MASTER'S DEGREE THESIS

Master of Science in Advanced Mathematics and Mathematical Engineering

Title: Optimal Mass Transport and Functional Inequalities

Author: Núria Pascual Miranda

Advisor: Fernando Charro Caballero

Department: Mathematics

Academic year: 2016-2017

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Universitat Politècnica de Catalunya
BarcelonaTech

Master's Degree Thesis

Optimal Mass Transport and Functional Inequalities

Author:
Núria Pascual Miranda

Advisor:
Dr. Fernando Charro Caballero

Master of Science in Advanced
Mathematics and Mathematical Engineering

Facultat de Matemàtiques i Estadística

June 2017

$\mathbb{F}$, $\mathbb{X}$,
$\mathcal{G}\mathcal{G}$,
and home

Abstract

In this work we show how to use optimal transport to prove functional inequalities such as the Gagliardo-Nirenbertg-Sobolev (GNS) and isoperimetric inequalities.

The optimal transportation problem was first formulated by Monge in 1781 and asks if, given two mass densities (an initial and target configuration), it is possible to transport one into the other with minimal cost. The full resolution of Monge's problem has taken more than 200 years, and in this work we review some major milestones, namely, the original setup by Monge, the reformulation of the problem by Kantorovich, and Brenier's theorem.

We provide two proofs of both the isoperimetric and GNS inequalities, one using classical analysis tools, and another one using optimal mass transport arguments. The optimal transport proof has the advantage of its efficiency and flexibility; without extra effort, it allows to prove the inequalities with sharp constants and determine all cases of equality, which is in general not trivial. Moreover, since the Euclidean structure of $\mathbb{R}^n$ plays no role in the mass transport proof of GNS, these inequalities, together with cases of equality, can be established for an arbitrary norm in $\mathbb{R}^n$.

We conclude showing that both, the GNS inequality and the isoperimetric inequality, are actually equivalent in a compact n-dimensional Riemannian manifold.

Keywords: Optimal transport, Monge problem, isoperimetric inequality, Sobolev inequality, Monge-Ampère.

AMS Subject Classification (2010): 35B45, 35J60, 49N90, 28A75, 49Q20

Contents

Preface

This is a “classic problem” meets “new approach” kind of story.

France, 1781. Monsieur Gaspar Monge publishes his paper *Mémoire sur la théorie des déblais et des remblais*, in which he addresses the problem that later on will adopt his name: How to transport a pile of sand or rubble to fill an excavation or a hole with minimal cost. With this paper, optimal mass transportation is born.

In modern terms, the problem can be stated as follows: Given two mass densities $f, g \geq 0$ in $\mathbb{R}^n$, with

$$\int_{\mathbb{R}^n} f(x)\,dx = \int_{\mathbb{R}^n} g(y)\,dy \qquad \text{(mass conservation condition)},$$

find a map $T : \mathbb{R}^n \to \mathbb{R}^n$ minimizing the total cost

$$I(T) = \int_{\mathbb{R}^n} c\big(x, T(x)\big) f(x)\,dx$$

among all maps T that push f onto g. The function $c(x, y)$ encodes the cost of transporting mass from x to y and in Monge's work it was proportional to the distance that the mass has travelled, i.e. $c(x, y) = |x - y|$. However, costs corresponding to other powers

$$c_p(x, y) = |x - y|^p, \qquad \text{for } p \geq 1$$

are also of great interest, as we shall see below.

By means of heuristic arguments, Monge showed (among other properties) that if a cost-minimizing map existed, then it had to be the gradient of some convex function. Yet, he did not really solve the problem, since he did not even address the question of the existence of a minimizer.

As it turned out, Monge's problem, happened to be very difficult to solve. Small advances were made over the years, but it was not until the 1940s that the Russian mathematician and economist, Nobel Prize winner Leonid Kantorovich, made the first breakthrough. He proposed a relaxation (or weak version) of Monge's problem, and it was finally possible to prove that a solution to the Monge-Kantorovich mass transportation problem existed, even if in a weak sense.

The question was, then, if the weak solution proposed by Kantorovich was actually a solution to Monge's original problem.

The person that really sealed the deal was Monsieur Yan Brenier in 1987, in the paper [2]. More than 200 years after Monge's original paper, Brenier provided a rigorous proof of existence and uniqueness of an optimal map characterized by a convex potential, as Monge correctly guessed, for the quadratic cost $c_2(x, y) = |x - y|^2$.

In fact, the results by Brenier could be easily adapted to other costs that are strictly convex functions of the difference $x - y$, such as $c_p(x, y)$ for $p > 1$.

The original cost considered by Monge was still an open question (that was later on closed by Caffarelli-Feldman-McCann [6] and, independently, Trudinger-Wang [17] via an approximation through strictly convex costs). However, with Brenier's paper optimal mass transportation was reborn and gained interest in many mathematical fields. It was already a well-known subject in probability theory, economics and optimization; but this paper opened the door to other applications in partial differential equations, fluid mechanics, geometry, functional analysis...

And it is about one of these applications that this Master's thesis is about: Functional inequalities such as Gagliardo-Nirenberg-Sobolev or the isoperimetric inequality meet optimal transport in [9], see also [19].

For the sake of comparison, we provide two proofs of both the isoperimetric and GNS inequalities, one using classical, elementary, analysis tools, and another one using optimal mass transport arguments. As we shall see, the optimal transport proof has the advantage of its efficiency and flexibility; without extra effort, it allows to prove the inequalities with sharp constants and determine all cases of equality, which is in general not trivial.

Moreover, since the Euclidean structure of $\mathbb{R}^n$ plays no role in the mass transport proof of GNS, these inequalities, together with cases of equality, can be established for an arbitrary norm in $\mathbb{R}^n$.

We conclude showing that both, the GNS inequality and the isoperimetric inequality, are actually equivalent in a compact n-dimensional Riemannian manifold.

Notation

- $B_r = \{x \in \mathbb{R}^n : \|x\| \leq r\}$ will denote the ball of radius r in $\mathbb{R}^n$ centered at the origin.
- $\mathcal{H}^{n-k}$ will denote the $(n-k)$-dimensional Hausdorff measure.
- Whenever Ω is a set, $\mathbb{1}_\Omega(x)$ denotes its indicator function. That is,

$$\mathbb{1}_\Omega(x) = \begin{cases} 1 \text{ if } x \in \Omega, \\ 0 \text{ otherwise.} \end{cases}$$

- The Dirac mass at a point x will be denoted by δ_x. That is,

$$\delta_x[\Omega] = \begin{cases} 1 \text{ if } x \in \Omega, \\ 0 \text{ otherwise.} \end{cases}$$

- We will write $\Delta\varphi(x)$ as the trace of the Hessian of the convex function $\varphi(x)$, i.e.

$$\Delta\varphi(x) = \mathrm{tr} D^2\varphi(x),$$

  in the almost everywhere sense.

Chapter 1

Notions on Optimal Transport

In this chapter we shall introduce some notions on optimal transport that we shall use in order to prove functional inequalities. There is much more on this topic than what we shall show here see, for instance, [19].

1.1 Formulation of the optimal transportation problem

Suppose that we are given a pile of sand, and a hole that we have to completely fill up with the sand. Since the sand fills and fits in the hole, the pile and the hole must have the same volume.

In fact, we could reformulate the problem as reshaping a pile of sand into another, so that they can be modelled by two non-negative distributions f, g in $\mathbb{R}^n$, that must fulfill the compatibility condition

$$\int_{\mathbb{R}^n} f(x)\ dx = \int_{\mathbb{R}^n} g(y)\ dy.$$

We can also normalize this quantities to 1. Then, we can model the pile and the hole by probability measures μ and ν, defined on $\mathbb{R}^n$. Whenever $A, B \subset \mathbb{R}^n$ are measurable subsets of $\mathbb{R}^n$, then $\mu[A]$ measures how much sand is located inside A, and $\nu[B]$ how much sand can be piled in B.

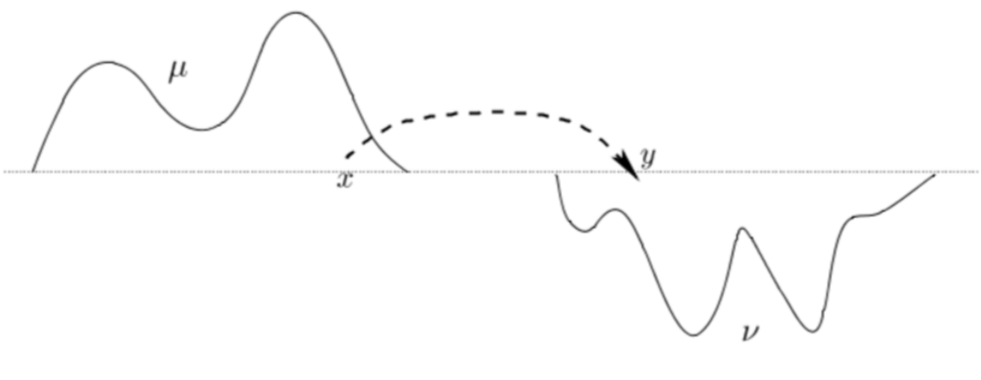

Figure 1.1: The mass transportation problem

Definition 1.1.1. If μ and ν are two probability measures on $\mathbb{R}^n$, then a Borel map $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to *push-forward* (or *transport*) μ onto ν if, whenever B is a Borel subset of $\mathbb{R}^n$, one has

$$\nu[B] = \mu\left[T^{-1}(B)\right],$$

or equivalently, for every non-negative Borel function $b : \mathbb{R}^n \to [0, \infty)$,

$$\int b(y)\, d\nu(y) = \int b(T(x))\, d\mu(x).$$

If T transports μ onto ν, we write

$$\nu = T_{\#}\mu.$$

This definition is also used for any couple of non-negative Borel measures on $\mathbb{R}^n$ with same total mass, even if it is not 1, that fulfill the conditions of Definition 1.1.1.

The effort of moving the sand around is modelled by a measurable *cost function*

$$c(x, y) : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty].$$

The function $c(x, y)$ tells how much does it cost to transport one unit of mass from location x to location y, and may take infinite values.

1.1.1 Monge's optimal transportation problem

Monge's problem consists on finding a transport map, or push-forward, that minimizes the total cost

$$I(T) = \int_{\mathbb{R}^n} c(x, T(x))\, d\mu(x),$$

over the set of all measurable maps T such that $T_{\#}\mu = \nu$.

Monge formulated this problem in 1781 [15] for the cost function $c(x, y) = |x - y|$. Later on, different cost functions have been considered. The most important ones are such as

$$c_p(x, y) = |x - y|^p,$$

for $p \geq 1$. Specially the case $p = 2$, known as the quadratic cost.

However, proving that this optimal transport map existed turned out to be very difficult, which is why on the 1940s Kantorovich presented a different approach [13].

1.1.2 Kantorovich's optimal transportation problem

So far we have been thinking about a transport map T such that $y = T(x)$ for all y, that is, we are transporting all the mass from x to a single point y. However, we may consider a more general *transport* or *transference plan*, such that the mass originally located in x can split and fill several possible destination y's. In the same way, not necessarily all the mass that arrives at a point y after the transference will have come from a single x.

We model this transference plans by probability measures π on the product space $\mathbb{R}^n \times \mathbb{R}^n$. The value $d\pi(x, y)$ measures the amount of mass transferred from location x to location y.

We shall say that a transference plan is admissible if all the mass taken from x coincides with $d\mu(x)$ and the total mass transferred to y coincides with $d\nu(y)$. This is called the marginal condition, namely

$$\int_{\mathbb{R}^n} d\pi(x, y) = d\mu(x) \text{ and } \int_{\mathbb{R}^n} d\pi(x, y) = d\nu(y).$$

More rigorously, we need that for all measurable subsets $A, B \subset \mathbb{R}^n$, we have

$$\pi[A \times \mathbb{R}^n] = \mu[A] \text{ and } \pi[\mathbb{R}^n \times B] = \nu[B],$$

or equivalently, for all functions b_x in $L^1(d\mu)$ and b_y in $L^1(d\nu)$,

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} [b_x(x) + b_y(y)] \, d\pi(x,y) = \int_{\mathbb{R}^n} b_x(x) \, d\mu(x) + \int_{\mathbb{R}^n} b_y(y) d\nu(y). \tag{1.1}$$

If these are satisfied, we say that μ and ν are *marginals* of π, an *admissible* transference plan. With this, we shall denote the set of such probability measures

$$\Pi(\mu, \nu) = \{\pi \text{ transference plan}; \pi \text{ is admissible}\}.$$

Remark 1.1.2. This set is nonempty, since $\mu \otimes \nu \in \Pi(\mu, \nu)$. This transportation plan distributes all pieces of sand, independently of its location, over the entire hole, proportionally to the depth.

Remark 1.1.3. As a matter of fact, the class of test functions in Equation (1.1) may vary. In general, they will be in L^1 or L^∞, but in some cases we shall focus on continuous bounded functions or continuous functions going to 0 at ∞.

With these notions, the Kantorovich's problems can be stated as follows.

Kantorovich's optimal transportation problem

Minimize the total transportation cost

$$I[\pi] = \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x,y) d\pi(x,y),$$

for $\pi \in \Pi(\mu, \nu)$.

Kantorovich's approach has the great advantage that it is easier to prove that an optimal transport plan exists in general. The question remains if by solving Kantorovich's problem, we are solving Monge's original problem.

1.1.3 Monge-Kantorovich problem

As we can see, Kantorovich's problem is a relaxation of Monge's problem, since the only difference between them is that Monge does not admit that the mass splits. That is, as we said earlier, to each location x it is associated a unique destination y.

In terms of transference plans, we are asking π, an optimal transference plan, to have the special form

$$d\pi(x,y) = d\mu(x)\delta[y = T(x)],$$

where T transports μ onto ν. We want the transference plan to be a transport map.

With this, "Monge-Kantorovich problem" will refer to either of the two minimization problems. Our questions will be whether there exists an optimal transference plan or not, and wheter it is a transport map or not.

We shall denote the optimal transport cost between μ and ν as

$$\mathcal{T}_c(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} I[\pi].$$

Example 1.1.4. *(Dirac mass)* If ν is a Dirac mass: $\nu = \delta_a$. Then, the set $\Pi(\mu, \nu)$ has a unique element: all the mass is transported to a. So,

$$\mathcal{T}_c(\mu, \delta_a) = \int_{\mathbb{R}^n} c(x, a)\ d\mu(x).$$

Example 1.1.5. *(The discrete case)* Suppose that we have goods at k different locations $x_1, \ldots, x_k$ in $\mathbb{R}^n$, that we want to transport to k new locations $y_1, \ldots, y_k$. We can express this situation as the densities with total mass k,

$$\mu = \sum_{i=1}^{k} \delta_{x_i} \text{ and } \nu = \sum_{j=1}^{k} \delta_{y_j}.$$

In this case, we want to minimize the total cost

$$I(y) = \sum_{i=1}^{k} c(x_i, y(x_i)),$$

where $y(x_i)$ denotes the destination as a map of the starting point. Since the number of locations is finite, we can check all permutations, and thus we know that this minimum exists, provided by an *optimal matching.*

This example is very useful in order to realize that the cost function that we pick has an effect on the choices of y. For instance, if we pick the linear cost $c(x, y) = |x - y|$, we know that the optimal map will not have any transport rays crossing. That is, the vectors $y(x_i) - x_i$ and $y(x_j) - x_j$ will not cross for any i, j different.

On the other hand, if we pick the quadratic cost $c(x, y) = |x - y|^2$, the transport rays may cross, but the map will be cyclically-monotone: the matching $y(x_i) = y_i$ is optimal if and only if for all permutations $\sigma \in \mathfrak{S}_k$,

$$\sum_{i=1}^{k} |x_i - y_i|^2 \leq \sum_{i=1}^{k} |x_i - y_{\sigma(i)}|^2.$$

For instance, without loss of generality assume that the matching $y(x_i) = y_i$ is optimal and consider the mapping $y(x_i) = y_{i+1}$, with the convention $y_{k+1} = y_1$. With this particular permutation,

$$\sum_{j=1}^{k} |x_j - y_j|^2 \leq \sum_{j=1}^{k} |x_j - y_{j+1}|^2,$$

which expanding each term is

$$\sum_{j=1}^{k} |x_j|^2 + |y_j|^2 - 2\langle x_j, y_j \rangle \leq \sum_{j=1}^{k} |x_j|^2 + |y_{j+1}|^2 - 2\langle x_j, y_{j+1} \rangle,$$

and canceling out the repeated terms, we get

$$\sum_{j=1}^{k} \langle x_j, y_j \rangle \geq \sum_{j=1}^{k} \langle x_j, y_{j+1} \rangle,$$

or

$$\sum_{j=1}^{k} \langle x_j, y_j - y_{j+1} \rangle \geq 0,$$

which is the condition for being cyclically monotone.

In conclusion, the cost function chosen has consequences on the characteristics of the resulting mappings.

1.2 Brenier's theorem and Monge-Ampère equation

Brenier's theorem is the most important tool in optimal transport for proving statements from very different areas. In [19] one can find both a proof for the original theorem and a proof for the refinement done by McCann, see [14].

Brenier's theorem closes Monge's problem for the quadratic cost, after being open for two centuries [2]. It is completely a game changer for optimal transport, and attracts researchers of several fields to start using these techniques.

Theorem 1.2.1. (Brenier) *If μ and ν are two probability measures on $\mathbb{R}^n$ and μ is absolutely continuous with respect the to Lebesgue measure, then there exists a convex function φ such that $\nabla\varphi$ transports μ onto ν. Furthermore, $\nabla\varphi$ is uniquely determined $d\mu-$almost everywhere, and uniquely minimizes Monge's problem for the quadratic cost.*

We shall refer to the mapping $\nabla\varphi$ as the *Brenier map* pushing μ forward to ν.

1.2.1 Fully-nonlinear Monge-Ampère type PDE

Let us consider $T_{\#}\mu = \nu$ and assume that $\mu = f(x)dx$ and $\nu = g(y)dy$ on $\mathbb{R}^n$. Then, if b is a non-negative Borel test function, it follows that

$$\int_{\mathbb{R}^n} b(T(x))f(x)dx = \int_{\mathbb{R}^n} b(y)g(y)dy. \tag{1.2}$$

If T were a diffeomorphism, we could apply the change of variables $y = T(x)$ to get

$$\int_{\mathbb{R}^n} b(T(x))f(x)dx = \int_{\mathbb{R}^n} b(T(x))g(T(x))|\det\nabla T(x)|dy,$$

and since this would be for any test function, we would conclude

$$f(x) = |\det\nabla T(x)|g(T(x)).$$

This is a fully nonlinear PDE, an equation of prescribed Jacobian. Actually, it can be seen that this equation holds $f(x)-$almost everywhere, see [14].

In the case of the Brenier map $T(x) = \nabla\varphi(x)$, we know that φ is convex, so the Jacobian is non-negative $\nabla T(x) = D^2\varphi(x) \geq 0$.

We may now recall Alexandrov's theorem, that states that if U is an open subset of $\mathbb{R}^n$ and $\phi : U \to \mathbb{R}^m$ is a convex function, then ϕ has a second derivative almost everywhere. Thus, the following proposition is valid.

Proposition 1.2.2. (Monge-Ampère equation for the Brenier map) *If $\nabla\varphi$ is the Brenier map transporting $\mu = f(x)dx$ onto $\nu = g(y)dy$, then we have*

$$f(x) = \det(D^2\varphi(x))g(\nabla\varphi(x)) \text{ almost everywhere.}$$

It is a fully nonlinear equation that prescribes the product of the eigenvalues of the Hessian of the solution u, in contrast with the "model" elliptic equation $\Delta u = f$ that prescribes their sum. However, in contrast with the Laplacian, which is always elliptic, Monge-Ampère equation is degenerate elliptic only when the solution is a convex function (see for instance [12, 16]).

Chapter 2

Isoperimetric Inequality

The first functional inequality that we shall discuss is probably one of the most important and famous: the isoperimetric inequality. It establishes that among all the subsets of $\mathbb{R}^n$ with the same volume, the ball will be the one with less surface area.

The isoperimetric inequality is specially well-known as its planar case, that claims that if L is the length of a closed curve and A the area of the region it encloses, the inequality $4\pi A \leq L^2$ holds, with equality if and only if the curve is a circle.

Ancient Greeks already discussed the problem to determine the plane figure with the largest area between those with the same perimeter (those who are *iso-perimetric*). They knew that the circle was the answer, even though they did not rigorously prove it.

Being such a classical problem, there are many different proofs using several mathematical techniques. For example, the standard proof of the isoperimetric problem [1] uses Steiner symmetrization; informally, if a domain is symmetric with respect to all hyperplanes passing through a point, then it must be a ball.

In this chapter, we present two proofs of the isoperimetric inequality in $\mathbb{R}^n$, one by induction on n and another one using an optimal transport approach using some of the concepts explained in Chapter 1.

Somehow related to the optimal transport method is the paper [3], where the author proves the isoperimetric inequality with best constant by means of the Alexandroff-Bakelman-Pucci (ABP) technique applied to a linear Neumann problem for the Laplacian. In addition, it shows easily that balls are the only smooth domains for which equality holds.

First of all, however, we need to revise some preliminary results that will be necessary in order to get to our goal and prove the isoperimetric inequality.

2.1 Preliminary results

2.1.1 Arithmetic Mean-Geometric Mean inequality

This inequality is probably one of the most elementary, but at the same time, one of the most useful.

Proposition 2.1.1. (Arithmetic Mean-Geometric Mean inequality) *For any list of n non-negative real numbers $x_1, x_2, \ldots, x_n$,*

$$\frac{x_1 + x_2 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \cdots x_n},$$

and equality holds if and only if $x_1 = x_2 = \cdots = x_n$.

Proof. By induction on n. For $n = 1$ the statement is trivially true with equality.

Assume now that the inequality holds for any n non-negative real numbers. Consider $x_1, \dots, x_{n+1} \geq 0$, without loss of generality assume $x_1 \geq x_2 \geq \cdots \geq x_{n+1}$, and let α be their arithmetic mean. Then,

$$\alpha(n+1) = x_1 + x_2 + \cdots + x_{n+1}.$$

If they are all equal, our statement is true with equality and we are done. On the other hand, if not all of them are equal, then we know that $x_1 > \alpha$ and $x_{n+1} < \alpha$. Then $x_1 - \alpha > 0$ and $\alpha - x_{n+1} > 0$, so $(x_1 - \alpha)(\alpha - x_{n+1}) > 0$.

Now, define $\beta := x_1 + x_{n+1} - \alpha \geq x_1 - \alpha > 0$, and since $\alpha(n+1) = x_1 + x_2 + \cdots + x_{n+1}$, we have

$$n\alpha = \beta + x_2 + \cdots + x_n,$$

so α is the arithmetic mean of the n non-negative numbers $\beta, x_2, \dots, x_n$. By induction hypothesis, we have

$$\alpha \geq \sqrt[n]{\beta x_2 \cdots x_n},$$

and thus,

$$\alpha^{n+1} \geq \alpha\beta x_2 \cdots x_n.$$

On the other hand, we have

$$0 < (x_1 - \alpha)(\alpha - x_{n+1}) = \alpha \underbrace{(x_1 + x_{n+1} - \alpha)}_{\beta} - x_1 x_{n+1},$$

hence $x_1 x_{n+1} < \alpha\beta$ and, in particular, $\alpha > 0$. Note that, if at least one of the numbers $x_2, \dots, x_n$ is equal to zero, we are done. Otherwise, we have $x_2 \cdots x_n \alpha\beta > x_1 x_2 \cdots x_n x_{n+1}$ and with this,

$$\alpha^{n+1} > x_1 x_2 \cdots x_n x_{n+1},$$

so

$$\frac{x_1 + x_2 + \cdots + x_{n+1}}{n+1} \geq \sqrt[n+1]{x_1 x_2 \cdots x_{n+1}}.$$

□

This inequality can be rewritten in terms of determinants and traces. As a matter of fact, this expression is the one that we shall use the most. For instance, it will be very useful as a bound for the Monge-Ampère operator (see Proposition 1.2.2).

Proposition 2.1.2. *Let* $M \in \mathbb{R}^{n \times n}$ *be a positive semi-definite symmetric matrix. Then,*

$$\frac{\mathrm{tr} M}{n} \geq \sqrt[n]{\det M}.$$

Proof. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of M. They are all non-negative since M is positive semi-definite. Therefore, using Proposition 2.1.1,

$$\frac{\mathrm{tr} M}{n} = \frac{\lambda_1 + \cdots + \lambda_n}{n} \geq \sqrt[n]{\lambda_1 \cdots \lambda_n} = \sqrt[n]{\det M}.$$

□

2.1.2 Integral results

This first result is one of the main theorems in calculus, and for that it is widely known. We shall need it when trying to relate the volume of a subset with the surface area of its boundary.

Theorem 2.1.3. (Divergence theorem) *Let $\Omega \subset \mathbb{R}^n$ compact with piece-wise smooth boundary. If F is a continuously differentiable vector field defined on a neighbourhood of Ω, then we have*

$$\int_\Omega \nabla \cdot F dx = \int_{\partial\Omega} F \cdot \hat{n} \; d\mathcal{H}^{n-1}(x).$$

Lastly, we shall also need Jensen's inequality. This result generalizes the idea that the secant line of a convex function lies *above* the graph of the function.

Proposition 2.1.4. (Jensen's inequality) *Let (Ω, Σ, μ) be a probability space, such that $\mu(\Omega) = 1$. If g is a real-valued function that is $\mu-$integrable, and φ is a convex function on the real line, then*

$$\varphi\left(\int_\Omega g d\mu\right) \leq \int_\Omega \varphi \circ g d\mu.$$

Proof. Since φ is convex, at each real number x we have a nonempty set of subderivatives: lines touching the graph of φ at x that lie below the graph of φ at all points. Now, define

$$x_0 := \int_\Omega g d\mu.$$

Then, we can choose a and b such that

$$ax + b \leq \varphi(x)$$

for all x, and such that $ax_0 + b = \varphi(x_0)$, i.e. the affine function $ax + b$ touches *varphi* from below at x_0. But then we have $(\varphi \circ g)(x) \geq ag(x) + b$ for all x. Now, since we have a probability measure, we can say that

$$\int_\Omega \varphi \circ g d\mu \geq \int_\Omega (ag + b) d\mu = a \int_\Omega g d\mu + b \int_\Omega d\mu = ax_0 + b = \varphi(x_0) = \varphi\left(\int_\Omega g d\mu\right).$$

□

2.2 A proof of the isoperimetric inequality using classical tools

With all these tools we are finally in position to state and prove the isoperimetric inequality. First, we shall see a more classical proof, by induction on the dimension of the space. This proof is extracted from [18].

Theorem 2.2.1. *For any bounded subset $\Omega \subset \mathbb{R}^n$, if $\mathrm{Vol}(\Omega) = \mathrm{Vol}(B_1)$, then $\mathcal{H}^{n-1}(\partial\Omega) \geq \mathcal{H}^{n-1}(\partial B_1)$.*

Proof. Proof by induction on n.

The base case $n = 1$ is trivial, with $\mathcal{H}^{n-1}(\partial\Omega)$ defined as the number of points of the border. Clearly, $\mathcal{H}^{n-1}(\partial\Omega) \geq 2 = \mathcal{H}^{n-1}(\partial B_1)$.

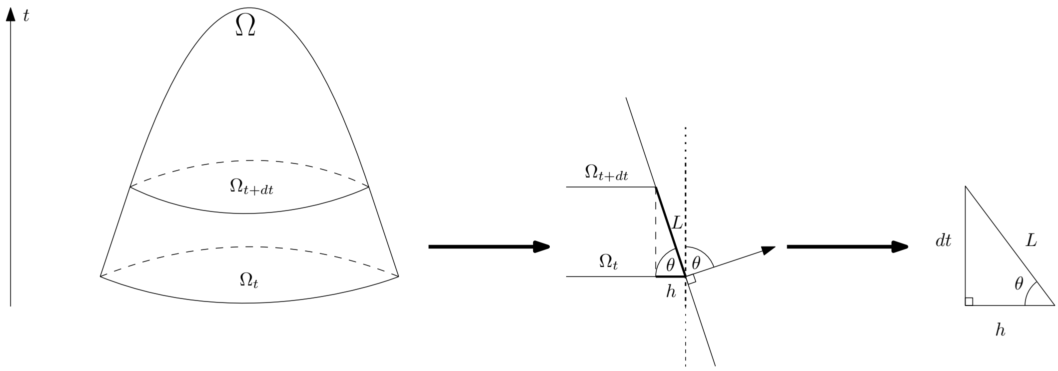

Figure 2.1: The set Ω and two close-ups.

Now assume that the inequality holds for dimension $n-1$. Using the same reasoning as in the proof of Corollary 2.4.2, we can see that this can be rewritten as: for all K bounded subset in $\mathbb{R}^{n-1}$, we have

$$\frac{|\partial K|}{|K|^{\frac{n-2}{n-1}}} \geq \frac{|\partial B_1^{n-1}|}{|B_1^{n-1}|^{\frac{n-2}{n-1}}}, \tag{2.1}$$

where B_1^{n-1} is the unit ball of dimension $(n-1)$.

We shall see that the theorem holds for dimension n. Define $\Omega_t = \Omega \cap \{x_n = t\}$ and $\partial\Omega_t = \partial\Omega \cap \{x_n = t\}$. Let $V(t) = \mathcal{H}^{n-1}(\Omega_t)$ and let $A(t) = \mathcal{H}^{n-2}(\partial\Omega_t)$. Let θ be the angle between the x_n-axis and the normal vector to $\partial\Omega$. Note that since Ω_t are parallel slices, we have

$$\int V(t)dt = \mathrm{Vol}(\Omega).$$

Consider now the diagram on figure Figure 2.1. Note that

$$V(t+dt) - V(t) \approx \int_{\partial\Omega_t} h \approx dt \int_{\partial\Omega_t} \frac{1}{\tan\theta},$$

and then,

$$V'(t) = \int_{\partial\Omega_t} \frac{1}{\tan\theta}.$$

On the other hand,

$$\begin{aligned}
\int_{\partial\Omega_t} \frac{1}{\sin\theta} &= \int_{\partial\Omega_t} \sqrt{1 + \frac{1}{\tan^2\theta}} && \text{using trigonometry} \\
&\geq \sqrt{A(t)^2 + \left(\int_{\partial\Omega_t} \frac{1}{\tan\theta}\right)^2} && \text{by Proposition 2.1.4 with} \varphi = \sqrt{1+\alpha^2} \\
&= \sqrt{A(t)^2 + V'(t)^2}.
\end{aligned}$$

On Figure 2.1, note that the surface between Ω_{t+dt} and Ω_t is approximately

$$\int_{\partial\Omega_t} L \, d\sigma = \int_{\partial\Omega_t} \frac{dt}{\sin\theta} \, d\sigma,$$

were σ is the revolution around the x_n-axis.

Therefore,

$$\mathcal{H}^{n-1}(\partial\Omega) = \int \left(\int_{\partial\Omega_t} \frac{1}{\sin\theta} \right) dt \geq \int \sqrt{A(t)^2 + V'(t)^2}\ dt. \tag{2.2}$$

Now, define the function $h(\tau)$ as the number in $[-1, 1]$ such that

$$\int_{-\infty}^{\tau} V(t)dt = \int_{-1}^{h(\tau)} \left|B_1^{n-1}\right| (1-s^2)^{\frac{n-1}{2}} ds.$$

The function $h(\tau)$ matches the volume of the ball up to height $h(\tau)$ with the volume of Ω up to height τ. It is well-defined because

$$\int V(t)dt = \mathrm{Vol}(B_1) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} = \frac{\pi^{(n-1)/2}}{\Gamma\left(\frac{n+1}{2}\right)} \cdot \frac{\pi^{1/2}\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)} = \left|B_1^{n-1}\right| \int_{-1}^{1} \left(1-s^2\right)^{\frac{n-1}{2}} ds.$$

Then, differentiating we get

$$V(t) = \left|B_1^{n-1}\right| (1-h(t)^2)^{\frac{n-1}{2}} h'(t).$$

Let

$$f(t) = V(t)^{-1/(n-1)} \left|B_1^{n-1}\right|^{1/(n-1)} (1-h(t)^2)^{1/2},$$

and observe that $f(t)^{n-1}h'(t) = 1$ and thus, $f(t)^{(n-1)/n}h'(t)^{1/n} = 1$. Note also that both $f(t)$ and $h'(t)$ are non-negative. Hence, by the arithmetic-geometric mean inequality,

$$1 \leq \frac{1}{n}\left((n-1)f(t) + h'(t)\right). \tag{2.3}$$

On the other hand, using Cauchy-Schwarz on the vectors $(A(t), V'(t))$ and $\left(\sqrt{1-h(t)^2}, -h\right)$, we have

$$\sqrt{A(t)^2 + V'(t)^2} \geq \sqrt{1-h(t)^2}A(t) - h(t)V'(t).$$

So, by the definition of f, we have

$$\sqrt{A(t)^2 + V'(t)^2} \geq \left|B_1^{n-1}\right|^{\frac{-1}{n-1}} f(t)V(t)^{\frac{1}{n-1}} A(t) - h(t)V'(t).$$

Using the induction hypothesis (Equation (2.1)), we have

$$\left|B_1^{n-1}\right|^{\frac{-1}{n-1}} V(t)^{\frac{1}{n-1}} A(t) \geq \frac{|\partial B_1^{n-1}|}{|B_1^{n-1}|} V(t) = (n-1)V(t),$$

and using Equation (2.3), we get to

$$\sqrt{A(t)^2 + V`(t)^2} \geq nV(t) - h'(t)V(t) - h(t)V'(t). \tag{2.4}$$

And putting Equation (2.2) and Equation (2.4) together,

$$\begin{aligned}
\mathcal{H}^{n-1}(\partial\Omega) &\geq \int \sqrt{A(t)^2 + V'(t)^2} dt \\
&\geq \int \left(nV(t) - (h(t)V(t))' \right)\ dt \\
&= n \int V(t)\ dt = \mathcal{H}^{n-1}(\partial B_1).
\end{aligned}$$

$\square$

2.3 Proof using optimal transport

Let us now proceed with the proof using optimal transport for the same theorem. The advantage of this approach is that it uses very few tools, and it is easily portable to other contexts. The ideas for this proof have been extracted from [14].

Theorem 2.3.1. *For any bounded subset $\Omega \subset \mathbb{R}^n$, if $\mathrm{Vol}(\Omega) = \mathrm{Vol}(B_1)$, then $\mathcal{H}^{n-1}(\partial\Omega) \geq \mathcal{H}^{n-1}(\partial B_1)$.*

Proof. Let $f(x) = \mathbb{1}_\Omega(x)$ and $g(y) = \mathbb{1}_{B_1}(y)$. Then, $\mu = f(x)dx$ and $\nu = g(y)dy$ are compactly supported measures. So Brenier's theorem (Theorem 1.2.1) gives us, for the quadratic cost, the existence of a unique transport map $T = \nabla\varphi$, for some φ convex.

Then, for any test function b defined on B_1,

$$\int_\Omega b\left(\nabla\varphi(x)\right) f(x)dx = \int_{B_1} b(y)g(y)dy,$$

and we can apply a change of variables, $y = T(x) = \nabla\varphi(x)$, and we get

$$\int_\Omega b\left(\nabla\varphi(x)\right) f(x)dx = \int_\Omega b\left(\nabla\varphi\left(x\right)\right) g\left(\nabla\varphi(x)\right) \left|\det\left(D^2\varphi(x)\right)\right| dx.$$

Thus, given that this equality holds for every test function b, we have that

$$f(x) = g(\nabla\varphi(x)) \left|\det(D^2\varphi(x))\right| \text{ a.e. in } \Omega.$$

Therefore, the convex function φ is a solution of the Monge-Ampère equation

$$\left|\det D^2\varphi\right| = \frac{f(x)}{g(\nabla\varphi(x))},$$

which by construction of f and g, we know that both of these quantities are equal to 1. In particular, we have $\det^{1/n}\left(D^2\varphi(x)\right) = 1$ almost everywhere in Ω.

Then, since we know that φ is convex, its Hessian matrix $D^2\varphi$ is positive semidefinite. Therefore, all its eigenvalues will be non-negative real numbers. So applying the inequality between the arithmetic and geometric means, we have

$$1 = \sqrt[n]{\det\left(D^2\varphi(x)\right)} \leq \frac{\mathrm{tr}\left(D^2\varphi(x)\right)}{n} = \frac{\Delta\varphi(x)}{n}.$$

Integrating over Ω gives us

$$\mathrm{Vol}(\Omega) = \int_\Omega 1dx \leq \frac{1}{n}\int_\Omega \Delta\varphi(x)dx.$$

By the Divergence theorem,

$$\int_\Omega \Delta\varphi dx = \int_\Omega \nabla\cdot\nabla\varphi dx = \int_{\partial\Omega} \nabla\varphi\cdot\hat{n}\; d\mathcal{H}^{n-1}(x),$$

and given that $\nabla\varphi(x) \in B_1$ if $x \in \Omega$, we know that for all $x \in \Omega$, since $\nabla\varphi$ transports Ω into B_1, we have $|\nabla\varphi(x)| \leq 1$, thus

$$\begin{aligned}
\mathrm{Vol}(B_1) = \mathrm{Vol}(\Omega) &\leq \frac{1}{n}\int_\Omega \Delta\varphi dx \\
&= \frac{1}{n}\int_{\partial\Omega} \nabla\varphi\cdot\hat{n}\; d\mathcal{H}^{n-1}(x) \\
&\leq \frac{1}{n}\int_{\partial\Omega} 1\; d\mathcal{H}^{n-1}(x) = \frac{1}{n}\mathcal{H}^{n-1}(\partial\Omega).
\end{aligned}$$

Now, since we know that $\text{Vol}(B_1) = \frac{1}{n}\mathcal{H}^{n-1}(\partial B_1)$, we get

$$\mathcal{H}^{n-1}(\partial B_1) \leq \mathcal{H}^{n-1}(\partial\Omega).$$

Furthermore, in the special case $\Omega = B_1$, Brenier's map coincides with the identity map and equality holds in all the steps. Thus, the inequality is tight with optimal constant. □

2.4 Consequences

There are several equivalent ways of stating the isoperimetric inequality. Below, we present some of them.

Corollary 2.4.1. *For any bounded subset $\Omega \subset \mathbb{R}^n$, if $\text{Vol}(\Omega) = \text{Vol}(B_r)$, for some $r > 0$, then we have that $\mathcal{H}^{n-1}(\partial\Omega) \geq \mathcal{H}^{n-1}(\partial B_r)$.*

Proof. If $\text{Vol}(\Omega) = \text{Vol}(B_r)$, then $\text{Vol}(\Omega/r) = \text{Vol}(B_1)$. By Theorem 2.3.1, we know that $\mathcal{H}^{n-1}(\partial(\Omega/r)) \geq \mathcal{H}^{n-1}(\partial B_1)$, thus $\mathcal{H}^{n-1}(\partial\Omega) \geq \mathcal{H}^{n-1}(\partial B_r)$. □

Corollary 2.4.2. *For any bounded subset $\Omega \subset \mathbb{R}^n$,*

$$\frac{\mathcal{H}^{n-1}(\partial\Omega)}{\text{Vol}(\Omega)^{\frac{n-1}{n}}} \geq \frac{\mathcal{H}^{n-1}(\partial B_1)}{\text{Vol}(B_1)^{\frac{n-1}{n}}}.$$

Proof. Choose $r > 0$ such that $\text{Vol}(\Omega) = \text{Vol}(B_r)$, since we know $\text{Vol}(B_r) = r^n\text{Vol}(B_1)$, that is

$$r = \left(\frac{\text{Vol}(\Omega)}{\text{Vol}(B_1)}\right)^{\frac{1}{n}}.$$

Since $\mathcal{H}^{n-1}(\partial B_r) = r^{n-1}\mathcal{H}^{n-1}(\partial B_1)$, by Corollary 2.4.1,

$$\mathcal{H}^{n-1}(\partial\Omega) \geq r^{n-1}\mathcal{H}^{n-1}(\partial B_1) = \left(\frac{\text{Vol}(\Omega)}{\text{Vol}(B_1)}\right)^{\frac{n-1}{n}} \mathcal{H}^{n-1}(\partial B_1).$$

Therefore,

$$\frac{\mathcal{H}^{n-1}(\partial\Omega)}{\text{Vol}(\Omega)^{\frac{n-1}{n}}} \geq \frac{\mathcal{H}^{n-1}(\partial B_1)}{\text{Vol}(B_1)^{\frac{n-1}{n}}}.$$

□

Note that each of these inequalities is tight, since we can take Ω equal to the (unit) ball, and all of them match an equality. As a matter of fact, sometimes, since the optimal subset is easily found, Corollary 2.4.2 is reformulated to the existence of a positive constant C such that for all $\Omega \subset \mathbb{R}^n$,

$$C\text{Vol}(\Omega)^{\frac{n-1}{n}} \leq \mathcal{H}^{n-1}(\partial\Omega).$$

Chapter 3

Sobolev Inequalities

Sobolev inequalities are well-known results in analysis. They provide one of the most important tools in the study of partial differential equations. They are a very flexible tool, useful in proofs of existence and regularity of solutions; in particular it plays a key role in Moser's iteration technique. Let us start with a definition.

Definition 3.0.1. Let $n \geq 1$ be an integer and $p \geq 1$ be a real number. We define the *Sobolev space* as

$$W^{1,p}(\mathbb{R}^n) := \{f \in L^p(\mathbb{R}^n); \nabla f \in L^p(\mathbb{R}^n)\}.$$

Here ∇f should be interpreted as the first weak-derivative of f, i.e., for all test functions $\phi \in C_c^\infty(\mathbb{R}^n)$, we have

$$\int_{\mathbb{R}^n} f \nabla \phi \, dx = -\int_{\mathbb{R}^n} \nabla f \phi \, dx.$$

Now, our question is the following: given an arbitrary function u, if u belongs to $W^{1,p}$, what can we say about its integrability?

3.1 Hölder's inequalities

In order to answer our question, however, we shall need to use Hölder's inequalities, which we revise here.

Definition 3.1.1. The *dual exponent*, or *Hölder conjugate*, of $p > 1$ is

$$q = \frac{p}{p-1}.$$

We shall need to use a general version of the well-known Hölder's inequality. In order to do that, below we recall the original version.

Let (Ω, Σ, μ) be a measure space and let $p, q \in [1, \infty]$ be Hölder conjugates. Then, for all measurable real-valued or complex-valued functions f and g on Ω,

$$\|fg\|_{L^1(\mu)} \leq \|f\|_{L^p(\mu)} \|g\|_{L^q(\mu)}.$$

Furthermore, if $p, q \neq 1$, and we have $f \in L^p(\mu)$ and $g \in L^q(\mu)$, then we have an equality if and only if there exist $\alpha, \beta \geq 0$, with at least one of them not zero, such that $\alpha|f|^p = \beta|g|^q$ μ-almost everywhere, i.e., the functions f and g are linearly dependent in $L^1(\mu)$.

Theorem 3.1.2. (Generalization of Hölder's inequality) *Let (Ω, Σ, μ) be a measure space and let $r \in [1, \infty)$ and $p_1, \dots, p_n \in (0, \infty]$ such that*

$$\sum_{k=1}^{n} \frac{1}{p_k} = \frac{1}{r}.$$

Then, for all measurable real-valued or complex-valued functions $f_1, \dots, f_n$ defined on Ω,

$$\left\| \prod_{k=1}^{n} f_k \right\|_{L^r(\mu)} \leq \prod_{k=1}^{n} \|f_k\|_{L^{p_k}(\mu)}.$$

In particular, if we have $f_k \in L^{p_k}(\mu)$ for all k, then $\prod_{k=1}^{n} f_k \in L^{p_k}(\mu)$.

Proof. By induction on n.

For $n = 1$, it is trivially true. Assume now that the result is true for $(n-1)$, we shall see that is also true for n. Without loss of generality, assume $p_1 \leq p_2 \leq \cdots \leq p_n$.

- If $p_n = \infty$, then
$$\sum_{k=1}^{n-1} \frac{1}{p_k} = \frac{1}{r}.$$
Thus, by induction hypothesis we have
$$\begin{aligned} \|f_1 \cdots f_n\|_{L^r} &\leq \|f_1 \cdots f_{n-1}\|_{L^r} \|f_n\|_{L^\infty} \\ &\leq \|f_1\|_{L^{p_1}} \cdots \|f_{n-1}\|_{L^{p_{n-1}}} \|f_n\|_{L^\infty}. \end{aligned}$$

- If $p_n < \infty$, then let $p, q \in (1, \infty)$ be
$$p = \frac{p_n}{p_n - r} \text{ and } q = \frac{p_n}{r}.$$
Note that $\frac{1}{p} + \frac{1}{q} = 1$, so we can apply Hölder's inequality as
$$\||f_1 \cdots f_{n-1}|^r |f_n|^r\|_{L^1} \leq \||f_1 \cdots f_{n-1}|^r\|_{L^p} \||f_n|^r\|_{L^q}.$$
Raising this inequality to the power $1/r$ and rewriting, we obtain
$$\|f_1 \cdots f_n\|_{L^r} \leq \|f_1 \cdots f_{n-1}\|_{L^{pr}} \|f_n\|_{L^{qr}}.$$
But since $qr = p_n$ and
$$\sum_{k=1}^{n-1} \frac{1}{p_k} = \frac{1}{r} - \frac{1}{p_n} = \frac{p_n - r}{r p_n} = \frac{1}{pr},$$
by induction hypothesis we have
$$\|f_1 \cdots f_n\|_{L^r} \leq \|f_1\|_{L^{p_1}} \cdots \|f_{n-1}\|_{L^{p_{n-1}}} \|f_n\|_{L^{p_n}}.$$

$\square$

3.2 Sobolev conjugate

Now we can return to our question, if we know that u belongs to $W^{1,p}$, what can we say about it? We will focus on the cases with p such that $1 \leq p < n$. We want to see if there is any relation we can find of the form

$$\|u\|_{L^r(\mathbb{R}^n)} \leq C\|\nabla u\|_{L^p(\mathbb{R}^n)}$$

for two constants independents of u, named $C > 0$ and $1 \leq r < \infty$ for all $u \in C_c^\infty(\mathbb{R}^n)$. Since $C_c^\infty(\mathbb{R}^n)$ is dense in $C^\infty(\mathbb{R}^n)$, that is dense in $W^{1,p}(\mathbb{R}^n)$, we shall establish this estimate for arbitrary functions in $W^{1,p}(\mathbb{R}^n)$.

The first thing we have to notice is that the value of the constant r cannot be arbitrary. In order to see this, consider a function $u \in C_c^\infty(\mathbb{R}^n)$, with $u \not\equiv 0$, and for $\lambda > 0$ define the rescaled function

$$u_\lambda(x) := u(\lambda x), \text{ for } x \in \mathbb{R}^n.$$

If there are C and r such that the previous inequality holds, applying it for u_λ we shall have $\|u_\lambda\|_{L^r(\mathbb{R}^n)} \leq C\|\nabla u_\lambda\|_{L^p(\mathbb{R}^n)}$. On the left hand side, applying the change of variables $y = \lambda x$, we have

$$\int_{\mathbb{R}^n} |u_\lambda|^r dx = \int_{\mathbb{R}^n} |u(\lambda x)|^r dx = \frac{1}{\lambda^n}\int_{\mathbb{R}^n} |u(y)|^r dy,$$

while on the right hand side we have

$$\int_{\mathbb{R}^n} |\nabla u_\lambda|^p dx = \lambda^p \int_{\mathbb{R}^n} |\nabla u(\lambda x)|^p dx = \frac{\lambda^p}{\lambda^n}\int_{\mathbb{R}^n} |\nabla u(y)|^p dy.$$

Thus,

$$\frac{1}{\lambda^{n/r}}\|u\|_{L^r(\mathbb{R}^n)} \leq C\frac{\lambda}{\lambda^{n/p}}\|\nabla u\|_{L^p(\mathbb{R}^n)},$$

and so

$$\|u\|_{L^r(\mathbb{R}^n)} \leq C\lambda^{1-\frac{n}{p}+\frac{n}{r}}\|\nabla u\|_{L^p(\mathbb{R}^n)}.$$

Now, if $1-\frac{n}{p}+\frac{n}{r} > 0$, if we choose λ close to 0 we reach contradiction, whereas if $1-\frac{n}{p}+\frac{n}{r} < 0$, choosing $\lambda \to \infty$ we get a contradiction as well. Therefore, our only possibility is that

$$1-\frac{n}{p}+\frac{n}{r} = 0,$$

that is, $r = \frac{np}{n-p}$.

Definition 3.2.1. Given $1 \leq p < n$, we define the *Sobolev conjugate of* p as

$$p^* := \frac{np}{n-p}.$$

Note that the dual exponent q defined on Definition 3.1.1 and the Sobolev conjugate p^* defined on Definition 3.2.1 are different and should not be confused.

And with this notion, we can introduce the homogeneous Sobolev spaces as follows.

Definition 3.2.2. Let $n \geq 1$ be an integer and $p \geq 1$ be a real number. We define the *homogeneous Sobolev space* as

$$\dot{W}^{1,p}(\mathbb{R}^n) := \left\{ f \in L^{p^*}(\mathbb{R}^n); \nabla f \in L^p(\mathbb{R}^n) \right\}.$$

3.3 Gagliardo-Nirenberg-Sobolev inequality

As we did with the isoperimetric inequality, we shall prove the GNS inequality with two different methods: a classic approach with iteration of integration and Hölder's inequality and a transport approach.

3.3.1 A proof of the GNS inequality using classical tools

We start with the classical approach, that can be found in [10].

Theorem 3.3.1. (Gagliardo-Nirenberg-Sobolev inequality) *Assume* $1 \leq p < n$. *There exists a constant* C*, depending only on* p *and* n*, such that, for all* $u \in C_c^1(\mathbb{R}^n)$,

$$\|u\|_{L^{p^*}(\mathbb{R}^n)} \leq C\|\nabla u\|_{L^p(\mathbb{R}^n)}.$$

Proof. Case $p = 1$. In this case, $p^* = \frac{n}{n-1}$.

Since u has compact support, for each $i = 1, \ldots, n$ and $x \in \mathbb{R}^n$ we can write

$$u(x) = \int_{-\infty}^{x_i} u_{x_i}(x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n)\, dy_i,$$

and thus

$$|u(x)| \leq \int_{-\infty}^{\infty} |\nabla u(x_1, \ldots, y_i, \ldots, x_n)|\, dy_i.$$

Therefore,

$$|u(x)|^{\frac{n}{n-1}} \leq \prod_{i=1}^{n} \left(\int_{-\infty}^{\infty} |\nabla u(x_1, \ldots, y_i, \ldots, x_n)|\, dy_i \right)^{\frac{1}{n-1}}.$$

Now, we integrate with respect to x_1:

$$\begin{aligned}
\int_{-\infty}^{\infty} |u|^{\frac{n}{n-1}} dx_1 &\leq \int_{-\infty}^{\infty} \prod_{i=1}^{n} \left(\int_{-\infty}^{\infty} |\nabla u| dy_i \right)^{\frac{1}{n-1}} dx_1 \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} |\nabla u| dy_1 \right)^{\frac{1}{n-1}} \prod_{i=2}^{n} \left(\int_{-\infty}^{\infty} |\nabla u| dy_i \right)^{\frac{1}{n-1}} dx_1 \\
&= \left(\int_{-\infty}^{\infty} |\nabla u| dy_1 \right)^{\frac{1}{n-1}} \int_{-\infty}^{\infty} \prod_{i=2}^{n} \left(\int_{-\infty}^{\infty} |\nabla u| dy_i \right)^{\frac{1}{n-1}} dx_1, \quad \text{and by Theorem 3.1.2,} \\
&\leq \left(\int_{-\infty}^{\infty} |\nabla u| dy_1 \right)^{\frac{1}{n-1}} \left(\prod_{i=2}^{n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\nabla u| dx_1 dy_i \right)^{\frac{1}{n-1}}.
\end{aligned}$$

Now, we integrate this with respect to x_2:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |u|^{\frac{n}{n-1}} dx_1 dx_2 \leq \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} |\nabla u| dy_1\right)^{\frac{1}{n-1}} \left(\prod_{i=2}^{n}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\nabla u| dx_1 dy_i\right)^{\frac{1}{n-1}} dx_2$$

$$= \int_{-\infty}^{\infty}\left(\iint_{\mathbb{R}^2} |\nabla u| dx_1 dy_2\right)^{\frac{1}{n-1}} \left(\int_{-\infty}^{\infty} |\nabla u| dy_1\right)^{\frac{1}{n-1}} \left(\prod_{i=3}^{n}\iint_{\mathbb{R}^2} |\nabla u| dx_1 dy_i\right)^{\frac{1}{n-1}} dx_2$$

$$= \left(\iint_{\mathbb{R}^2} |\nabla u| dx_1 dy_2\right)^{\frac{1}{n-1}} \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} |\nabla u| dy_1\right)^{\frac{1}{n-1}} \left(\prod_{i=3}^{n}\iint_{\mathbb{R}^2} |\nabla u| dx_1 dy_i\right)^{\frac{1}{n-1}} dx_2$$

$$\leq \left(\iint_{\mathbb{R}^2} |\nabla u| dx_1 dy_2\right)^{\frac{1}{n-1}} \left(\iint_{\mathbb{R}^2} |\nabla u| dy_1 dx_2\right)^{\frac{1}{n-1}} \prod_{i=3}^{n}\left(\iiint_{\mathbb{R}^3} |\nabla u| dx_1 dx_2 dy_i\right)^{\frac{1}{n-1}},$$

where in the last step we have used again Theorem 3.1.2.

Then we continue integrating with respect to each variable $x_3, \ldots, x_n$ and using Theorem 3.1.2. When we integrate with respect to x_k, we have

$$\int_{\mathbb{R}^k} |u|^{\frac{n}{n-1}} dx_1 \cdots dx_k \leq \left(\int_{\mathbb{R}^k} |\nabla u| dx_1 \cdots dx_k\right)^{\frac{k}{n-1}} \prod_{i=k+1}^{n}\left(\int_{\mathbb{R}^{k+1}} |\nabla u| dx_1 \cdots dx_k dy_i\right)^{\frac{1}{n-1}}.$$

Eventually, after integrating with respect to x_n, we find

$$\int_{\mathbb{R}^n} |u|^{\frac{n}{n-1}}\, dx \leq \left(\int_{\mathbb{R}^n} |\nabla u|\, dx\right)^{\frac{n}{n-1}}.$$

This can be rewritten as

$$\left(\int_{\mathbb{R}^n} |u|^{\frac{n}{n-1}}\right)^{\frac{n-1}{n}} \leq \int_{\mathbb{R}^n} |\nabla u| dx, \tag{3.1}$$

which is exactly $\|u\|_{L^{p^*}(\mathbb{R}^n)} \leq \|\nabla u\|_{L^p(\mathbb{R}^n)}$ for $p = 1$ and $p^* = \frac{n}{n-1}$.

Case $1 < p < n$.

We take the inequality Equation (3.1) and we apply it to $v := |u|^{\gamma}$, with $\gamma > 1$ yet to be defined. Then,

$$\begin{aligned}\left(\int_{\mathbb{R}^n} |u|^{\frac{\gamma n}{n-1}} dx\right)^{\frac{n-1}{n}} &\leq \int_{\mathbb{R}^n} |\nabla |u|^{\gamma}|\, dx = \gamma \int_{\mathbb{R}^n} |u|^{\gamma-1} |\nabla u| dx \\ &\leq \gamma \left(\int_{\mathbb{R}^n} |u|^{(\gamma-1)\frac{p}{p-1}} dx\right)^{\frac{p-1}{p}} \left(\int_{\mathbb{R}^n} |\nabla u|^p dx\right)^{\frac{1}{p}}.\end{aligned}$$

Now, choose γ such that the the exponents of $|u|$ are the same on both sides of the inequality, that is,

$$\frac{\gamma n}{n-1} = (\gamma - 1)\,\frac{p}{p-1} \implies \gamma = \frac{p(n-1)}{n-p} > 1,$$

and then

$$\frac{\gamma n}{n-1} = p^* = (\gamma - 1)\,\frac{p}{p-1}.$$

Hence, we have

$$\left(\int_{\mathbb{R}^n} |u|^{p^*}\, dx\right)^{\frac{1}{p^*}} \leq C \left(\int_{\mathbb{R}^n} |\nabla u|^p dx\right)^{\frac{1}{p}},$$

which is exactly

$$\|u\|_{L^{p^*}(\mathbb{R}^n)} \leq C\|\nabla u\|_{L^p(\mathbb{R}^n)}.$$

□

Remark 3.3.2. It is necessary that u has a compact support. For instance, for the case $u \equiv 1$, we cannot find any C fulfilling $\|u\|_{L^{p^*}(\mathbb{R}^n)} \leq C\|\nabla u\|_{L^p(\mathbb{R}^n)}$.

3.3.2 Proof using optimal transport

Here we shall proceed with the transport approach, following the ideas of [9]. With this approach we shall be able to easily prove our inequality for a general norm of $\mathbb{R}^n$, rather than restricting to Euclidean norms. To do so, however, we need to introduce a few concepts.

Let $E = (\mathbb{R}^n, \|\cdot\|)$, where $\|\cdot\|$ is an arbitrary norm on $\mathbb{R}^n$. Then the dual space is $E^* = (\mathbb{R}^n, \|\cdot\|_*)$ where, for $X \in E^*$,

$$\|X\|_* := \sup_{\|Y\|\leq 1} X \cdot Y = \sup_{\|Y\|\leq 1} \left(\sum_{i=1}^n X_i Y_i\right).$$

Theorem 3.3.3. (Young's inequality) *For all $X \in E^*$ and $Y \in E$, and for $\lambda > 0$,*

$$X \cdot Y \leq \frac{\lambda^{-p}}{p}\|X\|_*^p + \frac{\lambda^q}{q}\|Y\|^q,$$

where q is the dual exponent of p.

In this context, Hölder's inequality is rewritten as follows.

Theorem 3.3.4. (Hölder's inequality) *For $X : \mathbb{R}^n \to E^*$ in L^p and $Y : \mathbb{R}^n \to E$ in L^q, with p and q Hölder conjugates, we have*

$$\int X \cdot Y \leq \left(\int \|X\|_*^p\right)^{\frac{1}{p}} \left(\int \|Y\|^q\right)^{\frac{1}{q}}. \tag{3.2}$$

Proof. Integration of the inequality in Theorem 3.3.3 gives us

$$\int X \cdot Y \leq \frac{\lambda^{-p}}{p} \int \|X\|_*^p + \frac{\lambda^q}{q} \int \|Y\|^q, \tag{3.3}$$

for $\lambda > 0$. By optimizing with respect to λ, we get that the right-hand side has a minimum that is achieved at

$$\lambda = \left(\left(\int \|X\|_*^p\right)\left(\int \|Y\|^q\right)^{-1}\right)^{\frac{1}{pq}}.$$

Thus, substituting this λ into Equation (3.3), we obtain Equation (3.2). □

Note that we can extract from this inequality that the dual space of $L^p(\mathbb{R}^n)$ coincides with $L^q(\mathbb{R}^n)$.

Remark 3.3.5. Our arbitrary norm $\|\cdot\|$ is Lipschitz, thus, it is differentiable almost everywhere. If $x \in \mathbb{R}^n \setminus \{0\}$ is a point of differentiability, then the gradient of the norm at x is the unique vector $x^* = \nabla\left(\|\cdot\|\right)(x)$ such that $\|x^*\|_* = 1$ and

$$x \cdot x^* = \|x\| = \sup_{\|y\|_*=1} x \cdot y.$$

In fact, for the Euclidean norm $|\cdot|$, we have $x^* = \frac{x}{|x|}$.

Remark 3.3.6. If $f \in \dot{W}^{1,p}(\mathbb{R}^n)$, it is natural to consider the dual norm of ∇f. Then, we define

$$\|\nabla f\|_{L^p(\mathbb{R}^n)} := \left(\int \|\nabla f\|_*^p\right)^{\frac{1}{p}}.$$

For $p \in (1,n)$, we define the following function

$$h_p(x) := \frac{1}{\left(\sigma_p + \|x\|^q\right)^{\frac{n-p}{p}}},$$

where $\sigma_p > 0$ is determined by the condition $\|h_p\|_{L^{p^*}(\mathbb{R}^n)} = 1$.

For $p = 1$, however, we define

$$h_1(x) := \frac{\mathbb{1}_{B_1}(x)}{|B_1|^{\frac{n-1}{n}}}.$$

Remark 3.3.7. These h_p are chosen because, for almost every x, there is *equality* in Young's inequality (Theorem 3.3.3) when $X = -\nabla h_p(x)$, $Y = h_p^{p^*/q}(x)x$ and

$$\lambda = \lambda_p := \left(\frac{n-p}{p-1}\right)^{1/q}.$$

After some computations we can get

$$\left(\frac{n-p}{p-1}\right)\frac{\|x\|^q}{\left(\sigma_p + \|x\|^q\right)^n} = \frac{1}{p\lambda_p^p}\left(\frac{n-p}{p-1}\right)^p\frac{\|x\|^q}{\left(\sigma_p + \|x\|^q\right)^n} + \frac{\lambda_p^q}{q}\frac{\|x\|^q}{\left(\sigma_p + \|x\|^q\right)^n}.$$

As a consequence, the same choice of X and Y gives an equality in Theorem 3.3.4.

$$-\int \nabla h_p(x)\cdot\left[h_p^{p^*/q}(x)x\right]\,dx = \|\nabla h_p\|_{L^p}\left(\int \|x\|^1 h_p^{p^*}(x)\,dx\right)^{1/q}.$$

Now, we shall prove a result from which we shall deduce the GNS inequality.

Lemma 3.3.8. *Let $p \in (1,n)$ and q its dual exponent. Let $f \in \dot{W}^{1,p}(\mathbb{R}^n)$ and $g \in L^{p^*}(\mathbb{R}^n)$ be two functions such that $\|f\|_{L^{p^*}(\mathbb{R}^n)} = \|g\|_{L^{p^*}(\mathbb{R}^n)}$, then*

$$\frac{\int |g|^{p^*(1-1/n)}}{\left(\int \|y\|^q |g(y)|^{p^*}dy\right)^{1/q}} \leq \frac{p(n-1)}{n(n-p)}\|\nabla f\|_{L^p(\mathbb{R}^n)},$$

with equality if $f = h_p = g$.

Proof. First, note that f and $|f|$ have equal Sobolev norms, since $f \in \dot{W}^{1,p}(\mathbb{R}^n)$ and then $\nabla|f| = \pm\nabla f$ almost everywhere. Therefore without loss of generality, we shall assume that f and g are non-negative and $\|f\|_{L^{p^*}(\mathbb{R}^n)} = \|g\|_{L^{p^*}(\mathbb{R}^n)} = 1$. We can also assume f and g smooth and with compact support. The general case follows by density.

Now, let F, G two probability densities on $\mathbb{R}^n$,

$$F(x) = f^{p^*}(x) \text{ and } G(y) = g^{p^*}(y).$$

Now, by Brenier's theorem (Theorem 1.2.1) there is a convex function φ such that $\nabla\varphi$ transports $F(x)dx$ onto $G(y)dy$. Then, for $F(x)dx-$almost every $x \in \mathbb{R}^n$ we have the Monge-Ampère equation (Proposition 1.2.2)

$$F(x) = G\left(\nabla\varphi(x)\right) \det D^2\varphi(x).$$

Then,

$$\begin{aligned} G^{-1/n}(\nabla\varphi(x)) &= F^{-1/n}(x)(\det D^2\varphi(x))^{1/n} \\ &\leq F^{-1/n}(x)\frac{\Delta\varphi(x)}{n} \qquad \text{by the AM-GM inequality (Proposition 2.1.2).} \end{aligned}$$

Integrating both sides with respect to $F(x)dx$, we have

$$\int G^{-1/n}(\nabla\varphi(x))F(x)dx \leq \frac{1}{n}\int F(x)^{1-\frac{1}{n}}(x)\Delta\varphi(x)dx.$$

On the other hand, using the compatibility condition (1.2) with $b(y) = G^{-1/n}(y)$, we have

$$\int G^{1-\frac{1}{n}}(y)dy = \int G^{-1/n}(y)G(y)dy = \int G^{-1/n}(\nabla\varphi(x))F(x)dx.$$

Combining these two results, we get

$$\int G^{1-\frac{1}{n}} \leq \frac{1}{n}\int F^{1-\frac{1}{n}}\Delta\varphi.$$

We are using $\Delta\varphi = \mathrm{tr}D^2\varphi$ in the almost everywhere sense.

Since f and g are compactly supported by assumption, we know that $\nabla\varphi$ is bounded on the support of f, since $\nabla\varphi(\mathrm{supp}(f)) \subset \mathrm{supp}(g)$. Therefore, we can assume that $\mathrm{supp}(f)$ lies within an open set where φ is finite, extending the function if we need it. Hence, the Laplacian $\Delta\varphi$ can be bounded above by the distributional Laplacian $\Delta_{\mathcal{D}'}\varphi$, a non-negative measure on the set where φ is finite. With this, we can apply the integration by parts formula

$$\frac{1}{n}\int F^{1-\frac{1}{n}}\Delta\varphi \leq \frac{1}{n}\int F^{1-\frac{1}{n}}\Delta_{\mathcal{D}'}\varphi = -\frac{1}{n}\int \nabla\left(F^{1-\frac{1}{n}}\right)\cdot\nabla\varphi.$$

Recall that $F(x) = f^{p^*}(x)$ and $G(y) = g^{p^*}(y)$, so we have

$$\int g^{p^*(1-1/n)} \leq \frac{1}{n}\int \nabla\left(f^{p^*(1-\frac{1}{n})}\right)\cdot\nabla\varphi = -\frac{p(n-1)}{n(n-p)}\int f^{p^*/q}\nabla f\cdot\nabla\varphi.$$

On the other hand, if we apply Hölder's inequation (Theorem 3.3.4) with $X = -\nabla f$ and $Y = f^{p^*/q}\nabla\varphi$, we obtain

$$\begin{aligned} -\int f^{p^*/q}\nabla f\cdot\nabla\varphi &\leq \left(\int \|-\nabla f\|_*^p\right)^{1/p}\left(\int \|f^{p^*/q}\nabla\varphi\|^q\right)^{1/q} \\ &= \|\nabla f\|_{L^p(\mathbb{R}^n)}\left(\int f^{p^*}\|\nabla\varphi\|^q\right)^{1/q} \end{aligned}$$

Now, using again the compatibility condition (1.2) with $b(y) = \|y\|^q$, we get

$$\int f^{p^*} \|\nabla\varphi\|^q = \int \|y\|^q g^{p^*}(y)dy.$$

So we have

$$\int g^{p^*(1-1/n)} \leq \frac{p(n-1)}{n(n-p)} \|\nabla f\|_{L^p(\mathbb{R}^n)} \left(\int \|y\|^q g^{p^*}(y)dy \right)^{1/q},$$

which concludes the proof of the inequality.

Now, in the special case $f = h_p = g$, the Brenier map coincides with the identity map $\nabla\varphi(x) = x$, and with this and Remark 3.3.7, equality holds throughout the proof. □

An immediate consequence of this lemma is the duality principle

$$\sup_{\|g\|_{L^{p^*}(\mathbb{R}^n)}=1} \frac{\int |g|^{p^*(1-1/n)}}{\left(\int \|y\|^q |g(y)|^{p^*} dy\right)^{1/q}} = \frac{p(n-1)}{n(n-p)} \inf_{\|f\|_{L^{p^*}(\mathbb{R}^n)}=1} \|\nabla f\|_{L^p(\mathbb{R}^n)},$$

with h_p extremal in both variational problems.

Now we can prove the main result of this section.

Theorem 3.3.9. (Sharp Sobolev inequality) *Let $1 \leq p < n$. Then there exists a constant C such that*

$$\|f\|_{L^{p^*}(\mathbb{R}^n)} \leq C \|\nabla f\|_{L^p(\mathbb{R}^n)},$$

where:

(i) *If $p > 1$, the fuction $f \neq 0$ is any function in $\dot{W}^{1,p}(\mathbb{R}^n)$, and the optimal constant is $C = \left(\|\nabla h_p\|_{L^p(\mathbb{R}^n)}\right)^{-1}$.*

(ii) *If $p = 1$, the function $f \neq 0$ is any smooth compactly supported function. In this case the optimal constant is $C = n^{-1}|B_1|^{-1/n}$.*

Proof.

(i) Case $p > 1$.

If $f \neq 0$ lies in $\dot{W}^{1,p}(\mathbb{R}^n)$, so does $f/\|f\|_{L^{p^*}(\mathbb{R}^n)}$. Also, we have

$$\left\| \frac{f}{\|f\|_{L^{p^*}}} \right\|_{L^{p^*}(\mathbb{R}^n)} = 1 = \|h_p\|_{L^{p^*}(\mathbb{R}^n)}.$$

On the other hand, using that h_p is extremal in the duality principle of Lemma 3.3.8, we know that

$$\|\nabla h_p\|_{L^p(\mathbb{R}^n)} = \inf_{\|\tilde{f}\|_{L^{p^*}}=1} \|\nabla \tilde{f}\|_{L^p(\mathbb{R}^n)} \leq \left\| \nabla \left(\frac{f}{\|f\|_{L^{p^*}}} \right) \right\|_{L^p(\mathbb{R}^n)}.$$

Thus,

$$\|f\|_{L^{p^*}(\mathbb{R}^n)} \leq \left(\|\nabla h_p\|_{L^p(\mathbb{R}^n)}\right)^{-1} \|\nabla f\|_{L^p(\mathbb{R}^n)},$$

with $C = \left(\|\nabla h_p\|_{L^p(\mathbb{R}^n)}\right)^{-1}$ being the optimal constant since $f = h_p$ makes the equality hold.

(ii) Case $p = 1$.

This proof follows the same structure as the proof of Lemma 3.3.8.

Without loss of generality, assume f non-negative and $\|f\|_{L^{n/(n-1)}} = 1$. Let $\nabla\varphi$ be the Brenier map (see Theorem 1.2.1) pushing forward $F(x)dx = f^{n/(n-1)}(x)dx$ onto $G(y)dy = h_1^{n/(n-1)}(y)dy$. Using the same reasoning as before, we obtain

$$\int G^{(n-1)/n} \leq \frac{1}{n} \int F^{(n-1)/n} \Delta\varphi \leq -\frac{1}{n} \int \nabla \left(F^{(n-1)/n}\right) \cdot \nabla\varphi,$$

which in terms of f is

$$\int h_1 = |B_1|^{1/n} \leq \frac{1}{n} \int f \Delta\varphi \leq -\frac{1}{n} \int \nabla f \cdot \nabla\varphi.$$

Now, by definition of h_1, we have that $\nabla\varphi(x) \in B_1$ for almost every $x \in \operatorname{supp}(f)$. Thus,

$$-\nabla f \cdot \nabla\varphi \leq \sup_{x \in B_1} x \cdot \nabla f = \|\nabla f\|_*,$$

and therefore

$$n|B|^{1/n} \leq \int \|\nabla f\|_* = \|\nabla f\|_{L^1}.$$

In fact, the case $p = 1$ extends to functions of bounded variation, with equality if $f = h_1$, and therefore the constant is optimal.

$\square$

Chapter 4

Equivalence between the Sobolev and isoperimetric inequalities

The purpose of this chapter is to show that, in fact, the Sobolev inequality and the isoperimetric inequality are equivalent. We shall use the ideas provided in [7, 8].

Isoperimetric inequalities can also be studied for domains on manifolds. They are powerful analytical tools. For instance, every isoperimetric inequality leads, through the use of the coarea formula (see proposition 4.2.1 below) to sharp Sobolev inequalities also in the generality of Riemannian manifolds, see [7, 8].

Let us mention that, as an extension of the ABP method introduced in [3] and mentioned in the introduction to Chapter 2, in the recent paper [5] new isoperimetric and Sobolev inequalities with weights are stablished in convex cones of $\mathbb{R}^n$, see also the survey [4].

4.1 General statements

In this chapter we shall use more general statements for both inequalities, which we write below.

Definition 4.1.1. Let (M, g) be an n-dimensional Riemannian manifold, with $n \geq 2$. We say that it has the *isoperimetric property* if there exists a constant C such that for any relatively compact domain Ω with smooth boundary, we have

$$C\mathrm{Vol}\,(\Omega)^{\frac{n-1}{n}} \leq \mathcal{H}^{n-1}(\partial\Omega).$$

Recall that, in fact, thanks to Corollary 2.4.2 we know the optimal value for this constant C. However, in this case we just need the existence of such a C and we shall not discuss its optimality.

Theorem 4.1.2. (Sobolev inequality for $p = 1$) *Let (M, g) be a compact n-dimensional Riemannian manifold, possibly with boundary. Then there exists a constant C' such that*

$$C' \left(\int_M |u|^{\frac{n}{n-1}} \right)^{\frac{n-1}{n}} \leq \int_M |\nabla u|,$$

for any $u \in C_c^\infty(\mathbb{R}^n)$.

4.2 Preliminary results

The following propositions are well-known results that we shall need in the proof of the equivalence of the Sobolev and isoperimetric inequalities.

The first one is the Coarea formula, which can be found in [11].

Proposition 4.2.1. (Coarea formula) *Let $\Omega \subset \mathbb{R}^n$ be an open set and let u be a real-valued Lispchitz function in Ω. Then, for all $g \in L^1$,*

$$\int_\Omega g(x)\,|\nabla u(x)|\,dx = \int_{-\infty}^{\infty} \left(\int_{u^{-1}(t)} g(x) d\mathcal{H}^{n-1}(x) \right) dt.$$

Note that in the particular case $g \equiv 1$, what we obtain from the coarea formula is

$$\int_\Omega |\nabla u(x)| dx = \int_{-\infty}^{\infty} \mathcal{H}^{n-1}\left(u^{-1}(t)\right) dt.$$

As a matter of fact, the coarea formula can be generalized to Lipschitz functions u defined in $\Omega \subset \mathbb{R}^n$ and taking values in $\mathbb{R}^k$ as follows,

$$\int_\Omega g(x)|J_k u(x)|dx = \int_{\mathbb{R}^k} \left(\int_{u^{-1}(t)} f(x) d\mathcal{H}^{n-k}(x) \right) \, dt,$$

where $J_k u(x)$ is the k-dimensional Jacobian of u.

The second result that we shall need is the following version of the Cavalieri principle, also known as layer cake representation.

Proposition 4.2.2. (Layer cake representation) *Let $f : \mathbb{R}^n \to [0, \infty)$ measurable, then*

$$f(x) = \int_0^\infty \mathbb{1}_{\{y \in \mathbb{R}^n | f(y) \geq t\}}(x) \, dt \qquad \forall x \in \mathbb{R}^n.$$

Proof. For all $x \in \mathbb{R}^n$,

$$\int_0^\infty \mathbb{1}_{\{y \in \mathbb{R}^n | f(y) \geq t\}}(x) \, dt = \int_0^\infty \mathbb{1}_{[0, f(x)]}(t) \, dt = \int_0^{f(x)} 1 \, dt = f(x).$$

□

4.3 Equivalence

The following is the main result of the chapter.

Theorem 4.3.1. *The isoperimetric property is equivalent to the Sobolev inequality.*

Proof. First, assume that the Sobolev inequality holds, and let us deduce the isoperimetric property from it.

Let Ω be a relatively compact domain with smooth boundary.

For $\varepsilon > 0$, we define the function

$$u_\varepsilon(x) = \begin{cases} 1, & x \in \Omega, d(x, \partial\Omega) \geq \varepsilon, \\ \frac{d(x,\partial\Omega)}{\varepsilon}, & x \in \Omega, d(x, \partial\Omega) \leq \varepsilon, \\ 0, & x \notin \Omega. \end{cases}$$

If we apply the Sobolev inequality (Theorem 4.1.2) to u_ε, we obtain

$$C\left(\int_\Omega |u_\varepsilon|^{\frac{n}{n-1}}\ dx\right) \leq \int_\Omega |\nabla u_\varepsilon|\ dx \tag{4.1}$$

for some constant C. We know that $\nabla u_\varepsilon(x) = 0$ for $x \notin \Omega$ or $x \in \Omega$ with $d(x, \partial\Omega) \geq \varepsilon$. On the other hand, if $x \in \Omega$ with $d(x, \partial\Omega) \leq \varepsilon$, we shall have

$$|\nabla u_\varepsilon(x)| = \left|\nabla\left(\frac{d(x,\partial\Omega)}{\varepsilon}\right)\right| = \left|\frac{1}{\varepsilon}\nabla\left(d(x,\partial\Omega)\right)\right| \leq \frac{1}{\varepsilon},$$

as we saw in Remark 3.3.5. Now, taking $\varepsilon \to 0$ in eq. (4.1), we have

$$C\left(\int_\Omega dx\right)^{\frac{n-1}{n}} \leq \int_{\partial\Omega} d\mathcal{H}^{n-1},$$

which is

$$C\mathrm{Vol}(\Omega)^{\frac{n-1}{n}} \leq \mathcal{H}^{n-1}(\partial\Omega),$$

as we intended to prove.

Now, assume that M has the isoperimetric property, and let us show the Sobolev inequality. Without loss of generality, assume $u \geq 0$. Now,

$$\begin{aligned}
\int_M |u|^{\frac{n}{n-1}} dM &= \int_M \int_0^\infty \mathbb{1}_{\left\{y\in\mathbb{R}^n \middle| |u(y)|^{\frac{n}{n-1}} \geq\ t\right\}} dt dM && \text{by Proposition 4.2.2,}\\
&= \int_0^\infty \mathrm{Vol}\left(\left\{u^{\frac{n}{n-1}} \geq t\right\}\right)\ dt\\
&= \int_0^\infty \mathrm{Vol}\left(\left\{u \geq t^{\frac{n-1}{n}}\right\}\right)\ dt\\
&= \frac{n}{n-1}\int_0^\infty \mathrm{Vol}\left(\{u \geq s\}\right) s^{\frac{1}{n-1}} ds && \text{with a change of variables } s^{\frac{n}{n-1}} = t.
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
\int_M |\nabla u| &= \int_{-\infty}^\infty \mathcal{H}^{n-1}\left(u^{-1}(s)\right) ds && \text{by Proposition 4.2.1,}\\
&= \int_0^\infty \mathcal{H}^{n-1}\left(\{u = s\}\right)\ ds && \text{since } u \geq 0,\\
&\geq C' \int_0^\infty \mathrm{Vol}\left(\{u \geq s\}\right)^{\frac{n-1}{n}}\ ds, && \text{by the isoperimetric property.}
\end{aligned}$$

Now, if we want to prove that for some constant C

$$C\left(\int_M |u|^{\frac{n}{n-1}}\right)^{\frac{n-1}{n}} \leq \int_M |\nabla u|,$$

we just need to see that for some constant C'',

$$C''\int_0^\infty \mathrm{Vol}(\{u \geq s\}) s^{\frac{1}{n-1}} ds \leq \left(\int_0^\infty \mathrm{Vol}(\{u \geq s\})^{\frac{n-1}{n}} ds\right)^{\frac{n}{n-1}}.$$

If we consider the function $u \equiv 1$, even though it does not fulfill the conditions for the Sobolev inequality because it does not have a compact support, it seems natural to take $C'' = \frac{n}{n-1}$, because this is the value that would make the equality hold in this case. Since we are looking for the existence of a constant C'' and, as we shall see, this value works, we shall use this intuition.

Let

$$\alpha(t) = \left(\int_0^t \mathrm{Vol}\left(\{u \geq s\}\right)^{\frac{n-1}{n}} ds\right)^{\frac{n}{n-1}} \quad \text{and} \quad \beta(t) = \frac{n}{n-1}\int_0^t \mathrm{Vol}(\{u \geq s\}) s^{\frac{1}{n-1}}\, ds.$$

We want to show $\alpha(t) \geq \beta(t)$ for all t. Clearly, $\alpha(0) = 0 = \beta(0)$. Now, we have

$$\begin{aligned} \alpha'(t) &= \frac{n}{n-1}\mathrm{Vol}(\{u \geq t\})^{\frac{n-1}{n}} \left(\int_0^t \mathrm{Vol}(\{u \geq s\})^{\frac{n-1}{n}}\, ds\right)^{\frac{1}{n-1}}, \\ \beta'(t) &= \frac{n}{n-1}\mathrm{Vol}(\{u \geq t\}) t^{\frac{1}{n-1}}. \end{aligned}$$

Thus, $\alpha'(t) \geq \beta'(t)$ if and only if

$$\int_0^t \mathrm{Vol}(\{u \geq s\})^{\frac{n-1}{n}} ds \geq \mathrm{Vol}(\{u \geq t\})^{\frac{n-1}{n}} t,$$

but this is always true, since $\mathrm{Vol}(\{u \geq s\})$ is clearly non-increasing and

$$\int_0^t \mathrm{Vol}(\{u \geq s\})^{\frac{n-1}{n}} ds \geq \int_0^t \mathrm{Vol}(\{u \geq t\})^{\frac{n-1}{n}} ds = \mathrm{Vol}(\{u \geq t\})^{\frac{n-1}{n}} t.$$

Therefore, $\alpha(t) \geq \beta(t)$ and with $t \to \infty$, we are done. $\square$

Bibliography

[1] M. Breger. *Geometry I, II.* Springer-Verlag, 1987. Berlin.

[2] Y. Brenier. Décomposition polaire et réarrangement monotone des champs de vecteurs. *C. R. Acad. Sci. Paris, Série I*, 305:805–808, 1987.

[3] X. Cabré. Elliptic PDEs in Probability and Geometry. Symmetry and regularity of solutions. *Discrete Contin. Dyn. Syst.*, 20:425–457, 2008.

[4] X. Cabré. Isoperimetric, Sobolev, and eigen value inequalities via the ABP method: a survey. *To appear. Chin. An. Math.*, 38B(1), 2017.

[5] X. Cabré, X. Ros-Oton, and J. Serra. Sharp isoperimetric inequalities via the ABP method. *Journal of the European Mathematical Society*, 18:2971–2998, 2016.

[6] L. Caffarelli, M. Feldman, and R. J. McCann. Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs. *J. Amer. Math. Soc.*, 15:1–26, 2002.

[7] I. Chavel. *Riemannian Geometry, a Modern Introduction.* Cambridge University Press, 1993. Cambridge.

[8] I. Chavel. Isoperimetric inequalities. differential geometric and analytic perspectives. *Cambridge Tracts in Math., Cambridge University Press*, 145, 2001. Cambridge.

[9] D. Cordero-Erausquin, B. Nazaret, and C. Villani. A Mass-transportation approach to sharp Sobolev and Gagliardo–Nirenberg Inequalities. *Advances in Mathematics*, 182(2):307–332, 2004.

[10] L. C. Evans. Partial Differential Equations. *American Mathematical Society*, 19, 1998. Providence, RI.

[11] H. Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93(3):418–419, 1959.

[12] C. Gutierrez. The Monge-Ampère equation. *Progress in Nonlinear Differential Equations and their Applications*, 44:xii+127, 2001. Boston, MA.

[13] L. V. Kantorovich. On the translocation of masses. *C.R. (Dokl.) Acad. Sci. URSS*, 37:199–201, 1942.

[14] R. J. McCann and N. Guillen. Five lectures on optimal transportation: geometry, regularity and applications. *American Mathematical Society*, 56:145–180, 2010. Lecture Notes.

[15] G. Monge. Mémoire sur la théorie des déblais et des remblais. *Histoire de l'Académie Royake des Sciences de Paris*, pages 75–81, 1781.

[16] G. De Philippis and A. Figalli. The Monge-Ampère equation and its link to optimal transportation. *Bull. Amer. Math. Soc. (N.S.)*, 51(4):527–580, 2014.

[17] N. Trudinger and X.-J. Wang. On the Monge mass transfer problem. *Calculus of Variations and Partial Differential Equations*, 13:19–31, 2001.

[18] E. Tsukerman. Isoperimetric Inequalities and the Alexandrov Theorem. Master's thesis, Stanford University, 2013.

[19] C. Villani. *Topics in Optimal Transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, 2003. Providence, RI.