|f\|_{L^2(\sigma)} \|g \circ T\|_{L^2(\sigma)} = \|f\|_{L^2(\sigma)} \|g\|_{L^2(T_\#\sigma)} .$$

Using Riesz representation theorem we know the continuous linear functional F such that

$$F(g) = \langle T_\#(f\sigma), g\rangle = \langle hT_\#(\sigma), g\rangle ,$$

has norm $\|h\|_{L^2(T_\#\sigma)}$. This means

$$\|h\|_{L^2(T_\#\sigma)} \leq \|f\|_{L^2(\sigma)} .$$

■

The following result has an interesting physical interpretation, as the Wasserstein distance between two measures can be seen as the infimum of the energy needed to translate one density to the other.

Theorem 3.10 (Benamou-Brenier) *If μ_0, $\mu_1 \in \mathscr{P}_{2,AC}$, then we have the equality*

$$d_w^2(\mu_0, \mu_1) = \inf_{V_t, \mu_t \text{ admissible}} \int_0^1 \int |V_t|^2 \, d\mu_t dt.$$

Proof. Since we are assuming absolute continuity for μ_0 and μ_1, we know there is a convex funtion ψ such that $\nabla\psi_\#\mu_0 = \mu_1$ a.e. Let μ_t be the displacement interpolation function between μ_0 and μ_1.

For $0 \leq t \leq 1$, let

$$\mu_t = (T_t)_\#\mu_0 \text{ where}$$
$$T_t = (1-t)Id + t\nabla\psi.$$

So we define

$$V_t(x) := \frac{d}{dt} T_t(x) = \nabla\psi(x) - x$$

We claim that

$$V_t (T_t)_\# \mu_0 = V_t \mu_t = (T_t)_\# ((\nabla \psi - Id)\mu_0).$$

So using lemma 3.9 we have that

$$\|V_t\|_{L^2(T_{t\#}\mu)} \leq \|\nabla \psi - Id\|_{L^2(\mu)},$$

this means

$$\int |V_t|^2 \, d\mu_t \leq \int |x - \psi(x)|^2 \, d\mu_0 = d_w^2(\mu_0, \mu_1).$$

For the other inequality we take an admissible pairing, and first we suppose V_t is sufficiently regular so there exists a flow map T such that

$$\frac{dT_t}{dt}(x) = V_t(T(x)).$$
$$T_0(x) = x$$

We know the unique solution of the continuity equation is a displacement interpolation so

$$\mu_t = (T_t)_\# \mu_0.$$

We can compute

$$\int_0^1 \int |V_t|^2 \, d\mu_t dt = \int_0^1 \int |V_t(T_t(x))|^2 \, d\mu_0 dt \geq \int |T_1(x) - x|^2 \, d\mu_0 \geq d_w(\mu_0, \mu_1).$$

■

Chapter 4

Relaxation

4.1 Young measures

Young measures are an important tool in the Calculus of Variations and Optimal Control Theory. It gives a description of limits of minimizing sequences; most of the basic results can be found in L.C. Young's book [14]. In this chapter we will explain some work of Bernard [1], showing how he used the concept of Young measures to generalize Brenier's theory and prove some interesting results.

So far we have been working with measures that depend on time, instead of this we will define Young measures in $(I \times X)$, where $I = [a,b]$ with λ the normalized Lebesgue measure, and (X,d) is a complete and separable metric space.

Definition 4.1 *A Young Measure in $(I \times X)$, is a positive measure η on $(I \times X)$, such that for any measurable set $A \subset I, \eta(A \times X) = \lambda(A)$. We denote the set of Young measures as $\mathscr{Y}_1(I,X) \subset \mathscr{P}_1(I,X)$, and we endow the metric d_1 (see definition 2.8).*

Note that $\mathscr{Y}_1(I,X)$ is closed in $\mathscr{P}_1(I,X)$.

There is another way to express a *Young Measure* by using the disintegration theorem [9], as there exist is a family of measures $\{\eta_t\}_{t \in I}$ in X, such that

$$\int_{I \times X} f(t,x) d\eta = \int_I \int_X f(t,x) d\eta_t d\lambda. \tag{4.1}$$

Now we would like to study some properties of the map

$$\eta \mapsto \int_{I\times X} f(t,x)d\eta. \tag{4.2}$$

This map is continuous if $|f(t,x)|/(1+d(x_0,x))$ is bounded for some x_0 and f is continuous, but we can generalize this result. For this we need to define *Caratheodory* integrands , and remind the reader of some results.

Definition 4.2 *A Caratheodory integrand is a Borel-measurable function $f(t,x)$: $I\times X \mapsto \mathbb{R}$, which is continuous in the second variable. A normal integrand is a Borel function $f(t,x) : I\times X \mapsto (-\infty,\infty]$, which is lower semi-continuous in the second variable.*

Definition 4.3 *We say $Y \subset \mathscr{P}(X)$ has uniformly integrable first moment if for every $\varepsilon > 0$ there exists a ball $B \subset X$ such that*

$$\int_{X-B} d(x_0,x)d\mu \leq \varepsilon \qquad \forall \mu \in Y,$$

for one and hence for all $x_0 \in X$.We will use the following result in the proposition.

Definition 4.4 *A set $Y \subset \mathscr{P}(X)$ is called tight if for every $\varepsilon > 0$ $\exists$ K_ε compact such that*

$$\mu(X-K_\varepsilon) \leq \varepsilon \qquad \forall \mu \in Y.$$

Theorem 4.5 *The function $g(t,x) : I\times X \mapsto \mathbb{R}$ is a normal integrand if and only if $g = \sup_{n\in N} g_n(t,x)$, where g_n is a sequence of Caratheodory integrands.*

Proof. See Berliocchi, Lasry [7]. ■

Theorem 4.6 (Prokhorov) *Let $K \subset \mathscr{P}(X)$, K is tight if and only if it is relatively compact.*

Proof. See Ambrosio, Gigli, Savare [9]. ■

Theorem 4.7 *The following properties are equivalent*

· *The family Y is tight with uniformly integrable first moment.*

·There exists a function $f;X \mapsto [0,\infty]$ whose sub-levels are compact, a constant C and a point x_0 such that

$$\int_X (1+d(x_{0,}x)f(x)d\mu \leq C \qquad \forall \mu \in Y.$$

Proof. See Ambrosio, Gigli, Savaré [9] ■

Proposition 4.8 *The map* (4.2) *is continuous on $\mathscr{Y}_1(I\times X)$ if f is a* Caratheodory integrand *such that $|f(t,x)|/(1+d(x_0,x))$ is bounded for some $x_0 \in X$. It is lower semi-continuous if f is a normal integrand such that $|f(t,x)|/(1+d(x_0,x))$ is bounded from below for some $x_0 \in X$.*

Proof. Using the Scorza-Dragoni Theorem [12], we know there exists a sequence of compact sets $J_n \subset I$, such that f is continuous on $J_n \times X$, and $\lambda(J_n) \mapsto 1$ as $n \mapsto \infty$. For every set J_n we can extend the function f continuously to a function f_n with a bounded norm, so $|f_n(t,x)|/(1+d(x_0,x))$ is bounded for every n. This means $\eta \mapsto \int_{I\times X} f_n(t,x)d\eta$ is continuous, and converges uniformly to (4.2) , and therefore is continuous.

For the second part we define $g = f(t,x)/(1+d(x_0,x))$, then g is a normal integrand which is bounded from below. Using Theorem 4.5, we see $g = \sup_{n\in\mathbb{N}} g_n(t,x)$, where g_n have to be bounded Caratheodory integrands. So now we can see the map (4.2) as the increasing limit of the continuous maps

$$\eta \mapsto \int (1+d(x_0,x))g_n(t,x)d\eta.$$

Hence it is lower semi-continuous. ■

From this proposition we can conclude our first important result.

Theorem 4.9 *Let $f(t,x)$ be a normal integrand such that $f(t,x) \geq l(x)(1+d(x_0,x))+g(t)$, where $g: I \mapsto \mathbb{R}$, is an integrable function, and $l: X \mapsto [0,\infty)$ is a proper function. Then for each $C \in \mathbb{R}$, the set*

$$\left\{\eta \in \mathscr{Y}_1(I,X) \mid \int f d\eta \leq C\right\}$$

is compact.

Proof. The set

$$\left\{\eta \in \mathscr{Y}_1(I,X) \mid \int l(x)(1+d(x_0,x))d\eta \le C\right\} \supset \left\{\eta \in \mathscr{Y}_1(I,X) \mid \int f d\eta \le C\right\}$$

is closed and 1-tight by the equivalence results, hence it is compact. Since the map (4.2) is lower semi continuous the set $\{\eta \in \mathscr{Y}_1(I,X) \mid \int f d\eta \le C\}$ is closed. ■

4.2 Transport measures

In this section we will consider Young measures acting on $I \times TM$ where M is a complete Riemannian manifold without boundary, and d is a distance on TM such that the quotient

$$\frac{1+d((x_0,0),(x,v))}{1+\|v\|_x},$$

and its inverse are bounded for any point $x_0 \in X$.

If $\eta \in \mathscr{Y}_1(I,TM)$ is a Young measure, the image of η of the projection $I \times TM \mapsto I \times M$ will be denoted as μ. We can think of μ as a density in M. Using the disintegration theorem [9] with respect to this projection, we obtain the measurable family $\eta_{t,x}$ of probability measures on T_xM such that $\eta = \mu \otimes \eta_{t,x}$. We define the vector field $V(t,x) : I \times M \mapsto TM$ by the expression

$$V(t,x) = \int_{T_xM} v d\eta_{t,x}(v).$$

We note that $V(t,x)$ is a Borel vector field, that satisfies the integrability condition

$$\int \|V(t,x)\|_x d\mu(t,x) < \infty.$$

We would like to know wheter μ satisfies the continuity equation,

$$\partial_t \mu + \operatorname{div}(V\mu) = 0, \tag{4.3}$$

in the sense of distributions. We have the following characterization result.

Lemma 4.10 *The measure μ satisfies equation* (4.3) *if and only if*

$$\int_{I\times TM} [\partial_t g + \partial_x g \cdot v]\, d\eta(t,x,v) = 0 \tag{4.4}$$

for all smooth compactly supported test functions $g \in C_c^\infty((a,b)\times M)$.

Proof. If we disintegrate η, we have that

$$\int_{I\times TM} [\partial_t g + \partial_x g \cdot v]\, d\eta(t,x,v) = \int_{I\times TM} [\partial_t g + \partial_x g \cdot vd]\, \eta_{t,x}(v) d\mu(t,x),$$

for each test function. Considering the definition of V, we have the equality

$$\int_{TM} \partial_x g \cdot v d\eta_{t,x}(v) = \partial_x g \cdot V(t,x).$$

This means η satisfies equation (4.4) if and only if

$$\int_{I\times M} \partial_t g + \partial_x g \cdot V(t,x) d\mu(t,x) = 0.$$

Which is equivalent to say μ satisfies equation (4.3). ■

Any $\eta \in \mathscr{Y}_1(I,TM)$, that satisfies equation (4.4) will be called *transport measure,* and we will denote the space of transport measures as $\mathscr{T}(I,M)$.

For the boundary conditions we do the following, given two probability measures μ_i and μ_f on M, we say η is a transport measure between μ_i and μ_f, if in addition we have that

$$\int_{I\times TM} [\partial_t g + \partial_x g \cdot v]\, d\eta(t,x,v) = \int_M g_b(x) d\mu_f - \int_M g_a(x) d\mu_i,$$

for all $g : [a,b]\times M \mapsto \mathbb{R}$, smooth compactly supported function. We denote by $\mathscr{T}_{\mu_i}^{\mu_f}(I,M)$, the set of transport measures between μ_i and μ_f.

4.3 Generalized curves

A particular case of transport measures are generalized curves as studied by L.C. Young. The way he defined boundary points is equivalent as the way defined above for this particular case.

Definition 4.11 *A transport measure η is called a generalized curve if for each $t \in I$ we have that $\mu_t = \delta_{\gamma(t)}$, for a continuous curve $\gamma(t) : I \mapsto M$. We say η is a generalized curve over γ, and we denote them as $\mathscr{G}(I,M)$.*

The following result shows us some regularity we can obtain from our new continuity equation.

Lemma 4.12 *Let $\Gamma \in \mathscr{T}(I,M)$ be a generalized curve over γ, then γ is absolutely continuous.*

Proof. By the disintegration theorem, the measure Γ can be written in the from $d\Gamma = dt \otimes \delta_{\gamma(t)} \otimes d\Gamma_t$, with some measurable family $\{\Gamma_t\}$ of probability measure in $T_{\gamma(t)}M$. In other words

$$\int_{I\times TM} f(t,x,v)d\Gamma(t,x,v) = \int_0^1 \int_{T_{\gamma(t)}M} f(t,\gamma(t),v)d\Gamma_t(v)dt \qquad \forall f \in L^1(\Gamma).$$

Now, for each $f \in C_c^\infty(a,b)$ and $\varphi \in C_c^\infty(M)$, let´s apply the equation (4.4) to the function $g(t,x) = f(t)\varphi(x)$, to get

$$\begin{aligned} 0 &= \int_{I\times TM} \left[f'(t)\varphi(x) + f(t)d\varphi_x \cdot v\right] d\eta(t,x,v) \\ &= \int_0^1 f'(t)\varphi(\gamma(t))dt + \int_0^1 f(t)\int_{T_{\gamma(t)}M} d\varphi_{\gamma(t)} \cdot v d\Gamma_t(v)dt. \end{aligned}$$

This means, in the sense of distributions, that

$$(\varphi\circ\gamma)'(t) = \int_{T_{\gamma(t)}M} d\varphi_{\gamma(t)} \cdot v d\Gamma_t(v) \qquad \forall \varphi \in C_c^\infty(M).$$

Hence γ is absolutely continuous and

$$\int_{T_{\gamma(t)}M} v d\Gamma_t(v) = \dot{\gamma}(t). \tag{4.5}$$

■

Theorem 4.13 *The set $\mathscr{G}(I,M)$ is closed in $\mathscr{Y}_1(I,TM)$, and the map*

$$\Gamma \mapsto \gamma \tag{4.6}$$

is continuous.

Proof. Let Γ_n be a sequence of generalized curves converging to η in $\mathscr{P}_1(I,TM)$. The set $\{\Gamma_n\} \cup \eta$ is compact hence, it has uniformly integrable first moment, so if the Γ_n´s are generalized curves over γ_n , then the sequence γ_n is absolutely equicontinuous. Hence there exists a subsequence γ_{n_m} and a curve γ_0 absolutely continuous, such that $\gamma_{n_m} \to \gamma_0$. ■

4.4 Tonelli theorem

In this section we will prove the existence of minimizers of normal integrands L, by finding conditions for which sets of the type $\{\Gamma \mid \int L d\Gamma \leq C\}$ are compact. We will consider the space $AC_{x_i}^{x_f}$ of absolutely continuous curves $\gamma : I \mapsto M$, such that $\gamma(a) = x_i$ and $\gamma(b) = x_f$, and the set

$$\mathscr{G}_{x_i}^{x_f} = \mathscr{T}_{\mu_i}^{\mu_f}(I,M) \cap \mathscr{G}(I,M),$$

of generalized curves above elements of $AC_{x_i}^{x_f}$. We will notice convexity of L is not needed for the result in $\mathscr{G}_{x_i}^{x_f}$, but it is for $AC_{x_i}^{x_f}$, which is one of the advantages of working with generalized curves.

For the following results, we suppose $L : [a,b] \times TM \mapsto \mathbb{R} \cup \{+\infty\}$ is a normal integrand.

Definition 4.14 *We say L is fiber-wise convex if, the function $v \mapsto L(t,x,v)$, is convex on T_xM, for every $t \in [a,b]$, and $x \in M$.*

Definition 4.15 *We say L is uniformly super-linear over a compact K, if there exists a function $l : \mathbb{R}^+ \mapsto \mathbb{R}$, such that $\lim_{r \mapsto \infty} l(r)/r = \infty$ and such that $L(t,x,v) \geq l(\|v\|_x)$ for every $(t,x,v) \in [a,b] \times T_kM$.*

Lemma 4.16 *Let L be a fiber-wise convex normal integrand. If Γ is a generalized curve above γ, then*

$$\int_0^1 L(t,\gamma(t),\dot{\gamma}(t))dt \leq \int L d\Gamma.$$

Proof. Using equation (4.5) and Jensen´s inequality we have

$$L(t,\gamma(t),\dot{\gamma}(t)) = L(t,\gamma(t),\int_{T_{\gamma(t)}M} v d\Gamma_t(v)) \leq \int_{T_{\gamma(t)}M} L(t,\gamma(t),v) d\Gamma_t(v).$$

Hence

$$\int_0^1 L(t,\gamma(t),\dot{\gamma}(t))dt \leq \int_0^1 \int_{T_{\gamma(t)}M} L(t,\gamma(t),v)d\Gamma_t(v)dt = \int L d\Gamma.$$

■

Theorem 4.17 *Let L be a normal integrand such that the quotient*

$$\frac{L(t,x,v)}{1+\|v\|_x} \tag{4.7}$$

is bounded from below.

Conclusion: for each $C \in \mathbb{R}$, the set

$$\mathscr{A}_C^g := \left\{ \Gamma \in \mathscr{G}_{x_i}^{x_f} \mid \int L d\Gamma \leq C \right\}$$

is compact in $\mathscr{G}_{x_i}^{x_f}$. If L is fiber-wise convex, the set

$$\mathscr{A}_C := \left\{ \gamma \in AC_{x_i}^{x_f} \mid \int_a^b L(t,\gamma(t),\dot{\gamma}(t))dt \leq C \right\}$$

is compact in $AC_{x_i}^{x_f}$ for the uniform topology.

Proof. The compactness of $\mathscr{A}_C^g$ follows from theorem 4.9. If L is fiber-wise convex, using lemma 4.16 we know the image of $\mathscr{A}_C^g$ with the continuous map (4.6) is $\mathscr{A}_C$, hence it is compact. ■

A more general result is due originally to Tonelli.

Theorem 4.18 (Tonelli) *Let L be a normal integrand such that*

·L is uniformly super-linear over each compact subset of M.

·There exists a positive constant such that $L(t,x,v) \geq c\|v\|_x - 1$.

Then we have the same conclusion as in the last theorem.

Proof. If Γ is a generalized curve over γ such that $\int Ld\Gamma \leq C$, using $\|v\|_x \leq (L(t,x,v)+1)/c$

$$\int_a^b \left\|\dot{\gamma}(t)\right\|_{\gamma(t)} dt \leq \frac{C+b-a}{c}.$$

This means the curve γ lies in the ball $B(\frac{C+b-a}{c}, x_i)$, which is compact since M has finite dimension and d is complete. So if we define the convex integrand

$$L_B(t,x,v) = \begin{cases} L(t,x,v) & \text{if } x \in B(\frac{C+b-a}{c}, x_i) \\ \infty & \text{if } x \notin B(\frac{C+b-a}{c}, x_i) \end{cases},$$

we have that Γ satisfies $\int Ld\Gamma \leq C$ if and only if $\int L_B d\Gamma \leq C$. Using the fact L is uniformly super-linear on $B(\frac{C+b-a}{c}, x_i)$, we see that the quotient

$$\frac{L_B(t,x,v)}{1+\|v\|_x} \tag{4.8}$$

is bounded below. So we can use the previous theorem. ■

Chapter 5

Inequalities

Mass transport has already shown it is a powerful tool to prove known inequalities in sometimes remarkably simpler ways, for example one of the most simple inequalities one can prove using mass transportation techniques is the isoperimetric inequality. Using only the arithmetic-geometric inequality in the following sense

$$n(\det D^2\varphi)^{\frac{1}{n}} \leq tr(D^2\varphi) = \Delta\varphi,$$

we give a sketch, ignoring subtle analytic issues, of the original proof due to M. Gromov (see [13])

Theorem 5.1 *Let Ω be an open set, such that $|\Omega| = 1$, then we have that $|\partial\Omega| \geq |\partial B| = n$, where B is the ball with area one.*

Sketch. We take the unitary functions in Ω and B , 1_Ω and 1_B. Both are probability functions so we can take the optimal transport $\nabla\varphi$, from Ω onto B. Hence this function satisfies the Monge-Ampere equation

$$\det D^2\varphi = 1.$$

Since $|\nabla\varphi| \leq 1$, using Gauss theorem we can compute

$$|\partial\Omega| = \int_{\partial\Omega} 1 ds \geq \int_{\partial\Omega} \nabla\varphi \cdot \overrightarrow{n} ds = \int_{\Omega} \Delta\varphi dx \geq \int_{\Omega} n(\det D^2\varphi)^{\frac{1}{n}} = n|\Omega| = n.$$

■

It has been known that there is a relationship between the isoperimetric inequality and the Sobolev inequality. In fact the Sobolev inequality, can be proven using optimal transport in a similar spirit. There are several other applications like Brunn-Monkowski, HWI, Log-sobolev, and Gagliardo-Nirenberg. See [3],[2],[11]. Recently Agueh-Ghoussoub-Kang [10] showed that many of this inequalities actually belong to the same family of inequalities, in other words they are particular cases of the same general inequality. It is the purpose of this chapter to explain this result emphasizing on displacement convexity by proving first a general Sobolev inequality that can be used to obtain Log-Sobolev, Sobolev and Gagliardo-Nirenberg inequalities. Afterwards we will prove the Agueh-Ghoussoub-Kang's general inequality , and show that it generalizes the general Sobolev inequality as well as other general inequalities like the HWI and Gaussian inequalities .

5.1 General Sobolev inequality

In this section we will use the energy functional H^F (see definition 3.5)

In this chapter, T represents the optimal map from ρ_0 to ρ_1, and $\rho_t := ((1-t)I+T)_{\#}\rho_0$.

Lemma 5.2 *Suppose $F:[0,\infty)\to\mathbb{R}^n$ is differentiable with $F(0)=0$, and $x\mapsto x^n F(\frac{r}{x^n})$ is convex and non increasing for all $r>0$, then we have that*

$$H^F(\rho_1)-H^F(\rho_0)\geq\int_\Omega \rho_o(T-1)\cdot\nabla F'(\rho_0)dx,$$

for all $\rho_0,\rho_1\in\mathscr{P}_{2,AC}$.

Proof. *Since $H^F(\rho_t)$ is convex then we obtain*

$$\begin{aligned}\frac{H^F(\rho_1)-H^F(\rho_0)}{1}&\geq\left[\frac{d}{dt}H^F(\rho_t)\right]_{t=0}\\&=\left[\frac{d}{dt}\int_\Omega F(((1-t)I+tT)_{\#}\rho_0)dx\right]_{t=0}\\&=-\int_\Omega F'(\rho_0)div(\rho_0(T-I))dx=\int_\Omega\rho_o(T-1)\cdot\nabla F'(\rho_0)dx.\end{aligned}$$

■

Definition 5.3 *We will call a* Young function, *any strictly convex super-linear C^1-function $c : \mathbb{R}^n \to \mathbb{R}$, such that $c(0) = 0$, and we will denote by c^* its Legendre dual, as defined in remark 2.5.*

Theorem 5.4 (General Sobolev inequality) *Under the hypotheses of the previous lemma, let Ω be any open bounded convex set, then for any $\rho \in \mathscr{P}_{2,AC}$, satisfying supp$\rho \subset \Omega$ and $P_F(x) := xF'(x) - F(x) \in W^{1,\infty}(\Omega)$, we have that*

$$H^{F+nP_F}(\rho) \leq \int_\Omega c^*(-\nabla F'(\rho))\rho dx + K_c. \tag{5.1}$$

Proof. Using the previous lemma for $\rho_0 = \rho$, and $\rho_1 = \rho_c$, where $\rho_c \in \mathscr{P}_{2,AC}$ is a solution of

$$\nabla(F'(\rho_c) + c) = 0,$$

we get that

$$H^F(\rho_c) - H^F(\rho) \geq \int_\Omega \rho(Tx - x) \cdot \nabla F'(\rho).$$

We note that since $\rho \nabla(F'(\rho)) = \nabla(P_F(\rho))$ we have that

$$\int_\Omega \rho \nabla(F'(\rho)) \cdot x = \int_\Omega -nP_F(\rho) = H^{-nP_F}(\rho).$$

We obtain

$$\begin{aligned} H^F(\rho) - H^F(\rho_c) &\leq \int \rho(x - Tx) \cdot \nabla(F'(\rho)) \\ &\leq H^{-nP_F}(\rho) - \int_\Omega \rho \nabla(F'(\rho)) \cdot Tx dx. \end{aligned}$$

For the last term we can use the generalized Young's inequality to obtain that

$$-\nabla(F'(\rho)) \cdot Tx \leq c(Tx) + c^*(-\rho\nabla(F'(\rho)).$$

Integrating this to the inequality we have

$$\begin{aligned}
&H^F(\rho) - H^F(\rho_c) \\
&\leq H^{-nP_F}(\rho) + \int_\Omega c(Tx)\rho dx + \int_\Omega c^*(-\nabla F'(\rho))\rho dx \\
&= H^{-nP_F}(\rho) + \int_\Omega c(x)\rho_c dx + \int_\Omega c^*(-\nabla F'(\rho))\rho dx.
\end{aligned}$$

Finally we get

$$H^{F+nP_F}(\rho) \leq \int_\Omega c^*(-\nabla F'(\rho))\rho dx + \int_\Omega (c(x) + F'(\rho_c))\rho_c dx - H^{P_F}(\rho_c)$$

We name the constant $c(x) + F'(\rho_c) = K_c$, and we note that $H^{P_F}(\rho_c) \geq 0$ to conclude the proof. ■

In the following pages we will see that using different F's this inequality generalizes Log-Sobolev inequalities and Sobolev-Nirenberg-Gagliardo inequalities.

5.1.1 Euclidian Log Sobolev inequalities

The Log-Sobolev inequality was first introduced by L. Gross, see [6], here we prove it as a corollary of the previous inequality.

Corollary 5.5 *Let $\Omega \subset \mathbb{R}^n$ be an open bounded and convex set, and let c be a Young functional , such that c^* is p-homogeneous, for $p > 1$, we have that for all probability densities ρ, with supp(ρ) $\subset \Omega$, and $\rho \in W^{1,\infty}(\mathbb{R}^n)$*

$$\int_{\mathbb{R}^n} \rho \log \rho dx \leq \frac{n}{p} \log\left(\frac{p}{ne^{p-1}\sigma_c^{p/n}} \int_{\mathbb{R}^n} \rho c^*(-\frac{\nabla\rho}{\rho}) dx\right),$$

where $\sigma_c = \int_{\mathbb{R}^n} e^{-c(x)} dx$.

Proof. Let $F(x) = x\log(x)$, and $F(0) = 0$. We check that $x \mapsto x^n F(x^{-n}) = -n\log(x)$ is convex and non increasing. Considering that in this case $P_F(x) = x$, we get that for any probability measure ρ , $H^{P_F} = \int \rho = 1$. We take inequality (5.1),

$$H^{F+nP_F}(\rho) \leq \int_\Omega c^*(-\nabla F'(\rho))\rho dx + \int (F'(\rho_c) + c)\rho_c dx, \tag{5.2}$$

where ρ_c is a solution of the equation

$$\nabla(\log\rho_c + c) = 0,$$

which we take $\rho_c(x) = e^{-c(x)}/\sigma_c$, so we get

$$\int \rho\log\rho + n \leq \int c^*(-\frac{\nabla\rho}{\rho})\rho dx + \int \left(\log e^{-c(x)} - \log(\int_{\mathbb{R}^n} e^{-c(x)}dx) + c\right)\rho_c dx \tag{5.3}$$

$$= \int c^*(-\frac{\nabla\rho}{\rho})\rho dx - \log(\int_{\mathbb{R}^n} e^{-c(x)}dx). \tag{5.4}$$

Let $c_\lambda(x) := c(\lambda x)$, hence $c^*_\lambda(y) = c^*(\frac{y}{\lambda})$. If we apply the inequality to this Young function we get

$$\begin{aligned}\int \rho\log\rho + n &\leq \int c^*(-\frac{\nabla\rho}{\lambda\rho})\rho dx - \log(\int_{\mathbb{R}^n} e^{-c(\lambda x)}dx)\\ &= \int c^*(-\frac{\nabla\rho}{\lambda\rho})\rho dx - \log(\int_{\mathbb{R}^n} e^{-c(x)}dx) + n\log\lambda.\end{aligned}$$

Considering that $c^*(\frac{y}{\lambda}) = \frac{1}{\lambda^p}c^*(y)$, the infimum over λ is attained when

$$\lambda_o^p = \frac{p}{n}\int c^*(-\frac{\nabla\rho}{\rho})\rho dx.$$

So we get the inequality for all probability densities ρ, with $\mathrm{supp}(\rho) \subset \Omega$, and $\rho \in W^{1,\infty}(\mathbb{R}^n)$

$$\begin{aligned}\int \rho\log\rho &\leq \frac{n}{p\int c^*(-\frac{\nabla\rho}{\lambda\rho})\rho dx}\int c^*(-\frac{\nabla\rho}{\rho})\rho dx - \log(\sigma_c) + \frac{n}{p}\log\left(\frac{p}{n}\int c^*(-\frac{\nabla\rho}{\lambda\rho})\rho dx\right) - n\\ &\leq \frac{n}{p}\log\left(\frac{p}{n}\int c^*(-\frac{\nabla\rho}{\lambda\rho})\rho dx\right) - \log(\sigma_c) - n + \frac{n}{p}\\ &= \frac{n}{p}\log(\frac{p}{ne^{p-1}\sigma_c^{p/n}}\int_{\mathbb{R}^n}\rho c^*(-\frac{\nabla\rho}{\rho})dx).\end{aligned}$$

■

5.1.2 Sobolev and Gagliardo-Nirenberg inequalities

We will now derive the Gagliardo-Nirenberg inequality from the general Sobolev inequality. A classical proof can be found in [4], and a proof using mass-transport approach can be found in [2].

Corollary 5.6 (Gagliardo-Nirenberg) *Let $1 < p < n$ and $r \in (0, \frac{np}{n-p})$ such that $r \neq p$. We define $\gamma := \frac{1}{r} + \frac{1}{q}$, where $\frac{1}{p} + \frac{1}{q} = 1$. For any $f \in W^{1,p}(\mathbb{R}^n)$ we have that there exists θ such that*

$$\|f\|_r \leq C(p,r)\, \|\nabla f\|_p^{\theta}\, \|f\|_{r\gamma}^{1-\theta}.$$

Proof. We will use inequality (5.1) with $F(x) = \frac{x^\gamma}{\gamma-1}$. Since $r \neq p$ we have that $\gamma \neq 1$ and since $r \in (0, \frac{np}{n-p})$, we have that $1 > \gamma > 1 - \frac{1}{n}$. To use the inequality we check that $F(0) = 0$, and $x \mapsto x^n F(x^{-n}) = \frac{x^{n-n\gamma}}{\gamma-1}$ is convex and non increasing since $n - n\gamma < 1$ and $\gamma - 1 < 0$. Let $c(x) = \frac{r\gamma}{q}|x|^q$, so $c^*(x) = \frac{1}{p(r\gamma)^{p-1}}|x|^p$. Using inequality (5.1) we get

$$\int F(\rho) + n\rho F'(\rho) - nF(\rho)dx \leq \int_\Omega \rho \frac{1}{p(r\gamma)^{p-1}}(-\nabla F'(\rho))^p \rho dx + K_c. \quad (5.5)$$

Making the substitution of $F(x) = \frac{x^\gamma}{\gamma-1}$ we get

$$\frac{1}{\gamma-1}\int \rho^\gamma - n\rho^\gamma + n\gamma\rho^\gamma dx \leq \int_\Omega \rho \frac{1}{p(r\gamma)^{p-1}}(-\gamma\rho^{\gamma-2}\nabla\rho)^p \rho dx + K_c,$$

and rearranging the equation we get

$$(\frac{1}{\gamma-1} + n)\int \rho^\gamma dx \leq \int_\Omega \rho \frac{r\gamma}{p(r)^p}(\nabla\rho)^p \rho dx + K_c.$$

If we suppose that $\|f\|_r = 1$, we take $\rho = |f|^r$ to get

$$(\frac{1}{\gamma-1} + n)\int |f|^{r\gamma} dx \leq \int_\Omega \frac{r\gamma}{p}|\nabla f|^p \rho dx + K_c,$$

and for general f we get this inequality

$$\frac{r\gamma}{p}\frac{\|\nabla f\|_p^p}{\|f\|_r^p} - \left(\frac{1}{\gamma-1}+n\right)\frac{\|f\|_{r\gamma}^{r\gamma}}{\|f\|_r} \geq -K_c.$$

If we have the function $f_\lambda(x) = f(\lambda x)$, with a change of variables we get the following equalities

$$\begin{aligned}
\|f_\lambda\|_r^p &= \lambda^{-np/r}\|f\|_r^p,\\
\|f_\lambda\|_{r\gamma}^{r\gamma} &= \lambda^{-n}\|f\|_{r\gamma}^{r\gamma},\\
\|f_\lambda\|_r &= \lambda^{-n/r}\|f\|_r,\\
\|\nabla f_\lambda\|_p^p &= \|\lambda\nabla f(x\lambda)\|_p^p = \lambda^{p-n}\|\nabla f\|_p^p.
\end{aligned}$$

So the inequality becomes

$$\lambda^{p-n+np/r}\frac{r\gamma}{p}\frac{\|\nabla f\|_p^p}{\|f\|_r^p} - \lambda^{-n+n/r}\left(\frac{1}{\gamma-1}+n\right)\frac{\|f\|_{r\gamma}^{r\gamma}}{\|f\|_r} \geq -K_c.$$

We take $\lambda = \|\nabla f\|_p^a \|f\|_r^b \|f\|_{r\gamma}^c$, and we pick a, b, and c, so that the powers of the norms are the same in both terms, that is

$$\begin{aligned}
a &= \frac{pr}{pr+np-n},\\
b &= \frac{(p-1)r}{pr+np-n},\\
c &= \frac{r}{pr+np-n}.
\end{aligned}$$

So we obtain

$$\frac{1}{K_c}\left(-\frac{r\gamma}{p}+\frac{1}{\gamma-1}-n\right)\|\nabla f\|_p^{a'}\|f\|_{r\gamma}^{c'} \geq \|f\|_r^{b'},$$

where

$$a' = \frac{-npr + np}{pr + np - n}$$
$$b' = \frac{(p-1)r(-n+n/r)}{pr+np-n} - 1$$
$$c' = \frac{rp - nr + np}{pr + np - n}$$

■

Finally we note that if we take the limit as $r \to p^* = \frac{np}{n-p}$, we have that $c' \to 0$, and $a', b' \to \frac{np(n-np-p)}{(n-p)(pr+np-n)}$ so we get the Sobolev inequality

$$\|f\|_{p^*} \leq C(p,n)\, \|\nabla f\|_p .$$

5.2 General inequality

In this section we will generalize the previous result by showing an inequality that contains even more information, like HWI inequalities (see [10]). For this, inspired by the physics of interacting gases, we will define more energy functionals in $\mathscr{P}_{2,ac}$, and we will use the concept of semi-convexity.

Definition 5.7 *Let $F : [0,\infty) \to \mathbb{R}^n$ differentiable, and $V, W : \mathbb{R} \to [0,\infty)$, twice differentiable, then we can define the associated* Free Energy Functional *as*

$$H_V^{F,W}(\rho) := H^F(\rho) + H_V(\rho) + H^W(\rho).$$

Where we have

·Internal energy

$$H^F(\rho) := \int_\Omega F(\rho(x))dx, \, .$$

·Potential energy

$$H_V(\rho) =: \int_\Omega \rho(x)V(x)dx.$$

·Interaction energy

$$H^W(\rho) =: \frac{1}{2}\int_\Omega \rho(W * \rho),$$

where $*$ denotes the convolution product.

Furthermore we define the *relative energy of ρ_0 with respect to ρ_1* as

$$H_V^{F,W}(\rho_0 \mid \rho_1) := H_V^{F,W}(\rho_1) - H_V^{F,W}(\rho_0),$$

and the *relative entropy production of ρ with respect to ρ_V* as

$$\mathscr{I}_2(\rho \mid \rho_V) := \int_\Omega \left|\nabla(F'(\rho) + V + W * \rho)\right|^2 \rho dx.$$

So if ρ_V is a probability density that satisfies

$$\nabla(F'(\rho_V) + V + W * \rho) = 0,$$

then

$$\mathscr{I}_2(\rho \mid \rho_V) := \int_\Omega \left|\nabla(F'(\rho) - F'(\rho_V) + W * (\rho - \rho_V))\right|^2 \rho dx.$$

We will also work with non-quadratic versions of entropy, so we define the *generalized relative entropy production-type function of ρ with respect to ρ_V measured against c^** as

$$\mathscr{I}_{c^*}(\rho_0 \mid \rho_V) := \int_\Omega c^* \left(-\nabla(F'(\rho_0) + V + W * \rho_0)\right) \rho_0 dx,$$

where c^* is the Legendre conjugate of c.

Lemma 5.8 *Assume $V{:}\mathbb{R}^n \to \mathbb{R}$ satisfies that $D^2V \geq \lambda I$, for some $\lambda \in \mathbb{R}$, then we have that*

$$H_V(\rho_1) - H_V(\rho_0) \geq \int_\Omega \rho_o (T - 1) \cdot \nabla V dx + \frac{\lambda}{2} d_w^2(\rho_0, \rho_1),$$

for all $\rho_0, \rho_1 \in \mathscr{P}_{2,AC}$.

Proof. Expanding and using $D^2V \geq \lambda I$, we obtain

$$V(b) - V(a) \geq \nabla V(a) \cdot (b - a) + \frac{\lambda}{2} |a - b|^2.$$

This means that

$$V(Tx) - V(x) \geq \nabla V(x) \cdot (Tx - x) + \frac{\lambda}{2} |x - Tx|^2 .$$

Hence integrating we obtain

$$\begin{aligned} H_V(\rho_1) - H_V(\rho_0) &\geq \int V(Tx)\rho_0 - V(x)\rho_0 dx \\ &\geq \int_\Omega \nabla V(x) \cdot (Tx - x) + \frac{\lambda}{2} |x - Tx|^2 \rho_0 dx \\ &= \int_\Omega \rho_o (T - 1) \cdot \nabla V dx + \frac{\lambda}{2} d_w^2(\rho_0, \rho_1). \end{aligned}$$

■

Lemma 5.9 *Assume W:$\mathbb{R}^n \to \mathbb{R}$ is even and satisfies that $D^2 W \geq \nu I$, for some $\nu \in \mathbb{R}$, then we have that*

$$H^W(\rho_1) - H^W(\rho_0) \geq \int_\Omega \rho_o (T-1) \cdot \nabla(W * \rho_0) dx + \frac{\nu}{2}(d_w^2(\rho_0, \rho_1) - |b(\rho_0) - b(\rho_1)|^2),$$

for all $\rho_0, \rho_1 \in \mathscr{P}_{2,AC}$, and where b represents the centre of mass denoted by $b(\rho) = \int x\rho(x)dx$.

Proof. First we note that we can write the interaction energy as follows

$$\begin{aligned} H^W(\rho_1) &= \frac{1}{2} \int_{\Omega\times\Omega} W(x-y)\rho_1(x)\rho_1(y) dxdy \\ &= \frac{1}{2} \int_{\Omega\times\Omega} W(Tx - Ty)\rho_0(x)\rho_0(y) dxdy \\ &= \frac{1}{2} \int_{\Omega\times\Omega} W(x - y + (T-I)(x) - (T-I)(y))\rho_0(x)\rho_0(y) dxdy \end{aligned}$$

Since $D^2W \geq \nu I$ we obtain

$$
\begin{aligned}
H^W(\rho_1) \geq & \frac{1}{2}\int_{\Omega\times\Omega} [W(x-y)+\nabla W(x-y)\cdot((T-I)(x)-(T-I)(y))]\,\rho_0(x)\rho_0(y)dxdy \\
& + \frac{\nu}{4}\int_{\Omega\times\Omega} |(T-I)(x)-(T-I)(y)|^2\,\rho_0(x)\rho_0(y)dxdy \\
& = H^W(\rho_0) + \frac{1}{2}\int_{\Omega\times\Omega} \nabla W(x-y)\cdot((T-I)(x)-(T-I)(y))\rho_0(x)\rho_0(y)dxdy \\
& + \frac{\nu}{4}\int_{\Omega\times\Omega} |(T-I)(x)-(T-I)(y)|^2\,\rho_0(x)\rho_0(y)dxdy.
\end{aligned}
$$

Now we note the following equalities, for the last term

$$
\begin{aligned}
& \int_{\Omega\times\Omega} |(T-I)(x)-(T-I)(y)|^2\,\rho_0(x)\rho_0(y)dxdy \\
& = 2\int_{\Omega\times\Omega} |(T-I)(x)|^2\,\rho_0(x)dx - 2\left|\int_{\Omega\times\Omega}(T-I)(x)\rho_0(x)dx\right|^2 \\
& = 2\left[\int_{\Omega\times\Omega} |(T-I)(x)|^2\,\rho_0(x)dx - |b(\rho_1)-b(\rho_0)|^2\right].
\end{aligned}
$$

For the second term we consider that ∇W is odd

$$
\begin{aligned}
& \int_{\Omega\times\Omega} \nabla W(x-y)\cdot((T-I)(x)-(T-I)(y))\rho_0(x)\rho_0(y)dxdy \\
& = 2\int_{\Omega\times\Omega} \nabla W(x-y)\cdot((T-I)(x))\rho_0(x)\rho_0(y)dydx \\
& = 2\int_{\Omega\times\Omega} (\nabla W*\rho_0)\cdot(T-I)(x))\rho_0(x)dx.
\end{aligned}
$$

Using these two equalities we get

$$
H^W(\rho_1)-H^W(\rho_0) \geq \int_\Omega \rho_o(T-1)\cdot\nabla(W*\rho_0)dx + \frac{\nu}{2}(d_w^2(\rho_0,\rho_1)-|b(\rho_0)-b(\rho_1)|^2).
$$

■

Theorem 5.10 (Basic inequality) *Under the hypotheses of the three previous lemmas, let Ω be any open bounded convex set, then for any $\rho_0,\rho_1 \in \mathscr{P}_{2,AC}$, satisfying* *supp*$\rho_0 \subset \Omega$ *and* $P_F(x) := xF'(x) - F(x) \in W^{1,\infty}(\Omega)$ *, we have that*

$$H_{V+c}^{F,W}(\rho_0 \mid \rho_1) + \frac{\lambda+\nu}{2} d_w^2(\rho_0,\rho_1) - \frac{\nu}{2}|b(\rho_0 - b(\rho_1)|^2$$
$$\leq H_{c+\nabla V \cdot x}^{-nP_F, 2x\cdot\nabla W}(\rho_0) + \mathscr{I}_{c^*}(\rho_0 \mid \rho_V).$$

Proof. First we note that since $\rho_0 \nabla(F'(\rho_0)) = \nabla(P_F(\rho_0))$ we have that

$$\begin{aligned}
&\int_\Omega \rho_0 \nabla(F'(\rho_0) + V + W * \rho_0) \cdot x \\
&= \int_\Omega -nP_F(\rho_0) + \rho_0 \left[\nabla(V + \nabla W * \rho_0)\right] \cdot x \\
&= \int_\Omega -nP_F(\rho_0) + \rho_0 \nabla V \cdot x + \frac{1}{2}\rho_0(2x \cdot \nabla W * \rho_0) dx \\
&= H_{\nabla V \cdot x}^{-nP_F, 2x\cdot\nabla W}(\rho_0).
\end{aligned}$$

If we add the inequalities from the previous lemmas we get

$$\begin{aligned}
&H_V^{F,W}(\rho_1) - H_V^{F,W}(\rho_0) \\
&\geq \int_\Omega \rho_0(Tx - x) \cdot \nabla(F'(\rho_0) + V + W * \rho_0) dx + \frac{\lambda}{2} d_w^2(\rho_0, \rho_1) \\
&+ \frac{\nu}{2}(d_w^2(\rho_0, \rho_1) - |b(\rho_0) - b(\rho_1)|^2).
\end{aligned}$$

Rearranging and using the first inequality we have

$$\begin{aligned}
&H_V^{F,W}(\rho_0) - H_V^{F,W}(\rho_1) + \frac{\lambda+\nu}{2} d_w^2(\rho_0, \rho_1) - |b(\rho_0) - b(\rho_1)|^2) \\
&\leq \int_\Omega \rho_0(x - Tx) \cdot \nabla(F'(\rho_0) + V + W * \rho_0) \\
&\leq H_{\nabla V \cdot x}^{-nP_F, 2x\cdot\nabla W}(\rho_0) - \int_\Omega \rho_0 \nabla(F'(\rho_0) + V + W * \rho_0) \cdot Tx dx.
\end{aligned}$$

For the last term we can use the generalized Young's inequality to obtain that

$$\begin{aligned}
&-\nabla(F'(\rho_0) + V + W * \rho_0) \cdot Tx \\
&\leq c(Tx) + c^*(-\nabla(F'(\rho_0) + V + W * \rho_0).
\end{aligned}$$

Integrating this to the inequality we have

$$
\begin{aligned}
&H_V^{F,W}(\rho_0) - H_V^{F,W}(\rho_1) + \frac{\lambda+\nu}{2} d_w^2(\rho_0,\rho_1) - |b(\rho_0)-b(\rho_1)|^2) \\
&\leq H_{\nabla V\cdot x}^{-nP_F,2x\cdot\nabla W}(\rho_0) + \int_\Omega c(Tx)\rho_0 dx + \int_\Omega c^*(-\nabla(F'(\rho_0)+V+W*\rho_0)\rho_0 dx \\
&= H_{\nabla V\cdot x}^{-nP_F,2x\cdot\nabla W}(\rho_0) + \int_\Omega c(x)\rho_1 dx + \int_\Omega c^*(-\nabla(F'(\rho_0)+V+W*\rho_0)\rho_0 dx.
\end{aligned}
$$

This proves the inequality. ■

A simpler inequality is the one obtained when V and W are strictly convex hence $\nu,\lambda \geq 0$.

Lemma 5.11 *Under the same hypothesis as theorem 5.10, assume that V and W are also convex. Then for any Young function $c:\mathbb{R}^n \to \mathbb{R}$, we have*

$$
H_{V-\nabla V\cdot x}^{F+nP_F,W-2x\cdot\nabla W}(\rho) \leq -H^{P_F,W}(\rho_{V+c}) + \mathscr{I}_{c^*}(\rho \mid \rho_V) + K_{V+c}. \tag{5.6}
$$

Furthermore if we set $W=V=0$, since $H^{P_F}(\rho_c)\geq 0$, we obtain

$$
\begin{aligned}
H^{F+nP_F}(\rho) &\leq -H^{P_F}(\rho_c) + \mathscr{I}_{c^*}(\rho \mid \rho_V) + K_{V+c} && (5.7)\\
&\leq \int_\Omega c^*(-\nabla F'(\rho))\rho dx + K_{V+c}. && (5.8)
\end{aligned}
$$

Hence recovering inequality (5.1).

Proof. Let´s consider the inequality we just proved

$$
\begin{aligned}
H_{V+c}^{F,W}(\rho_0 \mid \rho_1) + \frac{\lambda+\nu}{2} d_w^2(\rho_0,\rho_1) - \frac{\nu}{2}|b(\rho_0 - b(\rho_1)|^2 && (5.9)\\
\leq H_{c+\nabla V\cdot x}^{-nP_F,2x\cdot\nabla W}(\rho_0) + \mathscr{I}_{c^*}(\rho_0 \mid \rho_V). && (5.10)
\end{aligned}
$$

In particular if we take $\rho_0=\rho$ and $\rho_1=\rho_{V+c}$, where ρ_{V+c} is a solution of

$$
\nabla(F'(\rho_{V+c}) + V + c + W*\rho_{V+c}) = 0.
$$

Hence we have that for any $\rho \in \mathscr{P}_c(\Omega)$, with supp $\rho \subset \Omega$, and $P_F(\rho) \in W^{1,\infty}(\Omega)$ we have that

$$H_{V-\nabla V\cdot x}^{F+nP_F,W-2x\cdot\nabla W}(\rho) + \frac{\lambda+\nu}{2} d_w^2(\rho,\rho_{V+c}) - \frac{\nu}{2} |b(\rho - b(\rho_1)|^2 \tag{5.11}$$

$$\leq -H^{P_F,W}(\rho_{V+c}) + \mathscr{I}_{c^*}(\rho \mid \rho_V) + \int \left(F'(\rho_{V+c}) + V + c + W * \rho_{V+c}\right)\rho_{V+c}. \tag{5.12}$$

Where we can define the constant $F'(\rho_V + c) + V + c + W * \rho := K_{V+c}$.

Since $\nu, \lambda \geq 0$ we get that

$$\frac{\lambda+\nu}{2} d_w^2(\rho,\rho_{V+c}) - \frac{\nu}{2} |b(\rho) - b(\rho_{V+c})|^2$$
$$= \frac{\lambda+\nu}{2} \int |Tx - x|^2 \rho_0(x)dx - \frac{\nu}{2} \left| \int (Tx - x)\rho_0(x)dx \right|^2 \geq 0.$$

So we can remove the terms involving ν and λ in the inequality to get the wanted inequality. ■

5.2.1 HWI inequalities

Now we proceed to get some corollaries when we apply a quadratic Young function.

Corollary 5.12 *Under the same hypothesis as theorem 5.10, let $\mu \in \mathbb{R}$, and U :$\mathbb{R}^n \to \mathbb{R}$ be a C^2 function such that $D^2U \geq \mu I$, then for any $\sigma > 0$ we have that*

$$H_U^F(\rho_0 \mid \rho_1) + \frac{1}{2}(\mu - \frac{1}{\sigma})W_2^2(\rho_0,\rho_1) \leq \frac{\sigma}{2} \int_\Omega \rho \left|\nabla(F' \circ \rho_0 + U\right|^2 dx.$$

Proof. If we take the basic inequality with $c(x) = \frac{1}{2\sigma}|x|^2$, $W = 0$, and we set, $V = U - c$. Hence we have that $c^*(p) = \frac{1}{2\sigma}|\sigma p|^2 = \frac{\sigma}{2}|p|^2$, so using the general inequality we get

$$H_U^F(\rho_0 \mid \rho_1) + \frac{(\mu - \sigma^{-1})}{2} d_w^2(\rho_0,\rho_1) \leq H_{c+\nabla(U-c)\cdot x}^{-nP_F}(\rho_0) + \frac{\sigma}{2} \int_\Omega \rho_0 \nabla(F' \circ \rho_0 + U - c)dx.$$

We can compute

$$
\begin{aligned}
&\frac{\sigma}{2}\int_{\Omega}\rho_0\left|\nabla(F'\circ\rho_0+U-c)\right|^2dx\\
&=\frac{\sigma}{2}\int_{\Omega}\rho\left|\nabla(F'\circ\rho_0+U\right|^2dx+\frac{1}{2\sigma}\int_{\Omega}\rho_0|x|^2dx-\int_{\Omega}x\rho_0\cdot\nabla(F'\circ\rho_0+U)dx,
\end{aligned}
$$

and

$$
H^{nP_F}_{c+\nabla(U-c)\cdot x}(\rho_0)=H^{nP_F}(\rho_0)-\int_{\Omega}\rho x\cdot\nabla U dx+\frac{1}{2\sigma}\int_{\Omega}|x|^2\rho_0 dx.
$$

By combining the two and using integration by parts we get that

$$
\begin{aligned}
&H^{-nP_F}_{c+\nabla(U-c)\cdot x}(\rho_0)+\frac{\sigma}{2}\int_{\Omega}\rho_0\left|\nabla(F'\circ\rho_0+U-c)\right|^2dx\\
&=\frac{\sigma}{2}\int_{\Omega}\rho\left|\nabla(F'\circ\rho_0+U\right|^2dx-\int_{\Omega}x\rho_0\cdot\nabla(F'\circ\rho_0)dx-H^{nP_F}(\rho_0)\\
&=\frac{\sigma}{2}\int_{\Omega}\rho\left|\nabla(F'\circ\rho_0+U\right|^2dx+\int_{\Omega}div(x\rho_0)\cdot(F'\circ\rho_0)dx-H^{nP_F}(\rho_0)\\
&=\frac{\sigma}{2}\int_{\Omega}\rho\left|\nabla(F'\circ\rho_0+U\right|^2dx+\int_{\Omega}n\rho_0\cdot(F'\circ\rho_0)dx+\int_{\Omega}x\cdot\nabla F(\rho_0)dx-H^{nP_F}(\rho_0)\\
&=\frac{\sigma}{2}\int_{\Omega}\rho\left|\nabla(F'\circ\rho_0+U\right|^2dx+\int_{\Omega}x\cdot\nabla F(\rho_0)dx+\int_{\Omega}n(F\circ\rho_0)dx\\
&=\frac{\sigma}{2}\int_{\Omega}\rho\left|\nabla(F'\circ\rho_0+U\right|^2dx.
\end{aligned}
$$

Returning to the first inequality we get that

$$
H^F_U(\rho_0\mid\rho_1)+\frac{(\mu-\sigma^{-1})}{2}d^2_w(\rho_{0,}\rho_1)\leq\frac{\sigma}{2}\int_{\Omega}\rho\left|\nabla(F'\circ\rho_0+U)\right|^2dx.
$$

■

Corollary 5.13 *Furthermore if we take* $\mu>0$, *that is, take* U *is uniformly convex, take* $\sigma=\frac{1}{\mu}$, *we can get the Generalized Log-Sobolev inequality:*

$$
H^F_U(\rho_0\mid\rho_1)\leq\frac{1}{2\mu}\int_{\Omega}\rho\left|\nabla(F'\circ\rho_0+U)\right|^2dx=\frac{1}{2\mu}\mathscr{I}_2(\rho_0\mid\rho_U).
$$

Corollary 5.14 (HWI) *Finally we can obtain the generalized HWI-inequality,* *which is originally due to Otto and Villani (see [5]).*

$$H_U^F(\rho_0 \mid \rho_1) + \frac{\mu}{2} d_w^2(\rho_0,\rho_1) \le \sqrt{\mathscr{I}_2(\rho_0 \mid \rho_U)}\, d_w(\rho_0,\rho_1).$$

Proof. If we write the inequality of the last corollary as

$$H_U^F(\rho_0 \mid \rho_1) + \frac{\mu}{2} d_w^2(\rho_0,\rho_1) \le \frac{\sigma}{2}\mathscr{I}_2(\rho_0 \mid \rho_U) + \frac{1}{2\sigma} d_w^2(\rho_0,\rho_1)$$

and minimize over σ, we obtain the minimum when $\sigma = \frac{d_w(\rho_0,\rho_1)}{\sqrt{\mathscr{I}_2(\rho_0|\rho_U)}}$, we can write the inequality as

$$H_U^F(\rho_0 \mid \rho_1) + \frac{\mu}{2} d_w^2(\rho_0,\rho_1) \le \sqrt{\mathscr{I}_2(\rho_0 \mid \rho_U)}\, d_w(\rho_0,\rho_1).$$

■

5.2.2 Gaussian inequalities

By taking a particular F we can prove Otto-Villani's HWI inequality.

Corollary 5.15 *Let $\mu \in \mathbb{R}$, and $U : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function such that $D^2U \ge \mu I$, then for any $\sigma > 0$, and any non-negative function f such that $f\rho_U \in W^{1,\infty}(\mathbb{R}^n)$ and $\int f\rho_U = 1$, we have that*

$$\int f\log(f)\rho_U + \frac{1}{2}\left(\mu - \frac{1}{\sigma}\right)W_2^2(\rho_0,\rho_1) \le \frac{\sigma}{2}\int_\Omega \rho_U \frac{|\nabla f|^2}{f}dx.$$

Where $\rho_U = e^{-U}/\int e^{-U}dx$.

Proof. The proof follows from corollary 5.12, taking $\rho_0 = \rho_U$, $\rho_1 = f\rho_U$, and $F(x) = x\log x$. So we compute

$$\begin{aligned} H_U^F(\rho_U) &= \int \rho_U \log \rho_U + U\rho_U dx = \int \left[\left(e^{-U}/\int e^{-U}dx\right)\log\left(e^{-U}/\int e^{-U}dx\right) + U\left(e^{-U}/\int e^{-U}dx\right)\right] \\ &= \frac{1}{\int e^{-U}dx}\int e^{-U}\left(-\log\int e^{-U}\right) = -\log\int e^{-U}, \text{ and} \end{aligned}$$

$$H_U^F(f\rho_U) = \int f\rho_U \log f\rho_U + U f\rho_U = \frac{1}{\int e^{-U} dx} \int e^{-U} f(\log f - \log \int e^{-U})$$
$$= \int f \log(f)\rho_U - \left(\log \int e^{-U}\right) \int f\rho_U$$
$$= \int f \log(f)\rho_U - \log \int e^{-U}$$

Hence $H_U^F(\rho_0 \mid \rho_1) = \int f \log(f)\rho_U$. ■

Furthermore if U is uniformly convex , we can consider $\mu > 0$, so we can simplify the inequality to get the original Log-Sobolev inequality of Gross

$$\int f \log(f)\rho_U \leq \frac{1}{\mu} \int_\Omega \rho_U \frac{|\nabla f|^2}{f} dx.$$

Bibliography

[1] P. Bernard. Young measures, supersposition and transport. *Indiana University Mathematics Journal*, 57(1):247–246, 2008. → pages 2, 17

[2] C. V. D. Cordero-Erausquin, B. Nazaret. A mass-transportation approach to sharp sobolev and gagliardo-nirenberg inequalities. *preprint*, 2002. → pages 2, 27, 31

[3] W. G. D. Cordero-Erausquin and C. Houdre. Inequalities for generalized entropy and optimal transport. *Proceedings of the workshop: Mass Transportation Methods in Kinetic Theory and Hydrodynamics*, 2003. → pages 2, 27

[4] L. Evans. *Partial Differential Equations*. AMS, 2005. → pages 31

[5] C. V. F. Otto. Generalization of an inequality by talagrand and links with the logarithmic sobolev inequality. *J. Funct. Anal.*, 173:361–400, 2000. → pages 41

[6] L. Gross. Logarithmic sobolev inequa. *Amer. J. Math*, 97:1061–1083, 1975. → pages 29

[7] J.-M. L. H. Berliocchi. Intgrands normales and mesures paramtres en calcul des variations. *Bull. Soc. Math. France*, 101:129–184, 1973. → pages 18

[8] Y. B. J.-D Benamou. A computational fluid mechics solution to the monge-kantorovich mass transfer problem. *Numerische Mathematik*, 84(3): 375–393, 2000. → pages 1, 10

[9] G. S. L. Ambrosio, N. Gigli. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Math. ETH Zurich, 2005. → pages 17, 18, 19, 20

[10] X. K. M. Agueh, N. Ghossoub. Geometric inequalities via a general comparison principle for interacting gases. *Geometric and Functional Analysis*, 14(1):215–244, 2004. → pages 2, 3, 27, 33

[11] R. McCann. A canvexity principle for interacting gases. *Adv. Math*, 128(1): 153–179, 1997. → pages 2, 27

[12] N.-S. P. S. Hu. *Handbook of Multivalued Analysis*. Kluwer Academic Publishers, 1975. → pages 19

[13] C. Villani. *Topics in Optimal Transportation*. AMS, 2003. → pages 1, 4, 6, 26

[14] L. Young. *Lectures on the Calculus of Variations and Optimal Control Theory*. Chelsea, 2005. → pages 17