Concentration of Measure And Ricci Curvature

Ryan Benty

A Thesis
in
The Department
of
Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements
For the Degree of Master of Arts at
Concordia University
Montreal, Quebec, Canada

June, 2012

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Ryan Benty

Entitled: Concentration of Measure and Ricci Curvature

and submitted in partial fulfillment of the requirements for the degree of

Master of Arts (Mathematics and Statistics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

Ronald J. Stern _______ Chair

Cody Hyndman _______ Examiner

_______________ Examiner

Alina Stancu _______ Supervisor

Approved by _______________
Graduate Program Director

____
Dean of Faculty

Date _______________

ABSTRACT

Concentration of Measure and Ricci Curvature

Ryan Benty

In 1917, Paul Lévy proved his classical isoperimetric inequality on the N-dimensional sphere. In the 1970's, Mikhail Gromov extended this inequality to all Riemannian manifolds with Ricci curvature bounded below by that of $\mathbb{S}^N$. Around the same time, the Concentration of Measure phenomenon was being put forth and studied by Vitali Milman. The relation between Concentration of Measure and Ricci curvature was realized shortly thereafter.

Elaborating on several articles, we begin by explicitly presenting a proof of the Concentration of Measure Inequality for $\mathbb{S}^N$, as the archetypical space of positive curvature, followed by a complete proof extending this result to all Riemannian manifolds with Ricci curvature bounded below by that of $\mathbb{S}^N$. In the process, we present a detailed technical proof of the Gromov-Lévy isoperimetric inequality.

Following Yann Ollivier, we note and prove a Concentration of Measure inequality on the discrete Hamming cube $\{0,1\}^N$, and discuss his extension of Ricci curvature to general metric spaces, particularly discrete metric measure spaces. We show that this coarse Ricci curvature on $\{0,1\}^N$ is positive, and present Ollivier's Concentration of Measure inequality for all spaces admitting positive coarse Ricci curvature. In addition, we calculate the coarse Ricci curvature for several discrete metric spaces.

Acknowledgments

First and foremost I would like to thank my supervisor, Dr. Alina Stancu, for the countless hours of instruction and advising provided these past two years. I would also like to thank all of my instructors for providing concise knowledge, and the staff of the Mathematics and Statistics department for having all of the answers.

TABLE OF CONTENTS

Chapter | Page

Chapter 1

Concentration of Measure and Isoperimetric Inequalities

Let $(X, \mathcal{M}, \nu, d)$ be a probability space, with Borel σ-algebra $\mathcal{M}$ and probability measure ν, endowed with a metric d. Consider the collection of sets $\mathcal{M}_a = \{A \in \mathcal{M};\ \nu(A) = a\}$, $a \in (0,1)$. One isoperimetric problem is minimizing the measure of the t-neighbourhoods of such sets: given $a \in (0,1)$, and given $t > 0$, find

$$\inf\{\nu(A_t);\ A \in \mathcal{M}_a\} \tag{1.1}$$

where $A_t = \{x \in X;\ d(x, A) \leq t\}$, and, if possible, characterize the minimizer(s). Explicit answers to this isoperimetric problem are often difficult to obtain. It is thus useful, and often in applications it suffices, to find a good lower bound for the measures of the sets A_t.

In the realm of concentration of measure, our equation is closely related to equation (1.1). Specifically, we are interested in the case $a = \frac{1}{2}$. However, instead of restricting ourselves to sets of measure exactly $\frac{1}{2}$, we look at this as a lower bound. Namely, we are looking for an upper bound of:

$$\sup\{1 - \nu(A_t) : \nu(A) \geq \frac{1}{2}, A \subset X, A \text{ Borel}\}, \tag{1.2}$$

as seen in [12]. Note that equation (1.2) allows this function to be defined in any probability metric space; it is an extension of equation (1.1).

A classic example is that of the N-sphere $\mathbb{S}^N$ in $\mathbb{R}^{N+1}$, where $N > 1$. More specifically, consider $(\mathbb{S}^N, B_{\mathbb{S}^N}, \nu, d)$, where $B_{\mathbb{S}^N}$ is the collection of Borel sets on $\mathbb{S}^N$ with the induced topology from $\mathbb{R}^{N+1}$, ν the *normalized* volume measure (so that $\nu(\mathbb{S}^N) = 1$), and d the geodesic metric in which distances are realized by curves which extend to great circles. Paul Lévy proved that among all Borel sets of measure $a \in (0,1)$, *spherical caps*, synonymous with a geodesic ball centred about a point, are those with neighbourhoods of least volume. As a specific example, consider any Borel set A on the unit sphere $\mathbb{S}^N$ such that $\nu(A) = 1/2$. Then, if B is a half sphere ($\nu(B) = 1/2$), we have that $\nu(A_t) \geq \nu(B_t)$. Since B is a half-sphere, it is much easier to obtain concentration lower bounds on $\nu(B_t)$ which will therefore hold for $\nu(A_t)$. If we take, for example, A such that $\nu(A) = 1/2$, then, as we will see shortly in the proof of Proposition 1.1, computing $\nu(B_t)$ above yields $\nu(A_t) \geq 1 - e^{-(N-1)t^2/2}$. This shows that for spheres in high enough dimension, almost all of the volume of the sphere lies within a small distance ϵ of any set containing at least half the volume of the sphere.

The results obtained on subsets of $\mathbb{S}^N$ may also be used to obtain concentration bounds on Lipschitz functions on $\mathbb{S}^N$. Specifically, if $f : \mathbb{S}^N \to \mathbb{R}$ is a 1-Lipschitz function, then for some m in $\mathbb{R}$:

$$\nu(\{x \in \mathbb{S}^N : |f(x) - m| \geq t\})$$

has an explicit upper bound (dependent on the dimension N and on the value t) which will be proven using Paul Lévy's isoperimetric inequality. Specifically, we will show that

$$\nu(\{x \in \mathbb{S}^N : |f(x) - m| \geq t\}) \leq 2e^{-(N-1)t^2/2}. \tag{1.3}$$

Concentration of measure may also be deduced for products of metric probability spaces. Along this idea, we will show that a concentration inequality exists for the Hamming Cube $\{0,1\}^N$. We consider $(\{0,1\}^N, \mathcal{P}(\{0,1\}^N), \nu, \delta)$, where ν is the uniform probability measure on the space, and the distance between two points is the normalized Hamming distance: $\delta(x,y) = \delta((x_1,\ldots,x_N),(y_1,\ldots,y_N)) = \dfrac{1}{N}\sum_{i=1}^{N}|x_i - y_i|$. we will prove the following later in this chapter: if f is a 1-Lipschitz function on this space, then there exists m such that

$$\nu(\{x : |f(x) - m| \geq t\}) \leq 2e^{-2Nt^2}. \tag{1.4}$$

1.1 Paul Lévy's Isoperimetric Theorem for $\mathbb{S}^N$

Theorem 1.1 (Paul Lévy's Isoperimetric Inequality)**.** *Let $\mathcal{S}_c$ denote the collection of all Borel sets in $\mathbb{S}^N$ with fixed normalized measure c, where $c \in (0,1)$. Then, for any $t > 0$ sufficiently small and any set $E \in \mathcal{S}_c$, we have that $\nu(E_t) \geq \nu(C_t)$, where C denotes a spherical cap of measure c.*

In preparation for the proof of the theorem for which we will follow Schechtman [11], we will start with a few preliminaries. Given a hyperplane H in $\mathbb{R}^{N+1}$ passing through the origin, we denote by $S_0 := \mathbb{S}^N \cap H$ and by S_+ and S_- the two open half spheres in H^c, where the sign of each will be later specified. Let $\sigma = \sigma_H$ be the reflection with respect to H. Note that σ preserves geometric distances of points on the sphere, and it is thus an isometry with respect to the sphere's geodesic metric. The following claim will be used:

Claim 1.1. *Suppose that a Borel set A of the sphere belongs to S_+ and that a point $x \in S_-$. Then $d(\sigma(x), A) \leq d(x, A)$, where $d(y,B) := \inf\{d(y,b); b \in B\}$.*

Proof of Claim. Note that we are working with Borel sets, which implies that $d(x,A) = d(x,\bar{A})$ where $\bar{A}$ denotes the closure of A. We may therefore assume that our set A

is closed, so that $d(x, A) = d(x, a)$ is realized by some $a \in A$. We assume $\sigma(x) \notin A$ or else the proof is trivial.

If x, $\sigma(x)$ and a all lie on the same great circle, then the geodesic segment $\overline{xa}$ contains $\sigma(x)$ which belongs to S_+, so $d(\sigma(x), a) \leq d(x, a)$ thus $d(\sigma(x), A) \leq d(x, A)$.

Suppose that they do not lie on the same great circle. Consider the geodesic lines $\overline{xa}$ and $\overline{\sigma(x)\sigma(a)}$. These geodesics intersect at some point $z \in S_0$, and the spherical (geodesic) triangles $\triangle_{\sigma(x)az}$ and $\triangle_{x\sigma(a)z}$ are congruent to one another. By the triangle inequality, and the congruency of the triangles, we conclude that $d(\sigma(x), a) \leq d(\sigma(x), z) + d(z, a) = d(x, z) + d(z, a) = d(x, a)$. Therefore $d(\sigma(x), A) \leq d(x, a) = d(x, A)$ which concludes the proof of the claim.

$\square$

Given $A \subset \mathbb{S}^N$, let us split this set into three disjoint subsets:

$$A = [A \cap (S_+ \cup S_0)] \sqcup [A \cap S_- \setminus (\sigma(A \cap S_+))] \sqcup [A \cap S_- \cap \sigma(A \cap S_+)],$$

which we will denote, respectively, by:

$$A = A_1 \sqcup A_2 \sqcup A_3.$$

Note that A_2 and A_3 belong to S_-. Let us, using σ, attempt to reflect as many of the elements of A as possible into S_+ while preserving the total measure of A. Since $\sigma(A_3) \subset A_1$, we cannot bring up A_3 by reflection into H without compromising the size of the set. However $\sigma(A_2) \cap A_1 = \emptyset$, so we define the *two point symmetrization of* A, denoted A^*, as follows:

$$A^* = A_1 \sqcup \sigma(A_2) \sqcup A_3,$$

by bringing up as many elements of A into S_+ as possible. Since these subsets are disjoint, and σ is an isometry, $\nu(A^*) = \nu(A)$, provided A is Borel. The following claim is that $(A^*)_t \subset (A_t)^*$; this will prove that $\nu((A^*)_t) \leq \nu((A_t)^*) = \nu(A_t)$. In other words, sets which have undergone two-point symmetrizations admit smaller t-neighbourhoods.

Lemma 1.1. $(A^*)_t \subseteq (A_t)^*$.

Proof of Lemma. Explicitly written out, these two sets look as follows:

$$(A^*)_t = (A_1)_t \cup (\sigma(A_2))_t \cup (A_3)_t$$
$$(A_t)^* = (A_t)_1 \cup \sigma((A_t)_2) \cup (A_t)_3.$$

We will show that each of the subsets of $(A^*)_t$ belong to $(A_t)^*$.

- $x \in (A_1)_t = [A \cap (S_+ \cup S_0)]_t$:

  First, note that $A_1 \subset A \Rightarrow (A_1)_t \subset A_t$, so either $x \in [A_t \cap (S_+ \cup S_0)]$ (ie. $x \in (A_t)_1$, so we are done) or $x \in (A_t \cap S_-)$; if $x \in S_-$, then $d(x, (A \cap S_+)) \leq t \Rightarrow d(\sigma(x), (A \cap S_+)) \leq t$ (it was shown in Claim 1.1 that if a set belongs to S_+ and x in S_-, the reflection of x into S_+ is closer to A than x was), so $x \in \sigma(A_t \cap S_+) \Rightarrow x \in [A_t \cap S_- \cap \sigma(A_t \cap S_-)] = (A_t)_3$. Therefore $(A_1)_t \subset (A_t)^*$.

- $x \in (\sigma(A_2))_t = [\sigma(A \cap S_- \setminus \sigma(A \cap S_+))]_t$:

  Suppose $x \in S_0$. Then $x = \sigma(x) \in [A \cap S_- \setminus \sigma(A \cap S_+)]_t$, which implies $x \in A_t \Rightarrow x \in [A_t \cap (S_+ \cup S_0)]$.

  Suppose $x \in S_+$. Then $\sigma(x) \in [A \cap S_- \setminus \sigma(A \cap S_+)]_t \Rightarrow \sigma(x) \in (A_t \cap S_-)$. If $\sigma(x) \notin \sigma(A_t \cap S_+)$, then $\sigma(x) \in [A_t \cap S_- \setminus \sigma(A_t \cap S_+)] \Rightarrow x \in [\sigma(A_t \cap S_- \setminus \sigma(A_t \cap S_+))]$; if $\sigma(x) \in \sigma(A_t \cap S_+) \Rightarrow x \in [A_t \cap (S_+ \cup S_0)]$.

  Suppose $x \in S_-$. We know $\sigma(x) \in [A \cap S_- \setminus \sigma(A \cap S_+)]_t \Rightarrow x \in [A \cap S_- \setminus \sigma(A \cap S_+)]_t$ (by the reflection reducing the metric distance, Claim 1.1) $\Rightarrow x \in A_t \cap S_-$.

Moreover, x must belong to $\sigma(A_t \cap S_+)$, otherwise $\sigma(x) \notin (A_t \cap S_+)$, but we have already seen that $\sigma(x) \in [A \cap S_- \setminus \sigma(A \cap S_+)]_t \subset A_t \Rightarrow \sigma(x) \in (A_t \cap S_+)$. Therefore $x \in \sigma(A_t \cap S_+) \Rightarrow x \in [A_t \cap S_- \cap \sigma(A_t \cap S_+)] = (A_t)_3$.

- $x \in (A_3)_t = [A \cap S_- \cap \sigma(A \cap S_+)]_t$:

  Suppose $x \in S_-$. Clearly $x \in A_t \Rightarrow x \in (A_t \cap S_-)$. We also know that $d(x, \sigma(A \cap S_+)) \leq t \Rightarrow d(\sigma(x), (A \cap S_+)) \leq t \Rightarrow \sigma(x) \in (A_t \cap S_+)$ (since $\sigma(x)$ certainly belongs in S_+), so $x \in \sigma(A_t \cap S_+) \Rightarrow x \in [A_t \cap S_- \cap \sigma(A_t \cap S_+)] = (A_t)_3$.

  Suppose $x \in S_0$. Then $x \in A_t \Rightarrow x \in [A_t \cap (S_+ \cup S_0)]$.

  Suppose $x \in S_+$. Then we know $d(x, \sigma(A \cap S_+)) \leq t \Rightarrow d(\sigma(x), (A \cap S_+)) \leq t \Rightarrow d(x, (A \cap S_+))$ (again by Claim 1.1) $\Rightarrow x \in (A_t \cap S_+) \Rightarrow x \in (A_t)_1$.

Therefore we have that $(A^*)_t \subseteq (A_t)^* \Rightarrow \nu((A^*)_t) \leq \nu((A_t)^*) = \nu(A_t)$.

□

Proof of Lévy's Theorem. Consider the metric space $\mathcal{C}$ consisting of all closed subsets of $\mathbb{S}^N$ with the Hausdorff metric. Fix $A \in \mathcal{C}$ and consider the collection of sets $B \in \mathcal{C}$ satisfying:

$$\nu(B) = \nu(A)$$
$$\nu(B_t) \leq \nu(A_t) \tag{1.5}$$

for all $t > 0$. Denote this collection of sets by $\mathcal{B}$. We will conclude the isoperimetric inequality by showing that a closed spherical cap whose measure is equal to $\nu(A)$ belongs to $\mathcal{B}$.

Note that $\mathcal{B}$ is closed: suppose B' is a limit point of $\mathcal{B}$ and consider a sequence $\{B_j\} \subset \mathcal{B}$ which converges to B'. For all $j \in \mathbb{N}$, $\nu(B_j) = \nu(A) \Rightarrow \lim_{j\to\infty} \nu(B_j) = \nu(B') = \nu(A)$. Also $\nu((B_j)_t) \leq \nu(A_t)$ for all j, thus $\lim_{j\to\infty} \nu((B_j)_t) = \nu(B'_t) \leq \nu(A_t)$.

Fix a point $x_0 \in \mathbb{S}^N$ and let C be the closed cap of measure $\nu(A)$ whose center is at x_0. If H is a hyperplane centered at the origin such that $x_0 \notin H$, denote by S_+ the open half-sphere with contains x_0. Consider the function from $\mathcal{C}$ to $\mathbb{R}$ which sends $B \in \mathcal{C}$ to $\nu(B \cap C)$. This function is upper semi-continuous:

Indeed, suppose that $B_j \to B$ in $\mathcal{C}$. For any $\epsilon > 0$ we may find a set B_j, for some $j \in \mathbb{N}$, such that $\nu(B_j \setminus B) < \epsilon$. Then $\nu(B_j \cap C) = \nu((B_j \cap B) \cap C) + \nu((B_j \setminus B) \cap C) < \nu(B \cap C) + \epsilon$. Since ϵ was arbitrary, we may conclude that $\lim_{j\to\infty} \nu(B_j \cap C) \leq \nu(B \cap C)$, which therefore proves the upper semi-continuity.

From this, we infer that the upper semi-continuous function $B \to \nu(B \cap C)$ will attain its maximum on the compact set $\mathcal{B}$, and we will denote this maximum by B. We shall show that $C \subset B$ which will certainly ensure that $C \in \mathcal{B}$. If C is not included in B, since they are both closed sets, there exist points of density in $B \setminus C$ and $C \setminus B$, respectively, such that $\nu(B \setminus C) > 0$ and $\nu(C \setminus B) > 0$. Let $x \in B \setminus C$, $y \in C \setminus B$ be points of density and let H be the hyperplane which is perpendicular to the geodesic line segment $[x, y]$ and crossing at the midpoint $(x+y)/2$. Let $B(x, r)$ be a ball centered at x such that $\nu(B(x, r) \cap (B \setminus C)) \approx \nu(B(x, r))$ (set r so that $B(x, r)$ is almost entirely contained in $(B \setminus C)$). Applying the two-point symmetrization to B will send most of $B(x, r)$ into $B(y, r)$ while any point in $B \cap C$ is sent to a point in C and therefore $B^* \cap C$. Therefore $\nu(B^* \cap C) > \nu(B \cap C)$, contradicting the fact that $B^* \in \mathcal{B}$. Therefore $C \subset B$. From the proof, it is not hard to see that only spherical caps are minimizers. Indeed, spherical caps are the sets which satisfy the property $C^* = C$ for all H, under the condition that $\nu(C \cap S_+) \geq \frac{1}{2}\nu(C)$ (ie. under the correction orientation), which allows us to say that $B \cap C$ is sent to $B^* \cap C$.

□

1.2 Concentration of Measure on $\mathbb{S}^N$

Using Paul Lévy's isoperimetric inequality on $\mathbb{S}^N$, we will now present a concentration of measure inequality on the sphere. Its proof will be a natural corollary of the following proposition:

Proposition 1.1. *[4] Let E_t be the t-neighborhood of a great circle of $\mathbb{S}^N$, $N \geq 2$. Then*

$$\nu(E_t^c) \leq 2 \exp\left[-(N-1)t^2/2\right],$$

where E_t^c denotes the complement of E_t relative to $\mathbb{S}^N$.

Proof. To evaluate the measure of these subsets of $\mathbb{S}^N$, we will use *hyper-spherical coordinates*:

$x_1 = \cos(\phi_1)$

$x_2 = \sin(\phi_1)\cos(\phi_2)$

$\vdots$

$x_N = \sin(\phi_1)\sin(\phi_2)\ldots\sin(\phi_{N-1})\cos(\phi_N)$

$x_{N+1} = \sin(\phi_1)\sin(\phi_2)\ldots\sin(\phi_{N-1})\sin(\phi_N)$,

where $\phi_i \in [0,\pi)$, $1 \leq i \leq N-1$, $\phi_N \in [0, 2\pi)$.

The volume form with respect to the coordinates $(\phi_1, \phi_2, \ldots, \phi_n)$ is

$$dV_{\mathbb{S}^N} = \sin^{N-1}(\phi_1)\sin^{N-2}(\phi_2)\ldots\sin^2(\phi_{N-2})\sin(\phi_{N-1})d\phi_1\ldots d\phi_N.$$

Note that calculating the normalized measure of $\mathbb{S}^N$ with respect to hyper-spherical coordinates yields:

$$
\begin{aligned}
1 &= \nu(\mathbb{S}^N) = \frac{1}{V(\mathbb{S}^N)} \int_{\mathbb{S}^N} dV_{\mathbb{S}^N} \\
&= \frac{1}{V(\mathbb{S}^N)} \int_0^{2\pi} \int_0^{\pi} ... \int_0^{\pi} \sin^{N-1}(\phi_1) \sin^{N-2}(\phi_2) ... \sin(\phi_{N-1}) d\phi_1 d\phi_2 ... d\phi_N \\
&= \frac{1}{V(\mathbb{S}^N)} V(\mathbb{S}^{N-1}) \int_0^{\pi} \sin^{N-1} d\phi_1 \\
&= s_N^{-1} \int_0^{\pi} \sin^{N-1}(\phi_1) d\phi_1,
\end{aligned}
$$

where $V(\cdot)$ represents the spherical measure induced from $\mathbb{R}^{N+1}$ and $s_N := \dfrac{V(\mathbb{S}^N)}{V(\mathbb{S}^{N-1})} = \displaystyle\int_0^{\pi} \sin^{N-1}(\theta)\, d\theta$.

Therefore if we wish to find the normalized volume of a spherical cap C of radius ϕ (in the case of half-sphere $\phi = \pi/2$), we may use the formula:

$$
\nu(C) = s_N^{-1} \int_0^{\phi} \sin^{N-1}(\theta)\, d\theta.
$$

Hence the measure of the t-neighborhood of a half-sphere A can be calculated as follows:

$$
\nu(A_t) = s_N^{-1} \int_0^{\pi/2+t} \sin^{N-1}(\theta)\, d\theta, \tag{1.6}
$$

and the measure of the complement of this set, A_t^c, is

$$
\begin{aligned}
1 - s_N^{-1} \int_0^{\pi/2+t} \sin^{N-1}(\theta)\, d\theta &= s_N^{-1} \int_{\pi/2+t}^{\pi} \sin^{N-1}(\theta)\, d\theta \\
&= s_N^{-1} \int_t^{\pi/2} \cos^{N-1}(\theta)\, d\theta.
\end{aligned}
$$

We seek an upper bound of A_t^c by finding an upper bound of the above integral. We will use the inequality $\cos(u) \leq e^{-u^2/2}$, $0 \leq \forall u \leq \pi/2$, and the fact that $\displaystyle\int_0^{\infty} e^{-u^2/2} du = \frac{\sqrt{\pi}}{\sqrt{2}}$. To deal with s_N, integration by parts gives us:

$$
s_N = \int_0^{\pi} \sin^{N-1}(\theta)\, d\theta = \frac{N-2}{N-1} \int_0^{\pi} \sin^{N-3}(\theta)\, d\theta = \frac{N-2}{N-1} s_{N-2}.
$$

Note that $(N-1)/(N-2) \geq \sqrt{N-3}/\sqrt{N-1} \Rightarrow \sqrt{N-1}s_N \geq \sqrt{N-3}s_{N-2} \geq 2$, so $s^N \geq 2/\sqrt{N-1}$. Using the above inequalities, along with the change of coordinates $\theta = \tau/\sqrt{N-1}$, yields:

$$\begin{aligned}
\int_t^{\pi/2} \cos^{N-1}(\theta)\, d\theta &= \frac{1}{\sqrt{N-1}} \int_{t\sqrt{N-1}}^{(\pi/2)\sqrt{N-1}} \cos^{N-1}(\frac{\tau}{\sqrt{N-1}}) d\tau \\
&\leq \frac{1}{\sqrt{N-1}} \int_{t\sqrt{N-1}}^{\infty} e^{-\tau^2/2}\, d\tau \\
&= \frac{1}{\sqrt{N-1}} \int_0^{\infty} e^{-(\tau+t\sqrt{N-1})^2/2}\, d\tau \\
&= \frac{1}{\sqrt{N-1}} \int_0^{\infty} e^{-(\tau^2+2\tau t\sqrt{N-1}+t^2(N-1))/2} d\tau \\
&\leq \frac{1}{\sqrt{N-1}} e^{-(N-1)t^2/2} \int_0^{\infty} e^{-\tau^2/2} d\tau \\
&= \frac{\sqrt{\pi}}{\sqrt{2(N-1)}} e^{-(N-1)t^2/2}.
\end{aligned}$$

Hence

$$\begin{aligned}
s_N^{-1} \int_t^{\pi/2} \cos^{N-1}(\theta)\, d\theta &\leq s_N^{-1} \frac{\sqrt{\pi}}{\sqrt{2(N-1)}} e^{-(N-1)t^2/2} \\
&\leq \frac{\sqrt{N-1}}{2} \frac{\sqrt{\pi}}{\sqrt{2(N-1)}} e^{-(N-1)t^2/2} \\
&\leq e^{-(N-1)t^2/2}.
\end{aligned}$$

We have shown that $\nu(A_t^c) \leq e^{-(N-1)t^2/2}$. Consider $B = \mathbb{S}^N \backslash A$, the complementary half-sphere. The symmetry of $\mathbb{S}^N$ allows us to conclude that $\nu(B^c) \leq e^{-(N-1)t^2/2}$, and therefore $\nu(A_t^c \cup B_t^c) \leq 2e^{-(N-1)t^2/2}$. But if E is the great circle corresponding to the boundary of A and B, and we consider E_t, then $E_t^c = (A_t \cap B_t)^c = A_t^c \cup B_t^c$, and therefore $\nu(E_t^c) \leq 2e^{-(N-1)t^2/2}$.

□

Corollary 1.1 (A Concentration of Measure Theorem on $\mathbb{S}^N$)**.** *[9] Let $\mathbb{S}^N \subset \mathbb{R}^{N+1}$ be the N-dimensional unit sphere of the Euclidean space and let $f : \mathbb{S}^N \to \mathbb{R}$ be a 1-Lipschitz function. Then there exists $m \in \mathbb{R}$ such that for any $t \geq 0$,*

$$\nu(\{x \in S^N, |f(x) - m| \geq t\}) \leq 2 \exp -\frac{t^2}{2D^2},$$

where $D = 1/\sqrt{N-1}$ *and* ν *is the natural volume measure on* $\mathbb{S}^N$*, normalized so that* $\nu(\mathbb{S}^N) = 1$.

Proof. We say that m is a median of f if $\nu(\{x : f(x) \leq m\}) \geq 1/2$ and $\nu(\{x : f(x) \geq m\}) \geq 1/2$. Let us denote by $A' := \{x : f(x) \geq m\}$ and $B' := \{x : f(x) \leq m\}$. Let A'_t and B'_t be the t-neighbourhoods of A' and B' respectively, and define the following sets

$A'_{f,t} := \{x : f(x) - m \geq -t\}$

$B'_{f,t} := \{x : f(x) - m \leq t\}$.

Claim 1.2. $A'_t \subset A'_{f,t}$ *and* $B'_t \subset B'_{f,t}$.

Proof of Claim. Suppose $x \in A'_t \cap A'^c$ (if x belongs in A', then we are done), which means $f(x) < m$, and $d(x, A') \leq t$. Because f is Lipschitz and therefore continuous, we may assume that A' is a Borel set, and we can assume that A' is closed so that $d(x, A')$ is realized by some value $y \in A'$. Then $0 \leq f(y) - f(x) \leq d(x, y) \leq t$, but $f(y) \geq m \Rightarrow m - f(x) \leq f(y) - f(x) \leq t$, so $x \in A'_{f,t}$. By a similar argument, $B'_t \subset B'_{f,t}$.

□

We define a set $C'_{f,t} := \{x : |f(x) - m| \leq t\} = A'_{f,t} \cap B'_{f,t}$. By the isoperimetric inequality, we know that $\nu(A_t) \leq \nu(A'_t) \leq \nu((A')_{f,t})$, where A is a half-sphere and A_t is its t-neighborhood. Therefore by Proposition 1.1, $\nu((A'_{f,t})^c) \leq \nu((A'_t)^c) \leq \nu((A_t)^c) \leq e^{-(N-1)t^2/2}$. Similarly, $\nu((B'_{f,t})^c) \leq e^{-(N-1)t^2/2}$, so

$$\nu(\{x \in \mathbb{S}^N; |f(x) - m| \geq t\}) = \nu((C'_{f,t})^c) = \nu((A'_{f,t})^c \cup (B'_{f,t})^c) \leq 2e^{-(N-1)t^2/2},$$

concluding the proof.

□

1.3 Concentration on the Hamming Cube

In this section, we prove a concentration of measure inequality on the Hamming Cube $\{0,1\}^N$. More precisely, we consider the space $(\{0,1\}^N, \mathcal{P}(\{0,1\}^N), \nu, \delta)$, where ν is a normalized uniform measure on this space and δ is the rescaled Hamming metric: $\delta(x,y) = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} |x_i - y_i|$, so that the maximal distance between points is 1 and the distance between neighbours (points which differ by exactly one value) is $\frac{1}{N}$. We will make use of the Laplace functional which we define below.

Definition 1.1. *Let (X,d) be a metric space, and let ν be a probability measure on the Borel sets of (X,d). The Laplace functional of ν is*

$$E(X,d,\nu)(\lambda) = \sup_f \int_X e^{\lambda f} d\nu, \tag{1.7}$$

where the supremum is taken after all 1-Lipschitz functions on X of mean zero.

In our context however, f is not restricted to have mean zero, so we consider instead

$$\sup_f \int e^{\lambda(f-\mathbb{E}f)} d\nu, \tag{1.8}$$

where f is a 1-Lipschitz function. The following lemma will help us prove a concentration inequality. Note the similarity between the inequality below and the analogous inequality on the sphere, Corollary 1.1.

Lemma 1.2. *[4]*

Let X be a metric space equipped with a probability measure ν. Let $f : X \to \mathbb{R}$ be a 1-Lipschitz function. Assume that there exists a $D > 0$ such that, for any $\lambda \in \mathbb{R}$, one has

$$\mathbb{E}e^{\lambda(f-\mathbb{E}f)} \leq e^{D^2\lambda^2/2} \tag{1.9}$$

where $\mathbb{E}$ *denotes the mean of* f *with respect to* ν*. Then, for any* $t > 0$*,*

$$\nu(\{x \in X,\ f(x) - \mathbb{E}f \geq t\}) \leq e^{-t^2/2D^2}.$$

Proof. The proof of this lemma makes use of the Chebyshev's inequality: Given a real non-negative measurable function f on a measure space $(X, \mathcal{M}, \mu)$, the following inequality holds for any positive number t

$$\mu(\{x \in X | f \geq t\}) \leq \frac{1}{t} \int_X f(x) d\mu.$$

Since the exponential function $x \to \exp(x)$ on $\mathbb{R}$ is increasing, by taking $t > 0$, and any real number $D > 0$, the following sets are equal:

$$\begin{aligned} &\{x \in X \quad | \quad f(x) - \mathbb{E}f \geq t\} \\ = &\{x \in X \quad | \quad \frac{t}{D^2}(f(x) - \mathbb{E}f) \geq \frac{t^2}{D^2}\} \\ = &\{x \in X \quad | \quad e^{\frac{t}{D^2}(f(x)-\mathbb{E}f)} \geq e^{\frac{t^2}{D^2}}\}. \end{aligned}$$

Applying Chebyshev's inequality to the latter set with respect to ν, and assuming (1.9), we obtain:

$$\begin{aligned} \nu\left(\{x \in X | e^{\frac{t}{D^2}(f(x)-\mathbb{E}f)} \geq e^{\frac{t^2}{D^2}}\}\right) \\ &\leq e^{\frac{-t^2}{D^2}} \mathbb{E}(e^{\frac{t}{D^2}(f-\mathbb{E}f)}) \\ &\leq e^{\frac{-t^2}{D^2}} e^{D^2(\frac{t}{D^2})^2/2} \\ &= e^{-t^2/2D^2}. \end{aligned}$$

□

Note that the result relies on the following assumption on the Laplace functional

$$E(X, d, \nu)(\lambda) = \sup_{f\text{1-Lipschitz}} \int e^{\lambda(f-\mathbb{E}f)} d\nu \leq e^{D^2\lambda^2/2}.$$

We will show that, in the case of the Hamming cube, such an estimate on the Laplace functional holds for $D = \dfrac{1}{2\sqrt{N}}$.

Consider the metric space $(\{0,1\}, d)$ with probability measure ν, and the metric d rescaled by $\dfrac{1}{N}$ $(d(1,0) = d(0,1) = \dfrac{1}{N})$. The rescaling is specific to the dimension N for which we wish to prove the inequality. Let us find an upper bound for the Laplace functional $E(\{0,1\}, d, \nu)$. If f is a 1-Lipschitz function on this space, then

$$\int_{\{0,1\}} e^{\lambda(f-\mathbb{E}f)} d\nu$$

$$= 1 + \int \lambda(f - \mathbb{E}f) d\nu + \int \frac{\lambda^2}{2!}(f - \mathbb{E}f)^2 d\nu + \int \frac{\lambda^3}{3!}(f - \mathbb{E}f)^3 d\nu + \cdots . \quad (1.10)$$

However, for any $k \geq 1$,

$$\int_{\{0,1\}} (f - \mathbb{E}f)^k d\nu = \int_{\{0\}} (f - \mathbb{E}f)^k d\nu + \int_{\{1\}} (f - \mathbb{E}f)^k d\nu$$

$$= \int_{\{0\}} \left(f(0) - \frac{f(0) + f(1)}{2} \right)^k d\nu + \int_{\{1\}} \left(f(1) - \frac{f(0) + f(1)}{2} \right)^k d\nu$$

$$= \frac{1}{2} \left[\frac{f(0) - f(1)}{2} \right]^k + \frac{1}{2} \left[\frac{f(1) - f(0)}{2} \right]^k,$$

so if k is odd, then $\int_{\{0,1\}} (f - \mathbb{E}f)^k d\nu = 0$. If k is even, then

$$\int_{\{0,1\}} (f - \mathbb{E}f)^k d\nu \leq \frac{1}{2} \left(\frac{1}{2N} \right)^k + \frac{1}{2} \left(\frac{1}{2N} \right)^k = \left(\frac{1}{2N} \right)^k, \quad (1.11)$$

since f is 1-Lipschitz and $d(0,1) = 1/N$. Applying this to (1.10), we obtain:

$$\int_{\{0,1\}} e^{\lambda(f-\mathbb{E}f)} d\nu \leq 1 + \frac{\lambda^2}{2!} \frac{1}{(2N)^2} + \frac{\lambda^4}{4!} \frac{1}{(2N)^4} + \cdots$$

$$= \sum_{i=0}^{\infty} \frac{\lambda^{2i}}{(2i)!(2N)^{2i}}$$

$$\leq \sum_{i=0}^{\infty} \frac{\frac{\lambda^2}{8N^2}}{i!} = e^{\lambda^2/8N^2}, \tag{1.12}$$

hence $E_{(\{0,1\},d,\mu)} \leq e^{\lambda^2/8N^2}$.

Note again that the metric d was rescaled as though we were working in $\{0,1\}^N$. This is due to viewing the Hamming Cube as a product space $\{0,1\} \times \{0,1\} \times \cdots \times \{0,1\}$ whose metric δ is the sum of metrics, so that $\delta(x,y) = d(x_1,y_1)+\cdots+d(x_n,y_n)$. The following proposition defines an upper bound of the Laplace Functional on a product space.

Proposition 1.2. *[4] Let (X,d,μ) and (Y,δ,ν) be metric spaces endowed with probability measures. Then $E_{(X\times Y,d+\delta,\mu\otimes\nu)} \leq E_{(X,d,\mu)}E_{(Y,\delta,\nu)}$.*

Proof. Suppose that F is a 1-Lipschitz function on $(X \times Y, d+\delta)$ with mean zero, and consider the functions $F^y(x) = F(x,y)$ and $G(y) = \int F^y d\mu$. These functions are 1-Lipschitz in their respective spaces, since $|F^y(x_1)-F^y(x_2)| = |F(x_1,y)-F(x_2,y)| \leq d(x_1,x_2) + \delta(y,y)$, and $|G(y_1) - G(y_2)| = |\int F^{y_1}(x) - F^{y_2}(x)d\mu| \leq |\int \delta(y_1,y_2)d\mu| = |\delta(y_1,y_2)\mu(X)| = \delta(y_1,y_2)$. Therefore

$$\int e^{\lambda F} d\mu \otimes \nu = \int e^{\lambda G(y)} \left(\int e^{\lambda[F^y(x) - \int F^y d\mu]} d\mu(x) \right) d\nu(y)$$

$$\leq E_{(X,d,\mu)}(\lambda) \int e^{\lambda G} d\nu \leq E_{(X,d,\mu)} E_{(Y,\delta,\nu)}. \tag{1.13}$$

□

Corollary 1.2. *If f is a 1-Lipschitz function on the N-dimensional Hamming Cube, then*

$$\int_{\{0,1\}^N} e^{\lambda(f-\mathbb{E}f)} d\nu \leq E_{(\{0,1\}^N,\delta,\nu)}$$

$$\leq (E_{(\{0,1\},d,\mu)})^N = e^{\lambda^2/8N} \tag{1.14}$$

We will now state and prove the Theorem of Concentration of Measure on the cube.

Theorem 1.2. *[9] Let $X = \{0,1\}^N$ be the Hamming Cube equipped with the uniform probability measure ν and the rescaled Hamming metric defined by $\delta(x,y) = \frac{1}{N}\sum_{i=1}^{N}|x_i - y_i|$, where $x = (x_1, x_2, \ldots, x_N), x_i \in \{0,1\}$. Let $f : X \to \mathbb{R}$ be a 1-Lipschitz function. Then there exists $m \in \mathbb{R}$ such that, for any $t \geq 0$,*

$$\nu(\{x \in X, |f(x) - m| \geq t\}) \leq 2e^{-t^2/2D^2}$$

where $D = 1/2\sqrt{N}$.

Proof. By Lemma 1.2 and Corollary 1.2, taking $m = \mathbb{E}f$, we have that $\nu(\{x \in X, f(x) - \mathbb{E}f(x) \geq t\}) \leq \exp^{-t^2N/8} = \exp^{-t^2D^2/2}$. Similarly, $\nu(\{x \in X, \mathbb{E}f(x) - f(x) \geq t\}) \leq \exp^{-t^2D^2/2}$, since $(\mathbb{E}f - f)$ is also a 1-Lipschitz function on X with mean zero. Therefore $\nu(\{x \in X, |f(x) - m| \geq t\}) = \nu(\{x \in X, f(x) - \mathbb{E}f(x) \geq t\} \cup \{x \in X, \mathbb{E}f(x) - f(x) \geq t\}) \leq 2\exp^{-t^2D^2/2}$ which concludes the proof.

□

Chapter 2

Ricci Curvature and Concentration inequalities

We will now turn our attention to finite dimensional Riemannian manifolds. The first part of this chapter will introduce the definitions and concepts of Riemannian geometry which we will need later. The second part of the chapter will deal with concentration inequalities on Riemannian manifolds with Ricci curvature bounded from below. This will be a natural, yet non-trivial, extension of the concentration of measure phenomenon on the sphere.

2.1 Riemannian Manifolds

A *Riemannian manifold* (M, g) is a differentiable manifold M such that, at each point $p \in M$, there exists a positive definite inner product $g_p : T_pM \times T_pM \to \mathbb{R}$. The family g of all such inner products g_p is known as the *Riemannian metric* tensor on M. The presence of an inner product on these tangent spaces allows for many convenient properties, such as angles between vectors, lengths of curves, and others. Most importantly, it allows us to view these manifolds as metric spaces, a fact which will be looked at shortly.

2.1.1 The Levi-Civita Connection, Christoffel Symbols, and the Riemann Curvature Tensor

Let M be a smooth manifold, and let $\mathcal{T}M$ denote the space of vector fields on the tangent bundle TM. A *linear connection*

$$\nabla : \mathcal{T}M \times \mathcal{T}M \to \mathcal{T}M$$

$$(X, Y) \mapsto \nabla_X Y$$

is a map which satisfies the following properties:

$$\nabla_{fX_1+gX_2} Y = f\nabla_{X_1} Y + g\nabla_{X_2} Y \quad \text{for } f, g \text{ in } C^\infty(M, \mathbb{R}) \tag{2.1}$$

$$\nabla_X (aY_1 + bY_2) = a\nabla_X Y_1 + b\nabla_X Y_2 \quad \text{for } a, b \text{ in } \mathbb{R} \tag{2.2}$$

$$\nabla_X fY = f\nabla_X Y + (Xf)Y. \tag{2.3}$$

where (2.3) is the known as the *product*, or the *Liebniz*, rule. The definition of a connection makes it apparent that they form a vast family of maps. We are fortunate to be working in Riemannian manifolds where connections with convenient properties exist.

If X is a vector field on M, we may view it as a differential operator on smooth functions on M. Given a local coordinate system x^i, $i = 1, 2, \ldots, N$, the tangent vectors $e_i = \dfrac{\partial}{\partial x^i}$ define a basis of the tangent space of M at each point. Here and for the rest of the chapter, unless otherwise stated, N is the dimension of the manifold M. If $f \in C^\infty(M, \mathbb{R})$ and $X = x^i e_i \in \mathcal{T}M$ is a vector field with local coordinates e_i, then $X(f) = x^i \dfrac{\partial f}{\partial x_i}$, using Einstein summation notation. Given two vector fields X and Y, we may define a new vector field denoted $[X, Y]$ which acts on smooth functions on M by

$$[X,Y](f) = X(Y(f)) - Y(X(f)).$$

This is known as the *Lie bracket* of X and Y. A linear connection ∇ is said to be *torsion free* if $[X,Y] = \nabla_X Y - \nabla_Y X$. Working with torsion free connections allow us to study the curvatures of manifolds with greater ease.

If (M,g) is a Riemannian Manifold, a linear connection ∇ is said to be *compatible* with g if

$$\nabla_X g(Y,Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$$

for all vector fields X, Y and Z in $\mathcal{T}M$. Compatibility implies that the metric tensor g is preserved under parallel transport which will be introduced later on.

If a connection on a Riemannian manifold is torsion free and compatible with its metric tensor, it is given a special name and it is in fact unique, by the following classical theorem whose proof can be found, for example, in [5]:

Theorem 2.1 (The Fundamental Theorem of Riemannian Geometry)**.** *Given a Riemannian manifold (M,g), there is a unique connection ∇ which is torsion free and compatible with g, called the* Levi-Civita connection *associated to g.*

On Riemannian manifolds, in light of the Levi-Civita connection, the *Christoffel symbols* often make an appearance. Given a local basis $e_i = \dfrac{\partial}{\partial x^i}$ of the tangent space of M, as earlier, the Christoffel symbols are the unique coefficients Γ^k_{ij} such that $\nabla_{e_i} e_j = \Gamma^k_{ij} e_k$. In terms of the coordinates of the metric tensor $g = [g_{ij}]$,

$$\Gamma^k_{ij} = \frac{1}{2} g^{il} \left(\frac{\partial g_{li}}{\partial x^j} + \frac{\partial g_{lj}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^l} \right),$$

where $[g^{ij}]$ denotes the inverse matrix of $[g_{ij}]$.

We will also require the use of the *Riemann curvature tensor*. In terms of the Levi-Civita connection it is the tri-linear map $R : \mathcal{T}M \times \mathcal{T}M \times \mathcal{T}M \to \mathcal{T}M$ given by

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z.$$

Intuitively, the curvature tensor measures the following [9]: suppose we are at a point $p \in M$ and we take three tangent vectors $U, V, W \in T_pM$. Parallel translate U along V and then along W, and parallel translate U along W and then along V. The difference between the endpoints of the two translations is given by the curvature tensor. In other words, the difference between the path VWU and WVU is given by $R(V,W)U$.

2.1.2 Geodesics and the Exponential Map

Now that we have introduced the Levi-Civita connection and the Christoffel symbols, we are ready to discuss *geodesics*, which can roughly be thought of as the equivalent of straight lines on Riemannian manifolds.

Suppose that $\gamma : [a,b] \to M$ is a smooth curve in M from $p = \gamma(a)$ to $q = \gamma(b)$. Then at each point $\gamma(t) \in M$, the derivative $\dot{\gamma} = \dfrac{d\gamma(t)}{dt}$ is an element of $T_{\gamma(t)}M$, so, by taking the inner product $g(\dot{\gamma}, \dot{\gamma}) = ||\dot{\gamma}||^2$, we may define the length of the curve as

$$l(\gamma) := \int_a^b ||\dot{\gamma}|| \; dt. \tag{2.4}$$

With this we may view M as a metric space by defining the distance between two points p and q as the infimum of the lengths of all such smooth curves $\gamma(t)$ such that $\gamma(a) = p$, $\gamma(b) = q$. These curves are realized by geodesics. However, geodesics are a larger family of curves which are in fact critical points of the length functional.

We say that a smooth curve γ is a geodesic if it satisfies the equation $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$ at each point along the curve. In the Euclidean space, with the standard metric, the ordinary differential equation amounts to the second derivative of γ being identically zero, whose solutions are precisely straight lines.

Let us observe that the coordinate functions of $\gamma(t) = (x^1(t), x^2(t), \ldots, x^n(t))$ are subject to $\nabla_{\dot\gamma}\dot\gamma = 0$ if and only if they satisfy

$$\ddot{x}^k(t) + \Gamma^k_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t) = 0,$$

where $\dot{x}^i(t) = \dfrac{dx^i}{dt}$ and the summation convention has been employed. This is known as the *geodesic equation*. The following theorem, for which we refer again to [5], will be relevant later.

Theorem 2.2 (Existence and Uniqueness of Geodesics)**.** *For any $p \in M$, any $v \in T_pM$, and $t_0 \in \mathbb{R}$, there exist an open interval $I \subset \mathbb{R}$ containing t_0 and a geodesic $\gamma : I \to M$ satisfying $\gamma(t_0) = p$, $\dot\gamma(t_0) = v$. Any two such geodesics agree on their common domain.*

The proof relies on the existence and uniqueness of solutions to second-order ODE systems (namely the geodesic equation) given initial conditions. Let the *maximal* geodesic associated to the given initial conditions be the geodesic curve $\gamma : I \to M$ such that $\gamma(0) = p$, $\dot\gamma(0) = v$, on some maximal interval I containing 0: allow I to be the union of all intervals which are domains of geodesics satisfying the initial conditions. This maximal geodesic will be useful in defining the exponential map.

For a Riemannian manifold (M, g), the *exponential map* $\exp : \mathcal{E} \subseteq TM \to M$ maps $(p, v) \mapsto \gamma_v(1)$, where $\gamma_v(t)$ is the maximal geodesic satisfying $\gamma_v(0) = p$, $\dot\gamma_v(0) = v$. For each $p \in M$, $\exp_p$ is the *restriction* of exp to T_pM.

We may define the entire geodesic $\gamma_v(t)$ using the exponential map, given by

$$\gamma_v(t) = \exp_p(tv)$$

for all t such that the geodesic is defined. The most important part of the exponential map is that it defines a local diffeomorphism between an open neighborhood of the origin of each tangent space T_pM and a neighborhood of p, as one can see, for example, in [5]:

Theorem 2.3 (Normal Neighborhood Lemma)**.** *For every $p \in M$, there is a neighborhood $\mathcal{V}$ of the origin of T_pM and a neighborhood $\mathcal{U}$ of p in M such that* $\exp_p : \mathcal{V} \to \mathcal{U}$ *is a diffeomorphism.*

As a useful consequence of the above lemma, we can define geodesic balls around any point p of M as follows. We know that there exists an $\epsilon > 0$ such that $B_\epsilon(0) \subset T_pM$ maps diffeomorphically to M, thus call $\exp_p(B_\epsilon(0)) \subset M$ the *geodesic ball* in M of radius ϵ, centered at p.

2.1.3 Variations and Jacobi Fields

Given a smooth curve $\gamma : [a, b] \to M$, a *variation* of γ is, in our context, a smooth map $\Gamma : (-\epsilon, \epsilon) \times [a, b] \to M$ such that $\Gamma_0(t) = \Gamma(0, t) = \gamma(t)$. The *variation field* V of Γ is the vector field $V(t) = \frac{\partial \Gamma}{\partial s}|_{(0,t)}$ which describes the way γ varies with respect to s at each point $\gamma(t)$. Suppose now that γ is a geodesic. Then we say that a variation Γ is a *variation through geodesics* if each of the curves $\Gamma_s(t) = \Gamma(s, t)$ is itself a geodesic.

Consider a geodesic $\gamma(t)$ and suppose that $J(t)$ is a vector field along γ. We will call $J(t)$ a *Jacobi field* if it satisfies the Jacobi equation:

$$\nabla_{\dot\gamma(t)} \nabla_{\dot\gamma(t)} J(t) + R(J(t), \dot\gamma(t))\dot\gamma(t) = 0.$$

Intuitively, Jacobi fields describe the difference between a geodesic γ and an infinitesimally close geodesic which is shifted according to $J(t)$. Indeed, variations through geodesics satisfy the Jacobi equation:

suppose that we have a variation of geodesics $\Gamma(s, t)$ such that $\Gamma(0, t) = \gamma(t)$ and $\Gamma_s(t) = \Gamma(s, t)$ are geodesics for all s. Denote by $Y(t)$ the variation field $\frac{\partial \Gamma(s,t)}{\partial s}|_{s=0}$ and $T(t) = \dot\gamma(t)$. Since γ is a geodesic, we know that $\nabla_T T = 0$. Because we may view s and t as local coordinates, we have by symmetry $\nabla_T Y = \nabla_Y T$. One consequence of symmetry is that $[Y, T] = 0$, so a direct calculation of the Riemann tensor shows us

$$R(Y,T)T = \nabla_Y \nabla_T T - \nabla_T \nabla_Y T - \nabla_{[Y,T]} T.$$

But the first and third terms are both zero, and by symmetry we have

$$R(Y,T)T = -\nabla_T \nabla_T Y,$$

which proves the earlier claim.

2.2 Curvature in Riemannian Manifolds

Sectional and Ricci curvatures are two ways to describe the curvature of a manifold, and offer up different information about the shape of the manifold. We will begin by describing the sectional curvature and then define the Ricci curvature, which we will be using frequently for the rest of this chapter. We will also introduce the shape operator and mean curvature of hypersurfaces of Riemannian manifolds, as they will play a major role in our generalization of Lévy's theorem.

2.2.1 Sectional Curvature

Take a point p in a Riemannian manifold M, and two linearly independent vectors v, $w \in T_pM$. The sectional curvature $K(v,w)$ is the Gauss curvature of the surface on M formed by the image of the plane spanned by v and w under $\exp_p$. It can be calculated by

$$K(v,w) = \frac{\langle R(v,w)v, w\rangle}{\langle v,v\rangle\langle w,w\rangle - \langle v,w\rangle^2},$$

where R is the Riemann curvature tensor. In the case that v and w are orthonormal, $\langle v,v\rangle = \langle w,w\rangle = 1$ and $\langle v,w\rangle = 0$, so

$$K(v,w) = \langle R(v,w)w, v\rangle.$$

Sectional curvature does not play a significant role in the remainder of this chapter, but we will see a more intuitive description in Theorem 3.1 when generalizing the Ricci curvature to arbitrary metric spaces.

2.2.2 Ricci Curvature

Like the sectional curvature, the Ricci curvature is obtained from the Riemann curvature tensor. Given a point p, and $v,\ w \in T_pM$, the *Ricci curvature tensor* $Ric(v,w)$ is defined by the trace of the map $u \mapsto R(u,v)w$. If $\{e_i\}_i$ is an orthonormal basis of T_pM, then $Ric(v,w) = \sum_i \langle R(e_i,v)w, e_i\rangle$. Suppose now that v is an arbitrary unit tangent vector. The *Ricci curvature* in the direction of v, denoted $Ric(v)$, is determined by $Ric(v,v) = \sum_i \langle R(e_i,v)v, e_i\rangle$. Furthermore, if $v = e_j$, then $Ric(e_j) = \sum_{i\neq j}\langle R(e_i,e_j)e_j, e_i\rangle = \sum_{i\neq j} K(e_i,e_j)$, so we may think of Ricci curvature in the direction of v as $N-1$ times the average of all sectional curvatures $K(v,.)$, where $N = \dim(M)$.

2.2.3 Hypersurfaces and the Shape Operator

Given an $(N+1)$-dimensional Riemannian manifold (M,g), a *hypersurface* (H,h) of M is an N-dimensional Riemmanian manifold embedded by the inclusion function $i : H \hookrightarrow M$ such that i preserves the metric h. That is, given $p \in H$ for $v,w \in T_pH$, $h(v,w) = g(i^*(v), i^*(w))$, where $i^* : T_pH \to T_pM$ is the push-forward of i. For the sake of simplicity, we will identify $i(p) \in M$ with p, as well as $i^*(v)$ with v. Given $p \in H$, there exists $\nu_p \in T_pM$ such that $g(\nu_p, v) = 0$ for all $v \in T_pH$. This vector ν_p is said to be *normal* to H. A normal vector field is a vector field on H consisting of vectors normal to H. The hypersurface H is said to be *orientable* if there exists a smooth, non-vanishing normal vector field ν defined on all of H. Suppose that we restrict our attention to unit normal vector fields. Then, in fact, only two such vector fields exist, depending on our "direction". Indeed, suppose $\nu_p \in T_pM$ is a unit normal vector. Then $-\nu_p$ is also a unit normal vector, but points in the opposite direction. These two vectors belong in two uniquely determined unit normal vector fields.

Fix a unit normal vector field ν. The *shape operator* S_ν at the point p is the map $S_\nu : T_pH \to T_pH$ mapping v to $\nabla_v \nu_p$. The shape operator describes the change of the normal unit vector field in the direction v at the point p, and therefore the curvature of H in M. Given an orthonormal basis $\{e_i\}_1^N$ of H, the *mean curvature* of H at p is $\frac{1}{N}\sum_1^N g(S_\nu(e_i), e_i)$, the trace of the shape operator normalized by N. We also require the definition of an umbilical hypersurface.

Definition 2.1. *A hypersurface $H \subset M$ is said to be* totally umbilical *if at each point $p \in H$, given an orthonormal basis $\{e_i\}_1^N$, $g(S_\nu(e_j), e_j) = g(S_\nu(e_k), e_k)$ for all j and k from 1 to N.*

2.3 Lévy-Gromov Isoperimetric Inequality

In this section, we will discuss Gromov's extension of Paul Lévy's isoperimetric inequality and its application to concentration of measure on Riemannian manifolds. We have seen Lévy's isoperimetric inequality on the sphere in the previous chapter. Gromov noticed that a similar inequality holds on any manifold whose Ricci curvature is greater than that of a sphere of equal dimension properly normalized. Therefore, concentration inequalities on manifolds of positive Ricci curvature are implied by the extension of the isoperimetric inequality. Let us first state the main theorem which we will prove in details not found in the literature. We will then present the results on concentration of measure as a corollary of this theorem.

Theorem 2.4. *[6] Let M be an $(N+1)$ - dimensional manifold, and let $Ric(M) := \inf_v Ric(v,v)$ where v runs over all unit tangent vectors in TM. Suppose $Ric(M) \geq N = Ric(\mathbb{S}^{N+1})$ where $\mathbb{S}^{N+1} \subset \mathbb{R}^{N+2}$ is the unit sphere. Let $M_0 \subset M$ be a domain with smooth boundary, and define $\alpha := \frac{Vol(M_0)}{Vol(M)}$. Let B_α be a ball in $\mathbb{S}^{N+1}$ such that $\frac{Vol(B_\alpha)}{Vol(\mathbb{S}^{N+1})} = \alpha$. Then*

$$\frac{Vol(\partial M_0)}{Vol(M)} \geq \frac{Vol(\partial B_\alpha)}{Vol(\mathbb{S}^{N+1})}, \tag{2.5}$$

where $Vol(\partial M_0)$ denotes the N-dimensional volume of the hypersurface which bounds M_0, and similarly for ∂B_α. Applying (2.5) to ϵ neighbourhoods $(M_0)_\epsilon$, and integrating *over ϵ, we also have*

$$\frac{Vol((M_0)_\epsilon)}{Vol(M)} \geq \frac{Vol((B_\alpha)_\epsilon)}{Vol(\mathbb{S}^{N+1})}. \tag{2.6}$$

The first step in proving this theorem is to compare hypersurfaces on M and $\mathbb{S}^{N+1}$, following the work of Ernst Heintze and Hermann Karcher, [7]. Note also that we will be simplifying notation for this section: as g is an inner product of T_pM for all $p \in M$, we will denote $g(\cdot,\cdot)$ as $\langle \cdot,\cdot \rangle$.

2.3.1 Heintze-Karcher Comparison Theorem

Suppose that M is an $(N+1)$-dimensional Riemannian manifold with Ricci curvature $\geq \delta N$ everywhere. Let $H \subset M$ be a smooth hypersurface with ν a smooth normal unit vector field. Suppose also that we have $\overline{M}$ and $\overline{H}$, a corresponding "model" pair such that $\dim(\overline{M}) = N+1$, $\overline{M}$ has constant sectional curvature δ, $\dim(\overline{H}) = N$, and $\overline{H}$ is totally umbilical with mean curvature η.

Consider a point $p \in H$ and a corresponding unit normal vector ν_p, $\overline{p} \in \overline{H}$ and a corresponding unit normal vector $\overline{\nu}_{\overline{p}}$ in the model pair, such that $tr(S_{\nu_p}) \leq tr(\overline{S}_{\overline{\nu}_{\overline{p}}})$, where S and $\overline{S}$ denote the shape operators of H and, respectively, $\overline{H}$. Let $\{e_i\}_{i=1}^N$ be an orthonormal basis of T_pH so that $\{\nu_p, e_1, \ldots, e_N\}$ forms an orthonormal basis of T_pM.

Our goal is to compare the volume "distortions" of H and $\overline{H}$, in the following sense: Take a point $p \in H$ and the corresponding unit normal $\nu_p \in T_pM$, part of the orthonormal basis $\{\nu_p, e_1, e_2, \ldots, e_N\}$. Given the geodesic $\gamma(t) = \exp_p(t\nu_p)$ in M, consider the vector fields $E_i(t)$, $1 \leq i \leq N$, along γ such that $E_i(0) = e_i$, and $\nabla_{\gamma'(t)} E_i(t) = 0$ (ie. the parallel vector field along $\gamma(t)$ corresponding to e_i). These vector fields $E_i(t)$ will form the basis of comparison for a specific type of Jacobi field called an H-Jacobi field, denoted $Y_i(t)$, which describes the change of H when shifted for a time t along γ.

Definition 2.2. *If H is a hypersurface of M and γ is a geodesic normal to H such that $\gamma(0) = p \in H$, an H-Jacobi field Y is a Jacobi field on M which also satisfies the following conditions:*

$$Y(0) \in T_pH \text{ and } Y'(0) = S_{\nu_p}(Y(0)),$$

where S_ν is the shape operator of H with respect to the unit normal vector field ν.

An H-Jacobi field thus describes a variation of geodesics normal to H, which are determined by the principal curvatures. Let Y_i therefore be the H-Jacobi fields with initial conditions

$$Y_i(0) = E_i(0) \text{ and } Y_i'(0) = S_{\nu_p}(E_i(0)).$$

We may analogously consider at a point $\overline{p} \in \overline{H}$ the "model" $\overline{H}$-Jacobi fields $\overline{Y_i}(t)$ along a geodesic $\overline{\gamma}(t)$ corresponding to $\overline{\nu}_{\overline{p}}$ and the parallel $\overline{E_i}(t)$ constructed in the same manner as above. Since we took $\overline{H}$ to be totally umbilical with mean curvature η, the initial conditions of $\overline{Y}_i(t)$ will be

$$\overline{Y}_i(0) = \overline{E}_i(0) \text{ and } \overline{Y}_i'(0) = \eta \overline{E}_i(0),$$

so the $\overline{H}$-Jacobi fields are quite easily constructed.

As mentioned, we are interested in volume distortions, and will therefore consider the maps

$$f(t) = \frac{|Y_1(t) \wedge Y_2(t) \wedge \cdots \wedge Y_N(t)|}{|E_1(t) \wedge E_2(t) \wedge \cdots \wedge E_N(t)|}$$

and

$$\overline{f}(t) = \frac{|\overline{Y}_1(t) \wedge \overline{Y}_2(t) \wedge \cdots \wedge \overline{Y}_N(t)|}{|\overline{E}_1(t) \wedge \overline{E}_2(t) \wedge \cdots \wedge \overline{E}_N(t)|}$$

so that we may calculate the volume form of the space spanned by the Y_i's (respectively the $\overline{Y}_i$'s) at a time t given a basis $E_i(t)$ (respectively $\overline{E}_i(t)$).

Claim 2.1 (Main Result of Heintze-Karcher Comparison Theorem)**.** $f(t) \leq \overline{f}(t)$.

Proof of Claim. Since $\lim_{t\to 0} f(t) = \lim_{t\to 0} \overline{f}(t) = 1$, $(\log f)' \leq (\log \overline{f})'$, $\forall t > 0$, implies $f \leq \overline{f}$ as $x \mapsto \log(x)$ is an increasing function. It therefore suffices to prove $(\log f)' \leq (\log \overline{f})'$. Note that $|E_1(t) \wedge E_2(t) \wedge \cdots \wedge E_N(t)|$ and $|\overline{E}_1(t) \wedge \overline{E}_2(t) \wedge \cdots \wedge \overline{E}_N(t)|$ are constant by construction, so it suffices to prove $(\log |Y_1(t) \wedge Y_2(t) \wedge \cdots \wedge Y_N(t)|)' \leq (\log |\overline{Y}_1(t) \wedge \overline{Y}_2(t) \wedge \cdots \wedge \overline{Y}_N(t)|)'$.

Fix r (smaller than the focal distances along γ and $\overline{\gamma}$). After appropriate linear combinations, assume that $Y_1(r), \ldots Y_N(r)$ are orthonormal, and similarly for $\overline{Y}_1(r), \ldots, \overline{Y}_N(r)$.

By the orthonormality of the basis $\{Y_i(r)\}$,

$$
\begin{aligned}
&(\log |Y_1 \wedge Y_2 \wedge \cdots \wedge Y_N|)'(r) \\
&= \frac{|Y_1 \wedge Y_2 \wedge \cdots \wedge Y_N|'}{|Y_1 \wedge Y_2 \wedge \cdots \wedge Y_N|}(r) \\
&= |Y_1 \wedge Y_2 \wedge \cdots \wedge Y_N|'(r) = \sum_1^N (\sqrt{\langle Y_i, Y_i\rangle})'(r) \\
&= \sum_1^N \frac{2\langle Y_i, Y_i'\rangle}{2\sqrt{\langle Y_i, Y_i\rangle}}(r) \\
&= \sum_1^N \langle Y_i, Y_i'\rangle(r),
\end{aligned}
$$

since $|Y_1 \wedge \ldots Y_N|(r) = \prod_1^N \sqrt{\langle Y_i, Y_i\rangle(r)}$.

Using integration by parts,

$$
\int_0^r \langle Y_i', Y_i'\rangle dt = \langle Y_i, Y_i'\rangle(r) - \langle Y_i, Y_i'\rangle(0) - \int_0^r \langle Y_i, Y_i''\rangle dt
$$

$$
\Rightarrow \sum_1^N \langle Y_i, Y_i'\rangle(r) = \sum_1^N \left(\langle Y_i, Y_i'\rangle(0) + \int_0^r (\langle Y_i', Y_i'\rangle + \langle Y_i, Y_i''\rangle) dt \right),
$$

but since for all i, Y_i is an H-Jacobi field, we conclude that

$$
\sum_1^N \langle Y_i, Y_i'\rangle(r) = \sum_1^N \left(\langle Y_i, S_{\gamma'(0)} Y_i\rangle(0) + \int_0^r (\langle Y_i', Y_i'\rangle - \langle Y_i, R(Y_i, \gamma')\gamma'\rangle) dt \right). \quad (2.7)
$$

Let I_r denote the index form of $\gamma|_{[0,r]}$ with respect to H. Equation (2.7) reduces to

$$\sum_{1}^{N} I_r(Y_i, Y_i),$$

which means we must compare the index forms of M and $\overline{M}$. An important property of this index form is the following: if Y is an H-Jacobi field, X any continuously differentiable vector field along γ such that $X(r) = Y(r)$ and $X(0) \in T_pH$, then $I_r(Y,Y) \leq I_r(X,X)$, with equality if and only if $X = Y$ (see Appendix B). We will use this property to prove the theorem; it will suffice to show that $I_r(Y_i, Y_i) \leq \overline{I}_r(\overline{Y}_i, \overline{Y}_i)$.

To compare these two index forms, choose a linear isometric injection

$$i_r : T_{\overline{p}}\overline{M} \to T_pM \text{ satisfying}$$

$$i_r(\overline{\gamma}'(0)) = \gamma'(0),$$

$$i_r(T_{\overline{p}}\overline{H}) = T_pH \text{ and}$$

$$i_r(\overline{V}_r) = V_r,$$

where V_r (respectively, $\overline{V}_r$) is the N-dimensional subspace of T_pM ($T_{\overline{p}}\overline{M}$) obtained by parallel translation along γ to $\gamma(0)$ ($\overline{\gamma}$ to $\overline{\gamma}(0)$) the span of $Y_i(r)$ ($\overline{Y}_i(r)$), $1 \leq i \leq N$. Define the vector fields W_i along γ as follows:

$$W_i(t) = P_t \circ i_r \circ \overline{P}_{-t} \circ \overline{Y}_i(t), \quad t \in [0, r],$$

where P_t denotes parallel translation along γ from $\gamma(0)$ to $\gamma(t)$ and $\overline{P}_{-t}$ denotes parallel translation along $\overline{\gamma}$ from $\overline{\gamma}(t)$ to $\overline{\gamma}(0)$.

As i_r is isometric, we have that

$$|W_i(t)| = |\overline{Y}_i(t)| \text{ and } |W_i'(t)| = |\overline{Y}_i'(t)|.$$

From the property $i_r(\overline{V}_r) = V_r$, a suitable linear combination will give us $Y_i(r) = W_i(r)$, for all i, since these two sets span the same space. Also, since $\overline{Y}_i(0) \in T_{\overline{p}}\overline{H}$, we have that $W_i(0) \in i_r(T_{\overline{p}}\overline{H}) = T_pH$. Therefore for all i, W_i is a vector field along γ satisfying $W_i(r) = Y_i(r)$, which implies that

$$\sum_1^N I_r(Y_i, Y_i) \leq \sum_1^N I_r(W_i, W_i).$$

If we prove that $\sum_1^N I_r(W_i, W_i) \leq \sum_1^N I_r(\overline{Y}_i, \overline{Y}_i)$, then we are done.

By construction, the $\overline{Y}_i$ are orthonormal, hence the W_i are as well. By the assumption $tr(S_{\nu_p}) \leq tr(\overline{S}_{\overline{\nu}_{\overline{p}}})$, we have that

$$\sum_1^N \langle W_i, SW_i \rangle(0) \leq \sum_1^N \langle \overline{Y}_i, \overline{SY}_i \rangle(0).$$

By the assumption that the Ricci curvature on M is everywhere $\geq \delta N$, while it is equal to δN at all points of $\overline{M}$, we may conclude that

$$\begin{aligned}
&\sum_1^N \left(\int_0^r (|W_i'|^2 - \langle W_i, R(W_i, \gamma')\gamma' \rangle dt) \right) \\
&= \int_0^r \left(\sum_1^N |W_i'|^2 - Ric(\gamma') \right) dt \\
&\leq \sum_1^N \left(\int_0^r (|\overline{Y}_i'|^2 - \langle \overline{Y}_i, R(\overline{Y}_i, \gamma')\gamma' \rangle dt) \right).
\end{aligned}$$

Hence, we obtain $I_r(W_i, W_i) \leq \overline{I}_r(\overline{Y}_i, \overline{Y}_i)$ concluding the proof of the claim.

□

The functions f and $\overline{f}$ are in fact the Jacobians of the corresponding *normal exponential map* defined as follows: given a hypersurface H of a Riemannian manifold M and a normal unit vector field ν, define the normal exponential map

$$\exp_H : H \times \mathbb{R} \to M$$

by $\exp_H(p, t) = \exp_p(t\nu_p)$. Given the vector fields E_i along $\gamma(t) = \exp_p(t\nu_p)$, the vector fields Y_i are simply the pushforward of E_i by the normal exponential map (ie.

$Y_i = d\exp_H E_i$). Therefore the Jacobian at p and at time t, denoted $J(p,t)$, is given as follows:

$$\frac{|Y_1(t) \wedge \cdots \wedge Y_N(t)|}{|E_1(t) \wedge \cdots \wedge E_N(t)|}.$$

So the Heintze-Karcher comparison theorem tells us that

$$|J(p,t)| \leq |\overline{J}(\overline{p},t)|,$$

as we have proven Claim 2.1.

Before proceeding, let us compute $|\overline{J}(\overline{p},t)|$. From now on, our model space $\overline{M}$ is the space of constant sectional curvature 1, ie. the $(N+1)$-dimensional unit sphere. Recall that the $\overline{H}$-Jacobi fields satisfy the initial conditions $\overline{Y}_i(0) = E_i(0) = e_i$ and $\overline{Y}'_i(0) = \eta E_i(0) = \eta e_i$, so, since we are in a space of constant curvature $\delta = 1$, the Jacobi equation for these fields reduces to $\overline{Y}''_i + \overline{Y}_i = 0$. Therefore solving this second order ODE with the given initial conditions as above, we find that $\overline{Y}_i(t) = (\cos(t) - \eta\sin(t))E_i(t)$, which implies that

$$|\overline{J}(p,t)| = (\cos(t) - \eta\sin(t))^N.$$

2.3.2 Proof of Theorem 2.4

Let us now turn our attention to hypersurfaces which divide our manifold M into two parts which we will denote M_0 and M_1. Fix $\alpha \in (0,1)$, and consider all hypersurfaces which divide M into two parts, M_0 and M_1, such that the relative volume, or $\dfrac{Vol(M_0)}{Vol(M)}$, is α and, respectively, $\dfrac{Vol(M_1)}{Vol(M)} = (1-\alpha)$. Among all such hypersurfaces there exists one with minimal N-dimensional volume. Such a hypersurface, which we will call H, has constant mean curvature. To see this, consider a (small) smooth function f on H, and consider the hypersurface H_f parametrized by

$$p \in H \mapsto \exp_p(f(p)\nu_p).$$

Conversely, any hypersurface $\tilde{H}$ close enough to H can be written as H_f for some small function f on H. If we consider the N-volume functional

$$\mathcal{A}(f) = \int_{H_f} dvol_{H_f},$$

then the differential of this map when $f = 0$ is described as (see [8]):

$$D\mathcal{A}|_{f=0}(V) = -\int_H S_\nu V dvol_H,$$

where S is the shape operator of H with respect to the normal unit vector field ν.

Now given any small function f on H, consider the decomposition $f = f^+ - f^-$, where $f^\pm := \max(\pm f, 0)$. Consider the $(N+1)$-dimensional spaces

$$B_{f^\pm} := \{\exp_p(t\nu_p) \; : \; \pm t \in (0, f^\pm)\},$$

which are precisely the domains between H and $H_{f^\pm}$, and define the $(N+1)$-volume functional

$$\mathcal{V}(f) := \int_{B_{f^+}} dvol_M - \int_{B_{f^-}} dvol_M,$$

so that volumes are counted positively when $f > 0$ and negatively when $f < 0$. The differential of this function at $f = 0$ is given by

$$D\mathcal{V}(V) = \int_H V dvol_H.$$

From the above equations, the goal is to minimize $\mathcal{A}$ while keeping $\mathcal{V}$ constant, a classical Lagrange multiplier problem. Therefore, we wish to find the critical points of

$$\mathcal{E} := \mathcal{A} + \lambda\mathcal{V},$$

but

$$D\mathcal{E}|_{f=0}(V) = \int_H (\lambda - S_\nu) V dvol_H,$$

and is thus critical when $S_\nu = \lambda$ or, equivalently, H has constant mean curvature. Further analysis of this problem can be found in [8].

With this in mind, we will prove that $\dfrac{Vol(H)}{Vol(M)} \geq \dfrac{Vol(\partial B_\alpha)}{Vol(\mathbb{S}^{N+1})}$, where H is the hypersurface which minimizes the N-dimensional volume. Consequently, for *any* hypersurface H' dividing M into two parts with relative volumes α and $(1-\alpha)$, we have $Vol(H') \geq Vol(H)$ as H has minimal N-dimensional volume among all such hypersurfaces, and so $\dfrac{Vol(H')}{Vol(M)} \geq \dfrac{Vol(\partial B_\alpha)}{Vol(\mathbb{S}^{N+1})}$, proving the theorem.

Let η denote the value of the constant mean curvature of H relative to the normal vector field which is in the direction of M_0. Since $|J(p,t)| \leq (\cos(t) - \eta\sin(t))^N$, and this latter function is independent of p, we may apply Fubini's theorem and integrate over our hypersurface H first, followed by the integration with respect to t, and conclude:

$$Vol(M_0) \leq Vol(H) \int_0^r (\cos(t) - \eta\sin(t))^N dt, \tag{2.8}$$

where r is the first zero of $\cos(t) - \eta\sin(t)$, as we only wish to take the domain on which this function is positive which corresponds to the domain on which the exponential map $\exp_p$ is a diffeomorphism onto its image. Note that the zero occurs when $\cos(r) - \eta\sin(r) = 0$, or $\eta = \cot(r)$.

It remains to calculate the value of the integral $\int_0^r (\cos(t) - \eta\sin(t))^N dt$ which is the object of the following lemma.

Lemma 2.1. *The geodesic ball $B(r)$ of radius r in $\mathbb{S}^{N+1}$ has boundary $\partial B(r)$ of constant mean curvature $\eta = \cot(r)$ and*

$$\frac{Vol(B(r))}{Vol(\partial B(r))} = \int_0^r (\cos(t) - \eta\sin(t))^N dt. \tag{2.9}$$

Proof. Consider a point $p \in \partial B(r)$ and an orthonormal basis $\{f_i\}_1^N$. Let ν_p be the unit vector normal to $\partial B(r)$ at p in the direction away from $B(r)$. Let $\{F_i(t)\}_1^N$ be the vector fields generated by parallel translation of f_i's along the geodesic γ given by $\gamma(0) = p$, $\gamma'(0) = \nu_p$. We wish to consider the $\partial B(r)$-Jacobi fields $\{Y_i\}_1^N$ such that $Y_i(0) = f_i$ and $Y_i'(0)$ changes according to the mean curvature. But $\partial B(r)$

is exactly an N-dimensional sphere of radius $\sin(r)$ embedded in $\mathbb{S}^{N+1}$, so setting $Y_i(t) = \frac{\sin(r+t)}{\sin(r)} F_i(t)$ will be the $\partial B(r)$-Jacobi fields. Therefore

$$Y_i(0) = \frac{\sin(r)}{\sin(r)} F_i(0) = f_i \text{ and } Y_i'(0) = \frac{\cos(r)}{\sin(r)} f_i = S_{\nu_p}(f_i),$$

so $S_{\nu_p}(f_i) = \cot(r) f_i$ for all i. Therefore $\partial B(r)$ is totally umbilical with mean curvature $\eta = \cot(r)$.

Taking this mean curvature value into account, equation (2.9) becomes

$$\begin{aligned} &\int_0^r (\cos(t) - \cot(r)\sin(t))^N dt \\ &= \int_0^r \frac{(\sin(r)\cos(t) - \cos(r)\sin(t))^N}{\sin^N(r)} dt \\ &= \int_0^r \frac{\sin^N(r-t)}{\sin^N(r)} dt, \end{aligned}$$

and after making a change of variable $t' = r - t$, this equation reduces to

$$\frac{1}{\sin^N(r)} \int_0^r \sin^N(t') dt'.$$

This is exactly the ratio $\dfrac{Vol(B(r))}{Vol(\partial B(r))}$, as is easily seen from the fact that these volumes are defined by:

$$Vol(B(r)) = Vol(\mathbb{S}^N) \int_0^r \sin^N(t) dt$$

and

$$Vol(\partial B(r)) = Vol(\mathbb{S}^N) \sin^N(r),$$

as was shown in our presentation of coordinates in Section 1.2.

□

Let us denote by $a(r) := \dfrac{Vol(\partial B(r))}{Vol(B(r))}$. By rearranging, and dividing by $Vol(M)$, we may rewrite inequality (2.8) as

$$\frac{Vol(H)}{Vol(M)} \geq \frac{a(r) Vol(M_0)}{Vol(M)}.$$

Note that if we will replace M_0 by M_1, and insured that the normal vector field at the separating boundary was pointing in the direction of M_1 now, we will obtain the same result, only the ratio would be $1-\alpha$ instead of α. Therefore by replacing M_0 with M_1, η with $-\eta$, and r with $\pi - r$, we also have

$$\frac{Vol(H)}{Vol(M)} \geq \frac{a(\pi - r)Vol(M_1)}{Vol(M)},$$

so

$$\frac{Vol(H)}{Vol(M)} \geq \max\{\alpha a(r), (1-\alpha)a(\pi - r)\}$$
$$\geq \inf_{x\in[0,\pi]} \max\{\alpha a(x), (1-\alpha)a(\pi - x)\}.$$

Claim 2.2. *On $[0,\pi]$, the function $a(x)$ is decreasing.*

Recall that $\frac{1}{a(x)} = \frac{Vol(B(x))}{Vol(\partial B(x))} = \int_0^x (\cos(t) - \cot(x)\sin(t))^N dt$. Let $f(x,t) = \cos(t) - \cot(x)\sin(t)$. Note that for $x \in [0,\pi]$, $t \in [0,x]$, $f(x,t) \geq 0$, and for $x' \geq x$, $f(x',t) \geq f(x,t)$. Therefore $f(x',t)^N \geq f(x,t)^N$, so integrating over t gives us

$$\int_0^{x'} (\cos(t) - \cot(x')\sin(t))^N dt \geq \int_0^{x} (\cos(t) - \cot(x')\sin(t))^N dt$$
$$\geq \int_0^{x} (\cos(t) - \cot(x)\sin(t))^N dt.$$

Therefore $\frac{1}{a(x)}$ is increasing and, consequently, $a(x)$ is decreasing.

□

If $a(x)$ is decreasing, then $a(\pi - x)$ is increasing so that the infimum of $\max\{\alpha a(x), (1-\alpha)a(\pi - x)\}$ is attained for some $x_0 \in [0,\pi]$ where

$$\alpha a(x_0) = (1-\alpha)a(\pi - x_0).$$

By symmetry, $Vol(\partial B(x_0)) = Vol(\partial B(x_0 - \pi))$, so that solving for α, we get

$$\alpha = \frac{Vol(B(x_0))}{Vol(\mathbb{S}^{N+1})},$$

so $B(x_0) = B_\alpha$, implying that $B(x_0)$ gives our desired inequality. Therefore we conclude

$$\frac{Vol(H)}{Vol(M)} \geq \frac{Vol(\partial B_\alpha)}{Vol(\mathbb{S}^{N+1})}.$$

As a consequence of this inequality, for any Borel subset $M_0 \subset M$ such that $\alpha = \frac{Vol(M_0)}{Vol(M)} = \frac{Vol(B_\alpha)}{Vol(\mathbb{S}^{N+1})}$, we have that $\frac{Vol(\partial M_0)}{Vol(M)} \geq \frac{Vol(\partial B_\alpha)}{Vol(\mathbb{S}^{N+1})}$. Fix $t > 0$ sufficiently small and consider the t-neighbourhoods of M_0 and B_α. We would again have the same inequality for the boundaries of $(M_0)_t$ and $(B_\alpha)_t$: $\frac{Vol(\partial((M_0)_t))}{Vol(M)} \geq \frac{Vol(\partial((B_\alpha)_t))}{Vol(S^{N+1})}$. Therefore, integrating with respect to t from 0 to ϵ, we conclude the inequality (2.6):

$$\frac{Vol((M_0)_\epsilon)}{Vol(M)} \geq \frac{Vol((B_\alpha)_\epsilon)}{Vol(\mathbb{S}^{N+1})}.$$

2.4 Concentration Results

As a Corollary to the Lévy-Gromov isoperimetric inequality, we have the following concentration result:

Corollary 2.1. *Let M be an N-dimensional manifold and suppose that its Ricci curvature is everywhere greater than that of $\mathbb{S}^N$. Let $f : M \to \mathbb{R}$ be a 1-Lipschitz function. Then there exists $m \in \mathbb{R}$ such that, for any $t \geq 0$,*

$$\nu(\{x \in M,\ |f(x) - m| \geq t\}) \leq 2e^{-\frac{t^2}{2D^2}},$$

where $D = \frac{1}{\sqrt{N-1}}$ and ν is the natural measure on M, normalized so that $\nu(M) = 1$.

Proof. Let m be a median of f, such that $\nu(\{x,\ f(x) \geq m\}) \geq \frac{1}{2}$ and $\nu(\{x,\ f(x) \leq m\}) \geq \frac{1}{2}$. Denote by M_0 the set $\{x,\ f(x) \geq m\}$ and by M_1 the set $\{x,\ f(x) \leq m\}$. As in the case of Chapter 1, because f is 1-Lipschitz, the set $\{x,\ f(x) - m \geq t\}$ contains $(M_0)_t$, where $(M_0)_t$ denotes the t-neighborhood of M_0. By the Gromov-Lévy inequality, we know that $\nu((M_0)_t) \geq \nu'(A_t)$, where ν' denotes in this case the normalized measure on $\mathbb{S}^N$ and A is a half-sphere. So $\nu((M_0)_t^c) \leq \nu'(A_t^c) \leq e^{-\frac{(N-1)t^2}{2}}$, by Proposition 1.1. Similarly, $\nu((M_1)_t^c) \leq e^{-\frac{(N-1)t^2}{2}}$. Combining the two, we obtain the desired result.

□

In conclusion, due to the Lévy-Gromov isoperimetric inequality, we have an extension of the concentration of measure phenomenon from the N-sphere to all Riemannian manifolds with Ricci curvature greater than that of the N-sphere. We will further investigate if such a phenomena can be extended to other metric measure spaces lacking the rigid Riemannian structure. In fact, we will go to the other extreme, when such spaces are discrete, showing that a generalized notion of Ricci curvature plays a similar role in the latter context, implying a concentration of measure.

Chapter 3

A Generalization of Ricci Curvature

The goal of this chapter is to introduce a generalization of Ricci curvature to metric spaces. For this chapter, we work under the assumption that (X, d, ν) is a separable and complete metric space with a probability measure ν. The main idea of this generalization is as follows: take a point $x \in X$ and a neighboring point $y \in X$. Imagine *shifting* a ball B_x of a fixed radius centered at x into a ball B'_y of the same radius centered at y. In $\mathbb{R}^N$, the average distance between the points of the two balls would be $d(x, y)$. In $\mathbb{S}^N$, however, the average distance between the points would in fact be *less* than $d(x, y)$. On a general Riemannian manifold, this average distance would vary according to the Ricci curvature. We note that we can consider average distances between balls on an arbitrary metric space, thus one can consider defining Ricci curvature in a more generalized setting by observing the effects on average distances between points in two nearby balls. The tools behind this generalization, however, are probabilistic in nature and therefore do not appear at first to always match that of the classical Ricci curvature. In an attempt to show that there is a good fit between the two, let us present first a *visual* geometric definition of the classical Ricci curvature after which we will proceed with the definition of the generalized Ricci curvature.

3.1 Alternative Definition of the Classical Sectional and Ricci Curvature

Suppose that (M, g) is an N dimensional Riemannian manifold, and suppose that $x \in M$. Take two orthogonal unit vectors $v, w \in T_xM$, and suppose that for some small $\delta > 0$, $y = exp_x(\delta v)$. Consider the parallel transport of w along v to y, and denote this vector $w' \in T_yM$. Now define two points: $x' = exp_x(\epsilon w)$, $y' = \exp_y(\epsilon w')$. In Euclidean space, the points x,y,x',y' would form a quadrilateral, and the distance between x' and y' would be $\delta = d(x, y)$. However, in the case of a Riemannian manifold this distance depends on the curvature of the space: in this case the sectional curvature K of the plane spanned by v and w. As seen in [9], $d(x', y')$ can be approximated as follows.

Theorem 3.1 (Sectional Curvature)**.** *Let (M, g) be a smooth complete Riemannian manifold. Let v, w be unit orthogonal tangent vectors at $x \in X$. Let ϵ, $\delta > 0$. Let $y = \exp_x(\delta v)$ and let w' be the tangent vector at y obtained by parallel transport of w along the geodesic from x to y. Then*

$$d(\exp_x(\epsilon w), \exp_y(\epsilon w')) = \delta(1 - \frac{\epsilon^2}{2}K(v, w) + \mathcal{O}(\epsilon^3))$$

as ϵ, $\delta \to 0$. Here $K(v, w)$ is the sectional curvature of the plane spanned by v, w.

Proof. To prove the theorem, we will construct an appropriate Jacobi field along the curve $c_w^{(x)}(t) = \exp_x(tw)$. Let Y be the Jacobi field along $c_w^{(x)}$ with the initial conditions:

$$Y(0) = v \quad \text{and} \quad Y'(0) = \nabla_T Y|_0 = 0$$

where T denotes the tangent to $c_w^{(x)}(t)$. We are basically taking the curve which connects x an y and parallel translating it along w for a time ϵ, which is why we take $Y'(0) = 0$.

Let us take the third order Taylor expansion of $||Y(t)|| = \langle Y(t), Y(t)\rangle^{1/2}$ about 0. The first two derivatives of the length of $Y(t)$ at $t = 0$ are:

$$(\langle Y(t), Y(t)\rangle^{1/2})'|_0 = \frac{\langle Y'(0), Y(0)\rangle}{\langle Y(0), Y(0)\rangle^{1/2}} = 0$$

$$\langle Y(t), Y(t)\rangle''|_0 = \frac{\langle Y''(0), Y(0)\rangle + \langle Y'(0), Y'(0)\rangle}{\langle Y(0), Y(0)\rangle^{1/2}} - \frac{\langle Y'(0), Y(0)\rangle^2}{\langle Y(0), Y(0)\rangle^{3/2}}$$
$$= \frac{\langle Y''(0), Y(0)\rangle}{\langle Y(0), Y(0)\rangle^{1/2}}.$$

It can easily be checked that after simplification, the only term of the third derivative $\langle Y(t), Y(t)\rangle'''|_0$ that does not vanish is

$$\frac{\langle Y'''(0), Y(0)\rangle}{\langle Y(0), Y(0)\rangle^{1/2}}.$$

As Y is a Jacobi field, it satisfies the Jacobi equation $Y'' = \nabla_T \nabla_T Y = -R(Y,T)T$, so we may rewrite the second derivative at zero as

$$\frac{-\langle R(Y,T)T, Y\rangle | t = 0}{\langle Y(0), Y(0)\rangle^{1/2}}.$$

Since $Y(0) = v$ and $T(0) = w$ are both of unit length, we have

$$\frac{\langle R(Y,T)Y, T\rangle}{\langle Y, Y\rangle^{1/2}|_0} = \frac{K(Y,T)|_0(||Y||_0 \cdot ||T||_0 - \langle Y, T\rangle^2|_0)}{||Y||_0}$$

$$= K(Y,T)|_0 ||T||_0$$

$$= K(v,w)||w||$$

$$= K(v,w),$$

where the orthonormality of v and w was used.

As for the third derivative, it suffices to see that

$$\frac{\langle Y'''(0), Y(0)\rangle}{\langle Y(0), Y(0)\rangle^{1/2}} = \frac{\nabla(\langle R(Y,T)T, Y\rangle)}{\langle Y(0), Y(0)\rangle^{1/2}}.$$

The Taylor expansion of $||Y(t)||$ will therefore look as follows:

$$\langle Y(t), Y(t)\rangle^{1/2}|_{t=0} = \langle Y(0), Y(0)\rangle^{1/2} + \frac{1}{2!}\frac{-\langle R(Y,T)T), Y\rangle|_0}{\langle Y(0), Y(0)\rangle^{1/2}}\,\epsilon^2 + \mathcal{O}(\epsilon^3)$$

$$= \langle v, v\rangle^{1/2} - \frac{1}{2!}\, K(v,w)\epsilon^2 + \mathcal{O}(\epsilon^3),$$

and integrating from zero to δ, we obtain

$$d(\exp_x(\epsilon w), \exp_y(\epsilon w')) = \delta\left(1 - \frac{\epsilon^2}{2}K(v,w) + \mathcal{O}(\epsilon^3)\right).$$

We can therefore see that for $K(v,w) > 0$, x' and y' are closer to one another than x and y. Indeed, imagine two points x and y on an equator of $\mathbb{S}^2$, and transporting these points along longitudinal lines. The further one travels along these lines (up to a distance $\pi/2$), the closer these points become. Also, if $K(v,w) < 0$, the points x' and y' are growing further apart.

□

Recall that the Ricci curvature in the direction of a unit tangent vector v at a point is the average of all sectional curvatures $K(v,w)$ after all tangent vectors w at the point orthonormal to v. Hence, similarly with Theorem 3.1, an analogous view can be obtained for the Ricci curvature in the following way: suppose that we have a point $x \in M$, $v \in T_xM$ a unit vector, $\epsilon, \delta > 0$ and $y = \exp_x(\delta v)$. Let S_x denote the sphere of radius ϵ in the tangent plane at x, and S_y the ϵ-sphere on T_yM. If we map the S_x to S_y via parallel transport, the points travel an average distance of

$$\delta\left(1 - \frac{\epsilon^2}{2(N-1)}\,\mathrm{Ric}(v) + \mathcal{O}(\epsilon^3)\right), \tag{3.1}$$

where N denotes the dimension of the manifold. This property of the Ricci curvature is what we will use to generalize it to an arbitrary metric space.

3.2 Markov Chains, Random Walks, and Transportation Distances

We will now present the probabilistic and measure-theoretic background necessary to define the "coarse" Ricci curvature. Let (X, d, ν) be a complete separable metric space endowed with a probability measure. A *random walk* $\{m_x\}_{x\in X}$ is defined by the following data: let m_x be a probability measure on X dependent, in a measurable way, on the point x, and define the generated Markov chain to be the jumps from x to a random point with probability corresponding to m_x. The n-step transition probability, or the probability of traveling from x to y in n steps, is given as follows:

$$dm_x^{*n}(y) := \int_{z\in X} dm_x^{*(n-1)}(z)dm_z(y),$$

where $m_x^{*1} = m_x$.

To give an example, suppose that $X = \{0,1\}^N$, d is the rescaled Hamming metric ($d(x,y) = \frac{1}{N}\sum_{i=1}^{N}|x_i - y_i|$), and that ν is the uniform measure on this space, such that $\nu(\{x\}) = \frac{1}{2^N}$ for all $x \in X$. We may define a random walk on this space as follows: each point $x \in X$ has N neighbors, implying that a ball $B_x(\frac{1}{N})$ would contain exactly $N+1$ elements. Define m_x to be the measure such that $m_x(y) = \frac{1}{N+1}$ for each $y \in B_x(\frac{1}{N})$, and $m_x(y) = 0$ otherwise. The corresponding random walk is the process of jumping from x to a point y in $B_x(\frac{1}{N})$ randomly, and then from y to a point in $B_y(\frac{1}{N})$ randomly, etc. Note that these measures m_x are proportional to ν: indeed, $m_x = \frac{\nu|_{B_x(\frac{1}{N})}}{\nu(B_x(\frac{1}{N}))}$; however, it is not always necessary that this be true. Indeed, one can consider the *lazy random walk* on the cube defined as follows: $m_x(x) = \frac{1}{2}$, $m_x(y) = \frac{1}{2N}$ for any neighbor y of x, and $m_x(z) = 0$ for all other $z \in X$.

This set of measures can be thought of as a natural replacement for balls centered at a point x. The advantage of the random walk is that we may define these measures m_x to suit specific needs. Indeed, the difference between the measures m_x corresponding to the uniform measure and the lazy random walk only can imply different properties for the metric space. We will return to the case of the discrete cube from this perspective at a later time.

3.2.1 Transportation Distances

The notion of transportation distances comes up in the problem of optimizing the transport of a given quantity from one location to another. The most popular example is that of moving a pile of sand from one place to another in the most cost-effective way. This is often modeled by a *transference plan* from a measure ν_1 to another measure ν_2 of equal mass. Formally, a transference plan is a measure π on $X \times X$ such that $\int_y d\pi(x,y) = d\nu_1(x)$ and $\int_x d\pi(x,y) = d\nu_2(y)$. Let $\Pi(\nu_1, \nu_2)$ be the set of all transference plans, and define the L_1 *transportation distance* $W_1(\nu_1, \nu_2)$ to be the optimal transference plan:

$$W_1(\nu_1, \nu_2) = \inf_{\pi \in \Pi(\nu_1, \nu_2)} \int \int d(x,y) d\pi(x,y).$$

We do not wish to delve deeper into the subject of transportation distances and optimal transference plans. We will only mention the *Kantorovich duality* property which is essential in our further discussion:

$$W_1(\nu_1, \nu_2) = \sup_{f: X \to \mathbb{R}} \left\{ \int f d\nu_1 - \int f d\nu_2 \, ; \, f \text{ 1-Lipschitz} \right\}.$$

Using Kantorovich duality, it is easy to prove a triangle inequality for the transportation distance.

Proposition 3.1. *Given three measures ν_1, ν_2, ν_3 of equal mass, then the following inequality holds*

$$W_1(\nu_1, \nu_2) \leq W_1(\nu_1, \nu_3) + W_1(\nu_3, \nu_2).$$

Proof. Indeed, by the definition of the supremum,

$$\sup_{f: X \to \mathbb{R}} \left\{ \int f d\nu_1 - \int f d\nu_2 \, ; \, f \text{ 1-Lipschitz} \right\}$$

$$= \sup_{f: X \to \mathbb{R}} \left\{ \left(\int f d\nu_1 - \int f d\nu_3 \right) + \left(\int f d\nu_3 - \int f d\nu_2 \right) \, ; \, f \text{ 1-Lipschitz} \right\}$$

$$\leq \sup_{g: X \to \mathbb{R}} \left\{ \int g d\nu_1 - \int g d\nu_3 \, ; \, g \text{ 1-Lipschitz} \right\} + \sup_{h: X \to \mathbb{R}} \left\{ \int h d\nu_3 - \int h d\nu_2 \, ; \, h \text{ 1-Lipschitz} \right\}$$

□

3.3 Coarse Ricci Curvature

Equipped with a set of measures $\{m_x\}_{x \in X}$ and transportation distances $W_1(m_x, m_y)$, we are now ready to define the coarse Ricci curvature on X.

Definition 3.1. *[10] Let (X, d) be a metric space with a set $\{m_x\}_{x \in X}$ of probability measures on X. Let x and y be points in X. The* coarse Ricci curvature *$\kappa(x, y)$ of X along xy is the value which satisfies:*

$$W_1(m_x, m_y) = (1 - \kappa(x, y))d(x, y).$$

Note the similarity with equation 3.1. Notice also that, by definition, $\kappa(x, y) \leq 1$ for all x and y. Let us look again at $X = \{0, 1\}^N$ with the set m_x of probability measures such that $m_x(y) = \dfrac{1}{N+1}$ for $y \in B_x(\frac{1}{N})$, 0 otherwise, and let us focus on the neighboring points $x = (0, 0, \ldots, 0)$, $y = (1, 0, \ldots, 0)$. The neighbors of x are the points $x_1 = y = (1, 0, \ldots, 0)$, $x_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $x_N = (0, 0, \ldots, 0, 1)$, while for y they are $y_1 = x = (0, 0, \ldots, 0)$, $y_2 = (1, 1, 0, \ldots, 0)$, $y_3 = (1, 0, 1, 0, \ldots, 0)$, $\ldots$, $y_N = (1, 0, \ldots, 1)$. Notice that for $2 \leq i \leq N$, $d(x_i, y_i) = \dfrac{1}{N}$, and $d(x_1, y) = d(x, y_1) = 0$. Let us now compute $W_1(m_x, m_y)$ via the Kantorovich duality.

First, note that if $f : X \to \mathbb{R}$, then

$$\int f dm_x = \frac{f(x) + f(x_1) + \cdots + f(x_N)}{N+1}$$

and

$$\int f dm_y = \frac{f(y) + f(y_1) + \cdots + f(y_N)}{N+1}.$$

After rearranging,

$$\int f dm_x - \int f dm_y$$
$$= \frac{(f(x) - f(y_1)) + (f(y) - f(x_1)) + (f(x_2) - f(y_2)) + \ldots (f(x_N) - f(y_N))}{N+1}$$
$$= \frac{(f(x_2) - f(y_2)) + (f(x_3) - f(y_3)) + \cdots + (f(x_N) - f(y_N))}{N+1}.$$

By taking the supremum after all such 1-Lipschitz functions f, we note that $\sup(f(x_i)-f(y_i)) = d(x_i, y_i) = \frac{1}{N}$ for all i, so that

$$\sup\left[\int f dm_x - \int f dm_y\right] = \frac{(N-1)/N}{N+1} = \frac{N-1}{N(N+1)},$$

which implies that $W_1(m_x, m_y) = \frac{N-1}{N(N+1)}$. Therefore $\frac{N-1}{N(N+1)} = (1-\kappa(x,y))d(x,y) = (1-\kappa(x,y))\frac{1}{N}$ and, consequently, the coarse Ricci curvature of X along xy is $\kappa(x,y) = 1 - \frac{N-1}{N+1} = \frac{2}{N+1} > 0$.

To give another example, let us look at Π^N, the symmetric group of permutations of $\{1,\ldots,N\}$, $N \geq 2$, equipped with the following normalized metric: for $x,\ y \in \Pi^N$, $d(x,y) = \frac{1}{N}\text{Card}\{i;\ x(i) \neq y(i)\}$. We will use the following notation to represent elements of Π^N: if x acts as follows: $x(1) = j_1,\ x(2) = j_2,\ \ldots, x(N) = j_N$, then $x = (j_1, j_2, \ldots, j_N)$, so that the identity element is $(1, 2, \ldots, N)$.

Given $x \in \Pi^N$, the minimal distance to any other element of the group is $\frac{2}{N}$; indeed, any other permutation y which differs from x will differ by the images of at least two of the objects $1, \ldots, N$, otherwise it will not be a permutation/bijection. Furthermore, any element x has exactly $\binom{N}{2} = \frac{N(N-1)}{2}$ neighbors of (minimal) distance $\frac{2}{N}$, so let us define m_x as the measure which assigns mass $\frac{1}{\frac{N(N-1)}{2}+1} = \frac{2}{N(N-1)+2}$ to each point in the ball $B_x(\frac{2}{N})$.

Before calculating the coarse Ricci curvature of this space, let us first analyze the structure of these neighbourhoods. We will concentrate on the neighborhood of the identity $(1, 2, 3, \ldots, N)$, and through symmetry, or the homogeneity of the space, we claim that the structure of the ball around the identity will be similarly throughout. The neighboring points of the identity will be the following: The first $N-1$ points will be permuting only 1 to 2, 1 to 3, etc.. The next $N-2$ will be permuting only 2 to 3, 2 to 4, etc.. So, the ball $B_x(\frac{2}{N})$ will consist of all the transpositions

$$(1, 2, 3, \ldots, N-1, N)$$

$$(2, 1, 3, \ldots, N-1, N)$$

$$\vdots$$

$$(N, 2, 3, \ldots, N-1, 1)$$

$$\vdots$$

$$(1, 2, 3, \ldots, N, N-1, N-2)$$

$$(1, 2, 3, \ldots, N-2, N, N-1).$$

Consider the neighboring point $y = (1, 2, 3, \ldots, N-2, N, N-1)$. The Table 3.1 reflects that all of the points in $B_x(\frac{2}{N})$ have neighbors in $B_y(\frac{2}{N})$.

Claim 3.1. *If x, $y \in \Pi^N$ such that $d(x, y) = \dfrac{2}{N}$, then $\kappa(x, y) = \dfrac{4}{N(N-1)+2}$.*

Proof of Claim. By symmetry, it is enough to prove it for the above points x and y. Let $n = \dfrac{N(N-1)}{2}$. We see from the table that $x_n = y$, $y_n = x$. Given $f : \Pi^N \to \mathbb{R}$,

$$\int f dm_x = \frac{(f(x) + f(y) + f(x_1) + \cdots + f(x_{n-1}))}{n+1}$$

and

$$\int f dm_y = \frac{(f(y) + f(x) + f(y_1) + \cdots + f(y_{n-1}))}{n+1},$$

Table 3.1: Neighbors of x Vs. Neighbors of y

i	x_i	y_i
-	(1,2,3,. . . , N-1,N)	(1,2,3,. . . ,N,N-1)
1	(2,1,3,. . . , N-1,N)	(2,1,3,. . . ,N,N-1)
$\vdots$	$\vdots$	$\vdots$
N-1	(N-1,2,3,. . . ,1,N)	(N,2,3,. . . ,1,N-1)
N	(N,2,3,. . . ,N-1,1)	(N-1,2,3,. . . ,N,1)
$\vdots$	$\vdots$	$\vdots$
$\frac{N(N-1)}{2} - 1$	(1,2,3,. . . ,N,N-1,N-2)	(1,2,3,. . . ,N-1,N,N-2)
$\frac{N(N-1)}{2}$	(1,2,3,. . . ,N,N-1)	(1,2,3,. . . ,N-1,N)

so

$$\int f dm_x - \int f dm_y = \frac{(f(x_1) - f(y_1)) + \cdots + (f(x_{n-1}) - f(y_{n-1}))}{n+1}.$$

For each pair x_i, y_i, $d(x_i, y_i) = \frac{2}{N}$, so that, by taking the supremum after all 1-Lipschitz functions, we will obtain that

$$W_1(x,y) = \left(\frac{n-1}{n+1}\right)\frac{2}{N} = \left(\frac{N(N-1)-2}{N(N-1)+2}\right)\frac{2}{N},$$

so

$$\left(\frac{N(N-1)-2}{N(N-1)+2}\right)\frac{2}{N} = (1-\kappa(x,y))\frac{2}{N} \Rightarrow \kappa(x,y) = \frac{4}{N(N-1)+2}.$$

□

Note that for both of these examples, we only computed the coarse Ricci curvature between neighbouring points. It turns out however that this is sufficient in a major class of examples, called ϵ-geodesic spaces:

Definition 3.2. *[10] A metric space (X, d) is said to be* ϵ-geodesic *if, for any two points x and $y \in X$, there exist an integer n and a sequence $x = x_0, x_1, \ldots, x_n = y$ such that $d(x,y) = \sum_0^{n-1} d(x_i, x_{i+1}) = d(x,y)$ and $d(x_i, x_{i+1}) \leq \epsilon$.*

Both of the examples we have seen are ϵ-geodesic. Indeed, by construction the Hamming cube with a rescaled metric is $\frac{1}{N}$-geodesic, and the symmetric group Π^N is $\frac{1}{2N}$-geodesic.

Proposition 3.2. *[10] Suppose (X, d) is ϵ-geodesic and $\kappa(x,y) \geq \kappa_0$ for any x,$y \in X$ such that $d(x,y) \leq \epsilon$. Then $\kappa(x,y) \geq \kappa_0$ for any pair $x, y \in X$.*

Proof. Fix x and y in the space and let $\{x_i\}$ be a sequence of points as in the definition of the ϵ-geodesic space. Then, by the triangle inequality for transportation distances (3.1),

$$(1-\kappa(x,y))d(x,y) = W_1(m_x, m_y) \leq \sum_0^{n-1} W_1(m_{x_i}, m_{x_{i+1}})$$

$$\leq \sum_{0}^{n-1}(1-\kappa_0)d(x_i, x_{i+1}) = (1-\kappa_0)d(x,y).$$

□

3.4 Coarse Ricci Curvature and Concentration of Measure

As we saw in the previous chapter, on smooth manifolds, positive Ricci curvature implies concentration of measure. We aim for a similar conclusion to hold for the coarse Ricci curvature on arbitrary metric measure spaces. To investigate this question, let us introduce some notation following Ollivier [10].

In this section, we let (X, d, ν) be a probability space with metric d, ν a probability measure and $\{m_x\}_{x\in X}$ be a random walk. We say that ν is *invariant* with respect to the random walk $\{m_x\}$ if for all $x \in X$, $d\nu(x) = \int_y d\nu(y)dm_y(x)$. In the examples seen so far, ν was invariant with respect to our random walks. Throughout this section, we assume that the coarse Ricci curvature $\kappa_0 > 0$ and that ν is an invariant measure. Define

$$D_x^2 := \frac{\sup\{Var_{m_x} f, f : Supp\, m_x \to \mathbb{R} \quad \text{1-Lipschitz}\}}{\kappa_0},$$

where $Var_{m_x} f := \displaystyle\int (f - \mathbb{E}_{m_x} f)^2 dm_x$, and set

$$D^2 := \mathbb{E}_\nu D_x^2.$$

Let $\sigma_\infty(x) := \dfrac{1}{2} \operatorname{diam} Supp\, m_x$ and $\sigma_\infty := \sup \sigma_\infty(x)$. Also, let the averaging operator M act on the space of L^2 functions of (X, d, ν) as follows:

$$Mf(x) := \int_y f(y)dm_x(y).$$

A very important property of the averaging operator is the following: for any $x \in X$,

$$\lim_{k\to\infty} M^k f(x) = \mathbb{E}_\nu(f).$$

Indeed,

$$\lim_{k\to\infty} M^k f(x) = \int_z \int_y f(y)dm_z(y)d\nu(z)$$
$$= \int_y f(y)d\nu(y)$$

by the invariance of ν.

We establish two lemmas before we state and prove the main theorem on concentration under assumptions on coarse Ricci curvature.

Lemma 3.1 (Lipschitz Contraction)**.** *[10] If the coarse Ricci curvature of X is at least κ_0, then for every k-Lipschitz function $f : X \to \mathbb{R}$, the function Mf is $k(1-\kappa_0)$-Lipschitz.*

Proof. Suppose that the coarse Ricci curvature is at least κ_0. Given a k-Lipschitz function f, we have

$$Mf(x) - Mf(y) = \int_z f(z)dm_x - \int_z f(z)dm_y$$
$$\leq k \cdot \sup_{g\ 1-\text{Lipschitz}} \int_z gd(m_x - m_y)$$
$$= kd(x,y)(1-\kappa_0(x,y)).$$

□

Lemma 3.2. *Let ϕ be an α-Lipschitz function with $\alpha \leq 1$. Assuming $\lambda \leq \dfrac{1}{3\sigma_\infty}$, we have, for $\forall x \in X$,*

$$Me^{\lambda\phi}(x) \leq e^{\lambda M\phi(x)+\lambda^2\alpha^2\kappa_0 D_x^2}.$$

Proof. The Taylor expansion of $g(\phi(y)) = e^{\lambda\phi(y)}$ about the point $M\phi(x) = \mathbb{E}_{m_x}\phi$ gives us

$$e^{\lambda\phi(y)} = e^{\lambda M\phi(x)} + \lambda e^{\lambda M\phi(x)}(\phi(y) - M\phi(x)) + \frac{\lambda^2}{2}e^{\lambda M\phi(x)}(\phi(y) - M\phi(x))^2 + \mathcal{O}(\phi(y)^3).$$

By integrating over m_x, and applying a Lagrange remainder, we obtain

$$Me^{\lambda\phi}(x) \le e^{\lambda M\phi(x)} + \frac{\lambda^2}{2}\left(\sup_{Supp\ m_x} e^{\lambda\phi}\right) Var_{m_x}\phi.$$

Since diam $Supp\, m_x \le 2\sigma_\infty$ and ϕ is α - Lipschitz ($\alpha \le 1$), thus $\sup_{Supp\ m_x} \phi \le \mathbb{E}_{m_x}\phi + 2\sigma_\infty$, we get

$$(Me^{\lambda\phi})(x) \le e^{\lambda M\phi(x)} + \frac{\lambda^2}{2} e^{\lambda M\phi(x) + 2\lambda\sigma_\infty} Var_{m_x}\phi.$$

By definition, $Var_{m_x}\phi \le \alpha^2\kappa_0 D_x^2$, and for $\lambda \le \dfrac{1}{3\sigma_\infty}$ we have $e^{2\lambda\sigma_\infty} \le 2$, so

$$Me^{\lambda\phi}(x) \le e^{\lambda M\phi(x)}(1 + \lambda^2\alpha^2\kappa_0 D_x^2) \le e^{\lambda M\phi(x) + \lambda^2\alpha^2\kappa_0 D_x^2}.$$

□

With these two lemmas in hand, we are ready to state and prove the concentration theorem.

Theorem 3.2. *[10] Suppose that the coarse Ricci curvature of X is at least $\kappa_0 > 0$ and that the function $x \mapsto D_x^2$ is C-Lipschitz. Set*

$$t_{max} := \min\left(\frac{8D^2}{9\sigma_\infty}, \frac{4D^2}{3C}\right).$$

Then for any 1-Lipschitz function f, and any $t \le t_{max}$, we have

$$\nu(\{x, |f(x) - \mathbb{E}_\nu f| \ge t\}) \le 2e^{-\frac{t^2}{6D^2}}.$$

Proof. Let f be a 1-Lipschitz function, and let $\lambda > 0$. Define by induction the sequence of functions $\{f_k\}_k$: $f_0 := f$, and $f_{k+1} := Mf_k(x) + \lambda\kappa_0 D_x^2(1 - \kappa_0/2)^{2k}$. Suppose that $\lambda \le 1/(2C)$. Then $\lambda\kappa_0 D_x^2$ is $\kappa_0/2$-Lipschitz.

Claim 3.2. *f_k is $(1 - \kappa_0/2)^k$-Lipschitz.*

Proof of claim. The function $f_0 = f$ is clearly 1-Lipschitz. Now, assume the property true for k and consider $f_{k+1} = Mf_k(x) + \lambda\kappa_0 D_x^2(1 - \kappa_0/2)^{2k}$. By hypothesis, and the Lipschitz contraction Lemma, $Mf_k(x)$ is $(1-\kappa_0/2)^k(1-\kappa_0)$-Lipschitz, and $\lambda\kappa_0 D_x^2(1-\kappa_0/2)^{2k}$ is $\kappa_0/2(1-\kappa_0/2)^{2k}$-Lipschitz, so that $f_{k+1}(x)$ is Lipschitz of constant $\left((1-\kappa_0/2)^k(1-\kappa_0)+\kappa_0/2(1-\kappa_0/2)^{2k}\right) = (1-\kappa_0/2)^k\left(1-\kappa_0+\kappa_0/2(1-\kappa_0/2)^k\right)$ $\leq (1-\kappa_0/2)^k(1-\kappa_0/2)$.

□

As a consequence of Lemma 3.2, and the fact that f_k is $(1-\kappa_0/2)^k$-Lipschitz, we have that

$$Me^{\lambda f_k}(x) \leq e^{\lambda Mf_k(x)+\lambda^2\kappa_0 D_x^2(1-\kappa_0/2)^{2k}} = e^{\lambda f_{k+1}(x)},$$

so that, by recursion, for all k and for all x,

$$(M^k e^{\lambda f})(x) \leq e^{\lambda f_k(x)}. \tag{3.2}$$

Call $g(x) := \kappa_0 D_x^2$. By definition, if we were to break down f_k, we would see that

$$f_k(x) = (M^k f(x)) + \lambda\sum_{i=1}^{k}(M^{k-i}g)(x)(1-\kappa_0/2)^{2(i-1)}.$$

As $k\to\infty$, $M^k f(x)\to\mathbb{E}_\nu f$. Since $\kappa_0\in(0,1)$, $\lim_{i\to\infty}(1-\kappa_0/2)^{2i}=0$, so $\lim_{k\to\infty}\sum_{i=1}^{k}(1-\kappa_0/2)^{2(i-1)} = \dfrac{1}{1-(1-\kappa_0/2)^2} \leq \dfrac{4}{3\kappa_0}$. Furthermore,

$$\sum_{i=1}^{k}(M^{k-i}g)(x)(1-\kappa_0/2)^{2(i-1)}$$

$$= \sum_{i=1}^{k/2}(M^{k-i}g)(x)(1-\kappa_0/2)^{2(i-1)} + \sum_{i=k/2+1}^{k}(M^{k-i}g)(x)(1-\kappa_0/2)^{2(i-1)}.$$

The sequence $M^j g(x)$ is bounded above by some number $\mathcal{M}$, and so

$$\sum_{i=k/2+1}^{k}(M^{k-i}g)(x)(1-\kappa_0/2)^{2(i-1)} \leq \sum_{i=k/2+1}^{k}\mathcal{M}(1-\kappa_0/2)^{2(i-1)}$$

$$= \mathcal{M}(1-\kappa_0/2)^k\frac{1-(1-\kappa_0/2)^{2(k-1)}}{1-(\kappa_0/2)^2},$$

whose limit as $k\to\infty$ is zero. Also,

$$\sum_{i=1}^{k/2}(M^{k-i}g)(x)(1-\kappa_0/2)^{2(i-1)} \to \sum_{i=1}^{\infty}\mathbb{E}_\nu g(1-\kappa_0/2)^{2(i-1)}.$$

Therefore

$$\lim_{k\to\infty} f_k(x) \le \mathbb{E}_\nu f + \lambda \mathbb{E}_\nu g \frac{4}{3\kappa_0}.$$

Clearly, $M^k(e^{\lambda f})$ tends to $\mathbb{E}_\nu e^{\lambda f}$. Therefore, by (3.2),

$$\mathbb{E}_\nu e^{\lambda f} \le \lim_{k\to\infty} e^{\lambda f_k} \le e^{\lambda \mathbb{E}_\nu f + \frac{4\lambda^2}{3\kappa_0}\kappa_0 \mathbb{E}_\nu D_x^2}.$$

Let us now end the argument using Chebyshev's inequality:

$$\begin{aligned} &\nu(\{x, f(x) \ge t + \mathbb{E}_\nu f\}) \\ &= \nu(\{x, e^{\lambda(f(x)-\mathbb{E}\nu(f))} \ge e^{\lambda t}\}) \\ &\le \frac{1}{e^{\lambda t}} \mathbb{E}_\nu e^{\lambda(f-\mathbb{E}\nu(f))} \\ &\le \frac{1}{e^{\lambda t}} e^{\lambda \mathbb{E}_\nu (f-\mathbb{E}(f)) + \frac{4\lambda^2}{3\kappa_0}\kappa_0 \mathbb{E}\nu D_x^2} \\ &= e^{\frac{4\lambda^2}{3}D^2 - \lambda t}. \end{aligned}$$

The minimum of this function with respect to λ occurs at $\lambda = \frac{3t}{8D^2}$, so

$$\nu(\{x, f(x) \ge t + \mathbb{E}_\nu f\}) \le e^{\frac{-3t^2}{16D^2}}$$

$$\le e^{\frac{-t^2}{6D^2}}.$$

Applying the same argument to $-f$ gives us the desired result. Note that the restrictions $\lambda \le 1/2C$ and $\lambda \le 1/3\sigma_\infty$, under which the reasoning was carried on, give us the value of $t_{\max}$. □

3.5 Other Examples of Random Walks on the Hamming Cube

3.5.1 The Lazy Random Walk

Let (X, d, ν) again be the discrete cube $\{0,1\}^N$ with the rescaled Hamming metric d and ν the invariant probability measure. Let the measures m_x be defined as follows:

$m_x(x) = \dfrac{1}{2}$, and for each point y which is a neighbor of x (i.e. $\{y,\ d(x,y) = \dfrac{1}{N}\}$), $m_x(y) = \dfrac{1}{2N}$. Let x and y be neighboring points. Let $x_1 = y, x_2, \ldots, x_N$ be neighbors of x and $y_1 = x, y_2, \ldots y_N$ be neighbors of y. Then

$$W_1(m_x, m_y) = \sup_{f\text{1-Lipschitz}} \left(\int f dm_x - \int f dm_y \right)$$

$$= \sup_{f\text{1-Lipschitz}} \left(\frac{1}{2} f(x) + \frac{1}{2N} f(y) - \frac{1}{2N} f(x) - \frac{1}{2} f(y) + \frac{1}{2N} \sum_2^N (f(x_i) - f(y_i)) \right)$$

$$= \sup_{f\text{1-Lipschitz}} \frac{N-1}{2N}(f(x) - f(y)) + \frac{1}{2N} \sum_2^N (f(x_i) - f(y_i)).$$

Assume that the x_i and y_i are enumerated such that they are neighbors (ie. $d(x_i, y_i) = \frac{1}{N}$ for all i). Then the supremum is attained when $f(x) - f(y) = \frac{1}{N}$ and $(f(x_i) - f(y_i)) = \frac{1}{N}$, so therefore

$$W_1(m_x, m_y) = \frac{N-1}{2N^2} + \frac{N-1}{2N^2} = \frac{N-1}{N^2},$$

so solving for κ in $W_1(m_x, m_y) = (1-\kappa) d(x,y)$ gives $\kappa = \dfrac{1}{N}$ for any two adjacent points.

3.5.2 The Biased Cube

Consider $(\{0,1\}^N, d, \nu)$ as before. We wish this time to consider a random walk $\{m_x\}$ with the following properties: the probability of transporting to a point closer to $\overline{0} := (0, 0, \ldots, 0)$ is p, the probability of transporting to a point closer to $\overline{1} := (1, 1, \ldots, 1)$ is q, and that of staying put $(1-p-q)$. The constraints in this case are that $p, q \in (0,1)$, $p + q < 1$. It can therefore be thought of as a random walk with a biased tendency to move either toward or away from the "origin", depending on the values of p and q.

The neighbors of an arbitrary point $x \in \{0,1\}^N$ generally consist of two types: those which are closer to $\overline{0}$ and those closer to $\overline{1}$. Let i denote the number of zeros in x. We will observe the transportation distance from x to a neighbor y which is closer to $\overline{1}$, which means we must consider three separate cases: i is between 2 and $N-1$, $i = N$ (so that $x = \overline{0}$), and $i = 1$ (so that $y = \overline{1}$). This will cover all cases; indeed, since $W_1(\nu_1, \nu_2) = W_1(\nu_2, \nu_1)$, we need only to transport in one direction, while getting the other distance for free.

Case 1: i between 2 and N-1

Now we must define the random walk $\{m_x\}$ for all points $x \in \{0,1\}^N$, given that i does not equal 0, 1 or N. Then given $x = (a_1, \ldots, a_N)$, where $a_j \in \{0, 1\}$, the neighbors of x consist of "switching" one of the a_j's with the opposite value, so that one of the i 0's are replaced with a 1, or one of the $N-i$ 1's replaced with a 0. Therefore x consists of i neighbors which are closer to 1 and $N-i$ neighbors which are closer to 0. Let $x_1, \ldots x_{N-i}$ denote the points which are closer to 0 and $x'_1, \ldots x'_i$ the points closer to 1. Then m_x is defined as follows: $m_x(x) = (1-p-q)$, $m_x(x_k) = \frac{p}{(N-i)}$ for all k between 1 and $(N-i)$, and $m_x(x'_j) = \frac{q}{i}$ for all j between 1 and i. This will ensure the desired probabilities.

Consider a neighbor y obtained by replacing one of the zeros in x with a 1, ordered so that $y = x'_i$. y therefore has $(N-i+1)$ closer to 0, which shall be named $y_0 = x, y_1, y_2, \ldots y_{N-i}$, and $(i-1)$ neighbors closer to 1, which we shall denote $y'_1, y'_2, \ldots y'_{i-1}$. Under the proper ordering, we can assure that for all k between 1 and $(N-i)$, x_k and y_k are in fact neighbors with one another, as well as x'_j and y'_j for all j between 1 and $(i-1)$. m_y is defined as follows: $m_y(y) = (1-p-q)$, $m_y(x) = m_y(y_k) = \frac{p}{(N-i+1)}$ for all k from 1 to $(N-i)$, and $m_y(y'_j) = \frac{q}{(i-1)}$ for all j from 1 to $(i-1)$.

Given a function $f : \{0,1\}^N \to \mathbb{R}$ and the above measures m_x and m_y, it is easy to see that

$$\int f dm_x = (1-p-q)f(x) + \frac{qf(y)}{i} + \sum_{k=1}^{N-i} \frac{pf(x_k)}{(N-i)} + \sum_{j=1}^{i-1} \frac{qf(x'_j)}{i}$$

and

$$\int f dm_y = (1-p-q)f(y) + \frac{pf(x)}{(N-i+1)} + \sum_{k=1}^{N-i} \frac{pf(y_k)}{(N-i+1)} + \sum_{j=1}^{i-1} \frac{qf(y'_j)}{(i-1)}.$$

Our goal is to subtract these two integrals, and consider the supremum of this value for all 1-Lipschitz functions. After some rearranging, note that

$$\int f dm_x - \int f dm_y$$

$$= (1-p-q)(f(x)-f(y)) + \frac{qf(y)}{i} - \frac{pf(x)}{(N-i+1)}$$

$$+\sum_{k=1}^{N-i}\left(\frac{pf(x_k)}{(N-i)} - \frac{pf(y_k)}{(N-i+1)}\right) + \sum_{j=1}^{i-1}\left(\frac{qf(x'_j)}{i} - \frac{qf(y'_j)}{(i-1)}\right).$$

Let us work for a moment with these sums.

$$\sum_{k=1}^{N-i}\left(\frac{pf(x_k)}{(N-i)} - \frac{pf(y_k)}{(N-i+1)}\right)$$

$$= \sum_{k=1}^{N-i} \frac{p(f(x_k)-f(y_k))}{(N-i+1)} + \sum_{k=1}^{N-i} \frac{pf(x_k)}{(N-i)(N-i+1)},$$

while

$$\sum_{j=1}^{i-1}\left(\frac{qf(x'_j)}{i} - \frac{qf(y'_j)}{(i-1)}\right)$$

$$= \sum_{j=1}^{i-1} \frac{q(f(x'_j)-f(y'_j))}{i} - \sum_{j=1}^{i-1} \frac{qf(y'_j)}{i(i-1)}.$$

Note that

$$\left(\sum_{k=1}^{N-i} \frac{pf(x_k)}{(N-i)(N-i+1)}\right) - \frac{pf(x)}{(N-i+1)} = \sum_{k=1}^{N-i} \frac{p(f(x_k)-f(x))}{(N-i)(N-i+1)},$$

simply by bringing $\frac{pf(x)}{(N-i+1)}$ into the sum. Similarly,

$$\frac{qf(y)}{i} - \sum_{j=1}^{i-1} \frac{qf(y'_j)}{i(i-1)} = \sum_{j=1}^{i-1} \frac{q(f(y)-f(y'_j))}{i(i-1)}.$$

Combining everything,

$$\int f dm_x - \int f dm_y$$

$$= (1-p-q)(f(x)-f(y)) + \sum_{k=1}^{N-i} \frac{p(f(x_k)-f(y_k))}{(N-i+1)} + \sum_{j=1}^{i-1} \frac{q(f(x'_j)-f(y'_j))}{i}$$

$$+ \sum_{k=1}^{N-i} \frac{p(f(x_k)-f(x))}{(N-i)(N-i+1)} + \sum_{j=1}^{i-1} \frac{q(f(y)-f(y'_j))}{i(i-1)}.$$

Note that given the proper arrangement, *all* function differences in the above equation are of points which are neighbors, telling us that the supremum of these differences are all identically $\frac{1}{N}$ as we are dealing with 1-Lipschitz functions. Therefore

$$W_1(m_x, m_y) = \frac{(1-p-q)}{N} + \frac{p(N-i)}{N(N-i+1)} + \frac{q(i-1)}{Ni} + \frac{p}{N(N-i+1)} + \frac{q}{Ni}$$

$$= \frac{1}{N}.$$

Solving for $\kappa(x,y)$ in the equation $W_1(m_x, m_y) = (1-\kappa(x,y))d(x,y)$ reveals $\kappa = 0$ among any two such neighbors.

Case 2: i = 1

The case where x has only 1 zero ensures that the only neighbor "closer" to $\overline{1}$ is the point $\overline{1}$ itself, while x has $(N-1)$ neighbors closer to $\overline{0}$, which I will again denote $x_1, \ldots x_{N-1}$. Therefore m_x is defined as follows: $m_x(x) = (1-p-q)$, $m_x(\overline{1}) = q$, $m_x(x_k) = \frac{p}{(N-1)}$ for all k between 1 and $(N-1)$. Meanwhile, let $x, y_1, \ldots, y_{N-1}$ denote the neighbors of $\overline{1}$, all of which are closer to $\overline{0}$; assume also that these neighbors are arranged so that x_k and y_k are neighbors for all k. In keeping with the text, the likelihood of staying put and traveling toward $\overline{1}$ should be combined in this case, so that $m_{\overline{1}}(\overline{1}) = (1-p-q)+q = (1-p)$, and $m_{\overline{1}}(x) = m_{\overline{1}}(y_k) = \frac{p}{N}$ for all k between 1 and $(N-1)$.

Let us now integrate with respect to f.

$$\int f dm_x = (1-p-q)f(x) + qf(\overline{1}) + \sum_{k=1}^{N-1} \frac{pf(x_i)}{(N-1)}$$

and

$$\int f dm_{\overline{1}} = \frac{pf(x)}{N} + (1-p)f(\overline{1}) + \sum_{k=1}^{N-1} \frac{pf(y_k)}{N}.$$

Therefore

$$\int dm_x - \int dm_{\overline{1}}$$

$$= (1-p-q)(f(x) - f(\overline{1})) - \frac{pf(x)}{N} + \sum_{k=1}^{N-1} \left(\frac{pf(x_k)}{(N-1)} - \frac{pf(y_k)}{N} \right).$$

By similar manipulation of summations as seen in Case 1, this becomes

$$(1-p-q)(f(x) - f(\overline{1})) + \sum_{k=1}^{N-1} \frac{p(f(x_k) - f(y_k))}{N} + \sum_{k=1}^{N-1} \frac{p(f(x_k) - f(x))}{N(N-1)}.$$

As all of these points are neighbors, the maximum of all 1-Lipschitz functions tells us that each of these function differences obtain maximum value $\frac{1}{N}$, so that

$$W_1(m_x, m_{\overline{1}}) = \frac{(1-p-q)}{N} + \frac{p(N-1)}{N^2} + \frac{p}{N^2} = \frac{1-q}{N},$$

which implies that $\kappa(x, \overline{1}) = q$.

Case 3: $\mathbf{i = N}$

The final case we shall consider is when we are moving from $\overline{0}$ to a neighbor y, which by default is closer to $\overline{1}$. Again, the likelihood of staying put and traveling toward $\overline{0}$ must be combined, so that $m_{\overline{0}}(\overline{0}) = (1-p-q) + p = (1-q)$, while for each of its neighbors, which I will denote $y, x'_1, \ldots, x'_{N-1}$, $m_{\overline{0}}(y) = m_{\overline{0}}(x'_j) = \frac{q}{N}$. Meanwhile, $m_y(\overline{0}) = p, m_y(y) = (1-p-q)$, and for the $N-1$ neighbors closer to $\overline{1}$, which I will denote $y'_1, \ldots, y'_{N-1}$, $m_y(y'_j) = \frac{q}{(N-1)}$. I also assume the arrangement such that x'_j and y'_j are neighboring points for all j between 1 and $(N-1)$.

Therefore given f,

$$\int f dm_{\overline{0}} = (1-q)f(\overline{0}) + \frac{qf(y)}{N} + \sum_{j=1}^{N-1} \frac{qf(x'_j)}{N}$$

and

$$\int f dm_y = pf(\overline{0}) + (1-p-q)f(y) + \sum_{j=1}^{N-1} \frac{qf(y'_j)}{(N-1)}.$$

subtracting these two and applying similar manipulations as in the previous cases, we see that

$$\int f dm_{\overline{0}} - \int f dm_y = (1-p-q)(f(\overline{0}) - f(y)) + \sum_{j=1}^{N-1} \frac{q(f(x'_j) - f(y'_j))}{N}$$

$$+ \sum_{j=1}^{N-1} \frac{q(f(y) - f(y'_j))}{N(N-1)}.$$

Once again, taking into account that each of these points are neighbors, the maximum difference for f being 1-Lipschitz, is $\frac{1}{N}$, so

$$W_1(m_{\overline{0}}, m_y) = \frac{(1-p-q)}{N} + \frac{q(N-1)}{N^2} + \frac{q}{N^2} = \frac{(1-p)}{N}.$$

Therefore $\kappa(\overline{0}, y) = p$.

We have computed an example for which the coarse Ricci curvature is almost everywhere 0, except for the "extreme" points. This example serves as a note that while the geometry of the space did not change, the chosen measurement served to "flatten" the normally binomial distribution of points on $\{0,1\}^N$.

Appendix A

A Proof of the Concentration on the Cube by Martingale Methods

As martingales are frequently used in discrete measure spaces as a tool to deal with concentration inequalities, we will include here a proof of the concentration inequality on the discrete hypercube (Theorem 1.2) using this method. We will recover the result obtained using product measure spaces and we will see that, in this instance, the two methods are nearly identical.

Definition A.1. *Let $(X, \mathcal{M}, \nu)$ be a probability space and let $\mathcal{M}_\alpha$ be a σ sub-algebra with respect to $\mathcal{M}$. Then given a function $f : X \to \mathbb{R}$ which is integrable on $(X, \mathcal{M})$ with respect to ν, the* conditional expectation *of f with respect to $\mathcal{M}_\alpha$, denoted $\mathbb{E}(f|\mathcal{M}_\alpha)$, is the unique (in the probability sense) function h, integrable on $(X, \mathcal{M}_\alpha)$ with respect to $\nu|_{\mathcal{M}_\alpha}$, such that*

$$\int_A h d\nu = \int_A f d\nu \text{ for all } A \in \mathcal{M}_\alpha.$$

Definition A.2. *Given a probability space $(X, \mathcal{M}, \nu)$, let $\mathcal{M}_0$, $\mathcal{M}_1$, $\dots \subset \mathcal{M}$ be a sequence, possibly infinite, of σ sub-algebras. A sequence of integrable functions f_0, f_1, $\dots$ is said to be a* martingale *with respect to this sequence of σ sub-algebras if $f_i = \mathbb{E}(f_j|\mathcal{M}_i)$, $\forall i \leq j$.*

Let $(\{0,1\}^N, \mathcal{M}, \nu, d)$ be the metric space of the N dimensional discrete cube with ν the uniform probability measure (so that $\nu(x) = \dfrac{1}{2^N}$ for all singletons x), $\mathcal{M}$ the σ-algebra consisting of all singletons of $\{0,1\}^N$, and d the rescaled Hamming metric $d(x,y) = \dfrac{1}{N}\sum_1^N |x_i - y_i|$.

Consider the following partition of this space:

$X_0 = \big\{\{0,1\}^N\big\}$

$X_1 = \big\{\{0,1\}^{N-1} \times \{0\}, \{0,1\}^{N-1} \times \{1\}\big\}$

$X_2 = \{\{0,1\}^{N-2} \times \{0\} \times \{0\}, \{0,1\}^{N-2} \times \{0\} \times \{1\}, \{0,1\}^{N-2} \times \{1\} \times \{0\},$

$\{0,1\}^{N-2} \times \{1\} \times \{1\}\},$

$\vdots$

$X_N = \{\{x\}_{x \in \{0,1\}^N}\},$

and denote by $\mathcal{M}_i$ the σ-algebra generated by X_i. Note that the number of elements of each partition X_i is 2^i. Let us denote by A^i_j, $0 \leq i \leq N$, $0 \leq j \leq 2^i$ the elements of the set X_i. For each partition i, the elements A^i_j split into two elements in the next, $(i+1)$, level (for example, $\{0,1\}^N$ splits into $\{0,1\}^{N-1} \times \{0\}$ and $\{0,1\}^{N-1} \times \{1\}$). If A^{i+1}_j, $A^{i+1}_k \subset A^i_l$, then there is a "distance minimizing" correspondence between these two sets, in the sense that there exists a bijective function $\phi : A^{i+1}_j \to A^{i+1}_k$ such that for any $x \in A^{i+1}_j$, $d(x, \phi(x)) \leq 1/N$. Indeed, the minimal distance between any two elements from A^{i+1}_j and A^{i+1}_k, respectively, is precisely $1/N$.

Let us now turn our attention to the σ-algebras generated by these partitions. Consider an integrable 1-Lipschitz real valued function f on $(\{0,1\}^N, \mathcal{M}, \nu, d)$. Denote by f_i the conditional expectation $\mathbb{E}(f|\mathcal{M}_i)$ of f with respect to $\mathcal{M}_i$.

It is therefore obvious that f_0 is precisely the constant function $\mathbb{E}(f)$, the usual expectation (mean) of f with respect to ν. Some other properties of these functions are:

- $\mathbb{E}(\mathbb{E}(f|\mathcal{M}_i)|\mathcal{M}_j) = \mathbb{E}(f|\mathcal{M}_j)$, whenever $j \leq i$ (which is in fact what makes this sequence of functions a martingale).

- If $g \in L_\infty(\{0,1\}^N, \mathcal{M}_i, \nu|_{\mathcal{M}_i})$, then $\mathbb{E}(fg|\mathcal{M}_i) = g\mathbb{E}(f|\mathcal{M}_i)$; in particular, $\mathbb{E}(af|\mathcal{M}_i) = a\mathbb{E}(f|\mathcal{M}_i)$ for any $a \in \mathbb{R}$.

Indeed, both of these properties can be seen directly from equation (1.1). Set $d_i := f_i - f_{i-1}$. We wish to find an upper bound for these functions (ie. $||d_i||_\infty$). Note by the above properties that $\mathbb{E}(d_i|\mathcal{M}_{i-1}) = 0$.

Proposition A.1. *With the above notations, we have that $||d_i||_\infty \leq 1/N$, for all i such that $1 \leq i \leq N$.*

Proof. Fix i between 1 and N, and consider $d_i(x) = f_i(x) - f_{i-1}(x)$ for some $x \in \{0,1\}^N$. As X_{i-1} is a partition of $\{0,1\}^N$, then there exists an A_l^{i-1} which contains x, for some l, $1 \leq l \leq 2^{i-1}$.

Let $B = A_j^i$ and $C = A_k^i$ be the two elements of X_i contained in A_l^{i-1}. Then f_i is constant on B and C, and

$$f_{i|C} = (\mathrm{Card(C)})^{-1} \sum_{\alpha \in C} f(\alpha) = (\mathrm{Card(B)})^{-1} \sum_{\beta \in B} f(\phi(\beta)). \tag{1.1}$$

Consequently, $|f_{i|C} - f_{i|B}| = |(\mathrm{Card(B)})^{-1} \sum_{\beta \in B} (f(\phi(\beta)) - f(\beta))| \leq d(\phi(x), x)$ for any $x \in B$, thus $|f_{i|C} - f_{i|B}| \leq 1/N$.

On the other hand, for $A := A_l^{i-1} = B \cup C$, we have $f_{i-1|A} = \frac{1}{2}(f_{i|B} + f_{i|C})$ which implies that $|f_{i|B} - f_{i-1|A}| \leq 1/N$. As x was arbitrary, we have that $||d_i||_\infty \leq 1/N$, $1 \leq i \leq N$.

□

The following lemma will then complete the proof of Theorem 1.2:

Lemma A.1. *Let $(\Omega, \mathcal{M}, \nu)$ be a probability space and let f be a function on Ω integrable with respect to ν. For every $r \geq 0$,*

$$\nu(\{|f - \mathbb{E}(f)| \geq r\}) \leq 2e^{-r^2/2D^2} \tag{1.2}$$

where $D^2 = \sum_{i=1}^N ||d_i||_\infty^2$.

Proof. As $\exp(x)$ is a convex function, we have, for any $\lambda \in \mathbb{R}$ and any $u \in [-1, 1]$,

$$e^{\lambda u} = e^{\frac{1+u}{2}\lambda + \frac{1-u}{2}(-\lambda)} \leq \frac{1+u}{2}e^{\lambda} + \frac{1-u}{2}e^{-\lambda}.$$

This is enough to imply the following inequality:

$$\mathbb{E}(e^{\lambda d_i}|\mathcal{M}_i) \leq \cosh(\lambda ||d_i||_\infty) \leq e^{\lambda^2 ||d_i||_\infty{}^2/2}. \tag{1.3}$$

Indeed,

$$\begin{aligned}
\mathbb{E}(e^{\lambda d_i}|\mathcal{M}_i) &= \mathbb{E}(e^{\lambda ||d_i||_\infty \frac{d_i}{||d_i||_\infty}}|\mathcal{M}_i) \\
&\leq \mathbb{E}(\frac{1 + \frac{d_i}{||d_i||_\infty}}{2}e^{\lambda ||d_i||_\infty} + \frac{1 - \frac{d_i}{||d_i||_\infty}}{2}e^{-\lambda ||d_i||_\infty}|\mathcal{M}_i)) \\
&= \mathbb{E}(\frac{1 + \frac{d_i}{||d_i||_\infty}}{2}e^{\lambda ||d_i||_\infty}|\mathcal{M}_i) + \mathbb{E}(\frac{1 - \frac{d_i}{||d_i||_\infty}}{2}e^{-\lambda ||d_i||_\infty}|\mathcal{M}_i) \\
&= \frac{e^{\lambda ||d_i||_\infty}}{2}\mathbb{E}(1 + \frac{d_i}{||d_i||_\infty}|\mathcal{M}_i) + \frac{e^{-\lambda ||d_i||_\infty}}{2}\mathbb{E}(1 - \frac{d_i}{||d_i||_\infty}|\mathcal{M}_i) \\
&= \frac{e^{\lambda ||d_i||_\infty}}{2}(1 + \frac{1}{||d_i||_\infty}\mathbb{E}(d_i|\mathcal{M}_i)) + \frac{e^{-\lambda ||d_i||_\infty}}{2}(1 - \frac{1}{||d_i||_\infty}\mathbb{E}(d_i|\mathcal{M}_i)) \\
&= \frac{e^{\lambda ||d_i||_\infty} + e^{-\lambda ||d_i||_\infty}}{2} = \cosh(\lambda ||d_i||_\infty),
\end{aligned}$$

using the fact that $\mathbb{E}(d_i|\mathcal{M}_i) = 0$. Using Taylor series, it is easy to see that $\cosh(x) \leq e^{x^2}$, hence we have further that

$$\mathbb{E}(e^{\lambda d_i}|\mathcal{M}_i) \leq e^{\lambda^2 ||d_i||_\infty^2/2}. \tag{1.4}$$

Let us now analyze $\mathbb{E}(e^{\lambda \sum_1^N d_i})$. The first thing to note is that $\sum_1^N d_i = f - \mathbb{E}(f)$ and thus $\mathbb{E}(e^{\lambda \sum_1^N d_i}) = \mathbb{E}(\mathbb{E}(e^{\lambda \sum_1^N d_i})|\mathcal{M}_N)$. Indeed, $\mathbb{E}(f) = \mathbb{E}(\mathbb{E}(f|\mathcal{M}_N))$, since f is integrable over the largest σ-algebra. So, $\mathbb{E}(e^{\lambda\sum_1^N d_i}) = \mathbb{E}(\mathbb{E}(e^{\lambda\sum_1^{N-1} d_i} e^{\lambda d_N})|\mathcal{M}_N) = \mathbb{E}(e^{\lambda d_N}\mathbb{E}(e^{\lambda\sum_1^{N-1} d_i})|\mathcal{M}_N)$ Therefore, we apply (1.3) to obtain

$$\mathbb{E}(e^{\lambda\sum_1^N d_i}) \leq \mathbb{E}(\mathbb{E}(e^{\lambda\sum_1^{N-1} d_i}|\mathcal{M}_{N-1})e^{\lambda^2||d_N||_\infty^2/2}.$$

Recall that $e^{\lambda d_i} \in L_\infty(\{0,1\}^N, \mathcal{M}_i, \nu|_{\mathcal{M}_i})$ and apply (1.3) successively to $\mathbb{E}(\cdot|\mathcal{M}_{N-1})$, $\mathbb{E}(\cdot|\mathcal{M}_{N-2})$ and so on. Since $\sum_1^N d_i = f - \mathbb{E}(f)$, we will see that

$$\mathbb{E}(e^{\lambda(f-\mathbb{E}(f))}) \leq e^{\lambda^2\sum_{i=1}^N ||d_i||_\infty^2/2}. \tag{1.5}$$

By applying Chebyshev's inequality, we obtain that, for any $r \geq 0$,

$$\begin{aligned} \nu(f - \mathbb{E}(f) \geq r) &= \nu(e^{\lambda(f-\mathbb{E}(f))} \geq e^{\lambda r}) \\ &\leq \frac{1}{e^{\lambda r}}\mathbb{E}(e^{(f-\mathbb{E}(f))}) \leq e^{-\lambda r+\lambda^2 D^2/2}, \end{aligned} \tag{1.6}$$

where $D^2 = \sum_{i=1}^N ||d_i||_\infty^2$. Applying the previous inequality to $-f$, we see that

$$\nu(|f - \mathbb{E}(f)| \geq r) \leq 2e^{-\lambda r+\lambda^2 D^2/2},$$

for all $\lambda \geq 0$.

Note that the function $\lambda \mapsto -\lambda r + \lambda^2 D^2/2$, defined for $\lambda \geq 0$, has a minimum value of $-r^2/2D^2$ which occurs at $\lambda = r/D^2$. Therefore, we conclude that:

$$\nu(|f - \mathbb{E}(f)| \geq r) \leq 2e^{-r^2/2D^2}, \quad \forall r \geq 0. \tag{1.7}$$

□

Appendix B

Index Forms of Vector Fields With Respect to a Hypersurface

Consider an $(N+1)$-dimensional Riemannian manifold (M, g) and a hypersurface H embedded in M with an oriented unit normal vector field ν. Choose a point $p \in H$ and a geodesic $\gamma : [0, r] \to M$ such that $\gamma(0) = p$ and $\gamma'(0) = \nu_p$. Consider a variation through geodesics $\Gamma : (-\epsilon, \epsilon) \times [0, r] \to M$ such that $\Gamma(s,t)_{|_{s=0}} = \gamma(t)$, and $Y(t) = \frac{\partial \Gamma}{\partial s}(0, t)$ is an H-Jacobi field, as defined earlier. Thus Y is a vector field along γ, orthogonal to it and we will call the index form $I_r(Y, Y)$ of Y

$$I_r(Y, Y) := \langle \nabla_Y Y, \gamma' \rangle|_0^r + \int_0^r (|Y'|^2 - \langle Y, R(Y, \gamma')\gamma' \rangle) dt,$$

where Y' denotes the derivative of Y with respect to $\dfrac{\partial}{\partial t}$. The variation is constructed so that $\Gamma(s, 0)$ stays on H and Y is an H-Jacobi field. So, we have that $\langle \nabla_Y Y, \gamma' \rangle(0) = \langle Y, S_{\gamma'(0)} Y \rangle(0)$. Moreover, in what follows, we will consider only variations for which $\langle \nabla_Y Y, \gamma' \rangle(s) = 0$ for $s > 0$. Thus, the index form of Y is entirely determined by the hypersurface H and the restrictions of the vector field Y to it:

$$I_r(Y, Y) = \langle Y, S_{\gamma'(0)} Y \rangle + \int_0^r (|Y'|^2 - \langle Y, R(Y, \gamma')\gamma' \rangle) dt. \tag{2.1}$$

Given a differentiable vector field Y along γ, we will call the index form of a vector field with respect to H at p the expression (2.1). As we noticed in the section on the Heintze-Karcher inequality, integration by parts yielded, for H-Jacobi fields, that $I_r(Y, Y) =< Y, Y' >_{|r}$.

The property that we use primarily in our thesis is that H-Jacobi fields minimize the quadratic form I_r among all vector fields along a geodesic. More precisely:

Theorem B.1. *[1] Suppose Y is an H-Jacobi field and that X is any differentiable vector field along γ such that $X(r) = Y(r)$ and $X(0) \in T_pN$. Then*

$$I_r(Y, Y) \leq I_r(X, X),$$

with equality if and only if $X = Y$.

The following lemma will prove useful in proving this theorem:

Lemma B.1. *If Y and X are any two H-Jacobi fields, then $\langle Y', X\rangle = \langle Y, X'\rangle$.*

Indeed, $\langle Y', X\rangle = \langle S_\nu Y, X\rangle = \langle Y, S_\nu X\rangle = \langle Y, X'\rangle$ by symmetry of the shape operator.

Proof of Theorem. Let $Y_1, \ldots, Y_N$ be a basis of the space of H-Jacobi fields along γ. We can then use these vectors to represent Y and X, so that

$$Y = y^i Y_i \text{ and } X = x^i Y_i,$$

using the summation notation. Because $\{Y_i\}$ forms a basis of the H-Jacobi fields, the functions y^i are constant for all i, whereas the functions x^i, for all i, generally are not. Computing the derivative of X we obtain $X' = A + B$, where $A = (x^i)'Y_i$ and $B = x^i Y_i'$.

With this in mind, we have

$$I(X, X) = \langle X, S_{\gamma'(0)}X\rangle + \int_0^r (\langle A, A\rangle + 2\langle A, B\rangle + \langle B, B\rangle - \langle X, R(X, \gamma')\gamma'\rangle)dt.$$

Since $\langle Y_i, Y_j'\rangle = \langle Y_i', Y_j\rangle$ from the lemma, it therefore follows that

$$\langle X, B\rangle' = (x^i)'x^j\langle Y_i, Y_j'\rangle + x^i x^j\langle Y_i', Y_j\rangle + x^i(x^j)'\langle Y_i, Y_j'\rangle + x^i x^j\langle Y_i, Y_j''\rangle$$

$$= 2\langle A, B\rangle + \langle B, B\rangle + x^i x^j\langle Y_i, Y_j''\rangle.$$

Replacing $2\langle A, B\rangle + \langle B, B\rangle$ with $\langle Y, B\rangle - x^i x^j \langle Y_i, Y_j''\rangle$ gives us:

$$I(X,X) = \langle X, S_{\gamma'(0)}X\rangle + \langle X, B\rangle|_0^r + \int_0^r \langle A, A\rangle dt + \int_0^r x^i x^j \langle Y_i, Y_j'' - R(Y_j, \gamma')\gamma'\rangle dt$$

$$= \langle X, S_{\gamma'(0)}X - B\rangle|_0 + \langle X, B\rangle|_r + \int_0^r \langle A, A\rangle dt,$$

where we used the fact that the Y_i's are H-Jacobi fields hence the previous integral term vanishes. Note also that $S_{\gamma'(0)}X - B = x^i(S_{\gamma'(0)}Y_i - Y_i') = 0$ and therefore $\langle X, S_{\gamma'(0)}X - B\rangle|_0 = 0$. Also, since $X(r) = Y(r)$, $x^i(r) = y^i(r)$, and so

$$\langle X, B\rangle|_r = \langle Y, x^i Y_i'\rangle = \langle Y, y^i Y_i'\rangle|_r = \langle Y, Y'\rangle|_r = I(Y,Y).$$

Therefore

$$I(X,X) = I(Y,Y) + \int_0^r \langle A, A\rangle dt,$$

so the inequality follows since $\langle A, A\rangle$ is positive definite. Moreover, we have equality between the two index forms if and only if $\langle A, A\rangle = 0$ or, equivalently, if $x^i(t) \equiv y^i(t)$ for all t.

$\square$

Bibliography

[1] Yu. D. Burago, V.A. Zalgaller, *Geometric Inequalities*, Springer-Verlag, 1988.

[2] Sylvestre Gallot, Dominique Hulin and Jaques Lafontaine, *Riemannian Geometry*, Springer-Verlag, 1987.

[3] Wolfgang Kühnel, *Differential Geometry: Curves-Surfaces-Manifolds,* Second Edition, American Mathematical Society, 2006.

[4] Michel Ledoux, *The Concentration of Measure Phenomenon*, American Mathematical Society, 2001.

[5] John M. Lee, *Riemannian Manifolds: An Introduction to Curvature*, Springer, 1997.

[6] Mikhail Gromov, *Paul Lévy's Isoperimetric Inequality*, Preprint IHÉS, 1979.

[7] Ernst Heintze, Hermann Karcher, *A General Comparison Theorem with Applications to Volume Estimates for Submanifolds*, Annales scientifiques de l'É.N.S. 4th série, tome 11, no. 4, pp. 451-470, 1978.

[8] Frank Pacard, *Constant Mean Curvature Hypersurfaces in Riemannian Manifolds*, Riv. Mat. Univ. Parma (7) 4, pp. 141-162, 2005.

[9] Yann Ollivier, *A Survey of Ricci Curvature for Metric Spaces and Markov Chains*, Probabilistic approach to geometry, Adv. Stud. Pure Math. 57, Math. Soc. Japan, 2010, pp. 343-381.

[10] Yann Ollivier, *Ricci Curvature of Markov Chains on Metric Spaces*, Journal of Functional Analysis 256, pp. 810-864, 2008.

[11] Gideon Schechtman, *Concentration, Results, and Applications*, Handbook of the geometry of Banach spaces, Vol. 2, pp. 1603-1634, North-Holland, Amsterdam, 2003.

[12] Michel Talagrand, *Concentration of Measure and Isoperimetric Inequalities in Product Spaces*, Publications mathématiques de I.H.E.S., pp. 73-205, 1995.