Interplay between curvature and isoperimetry: A relationship through functional inequalities

Joaquín Sánchez García

Born 13th of November 1992 in Mexico City, Mexico.

February 15, 2018

Master's Thesis Mathematics

Advisor: Dr. Matthias Erbar

Second Advisor: Prof. Dr. Andreas Eberle

Mathematical Institute

Mathematisch-Naturwissenschaftliche Fakultät der

Rheinischen Friedrich-Wilhelms-Universität Bonn

Contents

Chapter 0

Introduction and chapter breakdown

Recent papers [14](by M. Ledoux) and [16] by (E. Milman) explain the conditions on which one can recover isoperimetry from concentration inequalities under some curvature bounds, this implication is not trivial and requires care, this thesis aims to understand and explain this proposition.

The relationship between functional inequalities and isoperimetric type inequalities is well-known, this thesis explores several papers addressing it and also the statement: under certain conditions one can derive isoperimetry from concentration by using functional inequalities.
Evidently functional inequalities are highly related to curvature bounds, and therefore this work explores the interplay between isoperimetry and curvature.

The main idea is that curvature bounds allow us to obtain quadratic (Poincaré) and entropic (Log-Sobolev) inequalities. This entails that the geometry of the underlying space conditions the behaviour of our functions in terms of the semigroup. This inequalities are strong enough to imply some Lipschitz bounds on functions, which are fundamental tool with which we deal with concentration inequalities.
One of the keys for the topic is understanding that the functional inequalities explain how variance and entropy change with time. This idea (translated to Lipschitz bounds of functions) already establishes the behaviour of the concentration of measure. Concentration conditions relate to big enlargements of sets, while isoperimetry relates to small enlargements, therefore the results on Chapters 3 and 4 are very surprising.
The first chapter is dedicated to the background, making the next two chapters readable for everyone, it focuses on giving most definitions and a clear understanding of the tools needed to explore the topic. The content of this chapter is a compact way of exposing many of the ideas found in [3]. Also, this chapter summarizes the content of the lecture notes taken (by the author) in the course 'Adv. Topics in stochastic Analysis' taught on Summer Semester 2017 by Maria Gordina at the University of Bonn.

The second chapter exposes many of the known ideas of the first implication (how to use isoperimetry to get functional inequalities), it gives ideas of the common techniques and also details some not so known arguments, for example using the displacement of convexity to prove potential inequalities and use them to obtain Brunn-Minkowski.

The third chapter of the thesis focuses on the before mentioned paper [14] which uses the semigroup technique to almost obtain linear isoperimetry, nevertheless this is still an incomplete proof as to complete it one needs to take advantage of the geometric knowledge of the isoperimetric profile.
The fourth chapter is dedicated to understand the implications of the geometry of the underlying space on the isoperimetric profile, by using a generalized volume comparison theorem, we can understand properties weaker than the concavity of the isoperimetric profile but strong enough to finish the proof on chapter 3.

Chapter 1

Background

1.1 Basic definitions and Markov semigroups

In this chapter we write down all necessary tools for the development of the next chapters, focusing on definitions and not entirely on proofs. We talk about the elements needed to analyze diffusions on curved spaces.
Most of this definitions are found in [3]. We only include proofs that are usually not included in the literature. We start by defining semigroups and Markov semigroups which intuitively are just semigroups defined on the space of linear operators (the dual space in the case where such functions form a Banach space) of functions, meaning that they are functions on functions with an associative operation (composition).

As a first setting (will only change if stated) we start with a measurable space (X, S) and a set of real valued functions defined in this space, which will usually be either assumed to be a Banach space or just an algebra of functions. We denote this set by V.

Definition. *(Semigroup of operators)*

A family of operators $\{P_t\}_{t\geq 0}$ defined on some set V of real functions on (X, S) is called a semigroup if it satisfies:

- *For every t, P_t is a linear operator, sending bounded measurable functions to bounded measurable functions.*
- *$P_0 = I$ (Initial condition)*
- *$P_{t+s} = P_t \circ P_s$ (Semigroup property)*

The first condition is needed to ensure that things "don't explode", meaning that our operators are in fact reasonable operators. Preserving boundness is fundamental for the theory of operators as unbounded operators are studied with different techniques.
The second and third condition, resemble the fundamental properties of the exponential function ($e^0 = 1, e^{x+y} = e^x e^y$) this is of course no coincidence and will be clear later, nevertheless for intuition it is good to keep in mind this similarities.
This definition is still too general, so we can ask for more properties on $\{P_t\}$ and also on the set of functions, typically we put $V = L^2(X, \mu)$ where μ is a probability measure with specific properties defined throughout this chapter.

Definition. *(Positivity preserving)*
A semigroup as above is said to be positivity preserving if

$$f \geq 0 \text{ implies } P_t f \geq 0 \text{ for all } t.$$

Definition. *(Conservative)*
A semigroup as above is said to be conservative if $\mathbf{1} \in V$

$$P_t \mathbf{1} = \mathbf{1} \text{ for all } t.$$

where $\mathbf{1}$ *denotes* $\mathbf{1}(x) = 1\, \forall x$

Proposition 1.1.1. *(Jensen's inequality for semigroups)*
If P_t *is a* ***conservative*** *and* ***positivity preserving*** *semigroup, let* $\varphi \in V$ *be a convex function then*

$$\varphi(P_t f) \leq P_t \varphi(f)$$

We include the proof of the proposition because even though it is standard, because it is usually not done in the literature given the fact that it needs the concept of subderivative. For details of existence of subdifferentials, see [18].

Proof. For any x, define $x_0 = P_t(f)(x)$.
By convexity of φ and because we only consider finite functions, the subdifferential is non-empty everywhere, so we can choose a and b such that

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

$$ax_0 + b = \varphi(x_0)$$

observe that the first equation also reads

$$af(x) + b \leq \varphi(f)(x)$$

then because our semigroup is positivity preserving, in the first equation we get:

$$P_t(af(x) + b) \leq P_t(\varphi(f))(x)$$

and using linearity and using that it is conservative on the left hand-side we get:

$$\varphi(P_t(f(x)) = \varphi(x_0) = ax_0 + b = aP_t(f(x)) + b = aP_t(f(x)) + bP_t(\mathbf{1}) \leq P_t(\varphi(f))(x)$$

As x was arbitrary, we get

$$\varphi(P_t f) \leq P_t \varphi(f)$$

Before we go on to do analysis (which we'll need to do estimates) we shall introduce continuity concepts. Let us assume that V is a Banach space.

Definition. *(Strongly continuous)*
A semigroup of linear operators is called strongly continuous or C_0 *continuous if*

$$\lim_{t \downarrow 0} P_t f - f = 0, \text{ for every } f \in V$$

The definition of C_0 continuous uses the norm of V.
In other words, this means that $\forall f \in V$

$$||P_t f - f||_V \to 0 \text{ as } t \to 0^+$$

We state the usual definition of uniformly continuous which in contrast refers to the operator norm.

Observation 1.1.2. *Throughout this work, Markov Triples are assumed to be associated with positivity preserving, conservative strongly continuous semigroups.*

Definition. *(Uniformly Continuous)*
A semigroup of linear operators is called uniformly continuous if

$$\lim_{t \downarrow 0} ||P_t - I||_* = 0$$

where $|| \cdot ||_*$ *denotes the operator norm.*

Then every uniformly continuous semigroup is also strongly continuous.
Until now, we have used only a measurable space, from this moment on we will also talk about **measures**. In this sense we need a special measure, relating to our semigroup. We can use the idea of stationary distributions of stochastic processes in a more general sense.

Definition. *(Invariant measure)*
A measure μ*, is said to be invariant with respect to a semigroup of linear operators* $\{P_t\}_{t \geq 0}$ *if*

$$\int P_t f d\mu = \int f d\mu$$

for all bounded measurable f*.*

In the following we denote: $B_b = \{f : X \to \mathbb{R} | f \text{ is measurable and bounded}\}$

Definition. *(Symmetric measure)*
An invariant measure μ *for a semigroup* $\{P_t\}$ *is called symmetric if*

$$\int f P_t g \, d\mu = \int g P_t f \, d\mu \text{ for all } f, g \in L^2(X, \mu)$$

Observation 1.1.3. *Until now we have defined semigroups in some arbitrary set. Now that we have measures we would like to define them in* $L^2(X, \mu)$ *because it is a Hilbert space. So from now on we put* $V = L^2(X, \mu)$ *semigroups or semigroups defined on a dense set of* $L^2(X, \mu)$ *for which we use the notion of a core.*

Definition. *(Markov semigroup)*
A positivity-preserving, conservative semigroup on B_b *with invariant measure* μ *is said to be a Markov semigroup if it is strongly continuous on* $L^2(X, \mu)$

The first example of a Markov semigroup with respect to an invariant probability measure is the Ornstein-Uhlenbeck semigroup, for details [3] pg. 103.

1.2 Infinitesimal Generators

Definition. *(Infinitesimal Generator and it's domain)*
For a semigroup $\{P_t\}_{t\geq 0}$ over $L^2(X,\mu)$, we define the infinitesimal generator as the pair $(L,\mathcal{D}_L)$, where

$$\mathcal{D}_L = \left\{ f \in L^2(X,.\mu) : \lim_{t\downarrow 0} \frac{Ptf - f}{t} \text{ exists} \right\} \text{ and } L(f) = \lim_{t\downarrow 0} \frac{Ptf - f}{t} \text{ for } f \in \mathcal{D}_L$$

In this context it is necessary to define the infinitesimal generator L together with it's domain $\mathcal{D}_L$ because it is usual to occur that this limit is unbounded and therefore we can't treat it with our theory of operators. At this moment we know how to define operators (infinitesimal generators) given a semigroup, we ask ourselves if the reverse statement holds. When can we define semigroups from operators? The answer is the content of the next theorem.

Theorem 1.2.1. *(Hille-Yosida)*
Let L be a linear operator on a subspace $\mathcal{D}_L$ of a Banach space, then L is the generator of a contraction semigroup if and only if

- *$\mathcal{D}_L$ is a dense subspace of $L^2(X,\mu)$.*
- *every positive λ belongs to the resolvent of L and $||(\lambda I - L)^{-1}|| \leq \frac{1}{\lambda}$.*

The proof can be found in [21] pg. 237

This theorem is fundamental as it shows the generality of the concept, the two conditions on operators aren't difficult to meet, showing that we can find a wide spread of semigroups starting with these operators.

Observation 1.2.2. *By means of the Hille-Yosida theorem one can define a semigroup by first defining an operator L with a dense domain $\mathcal{D}_L$, this translates to the fact that P_t is uniquely characterized by $(L,\mathcal{D}_L$, for this matter we define a core.*

Definition. *(Core)*
A subset D_0 of the domain $\mathcal{D}_L$ of an operator L is said to be a core if it is dense (with respect to the topology of the domain) in the domain.

As $\mathcal{D}_L$ may be complicated to describe, we focus on understanding L in a core. In this sense we say that L is densely defined and use Hille-Yosida to extend it to the complete space. Typically smooth functions or compactly supported smooth functions are the chosen cores.

Properties 1.2.3. *(of infinitesimal generators) [3] pg. 19-21*
Let P_t a strongly continuous semigroup on $L^2(X,\mu)$, with infinitesimal generator $(L,\mathcal{D}_L)$ then

- *$LP_t = P_tL$ on $\mathcal{D}_L$*
- *$\frac{d}{dt}P_tf = LP_tf$ on $\mathcal{D}_L$*
- *L is a closed operator.*
- *The semigroup is uniquely determined by it's generator.*

Proof and further properties are also found on [3] pg. 19-21.

1.3 Carré du Champ

Let $\mathcal{A}$ be an algebra of $\mathcal{D}_L$ functions.

Definition. *(Carré du Champ operator)*
We define the Carré du Champ operator, $\Gamma : \mathcal{A} \times \mathcal{A} \to \mathbb{R}$, of a Markov semigroup $\{P_t\}$ as follows:

$$\Gamma(f,g) = \frac{1}{2}(L(fg) - fLg - gLf)$$

It is obvious from the definition that the operator is bilinear and symmetric.
Often we are interested in $\Gamma(f,f)$ which is usually denoted by $\Gamma(f)$.

Observation 1.3.1. *(Domain of Γ)*
Note that at this point the definition of Γ depends on the algebra $\mathcal{A}$, we need to specify the largest domain in which we can define the Carré Du Champ. We will require a lot of properties from this set so that the operator is useful.

Definition. *(Extended algebra)*
Let $\mathcal{A}$ be an algebra, we say that $\hat{\mathcal{A}}$ is an extended algebra of $\mathcal{A}$ if it satisfies

1. $f \in \hat{\mathcal{A}}, h \in \mathcal{A} \Rightarrow hf \in \mathcal{A}$
2. $f \in \hat{\mathcal{A}},\ \forall h \in \int fh d\mu \geq 0 \Rightarrow f \geq 0$
3. $\forall f \in \hat{\mathcal{A}}\ \ \forall \varphi \in \mathcal{C}^2, \varphi(0) = 0,\ \ \varphi \circ f \in \hat{\mathcal{A}}$
4. *L (and Γ) can be extended from $\mathcal{A}$ to $\hat{\mathcal{A}}$*
5. $\Gamma(f) \geq 0\ \forall f \in \hat{\mathcal{A}}$
6. *The diffusion property (1.1) holds for (L, Γ) in $\hat{\mathcal{A}}$*
7. *If $f \in \hat{\mathcal{A}}$ and $g \in \mathcal{A}$,* $\int \Gamma(f,g) d\mu = -\int fLg d\mu = -\int gLf d\mu$
8. $f \in \mathcal{A} \Rightarrow P_t f \in \hat{\mathcal{A}}$

Observation 1.3.2. *Basically, $\hat{\mathcal{A}}$ allows us to evaluate Γ and L not only in the original functions but also on smooth transformations of the functions and on the semigroup. Further we will need to estimate functions like $\Gamma(\sqrt{f})$ but $\sqrt{\cdot}$ is not differentiable at zero, in this cases we approximate by using Lipschitz functions like $\phi_\epsilon(x) = \sqrt{x+\epsilon}$ justifying being able to write terms of type $\Gamma(\sqrt{f})$. This type of arguments need to be totally detailed at least once, we do a similar argument in 3.1.*

Observation 1.3.3. *Note that we haven't (and we won't) require indicator functions to be in the extended algebra. This means that in general we can not evaluate Γ on indicator functions.*

Definition. *(Associated Dirichlet form)*
In the same setting, whenever we have defined Γ (the Carré du Champ) we define it's associated Dirichlet form to be the function $\mathcal{E} : \mathcal{A} \to \mathbb{R}$

$$\mathcal{E}(f) = \int \Gamma(f) d\mu$$

The Dirichlet form can be defined in a set larger than $\mathcal{A}$, this set is named the domain of the Dirichlet form

Definition. *(Domain of a Dirichlet form)*

$$\mathcal{D}(\mathcal{E}) = \left\{ f \in L^2(X,\mu) : \lim_{t\downarrow 0} \frac{1}{t} \int f(f - P_t f) d\mu \ \textit{exists} \right\}$$

Properties 1.3.4. *(Basic properties of Carré du Champ)*
$\{P_t\}$ a Markov semigroup with generator $(L, \mathcal{D}_L)$ and a Carré du champ defined in an algebra of functions $\mathcal{A} \subseteq \mathcal{D}_L$ then

- $\Gamma(f) \geq 0$ *for all* $f \in \mathcal{A}$
- $\Gamma(f,g)^2 \leq \Gamma(f)\Gamma(g)$ *for all* $f, g \in \mathcal{A}$
- $\mathbf{1} \in L^2 \Rightarrow L\mathbf{1} = 0$
- $f \in L^1 \cap \mathcal{D}_L$ *and* μ *finite, then* $\displaystyle\int Lf d\mu = 0$
- $\displaystyle\int \Gamma(f,g) d\mu = -\frac{1}{2}\int fLg + gLf d\mu$

The proofs are found on [3] chapter 1.

Observation 1.3.5. *This properties show the fundamental ideas behind Carré du Champ, they show that one obtains properties that resemble normed spaces (as the first two) while it also shows properties similar to the ones of differentiability (the last 3).*

Definition. *(Markov triple)*
Let $\{P_t\}$ be a Markov semigroup, with invariant measure μ and let Γ be it's associated Carré du Champ operator, then (X, μ, Γ) is called a Markov triple.

Definition. *(Intrinsic distance induced by a Markov Triple)*
For a Markov Triple (X, Γ, μ) we define it's induced intrinsic distance by $d : X \times X \to \mathbb{R}_+$

$$d(x,y) = \sup_{\Gamma(f)\leq 1} \{|f(x) - f(y)|\}$$

Definition. *(Diffusion operator)*
The pair (L, Γ) is considered a diffusion if $\forall f \in \hat{\mathcal{A}}$ $\forall \varphi \in \mathcal{C}^2, \varphi(0) = 0$,

$$L(\varphi \circ f) = \varphi'(f)L(f) + \varphi''(f)\Gamma(f) \tag{1.1}$$

Note that we are able to evaluate in this composition by property (3) of the extended algebra.
If the pair (L, Γ) is a diffusion, we say that the Markov Triple associated is a **Diffusion Markov Triple.**

Definition. *(Lipschitz with respect to Γ)*
In a Diffusion Markov Triple with extended algebra $\hat{\mathcal{A}}$ we say that a function f is Lipschitz with respect to Γ if $f \in \hat{\mathcal{A}}$ and $f \in L^\infty(\mu)$.
In this case we call $||\Gamma(f)||^{1/2}_{L^\infty(\mu)}$ it's Lipschitz constant.

Observation 1.3.6. *Note that for the concept of Diffusion Markov Triple to make sense, we need to be able to evaluate under $\mathcal{C}^2$ compositions. This means that we need to assume that if $f \in \mathcal{A}$ then $\forall \varphi \in \mathcal{C}^2$ such that $\varphi(0) = 0$ we have $\varphi \circ f \in \mathcal{A}$*

Definition. *(Ergodic)*
A Markov triple (X, Γ, μ) *is said to be ergodic if* $\forall f \in \mathcal{D}(L)$

$$Lf = 0 \Rightarrow f \text{ is constant}$$

Proposition 1.3.7. *(Chain rule)*
Let L *be a generator associated to a Carre du Champ operator* Γ *then if* $f, g \in \mathcal{A}$ *we have*

$$\Gamma(fg, h) = f\Gamma(g, h) + g\Gamma(f, h)$$

And furthermore,

$$\Gamma(\varphi(f), g) = (\varphi(f))'\Gamma(f, g)$$
$$\Gamma(\varphi(f)) = (\varphi f')^2\Gamma(f)$$

Proof. This follows from the diffusion property of (L, Γ), see [3] pg. 43-45.

Observation 1.3.8. *To interpret* Γ *correctly, we understand it as differential operator* ***in each variable****.*

Examples 1.3.9. *(Basic examples)*
Suppose that $f \in \mathcal{A}$*, and that* $P_t f \in \mathcal{A}$ $\forall t > 0$*, suppose also that* $\mathcal{A}$ *is stable under* $\mathcal{C}^2$ *compositions as explained in 1.3.6, then*

- $\Gamma(\sqrt{f}) = \left(\frac{1}{2\sqrt{f}}\right)^2 \Gamma(f) = \frac{1}{4}\frac{\Gamma(f)}{f}$
- $\Gamma\left(\log\left(\frac{1}{P_t f}\right)\right) = \frac{1}{\left(\frac{1}{P_t f}\right)^2}\Gamma\left(\frac{1}{P_t f}\right) = (P_t f)^2\Gamma\left(\frac{1}{P_t f}\right)$
- $\Gamma\left(\frac{1}{P_t f}\right) = \frac{1}{(P_t f)^4}\Gamma(P_t f)$

Observation 1.3.10. *We can combine the 3 basic examples to calculate an specific composition that will be needed in Chapter 4:*

$$\Gamma\left(\sqrt{\log\frac{1}{P_t f}}\right) = \frac{1}{4}\frac{1}{\log\frac{1}{P_t f}}\frac{(P_t f)^2}{(P_t f)^4}\Gamma(P_t f) = \frac{1}{4}\frac{1}{\log\frac{1}{P_t f}}\frac{1}{(P_t f)^2}\Gamma(P_t f)$$

1.3.1 Iteration of Carré du Champ

In this section, it will be clear how the Carré du Champ operator works. We start by rewriting the definition of the Carré du Champ.
With the same assumptions as before, **let** $\Gamma_0(f, g) = fg$ (the product of the functions), then the Carré du Champ Γ_1 is the operator:

$$\Gamma_1(f, g) = L(\Gamma_0(f, g)) - \Gamma_0(f, Lg) - \Gamma_0(Lf, g)$$

This rewriting teaches us how to generate more operators in the same space, by repeating this procedure, if we restrict our domain to be L-stable, meaning that if $f \in \hat{\mathcal{A}}$ then $Lf \in \hat{\mathcal{A}}$. For an integer k we write:

$$\Gamma_{k+1} = L\Gamma_k(f, g) - \Gamma_k(f, Lg) - \Gamma_k(Lf, g)$$

1.3.2 Generalized Carré du Champ

In many contexts, such as sub-Riemannian manifolds it is usual to define a general version of the Carré du Champ defined in terms of local coordinates or other functions. For example if $Z = [X, Y]$ (Lie bracket), one can define

$$\Gamma_0^Z(f, g) = (Zf)(Zg)$$

and consequently

$$\Gamma_{k+1}^Z(f, g) = L\Gamma_k^Z(f, g) - \Gamma_k^Z(f, Lg) - \Gamma_k^Z(Lf, g)$$

This idea makes clear the infinite possibilities of generalizations one can have. The one stated here is necessary to understand curvature bounds on the Heisenberg group on $\mathbb{R}^3$

1.4 Curvature Dimension Inequality

In this section we present the Curvature Dimension Inequality (sometimes referred to as curvature dimension condition), and explain it's origins. The use and impact of it will be clear as the reader goes through the next chapters.

Definition. *(CD(k, n))*
We say that the curvature dimension inequality is satisfied for the pair (k, n) (or simpler CD(k, n)) if

$$\Gamma_2(f) \geq k\Gamma_1(f) + \frac{1}{n}(Lf)^2$$

The name curvature dimension inequality comes from obtaining lower bounds for Γ_2 using Bochner-Weitzenböck identity, and relating the first part to a curvature bound relating the Ricci curvature and the dimension coming from a simple inequality in terms of the HS norm of the Hessian. For details the reference is again [3] pg 70-72.

Definition. *CD(k, ∞)*
As before, we say that $CD(k, \infty)$ is satisfied if

$$\Gamma_2(f) \geq k\Gamma_1(f)$$

Theorem 1.4.1. *(Bochner-Lichnerowicz formula)*
For any smooth function on a Riemannian manifold (M, g)

$$\frac{1}{2}\Delta_g(|\nabla f|^2) = \nabla f \cdot \nabla(\Delta_g f) + |\nabla\nabla f|^2 + \mathrm{Ricc}_g(\nabla f, \nabla f)$$

where Δ_g denotes the Laplace-Beltrami operator and Ricc_g denotes the Ricci curvature.

Observation 1.4.2. *A useful stronger formulation for the curvature dimension CD(k, ∞) inequality is*

$$4\Gamma(f)\left(\Gamma_2(f) - k\Gamma(f)\right) \geq \Gamma(\Gamma(f))$$

For reference, see [3] (C.6.4) pg 515.

1.5 Markov Triples on compact Riemannian Manifolds

This section is fundamental to the next chapters as it shows how the two apparently separated parts of this first chapter blend together.
Our setting will be a smooth connected Riemannian manifold M.

For $W \in \mathcal{C}^\infty(M)$, with finite integral, assume without loss of generality that

$$\int_M e^{-W} dV = 1$$

(change W by a constant if not).
We can use e^{-W} as a Radon-Nikodyn derivative and get a probability measure μ that is absolutely continuous with respect to the standard volume, i.e.

$$\mu(A) = \int_M \mathbf{1}_A e^{-W} dV$$

In the sense of 1.3, if M is a compact Riemannian manifold we write $\mathcal{A} = \mathcal{C}^\infty(M)$ but if M is assumed to be a Riemannian manifold (not nec. compact) we put $\mathcal{A} = \mathcal{C}_c^\infty(M)$, $\hat{\mathcal{A}} = \mathcal{C}^\infty(M)$ and we can write

$$\Gamma(f, g) = \langle \nabla f, \nabla g \rangle$$

.

Now it is clear how we can try to understand functional inequalities in Riemannian manifolds, by understanding the previously defined Carré du Champ. And throughout this thesis the connection between curvature and functional inequalities will mostly depend on Γ.

Definition. *(Density of a Riemannian manifold)*
In the setting as above we say that M has a smooth density e^{-W}

In many cases, as in chapter 4 we call $\Psi = e^{-W}$ and say that the density of the manifold is Ψ, sometimes manifolds have non-smooth densities but we won't focus on those cases.

Corollary 1.5.1. *of 1.4.1 In our setting, of compact Riemannian manifolds,*

$$\Gamma_2(f) = |\nabla\nabla f|^2 + \mathrm{Ricc}_g(\nabla f, \nabla f)$$

Observation 1.5.2. *This corollary shows clearly the relation between the geometry of the space and the Carré du Champ and it's iterations.*

1.6 Assumptions on a Markov Triple

Definition. *(Standard Markov Triple)*
A Markov triple (X, Γ, μ) is said to be standard if it is

1. *A diffusion markov triple in the sense of 1.1*
2. *Ergodic*
3. *Conservative*

Until now, the theory developed has been very general. The final framework will be the full markov triple:

Definition. *(Full Markov Triple)*
A full Markov Triple is a Standard Markov triple in which we additionally require:

- Γ *(and also L) are defined on an extended algebra, $\hat{\mathcal{A}}$ as in 1.3*
- *If $f \in \hat{\mathcal{A}}$ is such that $\Gamma(f) = 0$ then f must be constant.*
- *L (defined originally in $\mathcal{A}$ not in $\hat{\mathcal{A}}$) has a unique self-adjoint extension.*

From now on, when we mention Markov Triples we will be referring to Full Markov Triples.

Chapter 2

Functional inequalities

2.1 A non-exhausting list of functional inequalities

A functional inequality is exactly what it's name expresses, an inequality on functions. This "definition" may seem a bit vague but already shows that by the name "functional inequalities" we refer to an inequality on functions.
The problem with this idea is that almost every statement in the field of mathematical analysis can be understood as a functional inequality, therefore the aim of this part of the work is to make a list containing a lot of interesting inequalities that can be derived from curvature bounds or similar hypothesis. We aim to explore how the geometry of a space ensures that functions behave in a certain way.

Even though studying this property is interesting by itself, we dedicate a big part of this work to understanding how and where each inequality is relevant, this broadens the reach of this work.
In the same fashion, we analyze how specific geometric properties of a set conditions functions. One of the main tools for such analysis is the Carré du Champ operator presented in the background (1.3).

The study of ways to obtain functional inequalities is an on-going field of research, for example [20] shows how to use tools from optimal transport to generalize functional inequalities which are of high interest in physics. Also, understanding what one can derive from functional inequalities (done in [14] and [16]) is recent and explained in the next chapter.

2.1.1 Sobolev inequality

Theorem 2.1.1. *(Sobolev inequality in $\mathbb{R}^n$)*
Let f be a smooth, compactly supported function $f : \mathbb{R}^n \to \mathbb{R}$, $n \geq 3$ then it holds

$$\left(\int |f|^p dx\right)^{2/p} \leq c_n \int |\nabla f|^2 dx$$

where $p = \dfrac{2n}{n-2}$ *and the sharp constant* $c_n = \dfrac{1}{\pi n(n-2)} \left(\dfrac{\Gamma(n)}{\Gamma(\frac{n}{2})}\right)^{2/m}$

The sharp constant can be obtained with $f(x) = (a + b|x|^2)^{(1-n)/2}$.

2.1.2 Sobolev inequality in a Manifold

Definition. *(Sobolev inequality in a Riemannian Manifold of dimension $n \geq 2$)*
We say that the Sobolev inequality with constant c is satisfied if for all smooth f, compactly supported function $f : M \to \mathbb{R}$, then it holds

$$\left(\int_M |f|^{p^*}\right)^{1/p^*} \leq c\int_M |\nabla f|$$

where $p^* = \dfrac{n}{n-1}$ the integration is done with respect to the standard volume,

Applications of the Sobolev inequality

Sobolev type inequalities are the fundamental basis for the theory of PDE's. It allows to analyze regularity of PDE's and the L^p embeddings. This tools are crucial in the study of differential equations, as the Sobolev spaces are the starting point in this analysis, thus Sobolev inequalities are one of the main goals when studying differential equations.
The next section is devoted to the Log-Sobolev inequality which is a functional inequality derived from the Sobolev inequality.

2.1.3 Log-Sobolev inequality

Using the Sobolev inequality one can derive a different inequality, which is fundamental for information theory.

Starting from the Sobolev inequality

$$\left(\int |f|^p dx\right)^{2/p} \leq c_n \int |\nabla f|^2 dx$$

By taking logarithms we have,

$$\frac{2}{p}\log\left(\int |f|^p dx\right) \leq \log\left(c_n\int |\nabla f|^2 dx\right)$$

Let us assume $\int |f|^2 dx = 1$ (normalized variance), then by Jensen's inequality for the measure $f^2 dx$ we have

$$\log\left(\int |f|^p dx\right) = \log\left(\int |f|^{p-2} f^2 dx\right) \geq \int \log(|f|^{p-2}) f^2 dx = \frac{p-2}{2}\int f^2 \log f^2 dx$$

So, finally we get a first version of what we call Log-Sobolev inequality

$$\int f^2 \log f^2 dx \leq \frac{n}{2}\log\left(c_n \int |\nabla f|^2 dx\right), \int |f|^2 dx = 1$$

To express it in another form we write explicitly $d\mu = f^2 dx$ and so we get

$$\int \frac{d\mu}{dx}\log\left(\frac{d\mu}{dx}\right) dx = \int \log\left(\frac{d\mu}{dx}\right) d\mu \leq \frac{n}{2}\log\left(c_n\int \left|\nabla\sqrt{\frac{d\mu}{dx}}\right|^2 dx\right)$$

2.1.4 Rewriting of the Log-Sobolev inequality and Fisher's info

Functional inequalities can also be expressed in terms of probability measures. Because $f(x)dx$ is usually a measure. So for us to be able to use more advanced tools (such as optimal transport) we need to rewrite the functional inequalities in terms of measures. This rewriting allows us to understand the same inequalities in different spaces.
Let μ and ν be two probability measures absolutely continuous with respect to the standard volume.

Definition. *(Relative entropy)*
For μ and ν such measures we define the relative entropy as

$$H(\mu|\nu) = \int \log\left(\frac{d\mu}{d\nu}\right) d\mu$$

Observation 2.1.2. *The relative entropy, also called Kullback-Leibler divergence measures how much one measure differs from the other.*

Definition. *(Fisher's Information)*
For μ and ν measures as below, we define their Fisher's Information as

$$I(\mu|\nu) = \int \left|\nabla \log\left(\frac{d\mu}{d\nu}\right)\right|^2 d\mu = 4\int \left|\nabla\sqrt{\frac{d\mu}{d\nu}}\right|^2 d\nu$$

Intentionally the domain of integration is not explicitly written, in both definitions the integral is over the whole space, whether it is $\mathbb{R}^n$ or a Banach space X or a manifold M.

Observation 2.1.3. *This quantity is not defined if the Radon-Nikodym derivative is not differentiable, in this cases one must use approximations.*

Observation 2.1.4. *Fisher's information is one of the most important concepts in statistics, it is the reciprocal of the lower bound of the variance of the minimum variance unbiased estimators, by Cramer-Rao theorem, which are the fundamental blocks of statistics.*

Definition. *(Log-Sobolev inequality for measures)*
We say that the probability measure ν satisfies the Log-Sobolev inequality with constant ρ (LSI(ρ)) if ***for all*** *probability measures μ absolutely continuous with respect to ν it holds*

$$H(\mu|\nu) \leq \frac{1}{2\rho} I(\mu|\nu)$$

Now we turn our heads to the diffusions setting, in which we define

$$\text{Ent}_\mu(f) = \int f \log f d\mu - \int f d\mu \log\left(\int f d\mu\right)$$

Definition. *(Global Log-Sobolev inequality for Markov Triples)*
For a Markov Triple (X, Γ, μ) we say Log-Sobolev inequality holds with constant k if

$$\text{Ent}_\mu(f^2) \leq \frac{1}{2k}\mathcal{E}(f)$$

Definition. *(Local Log-Sobolev for Markov Triples)*
For a Markov Triple (X, Γ, μ) we say a local Log-Sobolev inequality holds with constant k if

$$P_t(f^2 \log f^2) - P_t(f^2) \log P_t(f^2) \leq 2\frac{1 - e^{-2kt}}{k} P_t(\Gamma(f)) \tag{2.1}$$

Observation 2.1.5. *Observe that $\rho = 0$ **is not** possible in the Log-Sobolev inequality, however as $e^{ct} \approx 1+ct$ near 0 by the Taylor expansion, we can rewrite a similar version of the Log-Sobolev inequality for Markov triples.*

Definition. *(Second version of local Log-Sobolev and reverse Log-Sobolev inequalities)*
For a Markov triple (X, Γ, μ) we say it satisfies Log-Sobolev and reverse Log-Sobolev inequalities if

$$t\frac{\Gamma(P_t f)}{P_t f} \leq P_t(f \log f) - P_t f \log P_t f \leq t P_t \left(\frac{\Gamma(f)}{f}\right)$$

Observation 2.1.6. *The log-Sobolev inequality immediately shows some Lipschitz properties of the function, namely if $0 \leq f \leq 1$ then:*

$$\Gamma(P_t f) \leq \frac{1}{t}(P_t f)^2 \log \frac{1}{P_t f} \leq \frac{1}{t} \log \frac{1}{P_t f}$$

To see the observation note that $P_t(f \log(f))$ is negative by positivity preserving property of P_t and so writing the reverse local Log-Sobolev inequality yields:

$$t\frac{\Gamma(P_t f)}{P_t f} \leq P_t f \log \frac{1}{P_t}$$

which is the observation.

Theorem 2.1.7. *(DE Bruijn's identity)*
In the setting for Markov triples, for every positive $f \in D(\mathcal{E})$ if one has $\int f|logf|d\mu < \infty$ then

$$\frac{d}{dt}\left(\int P_t f \log P_t f d\mu\right) = -\int \frac{\Gamma(P_t f)}{P_t f} d\mu$$

Observation 2.1.8. *This result is an essential part of the proof on chapter 3.*

Proof. Starting from $f \in \mathcal{A}$ and then extending to $\mathcal{D}(\mathcal{E})$ we have by hypothesis we have a majorant and we can apply dominated convergence to permute integral and derivative so we get

$$\frac{d}{dt}\left(\int P_t f \log P_t f d\mu\right) = \int (1 + \log P_t f) L P_t f d\mu$$

By integration by parts (property 7 of the extended algebra) we can rewrite the right-hand side and we have

$$\frac{d}{dt}\left(\int P_t f \log P_t f d\mu\right) = -\int \Gamma(P_t f, 1 + \log P_t f) d\mu = -\int \frac{\Gamma(P_t f)}{P_t f} d\mu$$

2.1.5 Talagrand's inequality

The first element one needs to define to get Talagrand's inequality is a way to measure distance between probability measures (aside from total variation).

Definition. *(Wasserstein distance)*
For a metric space (X, d), $p \in [1, \infty)$ and two probability measures μ, ν on X with finite order p we define the Wasserstein distance between μ and ν:

$$W_p(\mu, \nu) = \left(\inf_{\gamma \in \Gamma(\mu,\nu)} \left\{ \int_{X \times X} d(x,y)^p d\gamma \right\}\right)^{\frac{1}{p}}.$$

where $\Gamma(\mu,\nu) = \{\pi \in \mathcal{P}(X \times X : \pi_1(A) = \mu(A), \pi_2(B) = \nu(B)\}$, $\mathcal{P}$ *denotes the probability measures of a space and* π_1 *and* π_2 *are the marginals of* π.

Observation 2.1.9. *Wasserstein's distance is one of the fundamental elements of optimal transport, it shows the total cost of transport from* μ *to* ν *with respect to the distance function of the space. It is fundamental because it metrizes weak convergence.*

Definition. *(Talagrand's inequality)*
With probability measures μ *and* ν *as above, we say they satisfy Talagrand's inequality with constant* ρ, $T(\rho)$ *if*

$$W(\mu,\nu) \leq \sqrt{\frac{2H(\mu|\nu)}{\rho}}$$

By combining LSI(ρ) with $T(\rho)$ one can get another inequality.

Definition. *(LSI+T(ρ))*
For probabilities as above we say they satisfy (LSI+T)(ρ) if

$$W(\mu,\nu) \leq \frac{1}{\rho}\sqrt{I(\mu,\nu)}$$

Talagrand's inequality and it's generalizations are important to physics and the theory of optimal transport, in [20] (main theorem) it is shown that Log-Sobolev inequalities imply Talagrand's inequality.

2.1.6 Poincare inequality

Definition. *(Poincaré inequality in* $\mathbb{R}^n$*)*
Let $p \in [1,n)$, *we say that the Poincaré inequality is satisfied with constant* $C(n.p)$ *if* $\forall f \in W^{p,1}$ *and every ball* B,

$$\left(\int_B |f - f_B|^{np/(n-p)} dx\right)^{(n-p)/np} \leq C(n,p) \left(\int_B |\nabla f|^p dx\right)^{1/p}$$

where

$$f_B := \frac{1}{|B|}\int_B f dx$$

2.1.7 Poincaré inequality in terms of Markov triples

Now similarly as the Log-Sobolev inequality, here we define the variance with respect to μ instead of the entropy.

$$\mathrm{Var}_\mu(f) = \int f^2 d\mu - \left(\int f d\mu\right)^2$$

Definition. *(Global Poincaré inequality for Markov Triples)*
We say that a Markov triple (X, Γ, μ) *satisfies the Poincaré inequality with constant* k *if for every function* $f \in \mathcal{A}$

$$\mathrm{Var}_\mu(f) \leq k\mathcal{E}(f)$$

Definition. *(Local Poincaré inequality for Markov Triples)*
We say that a Markov triple satisfies the local Poincaré inequality with constant k if for every function $f \in \mathcal{A}$

$$P_t f^2 - (P_t f)^2 \leq \frac{1 - e^{-2kt}}{k} P_t(\Gamma f) \tag{2.2}$$

Definition. *(Local reverse Poincaré inequality for Markov Triples)*
We say that a Markov triple satisfies the reverse Poincaré inequality with constant k if for every function $f \in \mathcal{A}$

$$P_t f^2 - (P_t f)^2 \geq \frac{e^{2kt} - 1}{k} \Gamma(P_t f)$$

Observation 2.1.10. *In a similar fashion as 2.1.5, we can use the linear approximation of the exponential function to get another version of the Poincaré inequality when $k = 0$.*

Definition. *(Second version of the Poincaré inequality for Markov Triples)*
For a Markov triple $(X, \Gamma, , mu)$ we say it satisfies Poincaré and reverse Poincaré inequalities if

$$2t\Gamma(P_t f) \leq P_t(f^2) - (P_t f)^2 \leq 2t P_t(\Gamma(f))$$

Observation 2.1.11. *Note that directly from the reverse Poincaré inequality, one finds Lipschitz properties of the Carré du Champ operator, Γ, with respect to the semigroup: If $0 \leq f \leq 1$ then by the reverse Poincaré inequality (second version)*

$$2t\Gamma(P_t(f)) \leq P_t(f^2) - (P_t f)^2 \leq P_t(f^2) \leq P_t(f) \leq P_t(\mathbf{1})$$

And so, as the semigroup is conservative, we have $2tP_t(f) \leq 1$which of course can be rewritten in a way that it evokes Lipschitz properties:

$$\Gamma(P_t f) \leq \frac{1}{2t}$$

Proposition 2.1.12. *(Log-Sobolev inequalities imply Poincaré inequalities)*
Suppose that a Markov triple satisfies the global Log-Sobolev inequality with constant c then it satisfies the global Poincaré inequality with constant c.
In other words:

$$\text{LSI}(c) \Rightarrow \text{Poincaré}(c)$$

Sketch of proof Put $f_n = 1 + \frac{1}{n} f$ in the definition of global LSI(c) and take $n \to \infty$ considering the Taylor expansion of $\log(1 + x)$.

2.1.8 Gradient estimates

Gradient estimates as indicated by their name are ways to bound the norm of the gradient of a function in terms of the norm of the function, this means we are controlling the infinitesimal change of a function by the function.

2.1.9 Gradient estimate for Markov triples

Given a Markov triple (X, L, Γ) with L^2 domain we say that it satisfies a gradient estimate with constant k if $\forall f \in \hat{\mathcal{A}}$

$$\Gamma(P_t f) \leq e^{-2kt} P_t(\Gamma f) \tag{2.3}$$

Observation 2.1.13. *The name is understood as one realizes that in most cases Γ is similar to the norm of a gradient, for example $\Gamma(f, g) = \nabla f \cdot \nabla g$ in the Brownian motion and the O-U semigroup 1.1*

2.1.10 Stronger gradient estimate for Markov Triples

Given a Markov triple (X, L, Γ) we say that it satisfies a strong gradient estimate with constant k if $\forall f \in \hat{\mathcal{A}}$

$$\sqrt{\Gamma(P_t f)} \leq e^{-kt} P_t(\sqrt{\Gamma f}) \tag{2.4}$$

Observation 2.1.14. *Using the strong gradient estimate for $k = 0$ yields $\sqrt{\Gamma(P_t f)} \leq P_t(\sqrt{\Gamma(f)})$ which is a fundamental tool for the last chapter.*

We know introduce the concept of concentration, one of the main ingredients of the results we want to study. It explains how the probability measures enlargements of our sets.

2.1.11 Concentration inequality

Definition. *(Concentration function)*
For a metric probability space (X, d, μ) we define the concentration function as

$$\alpha_\mu(r) = \sup\left\{1 - \mu(A_r) : \mu(A) \geq \frac{1}{2}\right\}$$

Typically, concentration conditions like $\alpha_\mu(r) \to 0$ as $r \to \infty$ indicate the behaviour of our probability measure with respect to big enlargements of. The concentration function is highly related to the Lipschitz functions of the space, we see this in the following theorem.

Theorem 2.1.15. *If $f : X \to \mathbb{R}$ is a Lipschitz function with constant L then there exists M such that for all $\epsilon > 0$ one has*

$$\mu\left(\{x \in X : |f(x) - M| > \epsilon\}\right) \leq 2\alpha_\mu\left(\frac{\epsilon}{L}\right)$$

Proof. Let $A = \{x \in X : f(x) > M\}$, let us enlarge A.

If $x \in A_{\epsilon/L}$ then there exists $y \in A$ s.t. $d(x, y) < \dfrac{\epsilon}{L}$, by Lipschitz property $|f(x) - f(y)| \leq Ld(x, y) < \epsilon$. Also, $y \in A$ means $f(y) \geq M$ so we get $f(x) > M - \epsilon$.
This means that $x \in A_{\epsilon\ L}$ implies $f(x) > M - \epsilon$. In set notation this is

$$A_{\epsilon/L} \subseteq \{x : f(x) > M - \epsilon\} \text{ if and only if } \{x : f(x) \leq M - \epsilon\} \subseteq X \setminus A_{\epsilon/L}$$

So by monotonicity of μ and using that $\mu(X) = 1$,

$$\mu(\{x \in X : f(x) \leq M - \epsilon\}) \leq 1 - \mu(A_\epsilon) \leq \alpha_\mu\left(\frac{\epsilon}{L}\right)$$

Applying the same technique to $B = \{x \in X : f(x) > M\}$ we obtain

$$\mu(\{x \in X : f(x) \geq M + \epsilon\}) \leq \alpha_\mu\left(\frac{\epsilon}{L}\right)$$

But by adding these inequalities we obtain our result:

$$\mu(\{x \in X : |f(x) - M| \geq \epsilon\}) \leq 2\alpha_\mu\left(\frac{\epsilon}{L}\right)$$

Definition. *(Exponential concentration)*
We say that (X, Γ, μ) satisfies exponential concentration with constants c, C if for every integrable $1-$Lipschitz function f one has

$$\mu\left(f \geq \int f d\mu + r\right) \leq Ce^{-cr}$$

Observation 2.1.16. *We can interpret exponential concentration in a probabilistic way, it tells us according to the probability measure μ that the likeliness of f being away from it's mean decays exponentially in terms of this difference. Note also that in terms of our last theorem $M = \int f d\mu$, finding the correct median is necessary when dealing with concentration inequalities. We see how this plays a role in Chapter 3.*

Theorem 2.1.17. *(From Poincaré inequality to concentration)*
If (X, Γ, μ) satisfies the Poincaré inequality with constant C then every 1-Lipschitz function is exponentially integrable, i.e. if f is 1-Lipschitz and $s < \sqrt{\frac{4}{C}}$ then

$$\int e^{sf} d\mu < \infty$$

For a proof see [3] pg. 190, we do not do the proof here as it requires a lot of care and the focus of this work is to obtain the reverse implication.

2.2 Functional Inequalities from the curvature dimension inequality

This section explains how functional inequalities can be derived from the curvature dimension condition, this clearly shows how the geometry of the space determines if a space allows functional inequalities to hold. As the following theorem states, $CD(k, \infty)$ is equivalent to many functional inequalities, meaning that any manifold satisfying the curvature bound will have functional inequalities.

Theorem 2.2.1. *($CD(k, \infty)$ and functional inequalities) For a Markov triple (X, Γ, μ),*

a) Gradient estimates (2.3) $\Leftrightarrow CD(k, \infty)$

b) Local Poincaré (2.2) $\Leftrightarrow CD(k, \infty)$

c) Strong gradient estimates (2.4) $\Leftrightarrow CD(k, \infty)$

d) Local Log-Sobolev (2.1) $\Leftrightarrow CD(k, \infty)$

a) To see that the curvature dimension inequality implies the estimate, define $F(s) = e^{-2ks} P_s(\Gamma(P_{t-s}f))$ then by 1.3.4 we can explicitly compute the derivative, and use CD(k, ∞) to get,

$$F'(s) = 2e^{-2ks} \left(-kP_s(\Gamma(P_{t-s}f)) + P_s\Gamma_2(P_{t-s}f)\right) \geq 2e^{-2ks} \left(-kP_s(\Gamma(P_{t-s}f)) + P_s\Gamma(P_{t-s}f)\right) = 0$$

And therefore F is an increasing function. So $F(0) \leq F(t)$ which translates exactly to the gradient estimate.

To see that the gradient estimate implies the curvature dimension inequality, we need to repeat the procedure.
Define $G(t) = e^{-2kt} P_t(\Gamma f) - \Gamma(P_t f)$. Evaluating G in zero, shows $G(0) = \Gamma(f) - \Gamma(f) = 0$, our hypothesis is the gradient estimate so we know $G(t) \geq 0$ which means that $G'(0) \geq 0$, but

$$G'(0) = 2k\Gamma(f) + L\Gamma(f) - \Gamma(Lf) = \Gamma_2(f) - k\Gamma(f)$$

So we get the CD(K, ∞), $\Gamma_2(f) \geq \Gamma(f)$

b) Observe that we can express the left hand-side of the inequality in terms of an integral

$$P_t(f^2) - (P_t f)^2 = \int_0^t \frac{d}{ds} Ps((P_{t-s}f)^2) ds$$

Now it is our duty to understand the integrand. For a smooth function g

$$\frac{d}{ds}P_s(g(s)) = \lim_{h\to 0}\frac{P_{s+h}(g(s+h)) - g(s)}{h} = \lim_{h\to 0}\frac{P_{s+h}(g(s+h)) - P_{s+h}(g(s)) + P_{s+h}(g(s)) - g(s)}{h}$$

So by strong continuity of the Markov triple, we can get the limit inside the semigroup in the first part and get

$$\frac{d}{ds}P_s(g(s)) = \lim_{h\to 0}\frac{P_{s+h}(g(s+h)) - P_{s+h}(g(s))}{h} + \lim_{h\to 0}\frac{P_{s+h}(g(s)) - g(s)}{h} = P_s(g') + LP_s(g)$$

Now we use our computation, putting $g(s) = (P_{t-s}(f))^2$ by the chain rule,

$$g'(s) = -2(P_{t-s}(f))(LP_{t-s}(f))$$

And so, by using our explicit computation we have

$$\frac{d}{ds}Ps((P_{t-s}f)^2) = LP_s(P_{t-s}f)^2 - 2P_s((P_{t-s}f)LP_{t-s}f) = 2P_s(\Gamma(P_{t-s}f))$$

by definition of Γ. As we have proved $a)$ we can use the gradient estimate to get:

$$P_t(f^2) - (P_tf)^2 = 2\int_0^t P_s(\Gamma(P_{t-s}f))ds \leq 2\int_0^t P_s e^{-2k(t-s)}P_{t-s}(\Gamma f)ds$$
$$= 2\int_0^t e^{-2k(t-s)}P_t(\Gamma(f))ds = P_t(\Gamma(f))\frac{1-e^{-2kt}}{k}$$

which concludes the proof. The same reasoning is valid to prove the reverse local Poincaré inequality.

c) To prove that the curvature dimension inequality implies the strong estimate, using CD(k,∞) and the fact that L is a diffusion,

$$\Gamma_2(f) \geq k\Gamma(f) + \frac{\Gamma(\Gamma(f))}{4\Gamma(f)}$$

Define then $F(s) = P_s(\sqrt{\Gamma(P_{t-s}f)})$, this is applying the same idea as in the previous cases but with a big difference: the function is not differentiable at 0. This happens because $\sqrt{\cdot}$ is not differentiable at zero, therefore we need to approximate the function. Let $\phi_\epsilon(x) = \sqrt{x+\epsilon}$, applying the diffusion property to ϕ_ϵ we get

$$L\phi_\epsilon(g) = \frac{Lg}{2\phi_\epsilon(g)} - \frac{\Gamma(g)}{4\phi_\epsilon^3(g)}$$

Define, as before $F_\epsilon(s) = P_s\phi_\epsilon(e^{-2ks}\Gamma(P_{t-s}f))$, if we compute the derivative

$$F_\epsilon'(s) = e^{-2ks}P_s\left(\frac{\Gamma(P_{t-s}f) - k\Gamma(P_{t-s}f)}{\phi_\epsilon(e^{-2ks}\Gamma(P_{t-s}f))} - e^{-2ks}\frac{\Gamma(\Gamma(P_{t-s}f)}{4\phi_\epsilon^3(e^{-2ks}\Gamma(P_{t-s}f))}\right)$$

We can bound this from below by using the gradient estimate inequality for the last term and get

$$F_\epsilon'(s) \geq 0$$

And so F_ϵ' is increasing for $\epsilon > 0$, by letting $\epsilon \downarrow 0$ we get the result.

d) We want to prove that

$$P_t(f^2\log f^2) - P_t(f^2)\log P_t(f^2) \leq 2\frac{1-e^{-2kt}}{k}P_t(\Gamma(f))$$

We start by letting $g = f^2$, observe that we can write the left hand-side of the local Log-Sobolev inequality as an integral:

$$P_t(g \log g) - P_t(g) \log P_t(g) = \int_0^t \frac{d}{ds} P_s(P_{t-s}(g) \log P_{t-s}(g) ds$$

In the other hand, we can explicitly compute the derivative, using the diffusion property for $\varphi(x) = x \log x$, and get

$$\frac{d}{ds} P_s(P_{t-s}(g) \log P_{t-s}(g) ds = P_s \left(\frac{\Gamma(P_{t-s}(g))}{P_{t-s}(g)} \right)$$

To get the reverse statements, they are done similarly as the one in **a**).

Observation 2.2.2. *We can proof the same result, namely it's equivalence with the curvature condition, for the second versions of the inequalities (2.1.4 and 2.1.7). For example, to do the former note*

$$P_t(f^2) - (P_t f)^2 = -2 \int_0^t \frac{d}{ds} P_s((P_{t-s} f)^2) ds = 2 \int_0^t P_s(\Gamma(P_{t-s} f)) ds$$

and proceed in the same way.

2.3 The isoperimetric problem

The isoperimetric problem is well known across all fields of mathematics, in it's first version (the case of the plane) it states

"Among all the curves of fixed perimeter which (if any) is the one that encloses the biggest area?"

Even though the result is very intuitive (the circle) it is not trivial to prove. Nowadays we have many proofs and with different techniques, but recent work has shown that the isoperimetric problem is highly related to functional inequalities.

Furthermore, the problem is easily generalized to $\mathbb{R}^n$ by means of the Hausdorff measure.

"Among all the sets of fixed volume which (if any) is the one with biggest surface area?"

2.4 The exterior Minkowski boundary measure

Thee following definition is usual in analysis as a way to understand measure of boundaries. The idea is simple, take a set and ϵ-enlarge it. Compare the measure of this set with your original set and see how the rate changes.

Definition. *(Minkowski content)*
For a metric measure space (X, d, μ) we define the Minkowski boundary measure of a Borel set A as

$$\mu^+(A) = \liminf_{\epsilon \downarrow 0} \frac{\mu(A_\epsilon) - \mu(A)}{\epsilon}$$

where $A_\epsilon = \{x \in X : d(x, A) < \epsilon\}$

In a very general case we can understand the isoperimetric problem in terms of the isoperimetric profile:

2.4.1 Isoperimetric profile

Definition. *(Isoperimetric profile)*
In a measure metric space (X, d, μ), the ***isoperimetric profile*** *is the function $I : [0, \infty) \to \mathbb{R}$ given by*

$$I_\mu(x) = \inf_{A \in \mathcal{B}(X)} \{\mu^+(A) : \mu(A) = x\}$$

where $\mathcal{B}(X)$ is the Borel σ-algebra.

Observation 2.4.1. *In the case where μ is a probability measure, note that A and $X \setminus A$ are given the same boundary measure and the isoperimetric profile only takes values in $[0, 1]$, meaning that we can restrict the values to $[0, 1/2]$.*

Interpretation 2.4.2. *The Isoperimetric profile can be thought of as the minimum size of the boundary, for a set to be able to amount size x.*
From all the (Borel) measurable sets of measure x, $I_\mu(x)$ tells us what is the minimum size of it's boundary.

2.4.2 Isoperimetric type inequality

Definition. *(Isoperimetric type inequality)*
In a measure metric space (X, d, μ) an isoperimetric type inequality is said to hold with function f if

$$I_\mu(x) \geq f(x)$$

Observation 2.4.3. *In this notation, the true isoperimetry can be stated as I_μ having an actual minimizer. So there exists a set E such that $\mu^+(A) \geq I(\mu(E))$ for any set with $\mu(A) = \mu(E)$ so it can be rewritten as having $\mu^+(A) \geq I(\mu(A))$*

Example 2.4.4. *For the case when μ is the Gaussian measure in the $\mathbb{R}^n$, we have*

$$I_\mu(x) \geq \varphi(\Phi^{-1}(x))$$

where $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ and $\Phi(x) = \int_{-\infty}^{x} \varphi(s) ds$

For a proof see [3] pg. 416.

Linear isoperimetric bounds

In the case of 2.4.1 we say that **linear** isoperimetric bounds hold if there exists a constant c

$$I_\mu(x) \geq c \min\{x, 1-x\}$$

2.4.3 History

The isoperimetric problem has been one of the most intriguing problems in mathematics since the Greeks, one can find the isoperimetric problem in Greek literature where it is posed as a fencing problem. Given a fixed amount of fence how do I enclose the biggest area?. Even though many knew intuitively that the circle had to be the solution, it was not easy to get a readable proof.
It wasn't until Jakob Steiner came along as the first hero of the story.

Figure 2.1: Jakob Steiner (1796-1863)

Jakob Steiner (1796-1863) was a Swiss mathematician. Steiner is considered a leader in the field of geometry because his work was confined to this field, it is said in [7] that he even hated analysis, and thus tried to take the horizons of geometry without the support of analysis (which hadn't had it's climax yet). Even though he was born poor, he was able to attend to "Pestalozzi's school". Johann Henri Pestalozzi was a Swiss pedagogue known to have founded many education institutions in Switzerland. Steiner received an honorary degree from the university of Konigsberg thanks to Jacobi.
Most of Steiner's contributions to mathematics have been recollected in a book by the Berlin Academy.
According to [10] Steiner produced five proofs of the isoperimetric inequality, but they were disregarded by it's contemporaries due to the fact that he had assumed a solution existed. Steiner's proof was: given that there exists a closed curve maximizing the area, it must be a circle. So his proofs weren't completely accepted until more recent times.
Steiner's proofs rely on convexity, and a convexity lemma: *a maximizing curve must be the boundary of a convex set*. Today there are many ways to proof the isoperimetric inequality, one of them which we study here uses the Brunn-Minkowski's inequality.

2.4.4 Isoperimetry on $\mathbb{R}^n$ and Brunn-Minkowsky

One of the most important results on geometric measure theory is the Brunn-Minkowski's inequality and it's generalizations. In it's simplest form it shows some type of logarithmic concavity of the Lebesgue measure with respect to the Minkowski sum of sets.

Theorem 2.4.5. *(Brunn-Minkowski's inequality for the Lebesgue measure in $\mathbb{R}^n$)*
If $|\cdot|$ denotes the Lebesgue measure in $\mathbb{R}^n$, and X, Y are non-empty compact sets then

$$|X+Y|^{1/n} \geq |X|^{1/n} + |Y|^{1/n}$$

Observation 2.4.6. *The notation X and Y for the sets is done with the clear intention of resembling the usual definition of concavity of a function. Observe the* ***similarity*** *between this definition and a concave function f: $f(tx+(1-t)y) \geq tf(x)+(1-t)f(y)$, and then compare it to the log-concave one. The difficult part is that the sum operation in Brunn-Minkowski's is the Minkowski sum of set that doesn't behave as one would expect.*

The most direct way to proof Brunn-Minkowski's can be found on [4] pg. 225, and it follows from the fact that if u, v, w are Lebesgue integrable satisfying

$$w(tx+(1-t)y) \geq u(x)^t v(y)^{1-t} \qquad \forall x, y$$

then it holds

$$\int w(x)dx \geq \left(\int u(x)dx\right)^t \left(\int v(y)dy\right)^{1-t}$$

But as this proof is standard and can be found easily so we skip it and instead, show how tools from optimal transport can conclude the same results with methods that can be generalized to different spaces.

2.4.5 A way to obtain Brunn-Minkowski's inequality by means of optimal transport

What is Optimal Transport?

In 1781, french mathematician Gaspard Monge raised a problem that would intrigue the world of mathematics:

Given two subsets $U, V \subseteq \mathbb{R}^3$ with the same volume, find a volume preserving map between them, such that it minimizes some cost fixed function $c(x,y)$.

Little did everyone know this 'apparently easy' formulation would lead to centuries of work for many great scientists from a wide variety of backgrounds.

Monge's Problem

The before mentioned problem can be formulated in terms of measure spaces:
For measure spaces, (X, S_X, μ) and (Y, S_Y, ν) we aim to find:

$$\inf_{T_{\#}\mu=\nu} \left\{ \int_X c(x, Tx)d\mu \right\},$$

Where $T_{\#}\mu = \nu$ denotes that T is in fact what we will call a transport function: a *push forward* from μ to ν. This means that for any set A in S_Y, one has

$$T : X \to Y \text{ y } \mu(T^{-1}(A)) = \nu(A).$$

By this last equation one can use the change of variable formula for measures to formulate the problem in the space Y in terms of the inverse transport function T^{-1}, but in general this makes no difference.
The most interesting part of such problem is that even though it's simple formulation it turned out not to be so easy to solve. Almost 200 years passed without significant breakthroughs to Monge problem until V.N. Sudakov was able to find and prove existence conditions for the mapping T. **Why?** Mainly because the condition defining T is highly non-linear. Meaning that tools from linear analysis would not provide solutions.

Kantorovich's Problem

The real breakthrough came in the late twentieth century, when Leonid Kantorovich realized that he could formulate a different problem, easier to analyze but very related to the one of Monge.

$$\inf_{\gamma\in\Gamma(\mu,\nu)}\left\{\int_{X\times Y}c(x,y)d\gamma\right\}.$$

Where $\Gamma(\mu,\nu)=\{\gamma\in M_1(X\times Y,S):\pi_x(\gamma)=\mu,\pi_y(\gamma)=\nu\}$.
Here, M_1 denotes the set of all probability measures in the product space $X\times Y$ and π_x denotes the projection of the measure in the first coordinate.

How are these problems related? It turns out that Kantorovich's problem is a relaxation of Monge's.

Theorem. *Kantorovich's problem is a true relaxation of Monge's.*

Being **a true relaxation of a problem** means that whenever you have an element in the first problem, you get an element in the second problem; therefore the infimum on the second one must be smaller.

Proof. Let T such that $T_{\#}\mu=\nu$ for every set A in the product σ-algebra define

$$\gamma_T(A)=\mu(\{x\in X:(x,Tx)\in A\})$$

Then $\gamma_T\in\Gamma(\mu,\nu)$.

Convergence of measures and topologies in the space of measures

The breakthrough by Kantorovich relies on being able to put Monge's in a more manageable setting; namely, the space of probability measures on the product space. By this observation it is now obvious that we need to use the fundamental properties of the space of probability measures.
Now the problem has changed into analyzing how measures interact. We need to understand how measures interact with each other.

Theorem. *The set of signed probability measures can be turned into a Banach space using the total variation norm.*

This is a very well known result and by using the **variation norm** allows us to use Riesz's theorem to identify the dual space of the continuous functions of compact support with the set of positive Radon measures, which is a necessary condition to understand Kantorovich's Duality formula (next section).
While the second distance is highly related to isoperimetric type inequalities which is of course of great interest to us.

It is evident how this definition arises in the context of optimal transport, but the next result may appear utterly surprising:

Theorem. *The Wasserstein distance metrizes weak convergence of probability measures.*

For the proof see [19] pag. 212.

Kantorovich's Duality formula

Why is the Kantorovich problem understood to be 'easier' in some way to the one by Monge? Mainly because it was nice enough to give conditions for solutions.

Theorem. *Let X and Y compact spaces.*
Let μ be a probability measure on X and ν a probability measure on Y.
Let $c : X \times Y \to \mathbb{R} \cup \{\infty\}$ be non-negative.
Let $\Gamma(\mu, \nu)$ be the set of Borel measures on $X \times Y$ such that their projections on X and Y are μ and ν respectively
Let Φ_c be the set of measurable functions $(\phi, \psi) \in L_1(\mu) \times L_1(\nu)$ such that

$$\phi(x) + \psi(y) \leq c(x, y)$$

μ-a.e. in x y ν-a.e. in y.

Then

$$\inf_{\pi \in \Gamma(\mu,\nu)} \left\{ \int c(x, y) d\pi \right\} = \sup_{(\phi,\psi) \in \Phi_c} \left\{ \int \phi d\mu + \int \psi d\nu \right\}$$

Moreover, Kantorovich's problem admits a minimizer

Real case and the quadratic cost function

In the case were the space $X = \mathbb{R}^n$, and the cost function: $c(x, y) = \dfrac{||x - y||^2}{2}$.
Recall the definition of the **sub-differential** $\partial\phi$:

$$y \in \partial\phi(x) \Leftrightarrow \forall z \in \mathbb{R}^n \phi(z) \geq \phi(x) + \langle y, z - x\rangle$$

Theorem. *Optimal plans are supported in subdifferentials Let μ and ν be probability measures and $\pi \in \Gamma(\mu, \nu)$, then if π is optimal for c it must be supported in the subdifferential of a proper lower semi-continuous convex function.*

See [19] pg. 92.

Theorem. *Moreover if μ doesn't give mass to small sets (say a.c. with respect to Lebesgue measure) then there **exists** a convex function ϕ on $\mathbb{R}^n$ such that*

$$\nabla\phi_{\#}\mu = \nu$$

See [19] pg. 94. This result and the fact that Kantorovich's Duality formula is proved by means of convex analysis tells us of the real need of understanding forms of convexity to solve transport problems.

Time dependent formulation, displacement interpolation and displacement of convexity

We have defined the transport problem in a time independent setting. To analyze isoperimetric type inequalities we will need to reformulate the problem, so now for each x we obtain a trajectory $T_t(x)$ where t varies from 0 to 1 and of course it has associated a cost of displacement $C(T_t(x))$.
We need regularity in the trajectories so it is assumed that $t \to T_t(x)$ is continuous and piece wise C^1 for μ a.e. on x. So naturally the time dependent problem is formulated:

$$\inf \left\{ \int_X C(T_t(x)) d\mu; T_0 = Id, T_{1\#}\mu = \nu \right\}$$

, where $Id(x) = x$ is just the identity operator. For probability measures μ and ν on $\mathbb{R}^n$ that do not give mass to small sets, the function

$$[\mu, \nu]_t = [tId + (1-t)\nabla\phi]_{\#}\mu$$

is called the displacement interpolation of μ into ν, it can be seen as the most natural linear interpolation between the Identity operator and the solution to Kantorovich's problem.
A subset of the absolutely continuous (w.r.t Lebesgue) measures on $\mathbb{R}^n$ is said to be **displacement convex** if whenever μ and ν are in the set $[\mu, \nu]_t$ is a.c. and lies in the set for all $t \in [0.1]$.
Meaning $\mathbf{S} \subseteq P_{ac}(\mathbb{R}^n)$ is displacement convex iff $\mu, \nu \in S \Rightarrow [\mu, \nu]_t \in \mathbf{S}\ \forall t \in [0, 1]$ Similarly a functional F defined on a displacement convex set is said to be displacement convex if whenever $\rho_t = [\mu, \nu]_t$ is a displacement interpolation, the function $t \to F(\rho_t)$ is convex.

Definition. *(Internal Energy)*
Let $\mathbf{S}$ a displacement convex subset of $P_{ac}(\mathbb{R}^n)$ and $U : \mathbf{S} \to \mathbb{R} \cup \{\infty\}$ we define the internal energy of U, denoted $\mathcal{U}$ as

$$\mathcal{U}(\rho) = \int U(\rho(x))dx$$

Theorem. *Criteria for displacement convexity*
Let $\mathbf{S}$ be a displacement convex subset of $P_{ac}(\mathbb{R}^n)$. Let $U : \mathbf{S} \to \mathbb{R} \cup \{\infty\}$ then if

$$U(0) = 0 \ \textbf{\textit{and}}\ r \to r^n U(r^{-n}) \text{ is convex non increasing for positive } r$$

then $\mathcal{U}$ is displacement convex on $\mathbf{S}$.

See [19] pg. 153.

Brunn-Minkowski via Optimal Transport

Our aim is to understand the inequality

$$|X + Y|^{1/n} \geq |X|^{1/n} + |Y|^{1/n}$$

by means of optimal transport, dividing by 2,

$$\left|\frac{X+Y}{2}\right|^{1/n} \geq \frac{1}{2}\left(|X|^{1/n} + |Y|^{1/n}\right)$$

so it shows some type of midpoint concavity (not exactly) between the Minkowsky sum operator and the Lebesgue measure; to obtain Brunn-Minkowski's inequality **apply** displacement of convexity criteria to the functional:

$$U(\rho) = -\int_{\mathbb{R}^n} \rho(x)^{1-1/n} dx$$

and use that inequality to conclude Brunn-Minkowski, with the following lemma.

Lemma 2.4.7. *(Support of the interpolant) [19] pg.186*
If μ and ν are the uniform measures on compact sets X, Y and $\rho_t = [\mu, t]$ then $\operatorname{supp}(\rho_t) \subseteq tX + (1-t)Y$.

See [19] pg. 186.
To conclude Brunn-Minkowski, let $S_t. = \operatorname{supp}(\rho_t)$, then

$$\int_{S_t} U\left(\frac{d\rho_t}{dx}\right) dx \geq |S_t| U\left(\int_{S_t} d\rho_t\right) = -|S_t|^{\frac{1}{n}}$$

By the Lemma 2.4.7

$$-|S_t|^{\frac{1}{n}} \geq -|(1-t)X+tY|^{\frac{1}{n}}$$

But by the convexity theorem for displacement

$$\mathcal{U}(\rho_t) \leq (1-t)\mathcal{U}(\mu) + t\mathcal{U}(\nu)$$

which is by definition just

$$|(1-t)X+tY|^{\frac{1}{n}} \geq (1-t)|X|^{\frac{1}{n}} + t|Y|^{\frac{1}{n}}$$

which concludes the result.

2.4.6 From Brunn-Minkowski to isoperimetry

Let $A \subseteq \mathbb{R}^n$ be an open set with smooth boundary and if we denote $|\partial A| = \lim_{\epsilon \downarrow 0} \frac{|A+\epsilon B|-|A|}{\epsilon}$, and we consider a Ball such that $|B|=|A|$ then by Brunn-Minkowsky's

$$\frac{|A+\epsilon B|-|A|}{\epsilon} \geq \frac{(|A|^{\frac{1}{n}}+|\epsilon B|^{\frac{1}{n}})^n-|A|}{\epsilon} = \frac{(|B|^{\frac{1}{n}}+(\epsilon^n)^{\frac{1}{n}}|B|^{\frac{1}{n}})^n-|B|}{\epsilon} = \frac{(1+\epsilon)^n-1}{\epsilon}|B|$$

By taking the limit in both sides we obtain $|\partial A| \geq |\partial B|$, and therefore the result.

2.4.7 A direct proof of the isoperimetric inequality in $\mathbb{R}^n$ using optimal transport

This section is included as to understand one of the ways in which optimal transport helps to prove geometric inequalities (at least for the real case). This proof is a detailed version of the one on [1].

Define μ as the uniform measure in A and ν as the uniform measure in B. By optimality, there exists ϕ a convex function such that $\nabla\phi_{\#}\mu = \nu$ and so by the change of variable formula

$$\frac{1}{|A|} = \det(\nabla\phi)\frac{1}{|B|}$$

But ϕ is convex so $\nabla\phi(x)$ is a symmetric, non-negative matrix. By the arithmetic-geometric mean inequality one has

$$(\det\nabla\phi)^{\frac{1}{n}} \leq \frac{\operatorname{tr}\nabla\phi}{n}$$

Therefore, by definition of the divergence

$$(\det\nabla\phi)^{\frac{1}{n}} \leq \frac{\nabla\cdot\phi}{n}$$

Using this inequality in the change of variables we have

$$\frac{1}{|A|^{\frac{1}{n}}} \leq \frac{1}{|B|^{\frac{1}{n}}}\frac{\nabla\cdot\phi}{n}$$

Integrating both sides over A we get

$$|A|^{1-\frac{1}{n}} \leq \frac{1}{n|B|^{\frac{1}{n}}}\int_A \nabla\cdot\phi(x)dx = \frac{1}{n|B|^{\frac{1}{n}}}\int_{\partial A}\langle\phi(x),n(x)\rangle d\mathcal{H}^{n-1}(x) \leq \frac{1}{n|B|^{\frac{1}{n}}}\int_{\partial A} d\mathcal{H}^{n-1}(x)$$

where the last equality holds by the divergence theorem. But as $\phi(x)$ transports into B, $\phi(x) \in B$, meaning that $|\phi(x)| \leq 1$ which by duality gives us $\langle \phi(x), v(x) \rangle \leq 1$ and so

$$|A|^{1-\frac{1}{n}} \leq \frac{1}{n|B|^{\frac{1}{n}}} |\partial A|$$

So if we fix $|A| = |B|$ then we have $n|B| \leq |\partial A|$ but this hold with equality in the case of the Ball by definition of the surface measure.

Meaning that $n|B| = |\partial B| \leq |\partial A|$.

2.4.8 Other measure theoretic versions of Isoperimetry

In this section we show a general way to pose the isoperimetry problem for more general measures.

Definition. *(Convex measure)*
A Borel probability measure μ on $\mathbb{R}^n$ is called convex (or logarithmically concave) if for all non-empty Borel sets A, B and $t \in [0,1]$, it holds

$$\mu(tA + (1-t)B) \geq \mu(A)^t \mu(B)^{1-t}$$

Theorem 2.4.8. *(Representation of convex measures)*
A probability measure μ on $\mathbb{R}^n$ is convex if and only if there exists a convex function V, with domain $D \subseteq \mathbb{R}^n$ such that

$$\frac{d\mu}{dx} = \exp(-V) \text{ on } D$$
$$\frac{d\mu}{dx} = 0 \text{ on } \mathbb{R}^n \setminus D$$

The reference is [4] pg. 378.

Theorem 2.4.9. *(Regularity of densities for convex measures)*
If μ is a convex measure on $\mathbb{R}^n$ with density ρ then ρ is of bounded variation.
Furthermore, if $\rho > 0$ a.e. then $\rho \in W^{1,1}$.

The proof can be found in [4] pg. 377.

Observation 2.4.10. *Observe that this is exactly the definition of the setting we talked about in Markov Triples on Riemannian manifolds. Recall 1.5 to see that we would only need to understand convexity in this setting for this theorem to be generalized. This of course is no trivial task.*

Definition. *(Finite perimeter)*
A bounded measurable set E is said to have finite perimeter if it's indicator function $\mathbf{1}_E$ is of bounded variation.

Let E have finite perimeter and let $P(E) := ||D\mathbf{1}_E||$, we call $P(E)$ it's perimeter. [4] pg. 378.

Observation 2.4.11. *Note that this definition (which is not the one we will use) shows how to compute perimeters: as derivatives of indicator functions. To show the main statements of this work we will use this idea: to understand perimeters and the Minkowski content, we can approximate by differentiating indicator functions.*

Definition. *(Cacciapolli set)*
A set E is said to be a Cacciapolli set if it's intersection with each ball has finite perimeter.

Theorem 2.4.12. *(Isoperimetric inequality for Cacciapolli sets)*
If E is a bounded Cacciapolli set then

$$|E|^{(n-1)/n} \leq c_n P(E)$$

where c_n is the optimal constant of the Sobolev inequality in $\mathbb{R}^n$

For a reference one can use [4] pg 378.

2.5 The Co-area formula and the layer cake integral representation

2.5.1 Layer Cake integral

The Layer cake integral is a very easy equality that has a lot of applications, it is stated here just for the completeness of the work.

Lemma 2.5.1. *(Layer cake integral representation)*
In a measure space (X, μ) and for $f \in L^p(\mu)$ one has:

$$\int_X f^p d\mu = p \int_0^\infty \lambda^{p-1} \mu(\{f > \lambda\}) d\lambda$$

Proof. Note that $f^p(x) = p \int_0^{f(x)} \lambda^{p-1} d\lambda$. Substitute this in the left hand side and use Fubini's.:

$$\int_X f^p d\mu = p \int_X \int_0^{f(x)} \lambda^{p-1} d\lambda d\mu(x) = p \int_0^\infty \lambda^{p-1} \int_X \mathbf{1}_{f(x)>\lambda} d\mu(x) d\lambda = p \int_0^\infty \lambda^{p-1} \mu(\{f > \lambda\}) d\lambda$$

This is one of the most useful identities, as one can rewrite the p norm of a function in terms of the measure of a level set, for example it is used when proving Doob's inequality for martingales.

2.5.2 Area and Co-area

One of the main theorems of geometric measure theorem is the so-called area and co-area formulas. This formula shows how to express the integral of a function as the integral of the level sets of a different function.

For notation if $f : \mathbb{R}^n \to \mathbb{R}^k$ is a function in $W^{1,p}_{loc}(\mathbb{R}^n)$, denote by $|Jf|$ the absolute value of the k-th dimensional Jacobian of f.
Recall that $|Jf|$ is the volume of the parallelepiped generated by the vectors ∇f_i.
In this section we denote by H^α the α dimensional Hausdorff measure on $\mathbb{R}^n$ and $\mathrm{Card}(A)$ the cardinality of a set A.

Theorem 2.5.2. *(Area and Co-area formulas in $\mathbb{R}^n$)*
Let $f : \mathbb{R}^n \to \mathbb{R}^k$ be a Lipschitz map and A, B be measurable sets

- *If $n \leq k$*
$$\int_{A \cap f^{-1}(B)} |Jf(x)| dx = \int_B \mathrm{Card}(A \cap f^{-1}(y)) dH^n(y)$$
- *If $n > k$*
$$\int_A |Jf(x)| dx = \int_{\mathbb{R}^k} H^{n-k}(A \cap f^{-1}(y)) dy$$

Proof. The proof can be found in [4] pg. 380

Observation 2.5.3. *The co-area formula is more familiar if one remembers the way to do integration by spherical coordinates.*

Observation 2.5.4. *The co-area formula is the "bridge" between the isoperimetric inequality and the Sobolev inequality as we see in 2.6*

2.6 Equivalence: Sobolev-Inequality and Isoperimetry

The proof of this equivalence in Riemannian Manifolds is standard, and it is done by applying the Sobolev inequality to indicator functions and using the co-area formula for the converse statement, nevertheless there is a less-usual way to understand this situation **in terms of Markov Triples**. This technique is developed in [3] chap. 8, and is detailed and explained here. One of the novelties of this technique is it's wide applicability, it can be used not only for Sobolev inequalities but for many more as we will see later on this section.
A reader familiar with [3] chap. 8 will notice this is just a detailed version of the exposition found there and can skip this part. If the reader is not familiar with the technique, it's fundamental that this section is understood as it's the basis of our main results.

Definition. *(Capacity of a set)*
In the setting of a Markov triple (X, Γ, μ) with associated Dirichlet form $\mathcal{E}$, for a measurable set $A \subseteq X$ we define it's capacity with respect to μ, $\mathrm{Cap}_\mu(A)$ as

$$\mathrm{Cap}_\mu(A) = \inf_{\substack{\mathbf{1}_A \leq f \leq \mathbf{1} \\ f \in \mathcal{D}(\mathcal{E})}} \{\mathcal{E}(f)\}$$

If the integral defining $\mathcal{E}$ is restricted to another set B, the same limit is denoted $\mathrm{Cap}_\mu(A, B)$ and it's called the relative capacity of A with respect to B.

Definition. *(Modified Capacity of a set)*
If $\mathbf{1} \in \mathcal{D}(\mathcal{E})$ we define the modified capacity as

$$\mathrm{Cap}_\mu{}^*(A) = \inf \left\{ \mathrm{Cap}_\mu(A, B) : A \subseteq B, \mu(B) \leq \frac{1}{2} \right\}$$

In general, it is very complicated to compute capacities and modified capacities of sets.

2.6.1 Measure-Capacity inequalities and Orlicz spaces

Definition. *(Measure-Capacity inequality)*
A Markov triple is said to satisfy the measure-capacity inequality with growth function Φ if

$$\Phi(\mu(A)) \leq \mathrm{Cap}_\mu(A)$$

In the case the measure is finite, one writes

$$\Phi(\mu(A)) \leq \mathrm{Cap}_\mu{}^*(A)$$

The goal of this section is to relate this inequalities to the functional inequalities explained in the beginning of this chapter. For this we need a very specific framework. This framework is the Orlicz spaces, this concept is important as a simple way to generalize the $L^p(\mu)$ spaces without losing important embeddings such as the Sobolev embedding.

Definition. *(Orlicz pair) Let $\pi : [0,\infty) \to [0,\infty)$ be a continuous and increasing (wlog strictly) function such that $\pi(0) = 0$. Then,*

$$\Pi(r) = \int_0^r \pi(u)du, \qquad \Upsilon(r) = \int_0^r \pi^{-1}(u)du$$

Then we call (Π, Υ) an Orlicz pair.

Observation 2.6.1. *An Orlicz pair is Legendre dual, meaning that $\Pi^* = \Upsilon^{**}$, where $*$ means the Legendre transform.*

Definition. *(Orlicz space)*
Given an Orlicz pair (Π, Υ) we define the Orlicz space as the functions f satisfying $||f||_{L^\Pi} < \infty$ where

$$||f||_{L^\Pi} = \sup\left\{\int |f|g d\mu : g \geq 0, \int \Upsilon(g)d\mu \leq 1\right\}$$

Observation 2.6.2. *A useful observation is that one can compute the Orlicz norm of an indicator function with an easy formula:*

$$||\mathbf{1}_A||_{L^\Pi} = \mu(A)\Upsilon^{-1}\left(\frac{1}{\mu(A)}\right)$$

2.6.2 The Co-area inequality

Similar as 2.5.2, the co-area inequality finds a bound for the integral over level sets of a function.

Theorem 2.6.3. *(Co-area inequality for Markov Triples)*
For every $f \in \mathcal{D}(\mathcal{E})$ it holds

$$\int_0^\infty \mathrm{Cap}_\mu(\{|f| > \lambda\})\lambda d\lambda \leq 6\mathcal{E}(f)$$

Proof. We start by using a partition of the real line to organize our level sets

$$2\int_0^\infty \mathrm{Cap}_\mu(\{|f| > \lambda\})d\lambda = \sum_{k\in\mathbb{Z}} \int_{2^k}^{2^{k+1}} \mathrm{Cap}_\mu(\{|f| > \lambda\})d(\lambda^2)$$

As $\{|f| > a\} \subseteq \{|f| > b\}$ whenever $b < a$, we can bound each of this integrals to get

$$2\int_0^\infty \mathrm{Cap}_\mu(\{|f| > \lambda\})d\lambda \leq \sum_{k\in\mathbb{Z}} (2^{k+1} - 2^k)\,\mathrm{Cap}_\mu(\{|f| > 2^k\}) = 3\sum_{k\in\mathbb{Z}} 2^k\,\mathrm{Cap}_\mu(\{|f| > 2^k\})$$

As the definition of capacity depends on the indicator function of the set, a good idea is to find approximations in $\mathcal{D}(\mathcal{E})$, in this case we aim to approximate $\mathbf{1}_{\{|f|>\lambda\}}$, so we put

$$f_k = \frac{1}{2^k}\min\{(|f| - 2^k)^+, 2^k\}$$

This function resembles the idea of the non integer part of a function, observe that

$$2^k \leq |f(x)| \leq 2^{k+1} \text{ so } 0 \leq |f(x)| - 2^k \leq 2^k \text{ and } 0 \leq \frac{|f(x)| - 2^k}{2^k} \leq 1$$

So one gets that

$$\mathbf{1}_{\{|f|>2^{k+1}\}} \leq f_k \leq \mathbf{1}_{\{|f|>2^k\}}$$

which is a good sequence to approximate the indicators in our last equation. By definition of the capacity being an infimum we have

$$\mathrm{Cap}_\mu(\{|f| > 2^k\}) \leq \int \Gamma(f_{k-1})d\mu \leq 2^{-2k+2} \int_{\{|f|>2^k\}\setminus\{|f|>2^{k+1}\}} \Gamma(f)d\mu$$

where the last equality comes from the fact that Γ squares constants. So putting the two inequalities together we get

$$2\int_0^\infty \mathrm{Cap}_\mu(\{|f| > \lambda\})d\lambda \leq 3\sum_{k\in\mathbb{Z}} 2^k\,\mathrm{Cap}_\mu(\{|f| > 2^k\}) \leq 12\int \Gamma(f)d\mu$$

2.6.3 Proof of the equivalence

The usefulness and impact of the Orlicz spaces are summarized in the next proposition (8.2.1 in [3]). This is a general version of the equivalence between isoperimetry and the Sobolev inequality, and therefore extremely useful. A point is needed to be made, about the impact of this proposition, it gives a general technique to prove a fundamental theorem.

Theorem 2.6.4. *(Measure-capacity inequality and bounds for the Orlicz norm)*
The following statements are equivalent for a Markov triple (X, Γ, μ) with Dirichlet form $\mathcal{E}$

a) There is a constant C_1 such that if $\mu(A) < \infty$

$$||\mathbf{1}_A||_{L^\Pi} = \mu(A)\Upsilon^{-1}\left(\frac{1}{\mu(A)}\right) \leq C_1\,\mathrm{Cap}_\mu(A)$$

b) There is a constant C_2 such that if $f \in \mathcal{D}(\mathcal{E})$

$$||f^2||_{L^\Pi} \leq C_2\mathcal{E}(f)$$

Furthermore, $C_1 \leq C_2 \leq 12C_1$

Observation 2.6.5. *Before going in to depth in the proof, let us analyze a specific case, namely the Sobolev inequality. If we write $\pi(r) = r^{p-1}$ for $p > 1$ in the previous setting then*

$$\Pi(r) = \frac{r^p}{p} \qquad \Upsilon(r) = \int_0^r u^{\frac{1}{p-1}}du = \frac{p-1}{p}r^{\frac{p}{p-1}}$$

So by duality of the $L^p(\mu)$ spaces, the Orlicz norm $||f||_{L^{\frac{r^p}{p}}} = ||f||_{L^p}$.
Therefore by considering this inequality, we are really getting Sobolev's inequality. If the right hand-side of the equation, one has to remember that in this setting $\Gamma(f) = ||\nabla f||^2$. So we get the Sobolev inequality in the form of 2.1.1.
Note also that in this case one has:

$$||\mathbf{1}_A||_{L^{r^p/p}} = \mu(A)\Upsilon^{-1}\left(\frac{1}{\mu(A)}\right) = \left(\frac{p}{p-1}\right)^{(p-1)/p}\mu(A)^{(p+1)/(p-1)}$$

Also by taking Lipschitz approximations as we will do in Chapter 3, and using the Lebesgue measure the right hand-side converges to $|\partial A|$ giving us the isoperimetric inequality.

a) ⇒ b)

We start by computing the L^Π norm of f^2 and using the same technique ($f^2(x) = \int_0^{f^2(x)} rdr$) as in layer cake representation 2.5.1

$$||f^2||_{L^\Pi} = \sup\left\{\int f^2 gd\mu\right\} = 2\sup\left\{\int_0^\infty \left(\int_{\{|f|>r\}} gd\mu\right) rdr\right\} \leq 2\int_0^\infty \sup\left\{\int_{\{|f|>r\}} gd\mu\right\} rdr$$

where the last inequality is due to reverse Fatou's lemma. Notice now that the term in brackets, is the L^Π norm of $\mathbf{1}_{\{|f|>r\}}$, and by the observation 2.6.2 we know how to compute it, so we have

$$||f||_{L^\Pi} \leq 2\int_0^\infty \mu(\{|f|>r\})\Upsilon^{-1}\left(\frac{1}{\mu(\{|f|>r\})}\right) rdr$$

So know we can use our hypothesis **a)**, and we have

$$||f||_{L^\Pi} \leq 2C_1 \int_0^\infty \mathrm{Cap}_\mu(\{|f|>r\}) rdr$$

But this expression is already very familiar to us, as it is part of the co-area inequality 2.5.2, so we have

$$||f||_{L^\Pi} \leq 12C_1\mathcal{E}(f)$$

b) ⇒ a)

This is implication is simpler, note that if $A \subseteq X$ is such that $\mathbf{1}_A \leq f \leq \mathbf{1}$ then

$$||\mathbf{1}_A||_{L^\Pi} \leq ||f^2||_{L^\Pi} \leq C_2\mathcal{E}(f)$$

where the second inequality is our hypothesis **b)**, so by taking the infimum over all such f, we get the result as the right hand-side is the definition of the capacity, i.e.

$$\mu(A)\Upsilon^{-1}\left(\frac{1}{\mu(A)}\right) \leq C_2\,\mathrm{Cap}_\mu(A)$$

Observation 2.6.6. *This discussion was a more general proof of the Sobolev inequality, it is one of those examples where generalizing makes the proofs easier. Nevertheless, every element of the standard proof is there:*

- *The co-area formula.*
- *Approximations of indicator functions.*
- *Understanding the integrals of norms of indicator functions.*

2.6.4 Capacity and other functional inequalities

As one can imagine, many of the functional inequalities presented here have their capacity versions, adequate to use the Orlicz spaces to prove them. The chapter 8 on [3], not only covers the previous proofs but details the use of this technique in more functional inequalities.
Here we only state the theorem so the reader gets a good idea of how functional inequalities may be rewritten.

Capacity and Poincaré

Theorem 2.6.7. *(Capacity and Poincaré)*
For a Markov triple (X, Γ, μ) *suppose that for all sets* A *with* $\mu(A) \leq \frac{1}{2}$ *it holds*

$$\mu(A) \leq C\,\mathrm{Cap}_{\mu}{}^{*}(A)$$

then (X, Γ, μ) *satisfies a Poincaré inequality with constant* $12C$*. If Poincaré inequality holds for* $C/2$ *then the above inequality also holds.*

Capacity and Log-Sobolev

Theorem 2.6.8. *(Capacity and Log-Sobolev)*
Let (X, Γ, μ) *be a Markov triple, then if*

$$\mu(A)\log\left(1 + \frac{e^2}{\mu(A)}\right) \leq C\,\mathrm{Cap}_{\mu}{}^{*}(A)$$

then LSI($12C$*) holds. Furthermore if LSI*$\left(\frac{C}{8}\right)$ *then the above equation also holds.*

This section explains how capacity comes into play, capacity is the correct way to understand functional inequalities in terms of measures. This concept makes clear how functional inequalities and inequalities on sets and measures connect.

2.6.5 Co-area inequality for Minkowski Content

Similar as 2.5.2 one can find bounds of the level sets for the Minkowski content, this is fundamental to understanding isoperimetry. For this setting assume the "Important assumption" 3.1.

Theorem 2.6.9. *(Co-area inequality for Minkowski content)*
For a Markov triple (X, Γ, μ) *and the metric space induced by it's intrinsic distance, for every Lipschitz function* f *one has*

$$\int_{-\infty}^{\infty} \mu^{+}(\{x \in X : f(x) > r\})dr \leq \int \sqrt{\Gamma(f)}d\mu$$

Proof. The proof is a consequence of Fatou's lemma and the assumption 3.1, the details are omitted but again can be found in [3].

Theorem 2.6.10. *(Linear isoperimetry implies Poincaré)*
Suppose that for a Markov Triple (X, Γ, μ) *linear isoperimetry holds with constant* c *then the Poincaré inequality holds with constant* $\frac{4}{c^2}$*,*

Proof. We start by taking a positive Lipschitz function g and observe that

$$\int_0^{\infty} c\min\{\mu(\{x : g(x) > r\}), 1 - \mu(\{x : g(x) > r\})\}dr \leq \int_0^{\infty} I_{\mu}(\mu(\{x : g(x) > r\}))dr$$
$$\leq \int_0^{\infty} \mu^{+}(\{x : g(x) > r\})$$

where the first inequality is the hypothesis and the second one is the definition of I_μ as an infimum. Therefore by applying the Co-area formula for the Minkowski content 2.6.9 we have

$$\int_0^\infty c\min\{\mu(\{x : g(x) > r\}), 1 - \mu(\{x : g(x) > r\})\}dr \leq \int \sqrt{\Gamma(g)}d\mu \tag{2.5}$$

Now for a general Lipschitz function f, define f^+ and f^- as the positive and negative parts of $f - M$ where M is the median of f. By definition of M being the median,

$$\mu(\{x : (f^+)^2 > r\}) \leq \frac{1}{2} \text{ and } \mu(\{x : (f^-)^2 > r\}) \leq \frac{1}{2}$$

No we apply 2.5 to $(f^+)^2$ and $(f^-)^2$ so

$$c\int |f - M|^2 d\mu = c\int (f^+)^2 d\mu + c\int (f^-)^2 d\mu$$

But we can apply our usual layer cake representation 2.5.1 to get

$$c\int |f - M|^2 d\mu = c\int \mu(\{(f^+)^2 \geq r\})dr + c\int \mu(\{(f^-)^2 \geq r\})dr \leq c\int \sqrt{\Gamma(f^+)}d\mu + c\int \sqrt{\Gamma(f^-)}d\mu$$

By the chain rule for Γ and then by Cauchy-Schwartz we have

$$\int \sqrt{\Gamma((f^+)^2)}d\mu = 2\int f^+\sqrt{\Gamma(f^+)}d\mu \leq 2\left(\int (f^+)^2 d\mu\right)^{1/2}\left(\int \Gamma(f^+)d\mu\right)^{1/2}$$

By doing the same to f^- and adding both equations yield the estimate

$$c\int |f - M|^2 d\mu \leq 2\left(\int |f - M|^2 d\mu\right)^{1/2}\left(\int \Gamma(f)d\mu\right)^{1/2}$$

By taking the middle term to the left hand side and squaring both sides we have

$$\int |f - M|^2 d\mu \leq \frac{4}{c^2}\int \Gamma(f)d\mu$$

And as the variance can be shown to be the infimum of terms like the one on the left hand side we have

$$\mathrm{Var}_\mu(f) \leq \frac{4}{c^2}\int \Gamma(f)d\mu$$

which is of course the Poincaré inequality with constant $4/c^2$.

2.7 Levy-Gromov isoperimetry

In this section we discuss the situation of isoperimetry in manifolds, this results are due to Gromov, exposed in [12].

Theorem 2.7.1. *(Levi-Gromov's isoperimetry under curvature bound)*
Let M be a closed $(k+1)$-dimensional manifold and let A be a domain with smooth boundary.
If $\inf_\tau \mathrm{Ricc}(\tau, \tau) \geq k$ where the infimum is taken over all τ tangent unit vectors of the manifold then

$$\frac{\mathrm{Vol}(\partial A)}{\mathrm{Vol}(M)} \geq I_{Vol}\left(\frac{\mathrm{Vol}(A)}{\mathrm{Vol}(M)}\right)$$

The proof can be found in [12], and relies in geometric arguments and bounding the volume by the integral of the Jacobian of the exponential map which will be the main ideas we will use to prove 4.2.6.

2.8 A comment on the curvature dimension inequality

All the proofs in this work concern the curvature condition $\mathrm{CD}(k,\infty)$, nevertheless a more rigorous analysis is to be made when one needs to check $\mathrm{CD}(k,n)$ and it's relation with functional inequalities, this is no easy task but has been studied in [8]. Further studies for this project could involve the analysis of this stronger condition.

2.9 An example: Gabriel's Horn

We have seen that for general settings (wherever we can define Markov Triples) we have the equivalence between isoperimetry and the Sobolev inequality. We've seen this equivalence via 2.6.4.
So we know that **if** isoperimetry holds, then the Sobolev inequality will also hold. Similarly, **if** the Sobolev inequality holds, we instantly get isoperimetry.
In section 2.2.1 we showed that under $\mathrm{CD}(k,\infty)$, the Sobolev inequality holds, also 2.7 shows a geometric condition sufficient to have isoperimetry. But in general, both isoperimetry and this functional inequalities may not hold.
The technique in 2.4.5 doesn't work in other settings as it depends on several steps which are very special properties of $\mathbb{R}^n$.

Example 2.9.1. *(Gabriel's horn)*
Let us consider the solid of revolution generated by the map $x \to \dfrac{1}{x}$ *around the* x *axis in* $[1,\infty)$.

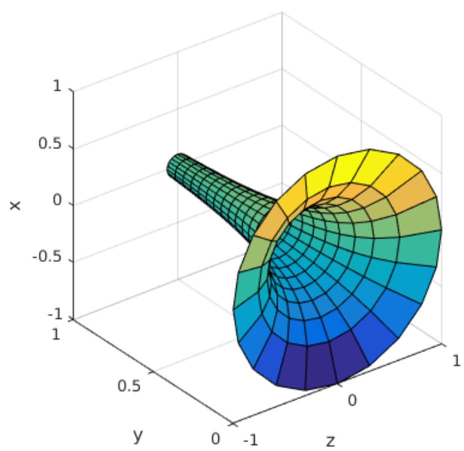

Figure 2.2: Gabriel's horn

Observe now that we can compute the volume V *and the surface area* A *of this object:*

$$V = \pi \int_1^\infty \left(\frac{1}{x}\right)^2 dx = \pi$$

$$A = 2\pi \int_1^\infty \frac{1}{x}\sqrt{1+\frac{1}{x^2}}dx = \infty$$

Now as the $A = \infty$ for every a we can find a set (say truncate the horn to the image of an interval) such that the area of this set is equal to a.
To produce a set with the same area but smaller volume one can just reproduce this construction to the right of the first set. This shows that there is no isoperimetry if we consider Gabriel's Horn as a manifold (inheriting the differential structure of $\mathbb{R}^3$) and by the results on this chapter, Log-Sobolev inequalities also fail in this setting. Via 2.2.1 one can now conclude geometric properties of Gabriel's horn.

So what we aim to see know is when can we get Isoperimetry or linear isoperimetry, this work depends on 2 principal papers: [14] and [16]. The latter gave geometric insight (and proofs) to understand the isoperimetric profile in general settings while the second one presented a semigroup proof (shown in the next chapter) that facilitates the understanding of Milman's geometric arguments.

2.10 A comment on extending the optimal transport technique

The technique used in 2.4.5 to prove isoperimetry was very nice. One can ask if such techniques can be generalized to use optimal transport to obtain every functional inequality. In general that is not true, nevertheless some results are achieved by generalizations of this technique, this is the content of Chapter 9 on [3]. Using optimal transport one can show that the log-Sobolev inequality already implies some inequalities regarding the quadratic cost of transport. Nevertheless it seems that at the moment, the technique developed in this chapter can take us a little bit further on understanding isoperimetry, as we see in the next chapter.

Chapter 3

Reversing the hierarchy: stochastic side

In the previous chapter we learned how one can derive concentration inequalities from functional inequalities, and how we can derive functional inequalities from isoperimetry but in that section we didn't talk about the reverse statement. In recent papers by Milman, he proved using geometric tools that concentration inequalities and curvature bounds are sufficient to obtain isoperimetry.
In this chapter we explain a proof found in [14], it is a semigroup proof concluding similarly to the results obtained in [16].
We are well suited to understand this proof with the tools developed in the first chapters.

3.1 Semigroup proof for linear Isoperimetry (Stochastic side)

Theorem 3.1.1. *(Milman) [3] [Theorem 8.7.1] Under $CD(0,\infty)$, if exponential concentration holds with constants c and C, then*

$$I_\mu(x) \geq c' \min\{x, 1-x\}, \;\; x \in [0,1]$$

Theorem 3.1.2. *(Milman) [14] [Theorem 1] Under $CD(0,\infty)$, if $\alpha_\mu(r) \to 0$ as $r \to \infty$ then*

$$I_\mu(x) \geq c_2^{'} \min\{x, 1-x\}, \;\; x \in [0,1]$$

Observation 3.1.3. *As we will see in the proof, the constant c does not depend on the dimension of the underlying manifold.*

Observation 3.1.4. *In this chapter we will not be able to prove 3.1.1 and 3.1.2 completely, as we need some geometric ideas to conclude. The whole purpose of the next chapter is to understand such ideas.*

Observation 3.1.5. *The key of this chapter is realizing that the proofs of both theorems are analogous, nevertheless doing only the proof of the latter is not as intuitive as doing the proof of the former, therefore we present the technique that proves both theorems, with the obvious distinctions in each case clearly stated.* ***We will focus on the details of 3.1.1 because they are more elaborate****. Whenever an equation has to be modified to be used for the second theorem* ***we denote it by writing an asterisk **** *next to it.*

We write observations 2.1.6 and 2.1.11 as a lemma:

Lemma 3.1.6. *(Poincaré and log-Sobolev inequalities imply Lipschitz properties)*
Let $0 \leq f \leq 1 \in \mathcal{A}$ then

- *If the second version of the Poincaré inequality (2.1.7) is satisfied then*

$$\Gamma(P_t f) \leq \frac{1}{2t}$$

- *If the second version of the log-Sobolev inequality (2.1.4) is satisfied then*

$$\Gamma(P_t f) \leq \frac{1}{t}(P_t f)^2 \log \frac{1}{P_t f} \leq \frac{1}{t} \log \frac{1}{P_t f}$$

Observation 3.1.7. *In this case it is helpful to study a larger class of functions, namely* $A_0^{\text{const}} = A + \mathbb{R}$ *where* $+$ *denotes the Minkowski sum.*

The proof presented here is a (way) more detailed version of the one found in [3].

Proof. Let $f \in A_0^{\text{const}}$ then

$$\int f \log f d\mu - \int P_t f \log P_t f d\mu = -\int_0^t \frac{d}{ds}\left(\int P_s f \log P_s f d\mu\right) ds$$

Now we can use De Bruijn's identity 2.1.7 and get

$$\int f \log f d\mu - \int P_t f \log P_t f d\mu = \int_0^t \int \frac{\Gamma(P_s f)}{P_s f} d\mu \tag{3.1}$$

For 3.1.2 we need

$$\int f^2 d\mu - \int (P_t f)^2 d\mu = 2\int_0^t \int \Gamma(P_s f) d\mu ds \tag{\textbf{3.1*}}$$

Observation 3.1.8. *It is clear even from this point that the details for 3.1.2 are a little easier.*

Now as $\mathrm{CD}(0, \infty)$ is part of our assumptions, by 2.2.1 local log-Sobolev inequalities hold in the sense of 2.1.4 and also 2.1.7. So we can use local Poincaré and the reverse local Poincaré inequalities. To use them, suppose further that $0 < \epsilon \leq f \leq 1$ then by the reverse Poincaré inequality

$$\frac{\Gamma(P_s f)}{P_s f} \leq \frac{1}{s}\left[P_s(f \log f) - P_s f \log P_s f\right]$$

But $1 \geq f \geq \epsilon$ means that $f \log f \leq 0$ and by positivity preserving property of P_s we have $P_s(f \log f) < 0$. We can use this in the last inequality to get a bound:

$$\frac{\Gamma(P_s f)}{P_s f} \leq \frac{1}{s}\left[-P_s f \log P_s f\right] = \frac{1}{s}\left[P_s f \log \frac{1}{P_s f}\right] \leq \frac{1}{s}[P_s f \log \left(\frac{1}{\epsilon}\right)]$$

And by simply taking $P_s f$ to the other side we get a bound for our Carré du champ evaluated in the semigroup:

$$\Gamma(P_s f) \leq \frac{1}{s} \log \left(\frac{1}{\epsilon}\right)(P_s f)^2$$

which by considering square roots can be rewritten as follows

$$\sqrt{\Gamma(P_s f)} \leq \sqrt{\frac{1}{s} \log \left(\frac{1}{\epsilon}\right)} P_s f$$

.

We can now use this bound,

$$\int \frac{\Gamma(P_s f)}{P_s f} d\mu = \int \frac{\sqrt{\Gamma(P_s f)}\sqrt{\Gamma(P_s f)}}{P_s f} d\mu \leq \sqrt{\frac{1}{s}\log\left(\frac{1}{\epsilon}\right)} \int \sqrt{\Gamma(P_s f)} d\mu \tag{3.2}$$

But under our assumption CD$(0,\infty)$ the strong gradient bounds hold (2.2.1), so we have $\sqrt{\Gamma(P_s f)} \leq P_s(\sqrt{\Gamma(f)})$, plugging the bound in 3.2 gives

$$\int \frac{\Gamma(P_s f)}{P_s f} d\mu \leq \sqrt{\frac{1}{s}\log\left(\frac{1}{\epsilon}\right)} \int \sqrt{\Gamma(f)} d\mu$$

To return to bounding in 3.1 we need to integrate with respect to s from 0 to t so we have

$$\int f \log f d\mu - \int P_t f \log P_t f d\mu \leq \int_0^t \sqrt{\frac{1}{s}\log\left(\frac{1}{\epsilon}\right)} \int \sqrt{\Gamma(f)} d\mu ds \tag{3.3}$$

For 3.1.2 we need the easier estimate:

$$\int f^2 d\mu - \int (P_t f)^2 d\mu \leq 2\sqrt{2t} \int \sqrt{\Gamma(f)} d\mu \tag{\textbf{3.4*}}$$

As only the first term depends on s on the bound, we can explicitly calculate the integral $\int_0^t \sqrt{(1/s)} ds = 2t^{1/2}$ so we get

$$\int f \log f d\mu - \int P_t f \log P_t f d\mu \leq 2\sqrt{t \log\left(\frac{1}{\epsilon}\right)} \int \sqrt{\Gamma(f)} d\mu \tag{3.4}$$

Observation 3.1.9. *In order to use the exponential concentration property, we need to use Lipschitz functions, therefore we can obtain the following lemma.*

Lemma 3.1.10. *Under CD$(0,\infty)$, $-\psi = -\sqrt{\log \frac{2}{P_t f}}$ is $\frac{1}{2\sqrt{t}}$ - Lipschitz with respect to Γ.*

Proof. Note that by the calculation done in 1.3.10 we know

$$\Gamma\left(\sqrt{\log \frac{1}{P_t f}}\right) = \frac{1}{4} \frac{1}{\log \frac{1}{P_t f}} \frac{1}{(P_t f)^2} \Gamma(P_t f)$$

By the Log-Sobolev inequality, we can bound

$$\Gamma\left(\sqrt{\log \frac{1}{\frac{1}{P_t f}}}\right) \leq -\frac{1}{4} \frac{1}{\log \frac{1}{P_t f}} \frac{1}{P_t f} \frac{2}{2t} P_t f \log P_t f$$

which by getting the sign inside the logarithm and cancelling the remaining terms is equivalent to

$$\Gamma\left(\sqrt{\log \frac{1}{P_t f}}\right) \leq \frac{1}{4t}$$

Now that we have a Lipschitz function, we can apply the hypothesis of exponential concentration, namely:

$$\mu\left(-\psi \geq \int -\psi d\mu + r\right) = \mu\left(\psi \leq \int \psi d\mu - r\right) \leq Ce^{-2cr\sqrt{t}}$$

Now we aim to use properties of ψ, namely the following lemma

Lemma 3.1.11. *The map* $u \to \sqrt{\log \frac{2}{u}}$ *is convex for* $u \in (0,1]$.

Proof. of the lemma: By differentiability it is enough to show that the second derivative is non-negative in $(0,1]$.

$$\frac{d}{du}\sqrt{\log\frac{2}{u}} = \frac{2\log\frac{2}{x}-1}{4x^2(\log\frac{2}{x})^{3/2}}$$

So the map is convex if $2\log\frac{2}{x} - 1 \geq 0$ which happens if $x < \frac{2}{\sqrt{e}}$. As $\frac{2}{\sqrt{e}} > 1$, we get that $u \to \sqrt{\log\frac{2}{u}}$ is convex for $u \in (0,1]$.

By Jensen's inequality, as we know by the lemma that the function is convex,

$$\sqrt{\log\frac{2}{\int f d\mu}} = \sqrt{\log\frac{2}{\int P_t f d\mu}} \leq \int \psi d\mu$$

And so we can bound,

$$\mu\left(\psi \leq \sqrt{\log\frac{2}{\int f d\mu}} - r\right) \leq \mu\left(\psi \leq \int \psi d\mu - r\right) \leq Ce^{-2cr\sqrt{t}}$$

Claim 3.1.12. *Let* $0 \leq r \leq \sqrt{\frac{1}{2}\log\frac{2}{\int f d\mu}}$ *If it holds that*

$$\mu\left(\psi \leq \sqrt{\log\frac{2}{\int f d\mu}} - r\right) \leq Ce^{-2cr\sqrt{t}}$$

then

$$\mu\left(P_t f \geq \sqrt{2\int f d\mu e^{r^2}}\right) \leq Ce^{-2cr\sqrt{t}}$$

Proof of claim. As we do not use specific properties of $P_t f$ and $\int f d\mu$, and the prove is only to work with the inequalities, we can write for easier understanding $y := P_t f$ and $x = \int f d\mu$, so our hypothesis are

$$\sqrt{\log\frac{2}{y}} \leq \sqrt{\log\frac{2}{x}} - r \text{ and also } 0 \leq r \leq \sqrt{\frac{1}{2}\log\frac{2}{x}} \tag{3.5}$$

and our aim is to conclude that

$$y \geq \sqrt{2x}e^{r^2} \tag{3.6}$$

Observe that

$$0 \leq (r - \frac{1}{2}\sqrt{\log\frac{2}{x}})^2 = r^2 - r\log\frac{2}{x} + \frac{1}{4}\log\frac{2}{x}$$

which by adding and subtracting $\frac{1}{2}\log\frac{2}{x}$ yields

$$\frac{1}{2}\log\frac{2}{x} - r^2 \leq \left(\sqrt{\log\frac{2}{x}} - r\right)^2 \tag{3.7}$$

As our hypothesis ensures the lower bound is positive so we can consider square roots on both sides and get

$$\sqrt{\frac{1}{2}\log\frac{2}{x} - r^2} \leq \sqrt{\log\frac{2}{x}} - r$$

To prove the claim, as we have sets for which the inequalities hold, it is enough to show that the hypothesis is necessary for the result, meaning that in order for 3.6 to hold we must have that 3.5 holds. So 3.6 holds if and only if

$$\log\frac{2}{y} \leq \frac{1}{2}\log\frac{2}{x} - r^2$$

so by using 3.7 we have

$$\sqrt{\log\frac{2}{y}} \leq \sqrt{\log\frac{2}{x}} - r$$

This means that whenever 3.6 holds 3.5 also holds. So by returning to our original problem we have

$$\mu\left(P_t f \geq \sqrt{2\int f d\mu e^{r^2}}\right) \leq \mu\left(\psi \leq \sqrt{\log\frac{2}{\int f d\mu}} - r\right) \leq Ce^{-2cr\sqrt{t}} \tag{3.8}$$

For 3.1.2 we need not to bound with the exponential of course, but with the concentration function so we have

$$\mu\left(P_t f \geq \sqrt{2\int f d\mu e^{r^2/4t}}\right) \leq \alpha_\mu(r) \tag{\textbf{3.8*}}$$

Returning to our original problem, suppose further that $0 < \delta \leq 1$ is such that $\delta \geq \sqrt{2\int f d\mu e^{r^2}}$ then we can use it to bound one of the terms in the Poincaré inequality, namely,

$$\int P_t f \log\frac{1}{P_t f} d\mu \geq \log\frac{1}{\delta}\int_{\{P_t \leq \delta\}} P_t f d\mu = \left(\int f d\mu - \int_{\{P_t f > \delta\}} P_t f d\mu\right)\log\frac{1}{\delta}$$

where the first inequality comes from considering the integral over a subset and the equality by writing $\{P_t f > \delta\} = X \setminus \{P_t f \leq \delta\}$ and using invariance.
As $\epsilon \leq f \leq \mathbf{1}$ by conservativeness $\epsilon \leq P_t f \leq \mathbf{1}$, meaning that $-P_t f \geq -\mathbf{1}$. So we can bound the last term as follows

$$\int P_t f \log\frac{1}{P_t f} d\mu \geq \left(\int f d\mu - \int_{\{P_t f > \delta\}} \mathbf{1} d\mu\right)\log\frac{1}{\delta} \geq \left(\int f d\mu - Ce^{-2cr\sqrt{t}}\right)\log\frac{1}{\delta} \tag{3.9}$$

where of course the last inequality comes from applying 3.8

KEY APPROXIMATION

Here we arrive to the core of the proof, until now even though a lot of inequalities and claims have been proved we haven't seen how this connects to isoperimetry. This step, fundamental to our analysis, shows the true relation between functional inequalities and isoperimetry. As we will see, if the functional inequality involves derivatives and we are able to approximate indicator functions, we will be able to understand isoperimetric type inequalities in such space.

Important assumption

In this section we adopt an additional assumption to the framework of Markov triples (X, Γ, μ) then for every $f \in \mathcal{D}(\mathcal{E})$

$$\sqrt{\Gamma(f)} = \limsup_{d(x,y)\to 0} \frac{|f(x) - f(y)|}{d(x,y)}$$

Notice that this is the case of Riemannian manifolds as discussed in (1.5). This assumptions is restrictive but it is good enough for our purposes because most (if not all) of our examples fulfill it. Note that $\sqrt{\Gamma(f)}$ is the *pointwise Lipschitz constant* of f at every x with respect to the absolute value.
For our theory, it is fundamental to understand the expression $\Gamma(\mathbf{1}_A)$ where A is some closed set. Nevertheless as this function may not be in $\mathcal{A}$ or in $\mathcal{D}(\mathcal{E})$ we need to approximate it. Let us attempt such approximation by putting

$$f_\epsilon(x) = \max\left\{1 - \frac{1}{\epsilon} d(x, A), 0\right\}$$

So let us analyze such limits, meaning we aim to compute

$$\sqrt{\Gamma(f_\epsilon)} = \limsup_{d(x,y)\to 0} \frac{|f_\epsilon(x) - f_\epsilon(y)|}{d(x,y)}$$

Just to have everything written explicitly, this is the same as

$$\sqrt{\Gamma(f_\epsilon)} = \limsup_{d(x,y)\to 0} \frac{\left|\max\left\{1 - \frac{1}{\epsilon}d(x,A), 0\right\} - \max\left\{1 - \frac{1}{\epsilon}d(y,A), 0\right\}\right|}{d(x,y)} \tag{3.10}$$

We solve this by analyzing cases of x, meaning that we see the sets to which x belongs and draw conclusions of the limit superior in each set.

Claim 3.1.13.

$$\sqrt{\Gamma(f_\epsilon)} \leq \left(\frac{1}{\epsilon}\right) \mathbf{1}_{A_\epsilon \setminus A}$$

Proof of claim. The relevant cases are determined by the right hand side of the claim so we can reduce our proof to $x \in A_\epsilon \setminus A$ and it's complement.

- If $x \in A$, as the limit superior is the infimum of the tail suprema, and the inner most term has an absolute value, we now that 0 is a lower bound for the suprema. But if $x \in A$ and y_n is a sequence such that $y_n \in A$ and $d(y_n, x) \to 0$ then we have $d(x, A) = 0 = d(y_n, A)$ so the numerator on 3.10 is 0. So the lower bound is attained and by the infimum property, $\sqrt{\Gamma(f_\epsilon)} = 0$.

- If $x \notin A_\epsilon$ then $\max\left\{1 - \dfrac{1}{\epsilon}d(x, A), 0\right\} = 0$ and so the numerator in 3.10 is only depending on y. So (by considering subsequences without loss of generality) we have two options, $y_n \in A_\epsilon$ or $y \notin A_\epsilon$, but the limit superior is the infimum of sequential limits, and as the latter produce 0 in the numerator, the lower bound 0 is achieved and this $\Gamma(f_\epsilon) = 0$.

- If $x \in A_\epsilon \setminus A$ then by using a sequence $y_n \in A_\epsilon \setminus A$ 3.10 becomes

$$\limsup_{d(x,y)\to 0} \frac{|\frac{1}{\epsilon}(d(x,A) - d(y,A))|}{d(x,y)}$$

But for every $z \in A$ we have the triangle inequality

$$d(x,z) \leq d(x,y) + d(y,z)$$

so considering the infimum over all $z \in A$ this becomes

$$d(x,A) \leq d(x,y) + d(y,A)$$

which by reversing the roles of x and y gives

$$|d(x,A) - d(y,A)| \leq d(x,y)$$

Using this we get

$$\sqrt{\Gamma(f_\epsilon)} \leq \limsup_{d(x,y)\to 0} \frac{1}{\epsilon}\frac{d(x,y)}{d(x,y)} = \frac{1}{\epsilon}$$

And the claim is proved.

As explained in 1.3.1 $\mathbf{1}_A$ may not be in the domain of Γ so we want to approximate what the value of $\sqrt{\Gamma(\mathbf{1}_A)}$ would be, recall that $f_\epsilon \to \mathbf{1}_A$ pointwise so let us write $1_A = \liminf_{\epsilon\to 0} f_\epsilon$.

$$\int \sqrt{\liminf_{\epsilon\to 0} \Gamma(f_\epsilon)} d\mu = \int \limsup_{d(x,y)\to 0} \liminf_{\epsilon\to 0} \frac{|f_\epsilon(x) - f_\epsilon(y)|}{d(x,y)} d\mu$$

But as the superior limit is an infimum and the inferior limit is a supremum we can bound this by the reversed limit and apply the claim, namely

$$\int \liminf_{\epsilon\to 0} \sqrt{\Gamma(f_\epsilon)} d\mu \leq \int \liminf_{\epsilon\to 0} \limsup_{d(x,y)\to 0} \frac{|f_\epsilon(x) - f_\epsilon(y)|}{d(x,y)} d\mu \leq \int \liminf_{\epsilon\to 0} \frac{1}{\epsilon} \mathbf{1}_{A_\epsilon \setminus A} d\mu$$

Now naturally in the last term we use Fatou's lemma and get

$$\int \liminf_{\epsilon\to 0} \sqrt{\Gamma(f_\epsilon)} d\mu \leq \liminf_{\epsilon\to 0} \frac{1}{\epsilon} \int \mathbf{1}_{A_\epsilon \setminus A} d\mu = \liminf_{\epsilon\to 0} \frac{1}{\epsilon}(\mu(A_\epsilon) - \mu(A)) = \mu^+(A)$$

Observation 3.1.14. *This bound is fundamental to our analysis. It is the connection between isoperimetric type inequalities and Markov semigroups.*

To summarize (with our key approximation) we have shown the following bound

$$\int \sqrt{\text{"}\Gamma(\mathbf{1}_A)\text{"}} d\mu \leq \mu^+(A) \tag{3.11}$$

where the quotes remind us that we are not evaluating on the indicator function but we have used our approximation. Now we are almost finished, we need to organize all the results and do a suitable parameter choice.

Summary. *We first proved 3.4, which says:*

$$\int f \log f d\mu + \int P_t f \log \frac{1}{P_t f} d\mu \leq 2\sqrt{t \log\left(\frac{1}{\epsilon}\right)} \int \sqrt{\Gamma(f)} d\mu$$

Then we showed that if δ and r satisfy some conditions, we have 3.9 for $f \in \mathcal{A}$

$$\left(\int f d\mu - Ce^{-2cr\sqrt{t}}\right) \log \frac{1}{\delta} \leq \int P_t f \log \frac{1}{P_t f} d\mu$$

and we then showed 3.11 which states

$$\int \sqrt{\Gamma(\mathbf{1}_A)}d\mu \leq \mu^+(A)$$

Now notice that $\epsilon \leq f \leq 1$ directly implies that

$$-\epsilon \log \frac{1}{\epsilon} \leq \int f \log f d\mu$$

Now we can put all of this bounds together, meaning that both terms in the left of 3.4 are bounded from below and the last term on the right is bounded from above so we get

$$-\epsilon \log \frac{1}{\epsilon} + \left(\mu(A) - Ce^{-2cr\sqrt{t}}\right) \log \frac{1}{\delta} \leq 2\sqrt{t \log\left(\frac{1}{\epsilon}\right)}\mu^+(A) \tag{3.12}$$

Furthermore, as $0 < \epsilon < 1$ we know $(1-\epsilon)\mu(A) \leq \mu(A)$ so we have

$$-\epsilon \log \frac{1}{\epsilon} + \left((1-\epsilon)\mu(A) - Ce^{-2cr\sqrt{t}}\right) \log \frac{1}{\delta} \leq 2\sqrt{t \log\left(\frac{1}{\epsilon}\right)}\mu^+(A) \tag{3.13}$$

Parameter choice

Now we can choose δ, r, ϵ as long as they satisfy the conditions in the lemmata.
Let us pick

- $\epsilon = \mu(A)^2$
- $r^2 = \frac{1}{4}\log\frac{1}{\mu(A)} = \frac{1}{8}\log\frac{1}{\epsilon}$
- $0 < \sqrt{2}\mu(A)^{1/4} \leq \delta \leq 2\mu(A)^{1/4}$.
  But to ensure that δ is smaller than 1 we need $2\mu(A)^{1/4} \leq 1$ i.e. $\mu(A) \leq 16$.

Let us first verify that this choice is valid in the sense that δ, r, ϵ meet their conditions.
For ϵ we just need $0 < \epsilon \leq 1$ which holds trivially as $\mu(X) = 1$.
Our condition for r is that $0 \leq r \leq \sqrt{\frac{1}{2}\log\frac{2}{\int f d\mu}} = \sqrt{\frac{1}{2}\log\frac{2}{\mu(A)}}$ but $\frac{1}{4}\log\frac{1}{\mu(A)} \leq \frac{1}{2}\log\frac{2}{\mu(A)}$ so this one also holds.
For δ we required $\delta \geq \sqrt{2\int f d\mu e^{r^2}} = \sqrt{2\mu(A)}e^{(1/4)\log\frac{1}{\mu(A)}} = \sqrt{2\mu(A)}\frac{1}{\mu(A)^{1/4}} = \sqrt{2}\mu(A)^{1/4}$
And so every condition is satisfied and we can plug in this values in 3.13 and we get

$$-2\mu(A)^2 \log\frac{1}{\mu(A)} + \log\frac{1}{\delta}\left[(1-\mu(A)^2)\mu(A) - Ce^{-2cr\sqrt{t}}\right] \leq 4\sqrt{2}r\sqrt{t}\mu(A)^+$$

And $\delta \leq 2\mu(A)^{1/4}$ implies $\log\frac{1}{\delta} \geq \frac{1}{4}\log\frac{1}{16\mu(A)}$ so we can further lower bound our term and get

$$-2\mu(A)^2 \log\frac{1}{\mu(A)} + \frac{1}{4}\log\frac{1}{16\mu(A)}\left[(1-\mu(A)^2)\mu(A) - Ce^{-2cr\sqrt{t}}\right] \leq 4\sqrt{2}r\sqrt{t}\mu(A)^+$$

But now as $\mu(A)^2 < \frac{1}{2}$ we have

$$-2\mu(A)^2 \log \frac{1}{\mu(A)} + \frac{1}{4} \log \frac{1}{16\mu(A)} \left[\frac{\mu(A)}{2} - Ce^{-2cr\sqrt{t}} \right] \leq 4\sqrt{2} r \sqrt{t} \mu(A)^+ \tag{3.14}$$

Similarly, for 3.1.2 we have

$$\mu(A) - \left(2\mu(A) + \frac{r}{\sqrt{2t}} \right)^2 - \alpha_\mu(r) \qquad \mathbf{(3.14^*)}$$

And finally we choose t such that $r\sqrt{t} = \frac{1}{2c} \log \frac{4C}{\mu(A)}$ and obtain

$$-2\mu(A)^2 \log \frac{1}{\mu(A)} + \frac{1}{16} \log \frac{1}{16\mu(A)} \mu(A) \leq 2\sqrt{2} \log \left(\frac{4C}{\mu(A)} \right) \mu^+(A)$$

From which we obtain that there exists c' such that if $0 < \mu(A) < c'$ we have

$$c' \mu(A) \leq \mu^+(A)$$

.

So we have proved that if $x = \mu(A)$ we have

$$c' x \leq I_\mu(x)$$

Proposition 3.1.15. *(At most linear growth of the isoperimetric profile)*
For the isoperimetric profile I_μ, if $\mathrm{Ricc}_\psi \geq 0$ we have

$$\frac{I_\mu(x)}{x} \text{ is a non-increasing function of } x$$

The proof is the main ingredient of the next chapter, namely 4.3.1

Now if $\frac{I_\mu(x)}{x}$ is non-increasing, suppose that $x > 1 - x$ then

$$c' \leq \frac{I_\mu(x)}{x} \leq \frac{I_\mu(x)}{1-x}$$

So finally we arrive to

$$c' \min\{x, 1-x\} \leq I_\mu(x)$$

This concludes the proof of theorem 3.1.1 as for the case $1 - x > x$ we just reverse the roles. We remark that almost the same proof works (by changing the exponential for α and picking suitable t) to prove theorem 3.1.2.

3.2 Semigroup proof for superlinear Isoperimetry

The next step is to try to achieve a different isoperimetric type inequality, this is: under curvature conditions we can bound the isoperimetric profile by a superlinear function. More, specifically,
When $\alpha_\mu \to 0$ as $r \to \infty$ denote by $r(x)$ the smallest r such that $\alpha_\mu(r) < x$

Theorem 3.2.1. *[14] [Theorem 2] Under $\Gamma_2 \geq 0$, if $\alpha_\mu(r) \to 0$ as $r \to \infty$ then*

$$I_\mu(x) \geq \frac{c}{r(x)} x \log \left(\frac{1}{x} \right)$$

Example 3.2.2. *The case of the Gaussian*
We use $f(x) \sim g(x)$ *as* $x \to a$ *to denote* $\lim_{x\to a} \frac{f(x)}{g(x)} = 1$*. We want to analyze Theorem 3.2.1 in terms of the Gaussian measure. Using the notation on 2.4.4, by integration by parts one can get the well known-bound:*

$$\frac{1}{\sqrt{2\pi}}\left(\frac{1}{x}-\frac{1}{x^3}\right)e^{-x^2/2} \leq 1-\Phi(x) \leq \frac{1}{\sqrt{2\pi}x}e^{-x^2/2}$$

So as $\Phi(-x) = 1 - \Phi(x)$*, we get*

$$\Phi(x) \sim \frac{1}{|x|}\varphi(x) \text{ as } x \to -\infty \tag{3.15}$$

Our goal is to understand the asymptotic behaviour of the isoperimetric function $\varphi \circ \Phi^{-1}$*. Therefore we need to analyze the asymptotics of* Φ^{-1}*.*
Using 3.15 for $\Phi\left((1+\epsilon)\sqrt{2\log 1/x}\right)$ *and* $\Phi\left((1-\epsilon)\sqrt{2\log 1/x}\right)$ *we obtain that* $\Phi^{-1}(x) \sim -\sqrt{2\log 1/x}$ *as* $x \to 0$*.*
Now if $x = \Phi(y)$ *then by 3.15*

$$x \sim \frac{1}{|y|}\varphi(y) = \frac{1}{\Phi^{-1}(x)}\varphi(\Phi^{-1}(x))$$

Therefore we get that the asymptotic of the isoperimetric function $\varphi \circ \Phi^{-1}$ *is*

$$\varphi \circ \Phi^{-1}(x) \sim x\Phi^{-1}(x) \sim -x\sqrt{2\log 1/x} \text{ as } x \to 0$$

As the right hand side is the one on 3.2.1, we conclude that the result is sharp for this case, showing the power of this theorem.

Before analyzing the details of the proof, let us compare it to the result proved by Milman in [16].

3.2.1 Comparison of the results

To be able to truly compare the results let us write another definition.

Definition. *(Log-concentration profile) For a metric probability space* (X, d, μ) *we define the log-concentration profile* $\mathcal{K} : [0, \infty) \to \mathbb{R}$ *to be defined as the maximal function such that*

$$1 - \mu(A_r) \leq \exp(-\mathcal{K}(r)) \;\; \forall A \in \mathcal{B}(X)$$

Theorem 3.2.3. *[16] [Theorem 1.1]*
Let β *be a non-decreasing function.*
Suppose $\Gamma_2 \geq 0$ *holds.*
If there exists M *such that* $\mathcal{K}(r) \geq \beta(r) \;\forall r > M$ *then*

$$I_\mu(x) \geq \min\left\{c\, x\, \gamma\left(\log\frac{1}{x}\right), c_\beta\right\}$$

where $\gamma(x) = \frac{x}{\beta^{-1}(x)}$ *and* c_β *is a positive constant depending only on* β*.*

Observation 3.2.4. *Note that the result by Milman 3.2.3 has a different concentration hypothesis. As soon as the log-concentration function is bounded from below (in the tails* $(r > M)$ *we get a stronger bound that combines the right handside of the previous results of the section. Milman is able to show that the superlinearity is obtained in terms of how fast the inverse of the lower bound* (β) *of* $\mathcal{K}$ *grows. This agrees* *with 3.2.1 as it provides the same bound if* $cx\gamma(\log\frac{1}{x}) > c_\alpha$ *but strengths it in the second case. Of course obtaining this stronger result requires the geometric content explained in the next chapter, as we did in 4.3.1, in this sense the result by Ledoux 3.2.1 is more specific.*

Sketch of Proof of Theorem 3.2.1 Actually as we developed theorem 3.1.1 we developed most of the needed tools, therefore we can skip the steps that are not changed and focus on the differences. Observe that the main difference is the appearance on the logarithmic term in the lower bound, so we can start already from 3.16 where we can replace the exponential concentration by α_μ so we have

$$-2\mu(A)^2\log\frac{1}{\mu(A)}+\frac{1}{4}\log\frac{1}{16\mu(A)}\left[\frac{\mu(A)}{2}-\alpha_\mu(r)\right]\leq 2\sqrt{2}r\mu(A)^+ \tag{3.16}$$

So let us write $r^* = \inf\{r > 0 : \alpha_\mu(r) \leq \frac{\mu(A)}{4}\}$ and so using this bound we have

$$-2\mu(A)^2\log\frac{1}{\mu(A)}+\frac{1}{4}\log\frac{1}{16\mu(A)}\left[\frac{\mu(A)}{2}-\frac{\mu(A)}{4}\right]\leq 2\sqrt{2}r^*\mu(A)^+$$

This means that the lower bound has two terms, one of the type $\mu(A)^2\log\frac{1}{\mu(A)}$ and one like $\mu(A)\log\frac{1}{\mu(A)}$. So close to zero the latter dominates, meaning that everything in the left-hand side is of the order of $\mu(A)\log\frac{1}{\mu(A)}$, and so there exists c such that

$$\frac{c}{r^*}\mu(A)\log\frac{1}{\mu(A)}\leq\mu^+(A)$$

so let us write $\mu(A) = x$, this means that if x is small enough, say $x \leq a$ we have

$$I_\mu(x)\geq\frac{c}{r(x)}x\log\frac{1}{x}\text{ whenever }x\leq a$$

Nevertheless, this concludes the theorem for $x \in (0, a)$, to complete the proof we must argue how to deal with $a \leq x \leq \frac{1}{2}$

Lemma 3.2.5. *(Of concentration)*
Let B be a measurable set with $\mu(B) \geq \rho > 0$ then for all $r > 0$ we have

$$1-\mu(B_{r+r_0})\leq\alpha_\mu(r)$$

where r_0 satisfies $\alpha_\mu(r_0) < \rho$

For $\tau > 0$ Let $B_\tau = \{x : P_t\mathbf{1}_A \leq (1+\tau)\mu(A)\}$ Note that $P_t\mathbf{1}_A$ is well defined as our semigroup is defined on all bounded measurable functions.
Find τ such that $\mu(B_\tau) \geq \frac{\tau}{1+\tau} = \rho$. A similar argument to the one in **3.14*** gives

$$2\sqrt{2t}\mu^+(A)\geq\mu(A)-\left((1+\tau)\mu(A)+\frac{r+r_0}{\sqrt{2t}}\right)^2-\alpha_\mu(r)$$

and finally a similar choice of parameters gives the result.

Chapter 4

The geometric side

4.1 Geometric side

In this chapter we aim to understand some of the ideas presented in [16], for understanding the isoperimetric profile. Even though the main goal is understanding further properties for the last proof of chapter 3, the chapters are separated because this one will have a more geometric feeling. It seems that to be able to completely reverse the hierarchy, one needs not only to understand the underlying semigroup but also the underlying geometry. Therefore this chapter presents the geometric aspects of the Isoperimetric profile, concluding with a result on [16].
The main idea is to get generalizations of Rauch's theorem, presented in chapter 1. The generalization of Rauch's theorem to a specific kind of manifolds will provide the essential tool to understanding the isoperimetric function, therefore we analyze with detail this generalizations.
Before we go into details we need some more definitions from Riemannian geometry.

4.2 Comparison theorems

In our objective of understanding the underlying geometry of the Isoperimetric profile in a manifold, we will encounter comparison theorem's for Jacobi fields. This theorems are generalization's of a theorem by Rauch.

4.2.1 Rauch's comparison theorem

Theorem 4.2.1. *(Rauch's Theorem)*
Let M be a submanifold of dimension n of a Riemannian manifold $\tilde{M}$ of dimension $n+k$. Let $\gamma : [0,\alpha] \to M$ and $\tilde{\gamma} : [0,\alpha] \to \tilde{M}$ be geodesics with the same velocity (i.e. $|\dot{\gamma}(t)|)|\dot{\tilde{\gamma}}(t)|)$.
Let J and $\tilde{J}$ be ***Jacobi fields*** *along γ and $\tilde{\gamma}$ such that*

- $J(0) = \tilde{J}(0) = 0$
- $|\dot{J}(0)| = |\dot{\tilde{J}}(0)|$
- $\langle \dot{J}(0), \dot{\gamma}(0)\rangle = \langle \dot{\tilde{J}}(0), \dot{\tilde{\gamma}}(0)\rangle$

Assume that $\tilde{\gamma}$ has ***no*** *conjugate points on $[0, \alpha]$ and assume that if K denotes the sectional curvature and for all $x \in T_{\gamma(t)}M, \tilde{x} \in T_{\tilde{\gamma}(t)}M$ one has*

$$\tilde{K}(\tilde{x}, \dot{\tilde{\gamma}}(t)) \geq K(x, \dot{\gamma}(t))$$

Then

$$|\tilde{J}| \leq |J|$$

Observation 4.2.2. *Rauch's theorem explains the interplay between sectional curvature and Jacobi fields, it shows how the spread of geodesics changes as sectional curvature changes.*

The proof can be found in [5] chapter 10 but relies heavily in the following Lemma which will be very useful for the generalizations.

4.2.2 The index Lemma

Definition. *(Index of a geodesic at a point)*
In a Riemannian manifold, for a geodesic γ and a piecewise differentiable vector field, we define the index of γ at $t_0 \in [0, a]$ as

$$I_{t_0}(V, V) = \int_0^{t_0} \langle V, V \rangle - \langle R(\dot{\gamma}, V)\dot{\gamma}, V \rangle dt$$

Lemma 4.2.3. *(Index Lemma)*
In a Riemannian manifold with a geodesic $\gamma : [0, a] \to M$ without conjugate points to $\gamma(0)$, let J be a Jacobi field with $\langle J, \dot{\gamma} \rangle = 0$.
Let V be a piecewise differentiable vector field, also parallel, i.e. $\langle V, \dot{\gamma} \rangle = 0$ then if J and V coincide in $t = 0$ and $t = t_0$ we have

$$I_{t_0}(J, J) \leq I_{t_0}(V, V)$$

For a proof see [5] pg 213.

Observation 4.2.4. *The index lemma states that Jacobi fields minimize the index. The index lemma is the key to proving 4.2.1 and also 4.2.6.*

Now we focus on a generalization, or a modified version of 4.2.1. We are now going to use Jacobi fields to get estimates on the volume of our manifold. This technique will be fully detailed and will be referred to as **Jacobi fields estimates**. The theorem presented is due to Heintze and Karcher and can be found in it's original presentation in [13].
For this theorem we need a different version of curvature assumptions, which we call curvature hypothesis (K), let N be isometrically immersed in M be a compact sub-manifold of the Riemannian manifold M.

Definition. *(Curvature hypothesis (K))*
We say that N satisfies curvature hypothesis (K) if one of the following holds:

- *If N is a hypersurface or a point, the Ricci curvatures of M for the tangent vectors of minimizing normal geodesics are bounded from below by some δ.*
- *If N has arbitrary codimension, the planes of M containing a tangent vector of a geodesic segment which minimizes distance to N have sectional curvatures bounded from below by some δ.*

To simplify the writing of the theorem we denote following [13]:

$$s_\delta(r) = \begin{cases} \sin(r\delta^{1/2})\delta^{-1/2}r & \delta > 0 \\ r & \delta = 0 \\ \sinh(r|\delta|^{1/2})|\delta|^{-1/2} & \delta < 0 \end{cases}$$

And also we write $c_\delta(r) = s'_\delta(r)$

Observation 4.2.5. *The reader must see that $s_\delta(r)$ is exactly the function of Jacobi fields in manifolds with constant curvature (go back to A.5.3), this is the fundamental idea behind the proof. To get estimates in the Jacobi fields we can understand locally a function of what Jacobi fields would be if we had constant curvature.*

4.2.3 Heintze-Karcher Comparison theorem

Theorem 4.2.6. *(Heintze-Karcher comparison theorem)*
Let N be isometrically immersed in M,N, M compact Riemannian manifolds and N satisfies the curvature hypothesis (K) then

$$\mathrm{vol}(M) \leq \int_N f_\delta(d(M), H) d\,\mathrm{vol}_N \leq \mathrm{vol}(N) f_\delta(d(M), \Lambda)$$

where

- *$H(p)$ denotes the length of the mean curvature normal of N at p.*
- *$d(M)$ denotes the diameter of M.*
- $f_\delta(d(M), H(p)) = \displaystyle\int_{S^{m-n-1}} \int_0^{\min\{d(M), z(\eta,\xi)\}} (c_\delta(r) - \langle \eta(p), \xi \rangle s_\delta(r))^n s_\delta(r)^{m-n-1} dr d\xi$
- *η is the mean curvature normal of N at p.*
- *$z(\eta, \xi)$ is the first zero of the integrand.*
- *Λ is an upper bound for H*

Before diving into the proof, we will need more geometric concepts, so let us define them before they appear.

Now we have enough tools to dive into the proof of the theorem 4.2.6.

The main part of the theorem is being able to understand distortions under the normal exponential map. Remember that $|d\exp|$ shows the rate at which geodesics spread apart, in the same fashion $|d\exp^\perp|$ measures how fast different normal geodesics to N separate. To understand how we can estimate volumes in terms of the volume of N we need to see how things 'diverge' from N to M, this is how normal geodesics behave.

Definition. *(Curvature assumption (K2))*
Let N be isometrically immersed in N as before, ξ be vertical to N at p, define $Y_i(s) = d\exp \perp U_i(s)$ be linearly independent, where $U(s)$ is defined in A.5.7.
Let $c(s) = \exp(s\xi)$ note that it is not the normal exponential.
Let $k(s)$ and $K(s)$ be the minimum and maximum of the sectional curvatures of planes containing $c'(s)$. Let $\lambda_1, \cdots, \lambda_n$ be the eigenvalues of the Weingarten map.
Consider the same assumptions for $\overline{N} \subseteq \overline{M}$. We say that the curvature assumptions (K2) are satisfied if one of the following is satisfied

a) $k(s) \geq \overline{K}(s), 2 \leq r+1 \leq \dim \overline{M} \leq \dim M, \dim N = \dim \overline{N} = 0$

b) $k(s) \geq \overline{K}(s), 2 \leq r+1 \leq \dim \overline{M} \leq \dim M, \operatorname{codim} N = \operatorname{codim} \overline{N} = 1, \max \lambda_i \leq \min \overline{\lambda_i}$

c) $k(s) \geq \overline{K}(s), r+1 = \dim M = \dim \overline{M}, \dim N = \dim \overline{N}, \lambda_i \leq \overline{\lambda_i}$

d) $\operatorname{Ricc}(c', c') \geq (m-1)\delta, \overline{N}$ *totally umbilic (every point is locally spherical i.e. constant positive curvature) and* $\operatorname{tr} S_\xi \leq \operatorname{tr} \overline{S}_\xi$

Observation 4.2.7. *This curvature assumptions can be understood as the immersion $\overline{N}$ in $\overline{M}$ being similar to the one of N in M but having smaller sectional curvatures. This means that the immersions are similar in terms of dimensions but different in terms of the curvatures, where the sectional curvatures of c' along N are bigger than the ones along $\overline{N}$.*

Theorem 4.2.8. *(Main inequality, distortions of the differential of exponential maps for N-Jacobi fields) Suppose $N, \overline{N}$ compact Riemannian manifolds isometrically immersed in M and $\overline{M}$, also Riemannian manifolds. Suppose that the curvature assumptions (K2) are satisfied, then for any s smaller than the first focal distance, we have*

$$\frac{|Y_1(s) \wedge \cdots Y_r(s)|}{|U_1(s) \wedge \cdots U_r(s)|} \leq \frac{|\overline{Y}_1(s) \wedge \cdots \overline{Y}_r(s)|}{|\overline{U}_1(s) \wedge \cdots \overline{U}_r(s)|}$$

The proof is long and elaborate, and it is not the focus of this work, it can be found in [13]. But for completeness we provide a sketch of the main ideas.

Sketch of proof

The main idea is to prove an estimate of the type:

$$(\log |Y_1(s) \wedge \cdots Y_r(s)|)' \leq (\log |\overline{Y}_1(s) \wedge \cdots \overline{Y}_r(s)|)'$$

for s smaller than the focal distances of N and $\overline{N}$.

For this purpose, one can show that for an N-Jacobi field

$$\langle Y, Y' \rangle(0) = \langle Y, S_{c'(0)} Y \rangle(0)$$

therefore one can show

$$(\log |Y_1(s) \wedge \cdots Y_r(s)|)'(s_1) = \sum_{i=1}^{r} \langle Y_i, Y_i' \rangle(s_i)$$

can be rewritten in terms of the Weingarten map according to A.5.7, to get:

$$(\log |Y_1(s) \wedge \cdots Y_r(s)|)'(s_1) = \sum_{i=1}^{r} \langle Y_i, S_{c'(0)} Y_i \rangle(0) + \int_0^{s_1} (\langle Y_i', Y_i' \rangle - \langle R(Y_i, c')c', Y_i \rangle ds$$

This term in the right is the index 4.2.2, so if $I_s(Y_i, Y_i)$ denotes the right hand side, then it is needed to show:

$$\sum_{i=1}^{r} I_s(Y_i, Y_i) \leq \sum_{i=1}^{r} I_s(\overline{Y}_i, \overline{Y}_i)$$

which is done by using the index lemma 4.2.2.

The next section presents a corollary for this comparison theorem, it is the main key that will allow us to prove 4.2.6.

Corollary 4.2.9. *(KEY COROLLARY)*
For the setting as above, if we apply case (d) of the curvature assumption (K2) and if s is smaller than the first focal distance of N in direction of unitary ξ then

$$|\det(d\exp^{\perp})(s\xi)|s^{m-n-1} \leq (c_\delta(s) - \langle \eta, \xi \rangle s_\delta(s))^n s_\delta(s)^{m-n-1} \tag{4.1}$$

The proof of this corollary is just the explicit computation of the distortion used in the last theorem when the second ones are assumed to be of constant sectional curvature. In this case the determinant of the wedge product of Y is by definition the one of the exponential functions, while

$$|U_1(s) \wedge \cdots \wedge U_r(s)| = Cs^{\operatorname{codim} N - 1}$$

where C is just a constant. This clearly explains why the term s^{m-n-1} appears, the right hand side comes from the example A.5.3 and this observation together.

Observation 4.2.10. *Note that as ξ is unitary, so we have $s = |s\xi|$.*

Let us return to the proof of 4.2.6, recall that we **want to show**:

$$\operatorname{vol}(M) \leq \int_N f_\delta(d(M), H) d\operatorname{vol}_N \leq \operatorname{vol}(N) f_\delta(d(M), \Lambda)$$

So let us start by estimating the volume of M, as $NB \to N$ is a Riemannian submersion, we use Fubini's theorem so restricting to U the set of al normal vectors ξ such that $|\xi| < \min\{d(M), s\}$ where s is the first focal distance of N (to satisfy the hypothesis of the last corollary) we have by integrating first over a fibre and then over the base N we get:

$$\operatorname{vol}(M) \leq \int_U |\det(d\exp^{\perp})| d\operatorname{vol}_{NB}$$

We can apply the corollary 4.2.9 and using the observation 4.2.10 if ξ is not unitary but $|\xi| = r$ then we have

$$\operatorname{vol}(M) \leq \int_U (c_\delta(r) - \langle \eta, \xi/r \rangle s_\delta(r))^n s_\delta(r)^{m-n-1} r^{-(m-n-1)} d\operatorname{vol}_{NB}$$

But the first zero of f_δ, say $z(\eta, \xi)$ (see [13] pg 459 for details) is an upper bound for the last term in 4.2.9, so we can restrict the integral to $[0, z(\eta, \xi)]$ giving us the exact expression of f_δ, so by applying the corollary 4.2.9 we get the desired inequality:

$$\operatorname{vol}(M) \leq \int_N f_\delta(d(M), H) d\operatorname{vol}_N$$

For the last step to be true we can show that $f_\delta(x, \cdot)$ is monotone increasing as a function of the second variable. This is done by using Newton's binomial for the terms $(x - Hy)^n + (x + Hy)^n$ and seeing that $d(M) = z(\eta, \xi)$ or $d(M) = z(\eta, -\xi)$, for details we address the original paper again. So suppose $H \leq \Lambda$, putting both results together we get the result:

$$\operatorname{vol}(M) \leq f_\delta(d(M), \Lambda) \operatorname{vol}(N)$$

Interpretation 4.2.11. *Now that we have proved the theorem it is important to take a moment to understand it's importance. We will focus on case (d) of the curvature assumption K2. This means that we are bounding the volume in terms of a lower bound of the Ricci curvature (δ) and the mean curvature. This is fundamental as we understand how to control volumes in terms of curvatures. In literature, the technique used for this proof is referred to as "integrating over infinitesimal wedges from N".*

The next theorem provides a generalization of Heintze-Karcher (4.2.6), to the setting of manifolds with densities discussed in 1.5 and it is due to Morgan [17]

Theorem 4.2.12. *(Generalized Heintze-Karcher to manifolds with densities)*
Let M be a smooth complete Riemannian manifold of dimension n with smooth density $\Psi = e^{\psi}$, assume that

$$\mathrm{Ricc} \geq (n-1)\delta \text{ and } -\mathrm{Hess}\,\psi \geq \gamma$$

Let S be a smooth, oriented, finite-area hypersurface in M with mean curvature H.
Let $V(r)$ be the volume of the region within distance r of S on the side of the unit normal. Then

$$V(r) \leq \int_S \int_0^{r*} [c_\delta(t) - H(s)s_\delta(t)]^{n-1} \exp\left(t\frac{d\psi}{dn}(s) - \frac{\gamma}{2}t^2\right) dtds$$

where ds is the weighted surface measure, and r^* is smaller than r and the first zero of the integrand.

Observation 4.2.13. *From 4.2.6 the first term is already familiar to us, the second term however changed, this is the influence of the density.*

Proof. If Ψ is a constant, indeed we have 4.2.6. Now, since $\mathrm{Hess}\,\psi \leq -\gamma$, by Taylor's theorem at a point distanced t to $s \in S$ in a geodesic normal to S, we have

$$\psi(s,t) \leq \psi(s,0) + t\frac{d\psi}{dn}(s,0) - \frac{\gamma}{2}t^2$$

By just considering the exponential in both sides, we get

$$\Psi(s,t) \leq \Psi(s,0) \exp\left(t\frac{d\psi}{dn}(s,0) - \frac{\gamma}{2}t^2\right)$$

So by following the same procedure as before but considering this influence of the density in f_δ the proof above would suffice.
The next theorem is similar but weaker, if we relax the hypothesis then we get a different estimate which is in some sense easier to follow.

Theorem 4.2.14. *(Relaxation of hypothesis in the generalization of Heintze-Karcher)*
Let M be a smooth complete Riemannian manifold of dimension n, with smooth density $\Psi = e^{\psi}$ satisfying

$$\mathrm{Ricc}_\psi \geq \gamma$$

Let S be a smooth oriented finite area hypersurface in M with generalized mean curvature:

$$H_\psi(s) = H(s) - \frac{1}{n-1}\frac{d\psi}{dn}$$

Let $V(r)$ denote the volume of the region within distance r of S on the side of the unit normal. Then

$$V(r) \leq \int_S \int_0^r \exp\left(-(n-1)H_\psi(s)t - \frac{\gamma}{2}t^2\right) dtds$$

where ds denotes the weighted surface area.

This theorem can be proved using the first and second variation formulas for weighted area if it is not understood within the generality of the last theorem. See [17].

Observation 4.2.15. *(Remark 3 in [17]) We can use this theorem for isoperimetric minimizers. Observe that when S is not smooth this theorem does not hold. Nevertheless it was observed by Gromov [12] and then Morgan [17] that the structure of the singularities of ∂A for an isoperimetric minimizer allows us to use this theorem.*

Now we apply this observation to see how this geometric estimates translate to our setup.

Corollary 4.2.16. *(Bringing 4.2.14 to our setup)*
By the last observation, 4.2.14 still holds while changing $V(r)$ for the measure of $A_r \setminus A$ so we get Let M be a smooth complete Riemannian manifold of dimension n, with smooth density $\Psi = e^{\psi}$ satisfying

$$\mathrm{Ricc}_{\psi} \geq \gamma$$

. Consider A an isoperimetric minimizer, let $H_{\mu}(A)$ denote the constant total curvature of the regular part of ∂A with respect to the outer unit normal vector then

$$\mu(A_r) - \mu(A) = \mu^+(A) \int_0^r \exp\left(H_{\mu}(A)t - \frac{\gamma}{2}t^2\right) dt$$

Observation 4.2.17. *It is usual to consider $\kappa = -\gamma$, to get in the notation:*

$$\mu(A_r) - \mu(A) = \mu^+(A) \int_0^r \exp\left(H_{\mu}(A)t + \frac{\kappa}{2}t^2\right) dt$$

4.3 Milman's use of the general Heintze-Karcher theorem

Finally we arrive to the goal of this section, namely we understand properties of the isoperimetric profile $I_{\mu}(v)$ by understanding the underlying geometry. For this purpose we use 4.2.14.

Theorem 4.3.1. *[16] [Proposition 3.1] If $\mathrm{Ricc}_{\psi} \geq 0$ the function $\frac{I_{\mu}(v)}{v}$ is non-increasing in $(0, 1)$.*

By means of more elaborate techniques, many have shown that I_{μ} is actually concave. But the main point of [16] is to show that such result is not the best as 4.3.1 is enough for most purposes.

Proof. Applying 4.2.14 to $B = X \setminus A$, and considering that $\gamma = 0$, by letting $r \to \infty$ we have $\mu(B_r) \to 1$ so

$$1 - \mu(B) \leq \mu^+(B) \int_0^{\infty} \exp(H_{\mu}(B)t)dt$$

But as $\mu^+(A) = \mu^+(B), H_{\mu}(B) = -H_{\mu}(A)$ we have

$$\frac{\mu(A)}{\mu^+(A)} \leq \int_0^{\infty} \exp\left(-H_{\mu}(A)\right) dt = \frac{1}{H_{\mu}(A)}$$

where to evaluate the integral we assumed wlog $H_{\mu}(A) > 0$. But this is the same as

$$H_{\mu}(A) \leq \frac{\mu^+(A)}{\mu(A)} = \frac{I_{\mu}(v)}{v} \tag{4.2}$$

Proposition 4.3.2. *(Lower bound on H)*
Let A be an isoperimetric minimizer of measure v then

$$\limsup_{\epsilon \to 0} \frac{I_{\mu}(v + \epsilon) - I_{\mu}(v)}{\epsilon} \leq H_{\mu}(A)$$

Lemma 4.3.3. *The regular part of the boundary of an isoperimetric minimizer must have constant total curvature.*

See [2] or [9] pg. 654, most of the ideas depend heavily in concepts from geometric measure theory and therefore are not included here.

Proof. of 4.3.2 Using the lemma, the definition of the graph of the isoperimetric minimizer cannot lie above $t \to (\mu(A_t), \mu^+(A_t))$, and as they touch at $t = 0$ they must be tangent at $(\mu(A), \mu^+(A))$ so

$$\limsup_{\epsilon \to 0} \frac{I_\mu(v+\epsilon) - I_\mu(v)}{\epsilon} \leq H_\mu(A) \tag{4.3}$$

Putting 4.2 together with 4.3 show

$$\limsup_{\epsilon \to 0} \frac{I_\mu(v+\epsilon) - I_\mu(v)}{\epsilon} \leq H_\mu(A) \leq \frac{I_\mu(v)}{v}$$

which implies $I_\mu(x)/x$ is not increasing.

Corollary 4.3.4. *If* $\mathrm{Ricc}_\psi \geq 0$ *then*

$$\inf_{v \in [0,1/2]} \min\{I_\mu(v), I_\mu(1-v)\} = 2I_\mu(1/2)$$

4.3.1 Discussion on existence of isoperimetric minimizers

For this proof to be totally complete, we would need to show that isoperimetric minimizers exist, this is no easy task. As seen in section 2.4.3, the existence of isoperimetric minimizers is a historic problem, nevertheless the methods of geometric measure theory provide the solution. The only problem is that in order to arrive to such conclusions one must first understand currents, which are a concept that requires a lot of care. Therefore we do not prove the existence of minimizers but refer the reader to [9] pg. 654 or 619.

Appendices

Appendix A

Background on Geometry

This thesis assumes a background in stochastics, nevertheless for the geometric concepts one can find a non-detailed explanation of all the results and concepts needed in this appendix.

A.1 Riemannian Manifolds

In this section we explain the basic definitions on Riemannian manifolds. As curvature bounds are one of the main topics of this thesis, we discuss the geometric tools necessary to understand the fundamentals of curvature. Also, we aim to understand diffusions of Markov processes in some Riemannian manifolds, therefore we need to have a clear image of the work space.

This introduction aims students with little to no knowledge in geometry, expert readers may easily skip this section.

Definition. *(Riemannian Manifolds)*
A second countable topological space M and a collection of open sets $\{U_i\}$ with functions $\phi_i : U_i \to \mathbb{R}^n$ are said to be an n-th dimensional Riemannian manifold if

- *Each ϕ_i is an homeomorphism (continuous with continuous inverse) into it's image.*
- *For every $p \in M$ there exists $U_p \in \{U_i\}$ such that $p \in U_i$*
- *If $U_i \cap U_j \neq \emptyset$ then $\phi_i \circ \phi_j^{-1} \in \mathcal{C}^\infty$*

A manifold is said to be closed if it is compact and has no boundary.
The functions ϕ_i are called charts, while the collection of charts is called an atlas.

Once the Riemannian manifolds are defined, we use their natural correspondence with the euclidean spaces (local coordinates) to understand smoothness.

Definition. *(Smoothness in Riemannian Manifolds)*
Let M and N be m and n Riemannian manifolds respectively, a function $f : M \to N$ is said to be smooth if

$$\phi_N \circ f \circ \phi_M^{-1} : \mathbb{R}^m \to \mathbb{R}^n \text{ is a smooth function}$$

Furthermore, let $\mathcal{C}^\infty(M, N)$ be the set of all smooth functions from M to N.

Usually when the second argument in $\mathcal{C}^\infty(\cdot,\cdot)$ is omitted it means $\mathbb{R}^m$. Now we use the definition of smoothness to obtain the most fundamental tool in smooth analysis: bump functions.

Lemma A.1.1. *(Extension Lemma)*
If $A \subseteq M$ is a closed set and $f : A \to \mathbb{R}$ is smooth then for an open set U, with $A \subseteq U$ there exists a smooth extension of f to all M such that

$$\hat{f}\Big|_A = f, \qquad \hat{f}\Big|_{M\setminus U} = 0$$

Definition. *(Tangent vector and tangent Space)*
For $p \in M$ a tangent vector at p is a map, $X : \mathcal{C}^\infty(M,\mathbb{R}) \to \mathbb{R}$ such that for any $f, g \in \mathcal{C}^\infty(M,\mathbb{R})$ one has

a) $X(af + g) = aX(f) + X(g)$

b) $X(fg) = f(p)X(f) + g(p)X(f)$

The set of all tangent vectors at p is denoted T_pM

Definition. *(Tangent Bundle)*
The disjoint union of all the tangent vectors is called the tangent bundle, namely

$$TM = \bigcup_{p\in M} T_pM$$

Definition. *(Tangent bundle)*
A vector field on M is a smooth map $X : M \to TM$ such that $X_p := X(p) \in T_pM$ for every $p \in M$. The set of all vector fields on M is denoted $\mathcal{T}M$

As for every p, T_pM is a finite dimensional vector space, then we can use it's dual for the following definition.

Definition. *(Cotangent bundle)*
We define the cotangent bundle as the union of the duals of the collections of tangent vectors, namely if $T_p^*M = (T_pM)^*$ the the cotangent bundle

$$T^*M = \bigcup_{p\in M} T_p^*M$$

Definition. *(Covector field)*
A covector field $\omega : M \to T^*M$ is a smooth map such that $\omega_p \in T_p^*M$ for every $p \in M$

Definition. *(Differential of a function)*
For a function $f : M \to \mathbb{R}$ the differential is defined as the covector field $df(X) = X(f)$

A.2 Tensor and Tensor bundle

In this section we discuss one of the most important tools to understand curvature: tensors, this idea arises as compact understanding of linear maps.

Definition. *((k,l) Tensor)*
A (k,l) tensor in a finite dimensional vector space V is a multilinear map $F : V^k \times (V^*)^l \to \mathbb{R}$. Furthermore, we denote by $T_l^k(V)$ the set of all (k,l) tensors of V.

The novel idea in this definition is that the space and it's dual are, in some way, working together.

Definition. *(Product tensor)*
Given a (k,l) tensor F and a (k',l') tensor G we define the product tensor $F \otimes G$ as the usual product of functions

$$F \otimes G(x_1, x_2, \ldots, x_{k+k'}, y_1, y_2, \ldots, y_{l+l'}) = F(x_1, \ldots, x_k, y_1, \ldots y_l) G(x_{k+1}, \ldots x_{k+k'}, y_{l+1}, \ldots, y_{l+l'})$$

Definition. *(Trace of a tensor)*

Let V be a finite dimensional vector space with a basis $e_1, \ldots, e_n$ and a dual basis $\phi^{(1)}, \ldots, \phi^{(n)}$ for a (k,l) tensor F we define the trace of F as a $(k-1, l-1)$ tensor

$$\operatorname{tr} F(X_1, \ldots, X_{k-1}, \omega^1, \ldots, \omega^{l-1}) = F(e_i, X_1, \ldots, X_{k-1}, \phi^{(i)}, \omega^1, \ldots, \omega^{l-1})$$

Definition. *(Tensor bundle)*
The tensor bundle is the union of all the spaces of tensors, namely $\bigcup_{p \in M} T_l^k(T_pM)$.

Definition. *(Riemannian metric)*
A Riemannian metric g on M is a $(2,0)$ tensor such that

- *$X, Y \in \mathcal{T}M$ then $g(X,Y) = g(Y,X)$*
- *For every $p \in M$, $X \in T_pM, X \neq 0$ we have $g(X,X) > 0$*

One should be familiar with these properties as they resemble the properties of an inner product, nevertheless the necessity of the bundles and tensors arises as we are not working on the space itself, but using the differential properties.
Note that an inner product in T_pM induces an isomorphism between T_pM and T_p^*M via $Y \to \langle \cdot, Y \rangle$.

A.3 Connections

A Riemannian connection on a Riemannian manifold is a map, $\nabla : \mathcal{T}M \times \mathcal{T}M \to \mathcal{T}M$, denoted $(X,Y) \to \nabla_X Y$ such that

- $\nabla_X Y$ is a $\mathcal{C}^\infty(M)$-linear map in the first coordinate: $\nabla_{fX_1 + gX_2} Y = f \nabla_{X_1} Y + g \nabla_{X_2} Y$
- $\mathbb{R}$-linear in the second coordinate: $\nabla_X(\alpha Y_1 + Y_2) = \alpha \nabla_X Y_1 + \nabla_X Y_2$
- Diffusion for $f \in \mathcal{C}^\infty(M)$: $\nabla_X(fY) = f \nabla_x Y + (Xf)Y$
- $X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle$
- $(\nabla_X Y - \nabla_Y X) f = X(Yf) - Y(Xf)$

The connection $\nabla_X Y$ can be understood as the derivative along X of Y.
While the first and the second properties seem straight-forward, a comment is to be made about the third and fifth ones. The third should remind the reader the diffusion property of the Carré du Champ operator, while the last one represents some kind of torsion along derivatives.
Then next theorem is a well-known result in the theory of Riemannian manifolds and shows the deterministic structure of this manifolds.

Theorem A.3.1. *(Levi-Cita/ Fundamental lemma of Riemannian Geometry)*
*Every Riemannian manifold admits a **unique** Riemannian connection.*

A nice proof can be found in [15] pg. 68.

Definition. *(Hessian)*
For a function $f \in \mathcal{C}^\infty(M)$ *is the* $(2,0)$ *tensor defined by* $\operatorname{Hess} f = \nabla^2 f = \nabla(\nabla f)$

Definition. *(Laplace-Beltrami) We define the Laplace Beltrami operator as the trace of the Hessian, in other words* $\triangle : \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$ *given by*

$$\triangle f = \operatorname{tr}(\operatorname{Hess} f)$$

A.4 Integration on manifolds

Definition. *(Integration for functions supported on charts)*
Let $f \in \mathcal{C}^\infty(M)$ *be such that* $\operatorname{supp}(f) \subset U_i$ *for some chart* U_i. *We define the integral as*

$$\int_M f dV = \int_U f \sqrt{|\det g|} dx$$

Even though it takes some care to see that this definition is enough to extend this integral to all functions, for the matters of this thesis we refer to [15] (pg. 29) and define the linear extension to be the integral.

A.5 Curvature

This section is the whole point of this geometric introduction as it will give us the definitions and intuitions necessary to comprehend the situation of Riemannian manifolds and curvature bounds.

Definition. *(Lie Bracket)*
For $X, Y \in \mathcal{T}M$, *the Lie Bracket or Jacobi-Lie bracket is a new vector field defined by*

$$[X, Y]f = X(Yf) - Y(Xf)$$

The Lie bracket can be thought as the derivative of Y along the **flow** generated by X.
Now we turn our heads to the concept of curvature, a great introduction to the topic can be found in [15], where many of this definitions come from. It provides the intuition and formality necessary.
Define a map $R : \mathcal{T}M \times \mathcal{T}M \times \mathcal{T}M \to \mathcal{T}M$ by the rule

$$R(X, Y, Z) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$$

Definition. *(Riemannian curvature endomorphism)*
The Riemannian curvature endomorphism is defined as the $(3,1)$ *tensor defined by*

$$R(X, Y, Z, \omega) = \omega(R(X, Y, Z))$$

This measures how much our manifold differs from a euclidean space from the view of ω.

Definition. *(Riemannian curvature tensor)*
The Riemannian curvature tensor is the $(4,0)$ *tensor*

$$R(X, Y, Z, W) = \langle R(X, Y, X), Z \rangle$$

Now we proceed to one of the fundamental ways to understand curvature.

Definition. *(Sectional Curvature)*
For X, Y linearly independent in T_pM we define the sectional curvature $K(X,Y)$ as

$$K(X,Y) = \frac{\langle R(X,Y)Y, X\rangle}{||X||^2||Y||^2 - \langle X,Y\rangle^2}$$

Definition. *(Ricci Curvature)*
The Ricci curvature is the $(2,0)$ tensor obtained by contracting the Riemannian curvature endomorphism, namely

$$\mathrm{Ricc}(X,Y) = R(e_i, X, Y, \phi^i)$$

One can also define the weighted Ricci curvature as

$$\mathrm{Ricc}_W(X,Y) = \mathrm{Ricc}(X,Y) + (\nabla^2 W)(X,Y)$$

Definition. *(Mean curvature)*
The mean curvature H is defined as

$$H = \frac{1}{n-1}\,\mathrm{tr}(\mathrm{I\!I})$$

where $\mathrm{I\!I}$ is the second fundamental form.

Observation A.5.1. *Curvature is meant to be interpreted as determining how fast geodesics spread apart.*

Definition. *(Geodesic)*
A parametrized curve $\gamma : [0,l] \to M$ is is called a geodesic if $\nabla_{\dot\gamma}\dot\gamma = 0$

Observation A.5.2. *Geodesics are locally (not globally) minimizing, meaning that if $\gamma : [0,t] \to M$ is a geodesic and $c : [0,t] \to M$ is a different curve, both starting at p, then there exists r such that $\ell(\gamma|_{[0,r]}) \leq \ell(c|_{[0,r]})$, where ℓ denotes the length.*

Definition. *(Differential of a map)*
Let $\varphi : V \to W$ be a differentiable map of an open set $V \subset M$ into W. For $p \in V$, $w \in T_pM$ we have that w is the velocity vector $\alpha'(0)$ of a differentiable parametrized curve α with $\alpha(0) = p$. Define $\beta = \varphi \circ \alpha$, so $\beta(0) = \varphi(p)$, therefore $\beta'(0) \in T_{\varphi(p)}W$.
Then $B'(0)$ does not depend on α and we call the differential of φ at p, the linear function $d\varphi_p : T_pM \to T_{\varphi(p)}W$ given by

$$d\varphi_p = \beta'(0)$$

Definition. *(Exponential map)*
Given $p \in M$ there exists a neighborhood V of p in M, $\epsilon > 0$ and a $\mathcal{C}^\infty(M)$ mapping $\gamma : (-2,2) \times \mathcal{V} \to M$ such that $t \to \gamma(t,q,w)$ is the unique geodesic that at instant $t = 0$ passes through q with velocity w, for all $q \in V$
where $\mathcal{V} = \{(q,w) \in TM | q \in V, w \in T_qM, |w| < \epsilon\}$.
We define the exponential map $\mathcal{V} \to M$ to be given by

$$\exp(q,v) := \exp_q(v) := \gamma(1,q,v)$$

The exponential map can be thought of starting a geodesic and see were it takes the initial point after a unit of time.
We now introduce the concept of Jacobi fields, which is the fundamental concept generalized in chapter 4. Jacobi fields are vector fields along geodesics.

Notation simplification ($\frac{D}{dt}$: Let M be a differentiable manifold, with affine connection ∇, let c be a curve, and let V be induced by c, i.e. $V(t) = Y(c(t))$ for some vector field Y, we now write

$$\frac{DV}{dt} = \nabla_{\dot{c}} Y$$

Definition. *(Jacobi field)*
In the same context as before, if $\exp_p$ *is defined at* $v \in T_pM$*, and* $w \in T_v(T_pM)$

$$(d\exp_p)_v(w) = \frac{\partial f}{\partial s}(1,0)$$

where f *is the parametrized surface given by* $f(t,s) = \exp_p tv(s)$*.*
Then f *satisfies* $\frac{D}{dt}\frac{\partial f}{\partial s} = 0$*.*
Let $J(t) = \frac{\partial f}{\partial s}(t,0)$*, we have*

$$\frac{D^2 J}{dt^2} + R(\dot{\gamma}(t), J(t))\dot{\gamma}(t) = 0$$

This last equation is called the Jacobi equation.
A Jacobi field J *is a vector field satisfying the Jacobi equation.*

The next example illustrates the concept of Jacobi fields but it's interesting as it makes the proofs on Chapter 4, more intuitive.

Example A.5.3. *(Jacobi fields in constant curvature)*
In a Riemannian manifold with constant curvature K*, let* $w(t), |w(t)| = 1$ *be a parallel field along* γ *(i.e.* $g(\dot{\gamma}(t), w(t)) = 0$ *then Jacobi fields are given by*

$$J(t) = \begin{cases} \sin(tK^{1/2})K^{-1/2}w(t) & K > 0 \\ tw(t) & K = 0 \\ \sinh(t|K|^{1/2})|K|^{-1/2}w(t) & K < 0 \end{cases}$$

For the proof see [5].

One of the main results of Jacobi fields is that they can be understood in terms of the differential of the exponential map, namely the following proposition.

Proposition A.5.4. *(Characterization of Jacobi fields)*
Let $\gamma : [0,a] \to M$ *be a geodesic, then a Jacobi field* J *along* γ *is given by*

$$J(t) = (d\exp_{t\gamma'(0)})(tJ'(0))$$

For the proof see the corollary 2.5 on pg 114 of [5]

Definition. *(Wedge product)*
A k*-linear function* f *is said to be alternating if for any permutation* σ *of* $\{1, \cdots, k\}$
$f(v_{\sigma(1)}, v_{\sigma(2)}, \cdots, v_{\sigma(k)}) = sign(\sigma) f(v_1, v_2, \cdots, v_k)$

For f *a* k*-linear function, the alternating version of* f *is defined as*

$$(Af)(v_1, v_2, \cdots, v_k) = \sum_{\sigma \in S_k} sign(\sigma) f(v_{\sigma(1)}, v_{\sigma(2)}, \cdots, v_{\sigma(k)})$$

where S_k *is the set of all permutations of* $\{1, \cdots, k\}$*.*

For f, a k-linear function and g an l-linear function, the tensor product is defined as

$$f \otimes g(v_1, v_2, \cdots v_{k+l}) = f(v_1, v_2, \cdots, v_k) g(v_{k+1}, v_{k+2}, \cdots, v_{k+l})$$

And finally, the wedge product between f and g, denoted $f \wedge g$ is defined to be the normalized alternating version of the product tensor. In symbols this is

$$f \wedge g = \frac{1}{k!l!} A(f \otimes g)$$

Observation A.5.5. *The wedge product of two vectors is interpreted as the whole area of the parallelepiped generated by the vectors, if the vectors are functions we have the surface generated by the both.*

Definition. *(Determinant)*
The determinant of a linear map T is the unique scalar $\det(T)$ such that for any basis $\{v_1, v_2, \cdots, v_n\}$ one has

$$T(v_1) \wedge T(v_2) \wedge \cdots \wedge T(v_n) = \det(T)(v_1 \wedge v_2 \wedge \cdots \wedge v_n)$$

Definition. *(Conjugate point)*
For a curve γ in a manifold with a connection, we say that $\gamma(t_0)$ is conjugate to $\gamma(0)$ along γ if there exists a Jacobi field J such that $J \neq 0$ and

$$J(0) = 0 = J(t_0)$$

The notion of conjugate point gives rise to the concept of focal point in a Riemannian manifold.

Definition. *(Focal point)*
Let N be a sub-manifold of a Riemannian manifold M, then $q \in M$ is called a focal point of N if there exists a geodesic $\gamma : [0, l] \to M$, $\gamma(0) = p \in N$, $\dot{\gamma}(0) \in (T_pN)^{\perp}, \gamma(l) = q$ and a non-zero Jacobi field J such that

- $J(l) = 0$
- $J(0) \in T_pN$
- $\dot{J}(0) + S_{\dot{\gamma}(0)}(J(0)) \in (T_pN)^{\perp}$

where S is the linear operator on T_pN of the second fundamental form of N.

Example A.5.6. *If $M = S^n$ and $N = S^{n-1}$ the equator of M, then $(0, 0, \cdots, 1)$ and $(0, 0, \cdots, -1)$ are the focal points of S^{n-1} in S^n*

Nevertheless this definition depends on the second fundamental form and the Jacobi equation, so it is very useful to have a different representation. For this we need a variation of the exponential map that works in the setting of manifolds and sub-manifold, meaning that we need a version of the exponential map that takes into account if we are seeing an object as a sub-manifold. For this fashion we define the normal exponential and for this we need the concept of normal bundle.

Definition. *(Normal bundle)*
Given an isometric immersion N in M for each $p \in N$ we have

$$T_pM = T_pN \oplus (TpN)^{\perp}$$

We call the normal bundle NB

$$NB = \bigcup_{p \in N} (T_pN)^{\perp}$$

Induced from the connection in M we get a normal connection $\nabla^{\perp}$ for NB. This is used to split the tangent vectors to the normal bundle:

$$T(NB) = \mathcal{H} + \mathcal{V}$$

where the elements in $\mathcal{H}$ are the ones $\nabla^{\perp}$ parallel, this means they are tangent to curves in NB. Each $\mathcal{V}_\xi$ is identified with a fibre and each $\mathcal{H}_\xi$ can be identified with a set of tangent vectors of N via the projection (say π_*), we get that we can have a canonical metric for $T_\xi NB$:

$$||u|| = ||\pi_* u||^2 + ||u - \pi_* u||^2 \text{ for } u \in T_\xi NB$$

This splitting is orthogonal. If we write $u = \dfrac{d\xi}{dt}|_{t=0}$ then $||u - \pi_* u|| = ||D^{\perp}/dt\xi||$ using the notation described in chapter 1.

Let $\pi : TM \to M$ be the projection of the tangent bundle onto M, denote this $TM \to M$, we also have $TN \to N$ and $NB \to N$. Now call $T(N)^{\perp} = \{(p, n) | p \in N, n \in (T_pN)^{\perp}\}$, we also get $T(N)^{\perp} \to N$. As we can think of the exponential map as a map from $TM \to M$ we can **restrict** it to $T(N)^{\perp}$

Definition. *(Exponential normal map)*
Following the discussion above, we define the exponential normal map, $\exp^{\perp} : T(N)^{\perp} \to M$ *as*

$$\exp^{\perp} = \exp|_{T(N)^{\perp}}$$

Definition. *(N-Jacobi Field)*
We say that a Jacobi field Y along a normal geodesic γ, $\gamma(s) = \exp^{\perp}(s\xi)$, is an N-Jacobi field if it is induced by a variation of geodesics normal to N, this is

$$Y(s) = \frac{d}{dt} \exp^{\perp}(s\xi(t))\bigg|_{t=0}$$

There are two important properties of N-Jacobi fields, which are proved in [13], so we state them as a lemma:

Lemma A.5.7. *(Properties of N-Jacobi fields)*

- *(Characterization of N-Jacobi Fields)*
  Y is an N-Jacobi field iff $Y(0) \in T_pN$ and $\dfrac{D}{ds}Y(0) - S_\xi Y(0) \perp T_pN$ where S_ξ is the Weingarten map.
- *(The mapping of the differential of N-Jacobi fields)*
  If $U(s) = \dfrac{d}{ds}(s\xi(t))\bigg|_{t=0} = A(s) + sB(s)$ then $Y(s) = d\exp^{\perp}(U(s))$ and

$$\pi_* A(s) = Y(0), |B(s)| = ||(D/dsY(0))^{\perp}||$$

See [6] to see the definition of the Weingarten equations (pg 155), proof of the proposition in [13]

Observation A.5.8. *This means that the N Jacobi field is the rate at which the geodesics normal to N change. This immediately characterizes the differential of* $\exp^{\perp}$*:*

$$d\exp^{\perp}(u) = Y(1)$$

Proposition A.5.9. *(Characterization of focal points)*
In the setting of above, $q \in M$ is a focal point of N if and only if q is a critical value of $\exp^{\perp}$

For the proof see [5] page 231.

By observation A.5.2 we know that there exists $r > 0$ such that a geodesic restricted to $[0, r]$ is length minimizing, note that if a curve is minimizing in $[0, r]$ then it is also minimizing in $[0, s]$ for all $s < r$. In the other hand, if a curve is not minimizing in $[0, t]$ then it cannot be minimizing in $[0, t_1]$ for all $t_1 > t$. Therefore by continuity we can define a unique value t_0 for which γ is minimizing in $[0, t_0]$ but not in $[0, t]$ if $t > t_0$.

Definition. *(Cut point of a point along a geodesic)*
We say that a geodesic $\gamma : [0, \alpha] \to M$ has a cut point at $\gamma(t_0)$ along γ if

$$t_0 = \sup\{r : \gamma|_{[0,r]} \textit{ is minimizing}\} = \inf\{t : \gamma|_{[0,t]} \textit{ is not minimizing }\}$$

The cut locus of p denoted $C_m(p)$ is the union of all cut points at p over all geodesics starting at p.

Proposition A.5.10. *(Minimizing distance to Cut locus)*
If $q \in M \setminus C_m(p)$ there is a unique minimizing geodesic joining p to q.

See [5] pg 271 for a proof.

Definition. *(Totally umbilic)*
A set N is said totally umbilic if every point looks locally like a sphere in the sense that the second fundamental form at the point is the mean curvature times the Riemannian tensor.

Bibliography

[1] L. Ambrosio and N. Giglia, *A User's guide to Optimal Transport*, Lecture notes of C.I.M.E. summer school, 2009.

[2] V. Bayle, *Propriétés de concavité du profil isopérimétrique et applications.*, Mathématiques. Université Joseph-Fourier - Grenoble I, 2003. Français

[3] D. Bakry and I. Gentil and M. Ledoux, *Analysis and Geometry of Markov Diffusion Operators*, Springler,Switzerland 2014.

[4] V.I. Bogachev, *Measure Theory* Volume 1,Springer-Verlag Berlin Heidelberg, 2000, USA. DOI: 10.1007/978-3-540-34514-5

[5] M. Do Carmo, *Riemannian geometry*, (translated by Flaherty, F.), Birkhauser Basel, 1992 Boston .

[6] Do Carmo, *Differential Geometry of curves and surfaces*, Prentice Hall, 1976 New Jersey.

[7] -, *Brittish Enciclopedia*, 15th edition, 2010.

[8] M. Erbar and K. Kuwada and K.T. Sturm, *On the Equivalence of the Entropic Curvature-Dimension Condition and Bochner's Inequality on Metric Measure Spaces*, arXiv:1303.4382

[9] H. Federer, *Geometric Measure Theory*, series Die Grundlehren der mathematischen Wissenschaften, Band 153, New York: Springer-Verlag New York Inc, 1969, Germany.

[10] R.J. Gardner, *The Brunn-Minkowski inequality*, Bulletin of American Mathematical Society (N.S), 39, 3 (2002), 355-405.

[11] A. Grigor'yan, *Heat Kernel analysis on manifolds*,s. AMS/IP Studies in Advanced Mathematics,47. American Mathematical Society, Providence, RI; International Press, Boston, MA, 2009.

[12] M. Gromov, *Paul Levy's isoperimetric inequality*, Appendix C of Metric structures for Riemannian and non-Riemannian spaces. Based on the 1981 French original, prepublications I.H.E.S. (1980)..

[13] E. Heintze and H.Karcher, *A general comparison theorem with applications to volume estimates for submanifolds.*, Annales scientifiques de l'École Normale Supérieure, Série 4 : Volume 11 (1978) no. 4 , p. 451-470

[14] M. Ledoux, *From Concentration to Isoperimetry: Semigroup Proofs*, Contemporary Mathematics Volume 545, 2011.

[15] J.M. Lee, *Manifolds and Differential Geometry*, American Mathematical Society, Graduate Studies in Mathematics Volume: 107; 2009 USA.

[16] E. Milman, *Isoperimetric and Concentration Inequalities - Equivalence under Curvature Lower Bound*, Duke Math. J. Volume 154, Number 2 (2010), 207-239.

[17] F. Morgan, *Manifolds with densities*, Notices of the AMS, Volume 52, Number 8, page 853.

[18] Rockafellar, Tyrell. *Convex Analysis*. Series:Princeton Landmarks in Mathematics and Physics. Princeton University Press, 1970, Second Ed, USA.

[19] C. Villani, *Topics in optimal transportation*, American Mathematical Society, 2003, USA.

[20] C. Villani and F. Otto *Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev*, Journal of Functional Analysis 173, 2 (2000), 361-400.

[21] Yosida, Kosaku. *Functional Analysis*, Springer-Verlag Berlin Heidelberg Springer, 1980 USA. DOI:10.1007/978-3-642-61859-8