UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE MATEMÁTICA

A geometrical point of view for the noncommutative Ergodic Theorems

Luís Miguel Neves Pedro Machado Sampaio

Mestrado em Matemática

Dissertação orientada por:
Professor Doutor Pedro Miguel Nunes da Rosa Dias Duarte

2018

To not be in Excavation

Acknowledgements

I'd like to start by thanking my advisor Prof. Pedro Duarte for all the time, patience, lifetime teachings, the brilliant classes on Differential Topology and Ergodic Theory and introducing me to Dynamical Systems. I'm also very grateful for all the fruitful conversations on multiple topics such as Measure Theory, Harmonic Analysis, Symplectic Geometry, Hyperbolic Geometry, Riemannian Geometry and others, some of which went way beyond the scope of this thesis.

I also find important to acknowledge the Department of Mathematics at FCUL for the opportunity to be a *Monitor* for two years in which I was allowed my joy in teaching and sharing Mathematics. Specially thankful for being assigned with the teaching of practical classes in Geometry during this last semester. In that regard I am very beholden to Prof. Maria Manuel Torres for guiding and advicing, Prof. Maria Cecília Ferreira for sharing and showcasing a huge love for geometry and finally to my colleague Ricardo Shimura for swapping classes so I could teach this course.

What would a Master's without great colleagues? I'd like to thank them all feeling the need to highlight two; Diogo for all the hours discussing mathematics and other topics, solving problems and helping with Functional Analysis questions and Florival for the inspiration of seeking knowledge and trying to learn more each day.

A little word is due to the people from the study room, namely Rui, Constança, Beatriz and Fabrice, for keeping it a lively and healthy study place where most of this text was written. Also taking this time to demonstrate my gratefulness to my students for baring with my somewhat rigid ways.

On the journey through mathematics I'd like to finish by thanking the professors at FCUL, namely Professors Gracinda Cunha and Carlota Gonçalves for cherishing their students, great lessons and the first for being an advisor on my seminar, Professors Nuno da Costa Pereira, Armando Machado and Mário Figueira for the extraordinary inspiring classes that always left me in awe and awaiting for the next one.

On a personal level I dedicate a few words to the special people who helped me through dark times, some of which without ever realizing and others by doing more than they were ever expected to. Of these people I feel the need to mention Prof. Carlos Florentino whose amazing classes on Algebraic Topology were detrimental to my continuation of pursuing mathematics and Annie for providing some needed controlled escapism.

For the company and support throughout these three years I finish by acknowledging Batista, David, Ana, Joanas, Carolina and Caldeira.

Resumo

Um dos resultados fundamentais na teoria de probabilidades é a Lei Forte dos Grandes Números, cujo conteúdo afirma que quanto mais repetirmos um experiência mais a média observada se aproxima do valor médio esperado. Formalmente, sendo X_i uma sequência de variáveis aleatórias independentes e identicamente distribuídas cujo valor esperado é finito, tem-se

$$\frac{1}{n}(X_1 + X_2 + \cdots X_n) \xrightarrow[n\to+\infty]{} \mathbb{E}(X_1).$$

Imagine-se agora um gás a mover-se livremente num contentor. O que podemos afirmar sobre o comportamento assimptótico deste movimento? Esta pergunta normalmente é dividida em questões menores - existe alguma periodicidade? Existem pontos de equilíbrio? Qual a natureza dos equilíbrios? Existe algum comportamento médio observável? Estas e outras questões são objecto de estudo da teoria de Sistemas Dinâmicos. Um dos primeiros resultados da teoria, devido a Poincaré, dá uma resposta inicial ao problema afirmando que, se aguardarmos tempo suficiente, o sistema volta arbitrariamente próximo da configuração inicial.

O movimento de gases e outros sistemas com grandes números de corpos e graus de liberdade é o objecto de estudo da Mecânica Estatística. No fim do século XIX, um dos pioneiros da área, Boltzmann, formulou uma hipótese, conhecida como Hipótese Ergódica, de teor semelhante à lei dos grandes números segundo a qual num sistema em equilíbrio as médias temporais convergem para a média espacial. Uma formulação matemática da hipótese ergódica não é clara no entanto, nos anos 30 surgiram dois resultados nessa direcção - os teoremas ergódicos "clássicos" de Birkhoff e von Neumann.

Tanto a lei dos grandes números como os teoremas de Birkhoff e von Neumann têm um aspecto importante em comum - a comutatividade da operação associada. Durante a década de 50, Furstenberg e Kesten questionaram-se quando a possíveis generalizações deste tipo de resultados a cenários mais gerais, nomeadamente casos em que a comutatividade falha. Como é comum, os espaços não comutativos paradigma são $GL(d,\mathbb{R})$ e os seus subgrupos. Levou até à década seguinte para os primeiros resultados nessa direcção serem exibidos. Primeiro o Teorema de Furstenberg - Kesten sobre normas de matrizes enquanto operadores que foi generalizado pouco depois pelo Teorema Ergódico Subaditivo de Kingman e finalmente o Teorema Ergódico Multiplicativo de Oseledets. É comum chamar-se aos teoremas ergódicos multiplicativos de teoremas ergódicos não comutativos.

No final da década de 80, Kaimanovich traduziu o teorema ergódico multiplicativo em linguagem geométrica, nomeadamente, utilizando a acção isométrica de $GL(d,\mathbb{R})$ sobre o espaço das matrizes simétricas definidas positivas. Assim, o teorema ergódico multiplicativo torna-se uma afirmação sobre isometrias e geodésicas. No final do século XX, Karlsson e Margulis repararam que todo o grupo com certas "boas características" actua sobre algum espaço métrico. O Teorema Ergódico Multiplicativo de Karlsson - Margulis é então uma generalização do Teorema Multiplicativo Ergódico de Oseledets que não só abrange grupos mais gerais como pode também ser aplicado a certos semigrupos.

O objectivo do presente texto é apresentar todos os teoremas mencionados anteriormente, as suas provas e a geometria subjacente ao seu estudo de forma auto-contida e simplificando notações sempre que possível. Apresentaremos também o Teorema de Karlsson - Ledrappier que generaliza o teorema de Karlsson e Margulis em vários casos. A importância do texto reside no facto de não se conhecer uma possível formalização da versão mais geral do Teorema Ergódico Multiplicativo e na apresentação certos resultados de uma forma mais detalhada.

No primeiro capítulo do texto, *Ergodic Theory*, apresentamos uma visão generalista do que são sistemas dinâmicos seguindo-se então o enunciado dos teoremas ergódicos clássicos, Furstenberg-Kesten, Kingman, Oseledets e a decomposição ergódica. Neste capítulo apresentamos a prova do teorema subaditivo ergódico de Kingman no caso ergódico que será um peça fundamental na prova dos teoremas multiplicativos. Finalizamos a tese com considerações sobre extremalidade e ergodicidade.

O segundo capítulo, *Geodesic Metric Spaces*, forma grande parte do texto e tem como objectivo principal apresentar a geometria necessária ao entendimento dos resultados. Como o título indica, estamos interessados em espaços métricos geodésicos e, para tal, teremos de definir curvas geodésicas em espaços métricos arbitrários. O estudo de geodésicas está fortemente relacionado com o estudo da curvatura, que terá um papel importante também no teorema de Karlsson - Margulis. Com este objectivo apresentamos e relacionamos três classes importantes de espaços métricos geodésicos:

- A primeira classe de espaços métricos geodésicos que introduzimos são os espaços CAT(k). Estes espaços já vêm equipados com uma noção de "curvatura limitada por k". A importância dos espaços CAT(k) reside na intuição geométrica sendo espaços em que as construções e definição são dadas à custa de triângulos. Em troca da intuição geométrica apresenta-se a dificuldade em manusear tais espaços preferindo-se então, sempre que possível, utilizar outros métodos para estudar estes espaços;

- A segunda classe que apresentamos são os espaços métricos convexos completos. Aqui faz sentido introduzir uma segunda noção de curvatura - curvatura não positiva no sentido de Busemann. Note–se que ao longo to texto estamos especialmente interessados no caso de curvatura não positiva, tais espaços têm propriedades importantes como a unicidade de geodésicas entres quaisquer dois pontos. É importante destacar que todo o espaço CAT(0) tem curvatura não positiva no sentido de Busemann;

- A última classe que estudamos, e a que mais aprofundamos, são as Variedades Riemannianas conexas e completas. Nesta secção percorremos as ideias e resultados principais da teoria: isometrias, conexões afim, derivada covariante, geodésicas, curvatura, o teorema de Hopf - Rinow, campos de Jacobi, o teorema de Cartan - Hadamard, e o teorema geométrico principal para o nosso texto que relaciona a curvatura seccional de uma variedade Riemanniana com a sua curvatura enquanto espaço métrico, nomeadamente, toda a variedade Riemannian completa, conexa, simplesmente conexa de curvatura seccional negativa é um espaço CAT(0). O capitulo apresenta um caracter mais completo nos resultados finais em que relacionamos a curvatura Riemanniana com as anteriores.

Seguidamente temos uma secção sobre Grupos de Lie e a sua relação com a geometria. A relevância desta secção encontra-se nas ideias originais de Kaimanovich e na apresentação da geometria hiperbólica plana a que se refere a secção seguinte, onde calculamos a conexão de Levi - Civita, as geodésicas, as isometrias e a distância do plano hiperbólico. Tendo estudado o plano hiperbólico podemos facilmente transportar os resultados para o modelo do disco de Poincaré no qual a visualização é facilitada. A geometria

hiperbólica será utilizada como meio privilegiado para apresentar exemplos. Na secção seguinte fazemos um estudo da geometria do espaço das matrizes simétricas definidas positivas que será fundamental à nossa apresentação do teorema ergódico multiplicativo, nomeadamente a sua curvatura. Terminamos este capítulo com o conceito de horofunção e a compactificação que lhe está associada. No caso em que trabalhamos com espaços CAT(0) as horofunções têm uma interpretação geométrica mais forte que exploramos de forma a mostrar que nestes espaços o teorema de Karlsson - Margulis é consequência do teorema de Karlsson - Ledrappier.

No último capítulo, *The Noncommutative Ergodic Theorems*, as construções dos capítulos anteriores finalmente se materializam quando demonstramos os Teoremas Ergódicos Multiplicativos de Karlsson - Ledrappier, Karlsson - Margulis e Oseledets. Terminamos a tese reinterpretando o teorema ergódico de Birkhoff e estudando a acção natural do grupo livre sobre o seu grafo de Caley. Na base dos teoremas de Karlsson-Ledrappier e Karlsson - Margulis encontra-se o teorema ergódico subaditivo assumindo assim a sua demonstração, feita no primeiro capítulo, especial relevância. O teorema de Karlsson - Ledrappier refere-se a horofunções e é independente da curvatura do espaço, sendo esta a vantagem do resultado em relação ao Teorema de Karlsson-Margulis, no qual a noção de curvatura não positiva no sentido de Busemann é fulcral.

Palavras-Chave: Teoria Ergódica, Teoremas Ergódicos não Comutativos, Karlsson-Margulis, Espaços Métricos Geodésicos, Curvatura não Positiva.

Abstract

The strong law of large numbers, surely a classical result in probability theory, says that the more an experiment is repeated the closer the sample mean is to the expected value. The attempts at bringing an analogue of this result to statistical mechanics, although hard to formulate from a mathematical point of view, gave rise to Ergodic Theory. Ergodic theory includes itself in the study of dynamical systems, namely it studies asymptotic behaviours of orbits from a measure theory viewpoint by looking at averages.

The first results in ergodic theory, von Neumann's Mean Ergodic Theorem and Birkhoff's Pointwise Ergodic Theorem as well as the strong law of large numbers all have an important aspect in common - the commutativity of the operation at hand. In the 50's Furstenberg and Kesten asked themselves how could they extend such results to more general scenarios, specifically the case in which we work with groups whose commutativity may fail. It took until the 60's for the first answers to such problems to be recorded. These were Furstenberg-Kesten Theorem, Kingman Subadditive Ergodic Theorem and Oseledets Multiplicative Ergodic Theorem.

This text aims to present the noncommutative ergodic theorems from a geometrical point of view. The first to notice the relationship between geometry and Oseledets theorem was Kaimanovich by looking at it as a consequence of the action of $GL(d,\mathbb{R})$ on the space of Symmetric Positive Definite Matrices. Later on, Karlsson and Margulis further extended the works of Kaimanovich to semigroups of semicontractions of more general spaces. This allows us to translate the problem into a geometric one on which we can use different machinery.

The thesis is comprised of three chapters with the goal of presenting all the results above. The first consists of the classical ergodic theory, the second is about the theory of geodesic metric spaces whilst the proof for the main theorems as well as some of the classic ones are presented in the third.

Key Words: Ergodic Theory, Noncommutative Ergodic Theorems, Karlsson-Margulis, Geodesic Metric Spaces, Nonpositive Curvature.

Contents

List of Figures

Introduction

A classical result in probability theory, the strong law of large numbers, asserts that given $\{X_i\}_{i\in\mathbb{N}}$ a sequence of independent identically distributed (i.i.d.) random variables taking values in $\mathbb{R}$ such that the mean value $\mathbb{E}(X_i)$ is finite, then

$$\frac{1}{n}(X_1 + X_2 + \cdots X_n) \xrightarrow[n\to+\infty]{} \mathbb{E}(X_1).$$

This result goes in the direction of frequentist probability. The more we repeat an experience the closer the sample mean is to the expected mean. There are two important aspects at play in this theorem. The more evident one is the i.i.d. property and the second is the commutativity of the sum operation.

Whilst working on statistical mechanics Boltzmann formulated an hypothesis of similar nature to the strong law of large numbers, according to which in a system at an equilibrium state the time averages converge to the space average. In other words, for large periods of times, the amount of time a particle spends in a certain region with the same energy is proportional to the volume of that region.

A solution to such hypothesis has still not been accomplished. Nonetheless, attempts at doing so in the scope of measure theory were responsible for the birth of Ergodic Theory, yielding important results such as the ergodic theorems of von Neumann's and Birkhoff's; the second of which greatly generalizes the strong law of large numbers.

When trying to generalize the results above to noncommutative cases we tend to look at the random variables as taking values in some group G. Although starting in the 50's, such approach only came to fruition in the 60's with the first result in this direction being due to Furstenberg and Kesten for the norm of maps taking values in $GL(d,\mathbb{R})$ which, together with it's subgroups, is the paradigmatic noncommutative group. Later on, this theorem was generalized by Kingman in his Subadditive Ergodic Theorem and finally by Oseledets in the Multiplicative Ergodic Theorem. It is common to call the multiplicative ergodic theorems noncommutative ergodic theorems.

In the 80's Kaimanovich [6] translated the multiplicative ergodic theorem of Oseledets into geometric language, namely, Kaimanovich used the fact that $GL(d,\mathbb{R})$ acts on the space of Symmetric Positive Definite Matrices. Whence the multiplicative ergodic theorem becomes a statement on isometries and geodesics. By the end of the 20th century, Karlsson and Margulis explored this direction by using the fact that every group with some "good properties" is a group of isometries. The Noncommutative Ergodic Theorem of Karlsson - Margulis is then a generalization of Oseledets theorem that not only accepts more general groups but can also be applied to some semigroups and reads as follows:

Theorem 0.1 (Karlsson-Margulis)**.** *Let S be a semigroup of semicontractions of some complete, uniformly convex, nonpositively curved in the sense of Busemann, metric space (X,d) with a marked point x_0. Let $(\Omega,\mathscr{B},\mu,T)$ be an ergodic measure preserving dynamical system and $g:\Omega\to S$ a measurable* *map. Given a right cocycle defined by g,* $Z_n(\omega) = g(\omega)g(T(\omega))\cdots g(T^{n-1}(\omega))$*. If*

$$\int_\Omega d(x_0, g(\omega)\cdot x_0)d\mu(\omega) < +\infty,$$

then, for μ*-a.e.* ω*, the following limit exists*

$$\lim_{n\to+\infty} \frac{1}{n} d(x_0, Z_n(\omega)\cdot x_0) = s.$$

Moreover, if $s > 0$*, then for* μ*-a.e.* ω *there is a unique geodesic ray in* X *starting at* x_0 *such that*

$$\lim_{n\to+\infty} \frac{1}{n} d(Z_n(\omega)\cdot x_0, \gamma_\omega(ns)) = 0.$$

As one can see right away, the theorem is highly geometric in essence. To fully grasp the content of the theorem it is important to understand the geometry it entails. We will also study another theorem in the same direction by Karlsson and Ledrappier [7, 8]. For this approach we need to introduce a compactification of our space on which the boundary elements will be functions, called horofunctions.

Theorem 0.2 (Karlsson-Ledrappier)**.** *Let* X *be a proper metric space and* Z_n *an integrable right cocycle taking values in the space of isometries of* X*,* $\mathrm{Isom}(X)$*. Then there is an almost everywhere defined mapping* $\omega \to D_\omega = D$*, where* D *is an horofunction, depending measurably on* ω*, such that*

$$\lim_{n\to\infty} -\frac{1}{n} D(Z_n(\omega)\cdot x_0) = s.$$

where

$$s = \lim_{n\to\infty} \frac{1}{n} d(Z_n(\omega)\cdot x_0, x_0) = \inf_{n\in\mathbb{N}} \frac{1}{n} \int_\Omega d(Z_n(\omega)\cdot x_0, x_0)d\mu(\omega).$$

The manuscript is divided into three major chapters, Ergodic Theory, Geodesic Metric Spaces and The Noncommutative Ergodic Theorems. In the first chapter we begin with a little introduction to Dynamical Systems and then proceed to present the classical ergodic theorems and ergodic decomposition. We also present a proof for the Subadditive Ergodic Theorem in the ergodic case whose steps we will need later in the text. We finish this chapter with the concept of extremality and its relationship with ergodicity.

The second chapter comprises most of the thesis introducing geodesic metric spaces and presenting three big classes in the study of such spaces:

- CAT(k) spaces which come equipped with a notion of curvature being bounded by k. These spaces are very important from an intuition point of view as we will be working and be interested in triangles. As a trade-off these spaces are hard to work with, whence, whenever possible, we use other methods to study them;
- Complete convex metric spaces which are specially important to study Banach spaces. In this section we will also introduce the notion of being nonpositively curved in the sense of Busemann.
- Riemannian manifolds whose study will be taken deeper than the other two. With that in mind we present the main ideas in the theory: Isometries, affine connections, geodesics, curvature, Hopf-Rinow theorem, Jacobi fields, Cartan-Hadamard theorem and our main results which relate sectional curvature with the other curvatures above.

A quick presentation of Lie Groups is also done to permit a further understanding on the geometry of homogeneous spaces. We then present the geometry of the Hyperbolic plane and the Space of Symmetric Positive Definite Matrices. This chapter finishes with a section on horofunctions and the associated compactification needed for the Karlsson and Ledrappier approach.

In the last chapter everything we've constructed so far comes together as we present the noncommutative ergodic theorems, often called multiplicative ergodic theorems, and their respective proofs. We end the thesis with some applications, namely by reinterpreting Birkhoff's Ergodic theorem and random walks on Cayley Graphs.

As a master's student it was my goal to make the text as complete and self-contained as possible. With that in mind, except for a small amount of results, every statement not addressed in a master's course I took is proven. The text supposes familiarity with Functional Analysis, Smooth Manifolds, General and Algebraic Topology and Metric spaces. As a last remark, some proofs contain a bibliographic reference at the start; this is done whenever I feel the text doesn't add enough to the proof to call it my own.

Chapter 1

Ergodic Theory

1.1 Dynamical Systems

The theory of Dynamical Systems is a big branch of mathematics that today finds applications in most areas of scientific knowledge, particularly in its deep-rooted connection to mechanics. The first steps in the theory are attributed to Henri Poincaré due to his finds whilst working on the three body problem at the turn of the 19th century. Since then, the contributions from multiple mathematicians vastly developed the theory, both in deepness and broadness; nonetheless, it is still a very active area of research in mathematics with a large array of open problems. In simplistic terms a dynamical systems is a mathematical structure consisting of a phase space, a notion of time and a law of evolution.

The *phase space*, X, displays the state of the system and its properties are related to the nature of the problem we are trying to solve. In topological dynamics, X is a topological space, usually metric and compact, whereas in ergodic theory it is often assumed to be a measurable space of finite measure. A third important case is that of differentiable dynamics, on which X is a manifold, often compact and riemannian.

Time is usually classified on a first instance by being either continuous or discrete and then by its invertibility. Discrete dynamical systems are presented by iterations of maps, whereas continuous ones often arise as flows of vector fields. Invertibility refers to whether we can predict both past and future or just one of them.

The *law of evolution* is the rule that determines the next state of the system, it is said to be deterministic if the law is unchangeable, that is, we can always know the next step from knowing the state we are in, and stochastic if there is more than one possibility for the next state. Such rule is asked to respect the category we chose to work in upon specifying the phase space.

To put it in purely mathematical terms, a dynamical system is an action $s : G \times X \to X$, where G is a semigroup (group in case of invertible dynamics). Studying a dynamical system consists of understanding the orbits of this action, namely its asymptotic properties. For the remainder of this text we will take $f : X \to X$, some map, and study its iterates. In essence, the action we consider is

$$\begin{aligned} s : \mathbb{N} \times X &\to X \\ (n,x) &\to f^n(x) \end{aligned}$$

where the power denotes composition of f.

1.2 Ergodic Theory

Ergodic Theory is a way to study dynamical systems based on some notion of measure preservation. Imagine an ideal gas moving freely inside some container, such movement can be described by an Hamiltonian system, which preserves Liouville measure. Sure the Hamilton equations display some existence of solutions, however, the problem lies in obtaining them. This difficulty together with preservation of Liouville measure brings us to employ Ergodic Theory when tackling this problem, namely we shall focus on how the particles behave on average. The movement of a gas was one of the original settings on which ergodic theoretical ideas came to be and, to this day, is still a great way to introduce it.

Let $(X,\mathscr{F},\mu)$ be a measure space, a map $f:X\to X$ is said to be *measure preserving* if it is measurable and $f_*\mu=\mu$, in other words, for every $E\in\mathscr{F}$, $f^{-1}(E)\in\mathscr{F}$ and $\mu(f^{-1}E)=\mu(E)$. If $\mu(X)=1$ then X is said to be a *probability space* and μ a *probability measure*. A measure space is said to be *complete* if every subset of a set with zero measure is measurable. A topological space is said to be *second countable* if it admits a countable basis for its topology; and *separable* if it contains some countable dense set. In a general topological space the first always implies the later whereas on a metric space the two are equivalent. A complete separable metric space is called a *Polish space*

The triple $(X,\mathscr{B},\mu)$ is said to be a *standard measure space* if X is a Polish space, $\mathscr{B}$ is the Borel σ-algebra and μ is some measure. For the remainder of this work, we will always assume that X is a metric separable probability space, $\mathscr{F}$ a σ-algebra containing the Borel σ-algebra, μ is a probability measure and f is some measure preserving transformation on X. The quadruple $(X,\mathscr{F},\mu,f)$ is called a *measure preserving dynamical system*, mpds for short.

Remark. Notice that any finite measure μ gives rise to a probability measure dividing it by $\mu(X)$. Finite measure in ergodic theory acts as an analogue to compactness in topological dynamics, it ensures the existence of some asymptotic structure. Completeness is not such a restrictive condition either as we can uniquely extend μ to some complete σ-algebra.

The first big result on Ergodic Theory is that of Poincaré, stating that a system preserving measure will return arbitrarily close to almost any given state. Philosophically the idea of ”eternal recurrence” pre-dates Poincaré, finding its roots in ancient civilizations and being detrimental to his contemporary Friedrich Nietzsche. Looking back at our gas, the theorem asserts that, at some point, the configuration of the phase space will be indistinguishable from the initial state we left it on.

Theorem 1.1 (Poincaré Recurrence)**.** *Let $(X,\mathscr{F},\mu,f)$ be a mpds such that $\mu(U)>0$ for every open set U. Then μ-a.e. x returns arbitrarily close to itself infinitely often. That is,*

$$\mu\big(\{x\in X \mid \textit{there is } n_k\to\infty \textit{ such that } f^{n_k}(x)\to x\}\big)=1.$$

Suppose $(X,\mathscr{F},\mu,f)$ is a mpds admitting some invariant set $E\in\mathscr{F}$, that is, $f^{-1}(E)=E$, then the dynamics can be decomposed into the action on E and $X\backslash E$. We say that $(X,\mathscr{F},\mu,f)$ is ergodic, if every invariant set E satisfies $\mu(E)=0$ or $\mu(E)=1$. We often write (μ,f) to say μ is an ergodic measure with respect to f or f is an ergodic map with respect to μ; it doesn't make sense to say one is ergodic without the other.

Proposition 1.2. *Let $(X,\mathscr{F},\mu,f)$ be an ergodic mpds and A,B be measurable sets such that $\mu(A)>0$. If $f^n(A)\subset B$ for every $n\geq 0$, then $\mu(B)=1$.*

Proof. Start by taking

$$\hat{A}:=\bigcap_{n\geq 0}\bigcup_{k\geq n} f^{-k}(B)$$

and noticing it is an invariant set

$$f^{-1}\Big(\bigcap_{n\geq 0}\bigcup_{k\geq n} f^{-k}(B)\Big) = \bigcap_{n\geq 0}\bigcup_{k\geq n} f^{-(k+1)}(B) = \hat{A}.$$

By ergodicity, since $A \subset \hat{A}$, $\mu(\hat{A}) = 1$. Finally, as $\hat{A} \subset B$, the result follows. □

Proving ergodicity of a system is still a difficult problem with various open questions. Finding ways to imply ergodicity is thus an important step in the theory.

Proposition 1.3. *Suppose $(X,\mathscr{F},\mu,f)$ is a mpds, (μ,f) is ergodic if and only if any measurable function $\varphi : X \to \mathbb{R}$ satisfying $\varphi \circ f = f$ μ-a.e. is constant μ-a.e.*

We say some mpds $(X,\mathscr{F},\mu,f)$ is *(strongly) mixing* if for all $E,F \in \mathscr{F}$,

$$\mu(E\cap f^{-k}F) \xrightarrow[k\to+\infty]{} \mu(E)\mu(F).$$

Suppose we are dealing with a mixing system and E is some invariant set, then $\mu(E\cap f^{-k}E) \xrightarrow[k\to+\infty]{} \mu(E)^2$, from which $\mu(E) = \mu(E)^2$. Therefore the system is ergodic.

Example 1.4. Consider $[0,1) \simeq \mathbb{S}^1$ with the Lebesgue measure and the circle rotation

$$\begin{aligned} R_\alpha : [0,1) &\to [0,1) \\ x &\to x+\alpha \ (\text{mod } 1). \end{aligned}$$

The set of all $E \in \mathscr{F}$ satisfying $m(R_\alpha^{-1}(E)) = m(E)$ is a monotone class containing the algebra of finite disjoint union of intervals. As such it is only needed to check the relation for intervals by monotone class theorem. For any interval I we have $m(R_\alpha^{-1}(I)) = m(I)$. Therefore R_α is measure preserving.

Suppose $\alpha = p/q$ with $p,q \in \mathbb{N}$. Then every point is q-periodic, so, for some $x \in [0,1)$, choose $\varepsilon > 0$ small enough so that the balls of centre $x+j\alpha$, $0 \leq j < q$, are disjoint and their union is not equal to $[0,1)$. Then this union is an invariant set whose measure is not zero nor one.

For the case $\alpha \notin \mathbb{Q}$, let E be some invariant set and consider φ its indicator function, in mathematical terms, $\varphi = \varphi \circ R_\alpha$. In $\mathbb{S}^1$ applying the Fourier transform and its inverse is equivalent to expanding into Fourier series,

$$\varphi(x) = \sum_{n\in\mathbb{Z}} \hat{\varphi}(n) e^{2\pi i n x}$$

where

$$\hat{\varphi}(n) = \int_0^1 \varphi(x) e^{2\pi i n x} dx.$$

However, for every $x \in X$,

$$\varphi(x) = \varphi \circ R_\alpha(x) = \sum_{n\in\mathbb{Z}} \hat{\varphi}(n) e^{2\pi i n\alpha} e^{2\pi i n x}.$$

Due to linear independence of $\{e^{2\pi i n x}\}_{n\in\mathbb{N}}$, $\hat{\varphi}(n) = \hat{\varphi}(n)e^{2\pi i n\alpha}$ for every $n \geq 0$. Since α is irrational we must have $\hat{\varphi}(n) = 0$ whenever $n \geq 1$, whence $\varphi = \hat{\varphi}(0) = m(E)$ m-a.e. Therefore $m(E) = 0$ or $m(E) = 1$, so (m,R_α) is ergodic.

We could extend the above to translation maps on $\mathbb{T}^n$ and classify the ergodicity based on the vector. Conceptually the challenge is similar, however we would need to present Fourier Analysis on locally compact abelian groups and Pontryagin Duality [10].

This simple case showcases some of the difficulties in proving ergodicity - there is no direct method to do it! Each problem must be considered independently and often requires broad knowledge in mathematics and the ability to relate various concepts.

We will now begin presenting some classical results known as the ergodic theorems. Firstly the Mean Ergodic Theorem proved by John von Neumann.

Theorem 1.5 (Mean Ergodic Theorem). *Let $(X,\mathscr{F},\mu,f)$ be a mpds, given $\varphi \in L^2(X,\mu)$, the following limit always exists*

$$\frac{1}{n}\sum_{j=0}^{n-1}\varphi\circ f^j \xrightarrow[n\to+\infty]{L^2} \varphi^*.$$

Moreover,

1. *$\varphi^*\circ f = \varphi^*$ in L^2;*
2. *$\int \varphi^* d\mu = \int \varphi d\mu$;*
3. *in case the system is ergodic, $\varphi^* = \int \varphi d\mu$.*

This theorem allows for a new classification of ergodicity, $(X,\mathscr{F},\mu,f)$ is ergodic if and only if for every $E,F\in\mathscr{F}$

$$\frac{1}{n}\sum_{k=0}^{n-1}\mu(E\cap f^{-k}F) \xrightarrow[n\to+\infty]{} \mu(E)\mu(F).$$

Shortly after the mean ergodic theorem came to public, Birkhoff presented his pointwise version. In the scope of our text, this theorem is the first generalization of the law of large numbers, as we shall discuss shortly.

Theorem 1.6 (Pointwise Ergodic Theorem). *Let $(X,\mathscr{F},\mu,f)$ be a mpds and $\varphi \in L^1(X,\mu)$. Then for μ-a.e. $x\in X$ the following limit exists*

$$\frac{1}{n}\sum_{j=0}^{n-1}\varphi\circ f^j(x) \xrightarrow[n\to+\infty]{} \varphi^*(x).$$

Moreover,

1. *$\varphi^*\in L^1$, $\varphi^*\circ f = \varphi^*$ in L^1;*
2. *$\int \varphi^* d\mu = \int \varphi d\mu$;*
3. *if the system is ergodic, $\varphi^* = \int \varphi d\mu$.*

Consider $(\mathbb{R},\mathscr{F},\mu)$ some probability space and $(\mathbb{R}^{\mathbb{N}},\mathscr{B},\nu)$ the infinite product space with the corresponding σ-algebra $\mathscr{B}$ and measure ν. Let $(\mathbb{R}^{\mathbb{N}},\mathscr{B},\nu,\tau)$ be the mpds, where τ denotes the shift, which is ergodic. If we consider φ to be the projection onto the first coordinate, Birkhoff's ergodic theorem states

$$\frac{1}{n}\sum_{j=0}^{n-1}\varphi\circ \tau^j(x) \xrightarrow[n\to+\infty]{} \int \varphi d\nu.$$

In other words, to meet the notation established in the introduction, given $X_i = \varphi\circ\tau^{i-1}$ a sequence of independent identically distributed random variables such that their mean values is finite, we have

$$\frac{1}{n}\Big(X_1+X_2+\cdots+X_n\Big) \xrightarrow[n\to+\infty]{} \mathbb{E}(X_1).$$

So we obtain the classical law of large numbers.

In this text we are interested in further generalizations, with that in mind, let $(X,\mathscr{F},\mu,f)$ be a mpds, G a locally compact Polish topological group, equipped with its normalized left invariant Haar measure μ_G, and $\varphi : X \to G$ some map in $L^1(X,G)$. Denote by $\mathscr{B}$ the Borel σ-algebra of G. The quadruple $(X\times G,\mathscr{F}\otimes\mathscr{B},\mu\otimes\mu_G,f_\varphi)$ where

$$\begin{aligned} f_\varphi : X\times G &\to X\times G \\ (x,g) &\to (f(x),\varphi(x)g). \end{aligned}$$

is a mpds called the *skew-product*. We call $f^n_\varphi(x,g) = (f^n(x),\varphi(f^{n-1}(x))\cdots\varphi(f(x))\varphi(x)g)$ a *cocycle* on G in general, or a *random walk* in the i.i.d. case. Often we are only interested on the position of the random walk starting at the identity, that is

$$A_n(x) := \varphi(f^{n-1}(x))\cdots\varphi(f(x))\varphi(x).$$

Some texts also call $A_n(x)$ a cocycle. The above defined is called a *left cocycle* as it naturally arises from a left action; we define the *right cocycles* analogously. To make it more visible we will denote left cocycles by A_n and right ones by Z_n. Note that the term cocycle is used quite loosely.

If we consider G to be $\mathbb{R}^n$ we obtain $Z_n(x) = \varphi(x)+\varphi(f(x))+...+\varphi(f^{n-1}(x))$ whose average, by Birkhoff's theorem, converges for almost every $x\in\mathbb{R}^n$. A more sophisticated example arrives when $f : M\to M$ is some diffeomorphism on a parallelizable manifold M, then, by chain rule, we have

$$d_xf^n(v) = d_{f^{n-1}(x)}f\circ d_{f^{n-2}(x)}f\circ\cdots\circ d_xf(v).$$

Doing the identification $\varphi(x) = d_xf\in GL(d,\mathbb{R})$ we have

$$A_n(x) = \varphi(f^{n-1}(x))\varphi(f^{n-2}(x))\cdots\varphi(f(x))\varphi(x).$$

The above is called the *derivative cocycle* and its importance lies in the study of nonuniformly hyperbolic diffeomorphisms in differentiable dynamical systems.

The behaviour of cocycles is rather convoluted, even for spaces of matrices as above. One of the first results in that direction is due to Furstenberg and Kesten. Consider $GL(d,\mathbb{R})$ with the operator norm, that is,

$$||A|| = \max_{v\in\mathbb{S}^{n-1}} ||Av||.$$

Theorem 1.7 (Furstenberg-Kesten). *Let $(X,\mathscr{F},\mu,f)$ be a mpds, suppose $A : X\to GL(d,\mathbb{R})$ is some measurable map such that $\log^+||A||\in L^1(X,\mu)$ and consider the cocycle $A_n(x) = A(f^{n-1}(x))\cdots A(x)$, then the limit*

$$\lim_{n\to\infty}\frac{1}{n}\log||A_n(x)||$$

exists μ-a.e.

In the 60s Kingman generalized both Birkhoff's and Furstenberg-Kesten's theorems with the subadditive ergodic theorem. Let $(X,\mathscr{F},\mu,f)$ be a mpds $a : \mathbb{N}\times X\to\mathbb{R}$ is a *subadditive cocycle* or *process* if for every natural m,n and x in X,

$$a(n+m,x)\le a(m,x)+a(n,f^m(x))$$

with the convention $a(0,x) = 0$ for every $x \in X$. Also define $a^+(1,x) = \max\{a(1,x), 0\}$. Notice we slightly change our notation when working with subadditive processes to highlight the fact it takes values in $\mathbb{R}$

Theorem 1.8 (Subadditive Ergodic Theorem)**.** *Let* $(X, \mathscr{F}, \mu, f)$ *be a mpds, and* $\{a : \mathbb{N} \times X \to \mathbb{R}\}$ *a subadditive process such that* $a^+(1,\cdot)$ *is in* $L^1(X,\mu)$*. Then*

$$A(x) = \lim_{n\to+\infty} \frac{1}{n} a(n,x)$$

exists almost surely. Moreover,

1. $A \circ f = A$ *μ-a.e;*
2. $\int_X A d\mu = \lim_{n\to+\infty} \frac{1}{n} \int_X a(n,x) d\mu(x) = \inf_{n\in\mathbb{N}} \int_X \frac{1}{n} a(n,x) d\mu(x)$*;*
3. $A^+(x) \in L^1(X,\mu)$*;*
4. *if the system is ergodic,* $A = \lim_{n\to+\infty} \frac{1}{n} \int_X a(n,x) d\mu(x) = \inf_{n\in\mathbb{N}} \frac{1}{n} \int_X a(n,x) d\mu(x)$

The last classical ergodic result we present is the multiplicative ergodic theorem of Oseledets whose proof is one of the goals for the text. Let $(X, \mathscr{F}, \mu, f)$ be a mpds and $A : X \to GL(n, \mathbb{R})$ some measurable transformation, define the left cocycle $A_n(x) := A(f^{n-1}(x)) \cdots A(x)$ and the real function $\log^+(z) := \max\{0, \log(z)\}$.

Theorem 1.9 (Multiplicative Ergodic Theorem)**.** *If* $\log^+ ||A(x)^{\pm 1}||$ *is in* $L^1(X,\mu)$*, then there are measurable functions* $k = k(x)$ *and* $-\infty < \chi_1(x) < \chi_2(x) < \cdots < \chi_k(x) < +\infty$ *and measurable filtration*

$$E_1(x) \subset E_2(x) \subset \cdots \subset E_k(x) \subset \mathbb{R}^n$$

such that:

1) *k and* χ_i *are* f*-invariant* ($K(f(x)) = k(x)$*,* $\chi_i(f(x)) = \chi_i(x)$)*;*
2) $A(x)(E_i(x)) = E_i(f(x))$ *for every* $1 \leq j \leq k$*;*
3) *for any vector v in* $E_i(x) \backslash E_{i-1}(x)$

$$\chi_i(x) = \lim_{n\to+\infty} \frac{1}{n} \log ||A_n(x)v||;$$

4) $\lim_{n\to+\infty} \frac{1}{n} \log|\det A_n(x)| = \sum_{i=1}^{k(x)} \chi_i(\dim E_i(x) - \dim E_{i-1}(x))$.

There also exists a multiplicative ergodic theorem for invertible systems which we will state but won't be the object of study.

Theorem 1.10 (Invertible version of the Multiplicative Ergodic Theorem)**.** *If* $f : X \to X$ *is invertible and* $\log^+ ||A(x)^{\pm 1}||$ *is in* $L^1(X,\mu)$ *then there are measurable functions* $k = k(x)$ *and* $-\infty < \chi_1(x) < \chi_2(x) < \cdots < \chi_k(x) < +\infty$ *and a direct sum decomposition* $\mathbb{R}^n = \bigoplus_{i=1}^k E_i(x)$ *into subspaces* $E_i(x)$ *depending measurably on x such that*

1) *k and* χ_i *are* f*-invariant* ($K(f(x)) = k(x)$*,* $\chi_i(f(x)) = \chi(x)$)*;*

2) $A(x)(E_i(x)) = E_i(f(x))$ *for every* $1 \leq j \leq k$*;*

3) *for any vector* v *in* $E_i(x)\backslash\{0\}$

$$\chi_i(x) = \lim_{n\to\pm\infty} \frac{1}{n} \log ||A_{\pm n}(x)v||;$$

4) $\lim_{n\to+\infty} \frac{1}{n} \log|\det A_n(x)| = \sum_{i=1}^{k(x)} \chi_i \dim E_i(x)$.

Where $A_{-n}(x) = A(f^{-n}(x))^{-1} \cdots A(f^{-1}(x))^{-1}$.

Due to the first item in both versions, if the system is ergodic, then K and χ_i are constant. The functions χ_i are known as the *Lyapunov exponents* and their study is an active area of research in itself.

As we've seen before ergodicity can be seen as a sort of measurable connectivity in the sense we can't break our space into smaller pieces to study their dynamics independently. It becomes quite natural to ask how do we go around building or finding those "ergodic components"? Or if they always exist?

Let $\mathscr{P} \subset \mathscr{F}$ be a partition of X_0 where $X_0 = X$ up to a zero measure set. Consider the natural quotient map

$$\begin{aligned} \pi : X_0 &\to \mathscr{P} \\ x &\to P \text{ if } x \in P, \end{aligned}$$

the σ-algebra $\hat{\mathscr{F}} = \{\mathscr{Q} \subset \mathscr{P} \ : \ \pi^{-1}(\mathscr{Q}) = \bigcup_{P\in\mathscr{Q}} P \ \in \mathscr{F}\}$ and the probability measure $\hat{\mu} = \pi_*\mu$, that is,

$$\begin{aligned} \hat{\mu} : \hat{\mathscr{F}} &\to [0,1] \\ \mathscr{Q} &\to \mu(\bigcup_{P\in\mathscr{Q}} P) = \sum_{P\in\mathscr{Q}} \mu(P). \end{aligned}$$

Theorem 1.11 (Ergodic Decomposition Theorem)**.** *Let* $(X,\mathscr{F},\mu,f)$ *be a mpds. Then there is* $X_0 \subset X$ *with* $\mu(X_0) = 1$*, a partition of* X_0*,* $\mathscr{P} \subset \mathscr{F}$ *and a family of probability measures* $\{\mu_p\}_{P\in\mathscr{P}}$ *such that*

1) *For every* $E \in \mathscr{F}$ *the function*

$$\begin{aligned} \mu^E : \mathscr{P} &\to [0,1] \\ P &\to \mu_P(E) \end{aligned}$$

is $\hat{\mathscr{F}}$*-measurable.*

2) *For every* $E \in \mathscr{F}$

$$\mu(E) = \int_{\mathscr{P}} \mu_P(E) d\hat{\mu}(P).$$

3) μ_p *are* f*-invariant and ergodic for* $\hat{\mu}$*-a.e.* $P \in \mathscr{P}$*.*

Having the ergodic decomposition tool at hand we often restrict our problems to the ergodic case. From this point on, if a result is stated in the general case and only proven for the ergodic one is because it follows immediately from the decomposition; this will be done without mention henceforth.

1.2.1 Subadditive Ergodic Theorem: Ergodic Case

We are only interested in proving the result in the ergodic case. There are multiple proofs for the subadditive ergodic theorem, we will follow very closely [9] whose proof steps we will need later. The other advantage to this proof is that we obtain Birkhoff's ergodic theorem along the way with no extra effort, hence getting two theorems for the labour of one.

Lemma 1.12 (Fekete)**.** *Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence of real numbers, possibly $-\infty$ such that $a_{m+n} \leq a_m + a_n$, then the following limit exists*

$$A = \lim_{n\to+\infty} \frac{1}{n} a_n = \inf_{n\in\mathbb{N}} \frac{1}{n} a_n.$$

Proof. If for some $n \in \mathbb{N}$, $a_n = -\infty$, then by subadditivity for every m greater than n $a_m = -\infty$ so the result follows. Let us now focus on the non-trivial case. For $n \geq p$, by division algorithm, $n = pq + r$ where $0 \leq r < p$,

$$\begin{aligned} \frac{a_n}{n} \leq & \frac{a_{pq+r}}{n} \leq \frac{q}{qp+r} a_p + \frac{a_r}{n} \\ & \leq \frac{a_p}{p} + \frac{a_r}{n}. \end{aligned}$$

Taking the limits

$$\limsup_{n\to+\infty} \frac{a_n}{n} \leq \inf \frac{a_p}{p} \leq \liminf_{n\to+\infty} \frac{a_n}{n},$$

so the result follows. □

Let $(X, \mathscr{F}, \mu, f)$ be a mpds, and $a : \mathbb{N} \times X \to \mathbb{R}$ a subadditive process satisfying the integrability condition $a^+(1,\cdot) \in L^1(X, \mu)$. Consider

$$a_n = \int_X a(n,x) d\mu(x).$$

The subadditive condition together with the measure preservation property implies $a_{n+m} \leq a_n + a_m$. Therefore the following limit exists

$$A := \lim \frac{1}{n} a_n$$

Let $c_1, \ldots, c_n$ be a finite sequence of real numbers. We call c_i a *leader* if at least one of the sums $c_i, c_i + c_{i+1}, \ldots, c_i + \cdots + c_n$ is negative.

Lemma 1.13. *The sum of all leaders is nonpositive.*

Proof. For a sequence of just one element the result is trivial as the only possible leader is c_1 which happens if and only if it is negative. With complete induction in view, suppose the result is true for some n and every $i < n$. Consider the sequence $c_2, \ldots, c_{n+1}$ whose sum of all leaders, B, is nonpositive. However all leaders of this sequence are leaders of $c_1, c_2, \ldots c_{n+1}$. In case c_1 is not a leader we are done. If c_1 is also a leader then there is a $k \leq n+1$ such that $c_1 + \ldots + c_k < 0$. Notice also that any c_i which is not a leader must be positive. Take $C \leq 0$ to be the sum of all leaders of $c_{k+1}, \ldots, c_{n+1}$, then

$$c_1 + B \leq c_1 + \ldots + c_k + C \leq 0.$$

□

Lemma 1.14. *For every positive ε there is a positive δ such that*

$$\int_C a^+(1,x)d\mu(x) < \varepsilon,$$

whenever $\mu(C) < \delta$.

Proof. If the result weren't true there would be a sequence of sets A_n such that $\lim \mu(A_n) = 0$ and

$$\int_{A_n} a^+(1,x)d\mu(x) > \varepsilon.$$

We know $0 \leq 1_{A_n} a^+(1,x) \leq a^+(1,x)$, so we can use the dominated convergence theorem to obtain

$$\lim \int_{A_n} a^+(1,x)d\mu(x) = \int_X \lim 1_{A_n} a^+(1,x)d\mu(x) = 0,$$

hence reaching an absurd. □

Lemma 1.15. *Suppose that $A > 0$. Let E be the set of $x \in X$ for which there are infinitely many n such that*

$$a(n,x) - a(n-k, f^k(x)) \geq 0$$

for all k, $1 \leq k \leq n$. Then $\mu(E) > 0$.

Proof. [9] For every positive natural number i define the function

$$b_i(x) = a(i,x) - a(i-1, f(x))$$

and the set

$$E_i = \{x \in X \mid \exists k : 0 \leq k \leq i \text{ and } a(i,x) - a(i-k, f^k(x)) < 0\}.$$

Clearly

$$\sum_{j=0}^{k-1} b_{n-j}(f^j(x)) = a(n,x) - a(n-k, f^k(x))$$

as the left side is a telescopic sum.

From definition, if $f^k(x) \in E_{n-k}$, then, for some t satisfying $1 \leq t \leq n-k$,

$$a(n-k, f^k(x)) - a(n-(k+t), f^{k+t}(x)) < 0.$$

Choosing $j = k+t-1$, we obtain $k \leq j \leq n$ and

$$b_{n-k}(f^k(x)) + \cdots + b_{n-j}(f^j(x)) = a(n-k, f^k(x)) - a(n-j-1, f^{j+1}(x)) < 0.$$

Hence, whenever $f^k(x) \in E_{n-k}$, $b_{n-k}(f^k(x))$ is a leader. Using lemma 1.13, for every x in X and n in $\mathbb{N}$

$$\sum_{\substack{k=0 \\ f^k(x) \in E_{n-k}}}^{n-1} b_{n-k}(f^k(x)) \leq 0.$$

Hence

$$
\begin{aligned}
\sum_{k=1}^{n}\int_{E_k} b_k(x)d\mu(x) &= \sum_{k=0}^{n-1}\int_{E_{n-k}} b_{n-k}(x)d\mu(x)\\
&= \sum_{k=0}^{n-1}\int_{f^{-k}E_{n-k}} b_{n-k}(f^k(x))d\mu(x)\\
&= \int_X \sum_{\substack{k=0\\ f^k(x)\in E_{n-k}}}^{n-1} b_{n-k}(f^k(x))d\mu(x)\\
&\leq 0.
\end{aligned}
$$

However, remembering our convention that $a(0,x)=0$,

$$
\begin{aligned}
a_n = \int_X a(n,x)d\mu(x) &= \int_X \sum_{k=0}^{n-1} b_{n-k}(f^k(x))d\mu(x)\\
&= \sum_{k=0}^{n-1}\int_X b_{n-k}(f^k(x))d\mu(x)\\
&= \sum_{k=0}^{n-1}\int_X b_{n-k}(x)df_*^k\mu(x)\\
&= \sum_{k=1}^{n}\int_X b_k(x)d\mu(x)
\end{aligned}
$$

From what we've seen so far and subadditivity, as $a_n/n \to A > 0$,

$$
\begin{aligned}
\sum_{k=1}^{n}\int_{X\setminus E_k} a^+(1,x)d\mu(x) &\geq \sum_{k=1}^{n}\int_{X\setminus E_k} a(1,x)d\mu(x)\\
&\geq \sum_{k=1}^{n}\int_{X\setminus E_k} a(k,x)-a(k-1,f(x))d\mu(x)\\
&= \sum_{k=1}^{n}\int_{X\setminus E_k} b_k(x)d\mu(x)\\
&> a_n > \frac{2A}{3}n.
\end{aligned}
$$

for all n greater than some N.

Consider $f_n = \sum_{k=1}^{n} 1_{X\setminus E_k}$ and denote by a_0 the integral $\int_X a^+(1,x)d\mu(x)$ and

$$
B_n := \left\{x \in X \mid \frac{A}{3a_0}n < f_n(x) \leq n\right\}
$$

the set of elements which are in $X\setminus E_i$ for at least $\frac{A}{3a_0}n$ choices of i. Whenever $n > N$,

$$
\begin{aligned}
\frac{2A}{3}n < \sum_{k=1}^{n}\int_{X\setminus E_k} a^+(1,x)d\mu(x) &= \int_X f_n(x)a^+(1,x)d\mu(x)\\
&= \int_{B_n} f_n(x)a^+(1,x)d\mu(x) + \int_{X\setminus B_n} f_n(x)a^+(1,x)d\mu(x)
\end{aligned}
$$

$$\begin{aligned}
&\leq n \int_{B_n} a^+(1,x)d\mu(x) + \frac{A}{3a_0} n \int_{X\setminus B_n} a^+(1,x)d\mu(x) \\
&\leq n \int_{B_n} a^+(1,x)d\mu(x) + \frac{A}{3} n
\end{aligned}$$

which in turn gives

$$\int_{B_n} a^+(1,x)d\mu(x) > \frac{A}{3}.$$

Using the previous lemma, there is a δ such that

$$\int_C a^+(1,x)d\mu(x) < \frac{A}{3},$$

whenever $\mu(C) < \delta$, so we must have $\mu(B_n) \geq \delta$.

To complete the argument, construct the sets

$$C_n = \{x \in X \mid x \in X\setminus E_i \text{ for at least } \frac{A}{3a_0} n \text{ positive integers } i\}$$

for which $B_n \subset C_n$ and $C_{n+1} \subset C_n$. Therefore,by monotonicity, the measure of $\bigcap_{n\geq 1} C_n$ is greater or equal than $\delta > 0$. The result follows as $\bigcap_{n\geq 1} C_n \subset E$. □

Proposition 1.16. *Suppose that f is ergodic and $A > -\infty$. For any $\varepsilon > 0$, let E_ε be the set of x in X for which there is an integer $K = K(x)$ and infinitely many n such that*

$$a(n,x) - a(n-k, f^k(x)) \geq (A-\varepsilon)k$$

for all k, $K \leq k \leq n$. Let $E = \bigcap_{\varepsilon>0} E_\varepsilon$, then $\mu(E) = 1$.

Proof. [9] For any $\varepsilon > 0$ let $c(n,x) = a(n,x) - (A-\varepsilon)n$

$$\begin{aligned}
c(n+m,x) =& a(n+m,x) - (A-\varepsilon)(n+m) \\
\leq& a(n, f^m(x)) - (A-\varepsilon)n + a(m,x) - (A-\varepsilon)m \\
=& c(n, f^m(x)) + c(m,x)
\end{aligned}$$

so c is a subadditive cocycle and by definition of A

$$\begin{aligned}
\lim_{n\to+\infty} \frac{1}{n} \int_X c(n,x)d\mu(x) =& \lim_{n\to+\infty} \frac{1}{n} \int_X a(n,x) - (A-\varepsilon)n d\mu(x) \\
=& A - (A-\varepsilon) \\
=& \varepsilon \geq 0
\end{aligned}$$

Notice that

$$\begin{aligned}
& c(n,x) - c(n-k, f^k(x)) \geq 0 \\
\Leftrightarrow\ & a(n,x) - (A-\varepsilon)n - a(n-k, f^k(x)) + (A-\varepsilon)(n-k) \geq 0 \\
\Leftrightarrow\ & a(n,x) - a(n-k, f^k(x)) \geq (A-\varepsilon)k.
\end{aligned}$$

Applying the previous lemma to $c(n,x)$ we obtain $\mu(E_\varepsilon) > 0$.

By subadditivity

$$a(n, f^j(x)) - a(n-k, f^{k+j}(x)) \geq a(n+j, x) - a(j, x) - a((n+j)-(k+j), f^{k+j}(x)).$$

If $f^j(x) \in E_\varepsilon$ there are infinite choices of n such that for all $K \leq k \leq n$

$$a(n+j, x) - a((n+j)-(k+j), f^{k+j}(x)) \geq (A-\varepsilon)k,$$

moreover, for k big enough, $a(j,k) \leq \varepsilon k$, so $f^j(E_\varepsilon) \subset E_{2\varepsilon}$. By proposition 1.2, $\mu(E_{2\varepsilon}) = 1$.

It is immediate from definition that $E_\varepsilon \subset E_\delta$ whenever $\varepsilon < \delta$, so we must have $\mu(E) = 1$ □

Theorem 1.17 (Subadditive Ergodic Theorem 1.8)**.** *If f is ergodic and $A > -\infty$. Then, for almost every x,*

$$\lim_{n\to\infty} \frac{1}{n} a(n,x) = A.$$

Proof. If $a(n,x)$ is an additive cocycle simply use the proposition above to both $a(n,x)$ and $-a(n,x)$ obtaining

$$A - \varepsilon \leq \liminf_{n\to\infty} \frac{1}{n} a_n \leq \limsup_{n\to+\infty} \frac{1}{n} a_n \leq A + \varepsilon$$

for every $\varepsilon > 0$. Thus the inferior and superior limit limits must coincide. In essence, we've proven the ergodic case of Birkhoff's ergodic theorem.

Let us focus on the subadditive case and notice that

$$\liminf_{n\to\infty} \frac{1}{n} a(n,x) \geq A - \varepsilon$$

remains valid by the previous proposition. Consider the new cocycle

$$\begin{aligned} b : \mathbb{N} \times X &\to \mathbb{R} \\ (n,x) &\to a(n,x) - \sum_{i=0}^{n-1} a(1, f^i(x)). \end{aligned}$$

We can easily verify

$$\begin{aligned} b(n+m,x) =& a(n+m,x) - \sum_{i=0}^{n+m-1} a(1, f^i(x)) \\ \leq& a(n,x) + a(m, f^n(x)) - \sum_{i=0}^{n-1} a(1, f^i(x)) - \sum_{i=n}^{n+m-1} a(1, f^i(x)) \\ =& a(n,x) - \sum_{i=0}^{n-1} a(1, f^i(x)) + a(m, f^n(x)) - \sum_{i=0}^{m-1} a(1, f^{i+n}(x)) \\ =& b(n,x) + b(m, f^n(x)), \end{aligned}$$

that is, b is subadditive. Moreover, $b(1,x) = 0$, so for every n we have $b(n,x) \leq 0$. This reduces the problem to nonpositive cocycles as $\sum_{i=0}^{n-1} a(1, f^i(x))$ forms an additive cocycle whose convergence of averages is assured from the first step of the proof.

Let $\varepsilon > 0$ and, since $a_n/n \to A$, take M such that

$$\frac{1}{M} \int_X a(M,x) d\mu(x) \leq A + \varepsilon$$

also define

$$
\begin{aligned}
a^M : \mathbb{N} \times X &\to \mathbb{R} \\
(n,x) &\to a(nM,x) - \sum_{i=0}^{n-1} a(M, f^{iM}(x)).
\end{aligned}
$$

We want to apply the previous proposition to a^M, with that in mind, let's verify the hypothesis with respect to f^M.

Subadditivity :

$$
\begin{aligned}
a^M(n+m,x) =& a((n+m)M,x) - \sum_{i=0}^{n+m-1} a(M, f^{iM}(x)) \\
\leq& a(nM,x) - \sum_{i=0}^{n-1} a(M, f^{iM}(x)) + a(mM, f^{nM}(x)) - \sum_{i=0}^{m-1} a(M, f^{iM+nM}(x)) \\
=& a^M(n,x) + a^M(m, f^{nM}(x)).
\end{aligned}
$$

Integrability hypothesis:

$$
(a^M)^+(1,x) = \big(a(M,x) - a(M,x)\big)^+ = 0.
$$

Calculating A_M:

$$
\begin{aligned}
A_M :=& \lim_{n\to\infty} \frac{1}{n} \int_X a^M(n,x) d\mu(x) \\
=& \lim_{n\to\infty} \frac{1}{n} \int_X a(nM,x) d\mu(x) - \sum_{i=0}^{n-1} a(M, f^{iM}(x)) \\
=& \lim_{n\to\infty} M \int_X \frac{a(nM,x)}{nM} d\mu(x) - \frac{1}{n} \sum_{i=0}^{n-1} \int_X a(M, f^{iM}(x)) d\mu(x) \\
=& MA - \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \int_X a(M,x) d\mu(x) \\
\geq& MA - (A+\varepsilon)M \\
=& -\varepsilon M.
\end{aligned}
$$

Ergodicity: notice (μ, f^M) may not be ergodic. In that case we apply the result to an ergodic decomposition of μ with respect to f^M.

We are now in condition to apply the proposition above with $k = n-1$, obtaining

$$
\begin{aligned}
0 \geq& \liminf_{n\to\infty} \frac{1}{nM} a^M(n,x) \\
\geq& \liminf_{n\to\infty} \frac{1}{nM} \big(a^M(1, (f^M)^{n-1}(x)) + (-\varepsilon M - \varepsilon)(n-1)\big) \\
=& -\varepsilon \frac{M+1}{M}.
\end{aligned}
$$

A classical consequence of Birkhoff's ergodic theorem (proven above) is that, for every integrable function g, $\lim_{n\to+\infty} g(f^n(x)) \to 0$ for μ-a.e. x. Due to the integrability hypothesis and the fact $A > -\infty$, every $a(n,x)$ is also integrable. Writing $n = q_nM + r_n = (q_n+1)M - s_n$ with $0 \le r_n < M$ and $0 < s_n \le M$, since s_n and r_n are limited we have $\limsup_{n\to+\infty} a(s_n, f^n(x))/n = \limsup_{n\to+\infty} a(r_n, f^{q_nM}(x))/n = 0$. Finally by subadditivity and the fact we are working with nonpositive cocycles,

$$\begin{aligned}
\limsup_{n\to+\infty}\frac{1}{n}a(n,x) - \liminf_{n\to+\infty}\frac{1}{n}a(n,x) \le& \limsup_{n\to+\infty}\frac{1}{q_nM}a(q_nM,x) + \limsup_{n\to+\infty}\frac{1}{n}a(r_n, f^{q_nM}(x)) \\
&- \liminf_{n\to+\infty}\frac{1}{(q_n+1)M}a((q_n+1)M,x) + \limsup_{n\to+\infty}\frac{1}{n}a(s_n, f^n(x)) \\
=& \limsup_{n\to+\infty}\frac{1}{q_nM}a(q_nM,x) - \liminf_{n\to+\infty}\frac{1}{(q_n+1)M}a((q_n+1)M,x) \\
=& \limsup_{q\to+\infty}\frac{1}{qM}a(qM,x) - \liminf_{q\to+\infty}\frac{1}{qM}a(qM,x) \\
=& \limsup_{q\to+\infty}\frac{1}{qM}a^M(q,x) - \liminf_{q\to+\infty}\frac{1}{qM}a^M(q,x) \\
\le& -\liminf_{q\to+\infty}\frac{1}{qM}a^M(q,x) \\
\le& \varepsilon\frac{M+1}{M}.
\end{aligned}$$

Since this holds for every $\varepsilon > 0$ and the quantity $(M+1)/M$ is bounded we obtain the result. □

1.2.2 Extremality and Ergodicity

In this subsection we present a fundamental tool in Ergodic Theory - extremality and how it relates to ergodicity. These results concern the existence of ergodic pairs (f,μ). For the following set of results $(\Omega, \mathscr{B}, \mu, T)$ is an ergodic mpds, X is a compact Polish space and f is a map from $\Omega \times X$ onto itself preserving the dynamics of Ω.

To set notation, given any measurable space $(Y, \mathscr{F})$ and a map $g : Y \to Y$, denote by $\mathrm{Prob}(Y)$ and $\mathrm{Prob}^g(Y)$ respectively the space of all probability measures on Y and the space of probability measures on Y that are g-invariant. Consider in these spaces the weak*-topology induced by the duality between the space of continuous functions $C(Y)$ and the space of measures $\mathscr{M}(Y)$. Recall that this duality comes from Riez Theorem. We will make the topology more precise when we use the results.

Definition 1.18. Let $K \subset V$ be a convex subset of some locally convex topological vector space. A point $x \in K$ is said to be an *extremal point* of K if for every $x_1 \neq x_2$ points in K and $t \in [0,1]$ if $x = tx_1 + (1-t)x_2$ then $t = 0$ or $t = 1$.

The are three central results in connecting extremality to ergodicity:

Proposition 1.19. *If Y is compact, then $Prob^g(Y)$ is a compact convex sets in $Prob(Y)$.*

Proposition 1.20. *A pair (g,θ) is ergodic if and only if θ is an extremal point in $Prob^g(Y)$.*

Theorem 1.21 (Krein-Milman)**.** *If K is compact, then it is the convex closure of its extremal points.*

Let $\varphi : V \to \mathbb{R}$ be a linear continuous function. Consider

$$s := \max_{\theta\in K}\varphi(\theta).$$

and the set $K_s = \{\theta \in K \mid \varphi(\theta) = s\}$.

Proposition 1.22. *If θ is an extremal point in K_s then it is extremal in K.*

Proof. Let $\theta = t\theta_0 + (1-t)\theta_1$, with $0 < t < 1$ and $\theta_0, \theta_1 \in K$. Suppose either θ_0 or θ_1 is not in K_s, that is, $\varphi(\theta_0) < s$ or $\varphi(\theta_1) < s$. In both cases

$$s = \varphi(\theta) = t\varphi(\theta_0) + (1-t)\varphi(\theta_1) < s.$$

Reaching this absurd we must have $\theta_0, \theta_1 \in K_s$. Since θ is extremal in K_s we have $\theta = \theta_0 = \theta_1$, thus we obtain extremality in K. $\square$

For our goals in this thesis we are specially interested in the space of measures which preserve μ, in other words,

$$\mathrm{Prob}^f_\mu(\Omega \times X) = \{\theta \in \mathrm{Prob}^f(\Omega \times X) \mid \pi_*\theta = \mu\},$$

where π denotes the projection onto Ω. Let us start by seeing how it relates to $\mathrm{Prob}^f(\Omega \times X)$.

Proposition 1.23. *Any extremal point of $Prob^f_\mu(\Omega \times X)$ is an extremal point of $Prob^f(\Omega \times X)$.*

Proof. Let θ be an extremal point in $\mathrm{Prob}^f_\mu(\Omega\times X)$. Suppose $\theta = t\theta_0 + (1-t)\theta_1$ with $\theta_0, \theta_1 \in \mathrm{Prob}^f(\Omega \times X)$ and $0 < t < 1$. Then

$$\mu = \pi_*\theta = t\pi_*\theta_0 + (1-t)\pi_*\theta_1.$$

However, μ is ergodic, hence extremal, thus $\pi_*\theta_0 = \mu$ or $\pi_*\theta_1 = \mu$ which yields $\pi_*\theta_0 = \pi_*\theta_1 = \mu$, that is, $\theta_0, \theta_1 \in \mathrm{Prob}^f_\mu(\Omega \times X)$. By extremality in $\mathrm{Prob}^f_\mu(\Omega \times X)$ we obtain $\theta_0 = \theta_1$. $\square$

Proposition 1.24. *The pair (f, θ) is ergodic whenever θ is an extremal point of $Prob^f(\Omega \times X)$.*

Proof. This is a direct consequence of 1.20, so this proof also gives the backward implication from this proposition. Suppose (f, θ) is not ergodic, then there are A and B invariant sets of positive measure such that $A \cup B = \Omega \times X$. This implies

$$\theta = \theta(A)\theta_A + \theta(B)\theta_B,$$

where θ_A and θ_B are the conditional measures given by

$$\theta_A(E) = \frac{\theta(E \cap A)}{\theta(A)}.$$

All that is left to prove is invariance

$$\begin{aligned}
\theta_A(f^{-1}(E)) =& \frac{\theta(f^{-1}(E) \cap A)}{\theta(A)} \\
=& \frac{\theta(f^{-1}(E) \cap f^{-1}(A))}{\theta(A)} \\
=& \frac{\theta(f^{-1}(E \cap A))}{\theta(A)} \\
=& \frac{\theta(E \cap A)}{\theta(A)} \\
=& \theta_A(E).
\end{aligned}$$

Hence θ is not extremal and we obtain the assertion. $\square$

Chapter 2

Geodesic Metric Spaces

In this chapter we briefly present the theory of metric spaces on which every two points are joined by some notion of line segment, called a geodesic. Just like when defining geodesics in Riemannian manifolds we try to extend our natural ideas from $\mathbb{R}^n$. In Riemannian manifolds we define affine connections as a way to think of second derivatives in order do introduce geodesics as curves with no acceleration. In a general metric space we may not have access to a differentiable structure so we must make do with the metric structure alone.

Throughout the first three sections we will present some classes of geodesic metric spaces and relate them. In particular, we shall see Riemannian manifolds, under particular conditions, give rise to metric spaces maintaining certain properties. From that point on we navigate through these notions using what suits the problem better. Namely, we will use Riemannian geometry to study Hyperbolic Geometry and the space of Symmetric Positive Definite Matrices and the theory of metric spaces to introduce Horofunctions.

2.1 CAT(k) Spaces

Let X be a metric space. A *geodesic path* joining a to b in X is a curve $\gamma : [0,l] \to X$ such that $\gamma(0) = a$, $\gamma(l) = b$ and, for every $t,s \in [0,l]$, $d(\gamma(t),\gamma(s)) = |t-s|$. The image of γ in X is called a *geodesic segment* between a and b, often denoted $[ab]$. A *geodesic ray* is a curve $\gamma : [0,+\infty] \to X$ such that $d(\gamma(t),\gamma(s)) = |t-s|$ for every $t,s \geq 0$.

A metric space in which every two points are joined by some geodesic segment is called a *geodesic space*. If there is a unique geodesic segment joining any two points the space is said to be *uniquely geodesic*.

We say that map $f : X \to X$ is an *isometry* if, for every x and y in X, $d(f(x), f(y)) = d(x,y)$ and a *semicontraction* if $d(f(x), f(y)) \leq d(x,y)$.

Example 2.1. The classical example is $\mathbb{R}^n$ with the euclidean metric, that is, the metric induced by the euclidean inner product

$$\langle x,y\rangle = \textstyle\sum_{i=1}^n x_i y_i \text{ for } x,y \in \mathbb{R}^n.$$

It is well known that the euclidean space is geodesic with geodesic segments of the form $\{at + (1-t)b : t \in [0,1]\}$.

Another example is the n-dimensional sphere $\mathbb{S}^n$ which, being a subspace of $\mathbb{R}^{n+1}$ carries a natural metric. We will however use a different one; let x and y be two points on the sphere $\mathbb{S}^n$ we will consider the metric which returns the angle between the lines containing each of the points and the origin, that is, the unique real number in $[0,\pi]$ such that

$$\cos(d_{\mathbb{S}}(x,y)) = \langle x,y \rangle.$$

With this construction $(\mathbb{S}^n, d_{\mathbb{S}})$ is a geodesic space with the geodesic lines being given by the great arcs which arise as intersections of $\mathbb{S}^n$ with 2-dimensional linear subspaces of $\mathbb{R}^{n+1}$, moreover, if $d_{\mathbb{S}}(x,y) < \pi$ there is a unique geodesic joining x to y.

For our last example begin by considering in $\mathbb{R}^{n+1}$ the Lorentzian quadratic form

$$(x,y) = -x_{n+1}y_{n+1} + \sum_{i=1}^{n} x_i y_i.$$

The hyperbolic n-space $\mathbb{H}^n$ is the upper sheet of the hyperboloid $\{x \in \mathbb{R}^n : (x,x) = -1\}$ together with the metric

$$\cosh(d_{\mathbb{H}}(x,y)) = -(x,y).$$

The space $(\mathbb{H}^n, d_{\mathbb{H}})$ is geodesic being the geodesic segments uniquely defined. Just as in the case of the sphere geodesic lines are determined by intersections of $\mathbb{H}^n$ with 2-dimensional linear subspaces of $\mathbb{R}^{n+1}$.

From these examples we shall now construct the model spaces for the theory of geodesic metric spaces.

Definition 2.2. Given a real number k, denote by M_k^n the following metric spaces:

1. if $k = 0$ then (M_0^n, d_0) is the euclidean space;
2. if $k > 0$ then (M_k^n, d_k) is the space $(\mathbb{S}^n, d_{\mathbb{S}}/\sqrt{k})$;
3. if $k < 0$ then (M_k^n, d_k) is the space $(\mathbb{H}^n, d_{\mathbb{H}}/\sqrt{-k})$.

The spaces defined above are called model spaces as the local study of geodesic spaces is mostly done by comparing the space at hand to some M_k^n. One of the ways this can be done is by using the theory of CAT(k) spaces.

Let x,y,z be three points on a metric space X. A *geodesic triangle* Δxyz is the result of joining the *vertices* x,y,z by a choice of geodesic segments. Notice that such choice may not exist or even be unique. A *comparison triangle* in (M_k^2, d_k) for the triplet (x,y,z) is a geodesic triangle $\bar{\Delta}\bar{x}\bar{y}\bar{z}$ such that $d(x,y) = d_k(\bar{x},\bar{y})$, $d(x,z) = d_k(\bar{x},\bar{z})$ and $d(y,z) = d_k(\bar{y},\bar{z})$. If we can form a geodesic triangle Δxyz we say that $\bar{\Delta}$ is a comparison triangle for the triangle Δ.

Proposition 2.3. *Given three points x,y,z in a metric space X, there is a comparison triangle in M_k^2 if $d(x,y)+d(y,z)+d(x,z) < 2D_k$, where D_k denotes the diameter of M_k^2 which is ∞ for $k \le 0$ and $\pi/\sqrt{k}$ for $k > 0$. Moreover, such triangles are unique up to isometry of M_k^2.*

A point $\bar{a}$ on the geodesic $[\bar{x}\bar{y}]$ is called a *comparison point* for a in $[xy]$ if $d(x,a) = d_k(\bar{x},\bar{a})$.

Definition 2.4. Let (X,d) be a metric space and k a real number. Let Δ be a geodesic triangle in X with perimeter less than twice the diameter and $\bar{\Delta}$ a comparison triangle for Δ in M_k^2. We say that Δ satisfies the CAT(k) *inequality* if for any $a,b \in \Delta$ and $\bar{a},\bar{b}$ their respective comparison points we have $d(a,b) \le d_k(\bar{a},\bar{b})$.

If $k \leq 0$, then X is said to be a CAT(k) space if X is a geodesic space in which geodesic triangles satisfy the CAT(k) inequality.

If $k > 0$, then X is called a CAT(k) space if there is a geodesic joining x to y whenever $d(x,y) < D_k$ and any triangle in X, whose perimeter is less than twice the diameter, satisfies the CAT(k) inequality.

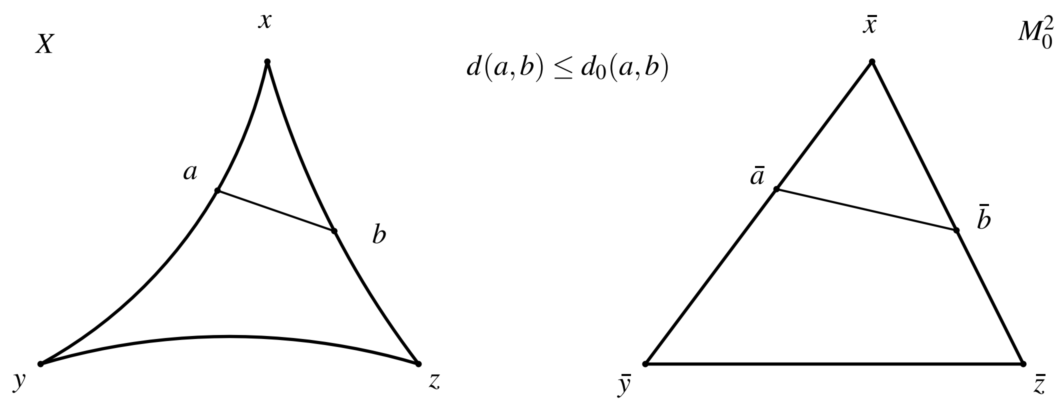

Figure 2.1: Representation of X being a CAT(0) space ($d(x,y) = d(\bar{x},\bar{y})$, $d(x,z) = d(\bar{x},\bar{z})$, $d(y,z) = d(\bar{y},\bar{z})$, $d(x,a) = d(\bar{x},\bar{a})$ and $d(x,b) = d(\bar{x},\bar{b})$).

The definition of a CAT(k), although intuitive from a geometric point of view, may be hard to work with. At this point, the least we can do is state that M_k^n is a CAT(k) space. To delve further we need to understand better the behaviour of triangles in M_k^n as that is what we will be comparing with. The most fundamental result in this direction is the law of cosines.

To introduce the law of cosines we need to work with some notion of angle on a metric space, for example, the Alexandrov angle. Here we shall look at M_k^n as Riemannian manifolds (see Section 2.3) and define the angle between two geodesics by the angle between their tangent vectors. For a discussion on the law of cosines and Alexandrov angles we refer to [3].

Theorem 2.5 (k-Law of cosines)**.** *Given a geodesic triangle ΔABC in M_k^n with sides of positive length a,b,c opposite to vertices A,B,C respectively and angle γ at C we have:*

1. *for* $k = 0$, $c^2 = a^2 + b^2 - 2ab\cos(\gamma)$*;*
2. *for* $k > 0$, $\cos(\sqrt{k}c) = \cos(\sqrt{k}a)\cos(\sqrt{k}b) + \sin(\sqrt{k}a)\sin(\sqrt{k}b)\cos(\gamma)$*;*
3. *for* $k < 0$, $\cosh(\sqrt{-k}c) = \cosh(\sqrt{-k}a)\cosh(\sqrt{-k}b) - \sinh(\sqrt{-k}a)\sinh(\sqrt{-k}b)\cos(\gamma)$.

Namely, c increases with γ.

We will finish this section with the notion of curvature of a metric space which later we shall relate to the sectional curvature of a Riemannian manifold. By Hopf-Rinow Theorem complete connected Riemannian manifolds are also metric spaces so we obtain an improved way of doing calculations whenever we are able to find a suitable Riemannian structure.

Definition 2.6. A metric space is said to have its *curvature* bounded by k if for every x in X there is an open neighbourhood which is a CAT(k) space for the induced metric; using the usual denomination from topology this is also called *locally* CAT(k).

2.2 Convex Metric Spaces

Let (X,d) be a metric space, a point z is called a *midpoint*, also denoted m_{xy}, of x and y if

$$d(z,x) = d(z,y) = \frac{1}{2}d(x,y).$$

If any two points in X admit a midpoint we say the space is *convex*. The first step is clearly showing how convex metric spaces insert themselves inside the study of geodesic spaces.

Proposition 2.7. *Every complete convex metric space is geodesic.*

Proof. We shall give a sketch proof for this result as a complete one goes beyond the scope of this text, moreover, it is a quite natural argument. Let $x,y \in X$ and let us define $c:[0,1] \to X$ by starting with $c(0) = x$ and $c(1) = y$. Now proceed inductively doing

$$c\Big(\frac{a+b}{2}\Big) = m_{c(a)c(b)}.$$

This defines c on the dense set of the dyadic rational numbers on $[0,1]$. By completeness, c is defined on the whole interval. Finally define the geodesic between x and y as

$$\begin{aligned} \gamma : [0,d(x,y)] &\to X \\ t &\to c\Big(\frac{t}{d(x,y)}\Big). \end{aligned}$$

□

Definition 2.8. A metric space is called *uniformly convex* if it is convex and there is a strictly decreasing continuous function $g:[0,1] \to [0,1]$ such that, for every non coincident points $x,y,z \in X$,

$$\frac{d(m_{xy},z)}{R} \leq g\Big(\frac{d(x,y)}{2R}\Big),$$

where R is the maximum between $d(x,z)$ and $d(y,z)$. In particular, $d(m_{xy},z) \leq R$.

Proposition 2.9. *On any uniformly convex metric space midpoints are unique.*

Proof. Let x,y be two points having at least two distinct midpoints, m_1 and m_2. By triangle inequality

$$d(m_1,m_2) \leq d(m_1,x) + d(m_2,y) = d(x,y).$$

Consider m_k the midpoint between m_1 and m_2. Then, by uniform convexity

$$\frac{d(x,m_k)}{R} \;\leq\; g\Big(\frac{d(m_1,m_2)}{2R}\Big) \;=\; g\Big(\frac{d(m_1,m_2)}{d(x,y)}\Big) \;<\; 1$$

Therefore,

$$d(x,y) \;\leq\; d(x,m_k) \;+\; d(m_k,y) \;<\; 2R \;=\; d(x,y),$$

from which we derive our absurd. □

The following is an immediate corollary to this proposition due to the way we constructed geodesics.

Corollary 2.10. *Every complete uniformly convex metric space is uniquely geodesic.*

Example 2.11. 1. The Euclidean space M_0^n is uniformly convex considering the function $g:[0,1]\to[0,1]$ given by $g(x)=(1-x^2)^{1/2}$. Let $x,y,z\in M_0^n$ and suppose without loss of generality that $d(x,z)\geq d(y,z)$. Then take the triangle $\Delta x m_{xy} z$ and denote by $\theta\geq\pi/2$ the angle at m_{xy}. Applying law of cosines

$$\begin{aligned}
&d(z,x)^2=d(z,m_{xy})^2+d(x,m_{xy})^2-2cos(\theta)d(z,m_{xy})d(x,m_{xy})\\
\Leftrightarrow\ &d(z,m_{xy})^2=d(z,x)^2-\Big(\frac{d(x,y)}{2}\Big)^2+cos(\theta)d(z,m_{xy})d(x,y)\\
\Leftrightarrow\ &\frac{d(z,m_{xy})}{R}\leq\Big(1-\Big(\frac{d(x,y)}{2R}\Big)^2\Big)^{\frac{1}{2}}.
\end{aligned}$$

2. The same function we've used above works for CAT(0) spaces due to the comparison being made with the Euclidean space. Let Δxyz be a geodesic triangle in some CAT(0) space and its Euclidean comparison triangle $\Delta\bar{x}\bar{y}\bar{z}$. Then, considering $R=\max\{d(x,z),d(y,z)\}$ and $\bar{R}=\max\{d(\bar{y},\bar{z}),d(\bar{x},\bar{z})\}$,

$$\frac{d(m_{xy},z)}{R}\ \leq\ \frac{d(\overline{m_{xy}},\bar{z})}{\bar{R}}\ \leq\ g\Big(\frac{d(\bar{x},\bar{y})}{2\bar{R}}\Big)\ =\ g\Big(\frac{d(x,y)}{2R}\Big).$$

3. Recall the Clarkson's inequality: in L^p with $1<p<+\infty$,

$$\Big|\Big|\frac{\hat{f}+\hat{g}}{2}\Big|\Big|^p+\Big|\Big|\frac{\hat{f}-\hat{g}}{2}\Big|\Big|^p\leq\frac{1}{2}\Big(||\hat{f}||^p+||\hat{g}||^p\Big),$$

for every $\hat{f}$ and $\hat{g}$ in L^p. Let $f,g,h\in L^p$, $(f+g)/2$ is the midpoint of f and g. Using Clarkson's inequality with $\hat{f}=f-h$ and $\hat{g}=g-h$ and $R=\max\{||f-h||,||g-h||\}$ we have

$$\begin{aligned}
&\Big|\Big|\frac{f+g-2h}{2}\Big|\Big|^p+\Big|\Big|\frac{f-g}{2}\Big|\Big|^p\leq\frac{1}{2}(||g-h||^p+||f-h||^p)\leq R^p\\
\Rightarrow\ &\Big|\Big|\frac{f}{2}+\frac{g}{2}-h\Big|\Big|^p\leq R^p-\frac{||f-g||^p}{2^p}\\
\Leftrightarrow\ &\Big(\frac{||{}^f/_2+{}^g/_2-h||}{R}\Big)^p\leq 1-\Big(\frac{||f-g||}{2R}\Big)^p\\
\Leftrightarrow\ &\frac{||{}^f/_2+{}^g/_2-h||}{R}\leq\Big(1-\Big(\frac{||f-g||}{2R}\Big)^p\Big)^{\frac{1}{p}}.
\end{aligned}$$

So, for $1<p<\infty$ every L^p space is uniformly convex considering the function $g:[0,1]\to[0,1]$ given by $g(x)=(1-x^p)^{1/p}$.

The next result is one we will need later on in the thesis.

Lemma 2.12. *Let (X,d) be a uniformly convex metric space. Let $x,y,z\in X$ and assume*

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

where $\delta\in[0,1]$. Consider w, the point on the geodesic between y and z such that $d(y,w)=d(y,x)$, then there is some f and function such that $f(s)\to 0$ as $s\to 0$ and

$$d(w,x)\leq f(\delta)d(y,x).$$

Proof. [9] By uniform convexity

$$d(m_{xw},z)\leq\max\{d(w,z),d(x,z)\}.$$

Applying the hypothesis we can replace this maximum by a more manageable quantity

$$\begin{aligned} d(w,z) &= d(y,z) - d(y,w) = d(y,z) - d(y,x) \\ d(x,z) &\leq d(y,z) - d(y,x) + \delta d(y,x). \end{aligned}$$

Hence

$$d(m_{xw},z) \leq d(y,z) - d(y,x) + \delta d(y,x).$$

which is equivalent to

$$d(y,z) - d(m_{xw},z) \geq d(y,x) - \delta d(y,x).$$

Applying the triangle inequality,

$$d(y,m_{xw}) \geq (1-\delta)d(y,x).$$

Using uniform convexity again, we know there is some strictly decreasing function $g : [0,1] \to \mathbb{R}$ such that, for $R = \max\{d(x,y), d(y,w)\} = d(x,y)$,

$$(1-\delta) \leq \frac{d(m_{xw},y)}{d(x,y)} \leq g\Big(\frac{d(w,x)}{2d(x,y)}\Big)$$

due to g being strictly decreasing we have

$$g^{-1}(1-\delta) \geq \frac{d(w,x)}{2d(x,y)}$$

which yields the result for $f(\delta) = 2g^{-1}(1-\delta)$. □

We finish this section with an extension to the notion of being CAT(0). We call *nonpositively curved (space) in the sense of Busemann* any geodesic metric space (X,d) for which, given two geodesics γ_1, γ_2 such that $\gamma_1(0) = \gamma_2(0)$, the function

$$\begin{aligned} g : \mathbb{R} &\to \mathbb{R}^+ \\ t &\to \frac{1}{t} d(\gamma_1(t), \gamma_2(t)) \end{aligned}$$

is increasing.

Example 2.13. Let $\gamma_1, \gamma_2 : \mathbb{R} \to X$ be two geodesics on a CAT(0) space coinciding on 0, let $s < t \in \mathbb{R}$ and consider the points $x = \gamma_1(0)$, $y = \gamma_1(t)$, $z = \gamma_2(t)$, $a = \gamma_1(s)$ and $b = \gamma_2(s)$. Take the respective euclidean comparison triangle $\Delta\bar{x}\bar{y}\bar{z}$ and comparison points $\bar{a}, \bar{b}$. The triangles $\Delta\bar{x}\bar{a}\bar{b}$ and $\Delta\bar{x}\bar{y}\bar{z}$ are similar with ratio s/t. Therefore

$$\frac{1}{s}d(a,b) \;\leq\; \frac{1}{s}d(\bar{a},\bar{b}) \;=\; \frac{1}{s}\frac{s}{t}d(\bar{y},\bar{z}) \;=\; \frac{1}{t}d(y,z).$$

Another important class of examples is that of uniformly convex Banach spaces. We will introduce a last example in the form of a subsection since it has more things to be said.

2.2.1 Tree Graphs

A graph is a space X obtained from a discrete countable set V, whose elements are called vertices, by joining some pairs of points a, b by the interval $I_{ab} = (0,1)$, in essence a 1-dimensional CW-complex. From here on we suppose a and b are also in I_{ab} to simplify notation. We denote by E the set of all intervals I_{ab} which we call edges. We say that there is a *path* between vertices a and b if there is a the set of vertices $a = v_1, v_2, ..., v_n = b$ such that $I_{v_i v_{i+1}} \in E$ for every $0 \leq i < n$. We say a graph is *connected* if there is a path between any two points. Given a connected graph X define the distance between vertices to be

$$d(a,b) = \min_{\substack{a=v_1,v_2,...,v_n=b \\ I_{v_i v_{i+1}} \in E}} \#\{v_1, v_2, \cdots v_n\} - 1,$$

Equivalently, attribute to each edge length one. For general points $x \in I_{ab}$ and $y \in I_{cd}$ denote by $\hat{x}, \hat{y}$ the distance of x and y to a and c respectively.

$$d(x,y) = \begin{cases} |\hat{x} - \hat{y}| & \text{if } I_{ab} = I_{cd}, \\ \min\{d(a,c) + \hat{x} + \hat{y}\ ,\ d(a,d) + \hat{x} + 1 - \hat{y}\ , & \\ \qquad d(b,c) + 1 - \hat{x} + \hat{y}\ ,\ d(b,d) + 2 - \hat{x} - \hat{y}\} & \text{ortherwise.} \end{cases}$$

Define the degree of a vertex a, $\deg(a)$ to be the number of edges at a. For the remainder of this section all vertices of all graphs are supposed to have finite degree greater than one.

Before studying some of the geometry of connected graphs let's present a useful proposition which will simplify notation and we will use henceforth without mention. A metric space is said to be *proper* if every closed limited set is compact.

Proposition 2.14. *Every proper metric space is complete.*

Proof. Let $\{x_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence. Given $\delta > 0$ there is an order p such that for all m, n greater than p we have $d(x_n, x_m) < \delta$. For every $n \geq p$ we know $\{x_n\} \in \overline{B_\delta(x_p)}$ which is closed and bounded, hence compact. Then $\{x_n\}_{n\geq p}$ has a convergent subsequence which is also a converging subsequence of $\{x_n\}_{n\in\mathbb{N}}$.

Let x_{n_k} be the converging subsequence and x its limit. Let $\varepsilon > 0$, then there is N such that $d(x_{n_k}, x) < \varepsilon/2$ and $d(x_n, x_{n_k}) < \varepsilon/2$ for every $k, n > N$. By triangle inequality

$$d(x_n, x) \leq d(x_{n_k}, x_n) + d(x_{n_k}, x) < \varepsilon,$$

whenever $k, n > N$. Hence the space is complete. □

Proposition 2.15. *A graph (X, d) is a proper convex metric space.*

Proof. We will only prove properness as the other two statements are usually part of topological graph theory. Let A be a closed bounded set in X. Then it contains a finite number of edges together with their end points or portions of edges, which are given by $[0,1]$ or by closed subsets of it. In either case A is a finite union of compact sets whence compact. □

As we've seen before every convex complete metric space is a geodesic metric space, this inserts connected graphs in the study of geodesic metric spaces. We say that a graph is a *tree* if it is simply connected. A quick consequence of being simply connected is the uniqueness of paths minimizing the distance between any two vertices, whence uniqueness of geodesics.

Proposition 2.16. *Every tree is uniformly convex.*

Proof. Let $x, y, z \in X$ and m_{xy} the unique midpoint of $[xy]$. Since we are working on a tree $m_{xy} \in [xz] \cup [yz]$. Suppose $m_{xy} \in [xz]$, we have $R = d(x,z)$ and

$$d(z, m_{xy}) = d(x,z) - d(x, m_{xy})$$

$$\Leftrightarrow \frac{d(z,m_{xy})}{R} = 1 - \frac{d(x,y)}{2R}.$$

Hence we obtain uniform convexity with $g(\delta) = 1-\delta$. □

Proposition 2.17. *Every tree is nonpositively curved in the sense of Busemann.*

Proof. Let γ_1 and γ_2 be two geodesic rays with the same origin on some tree. Take

$$t_0 = \sup\{t : \gamma_1(t) = \gamma_2(t)\}.$$

Then

$$d(\gamma_1(t),\gamma_2(t)) = \begin{cases} 0 & \text{if } t \le t_0, \\ 2t-2t_0 & \text{if } t > t_0. \end{cases}$$

Finally

$$\frac{1}{t}d(\gamma_1(t),\gamma_2(t)) = \begin{cases} 0 & \text{if } t \le t_0, \\ 2-2\frac{t_0}{t} & \text{if } t > t_0, \end{cases}$$

is an increasing function so we obtain the result. □

2.3 Riemannian Geometry

A *Riemannian manifold* is a pair (M,g) in which M is a differentiable manifold and g, usually called a Riemannian structure, is some covariant 2-tensor such that g_p is nondegenerate, symmetric and positive definite for every p in M. In other terms, g_p is a choice of inner product in T_pM which varies smoothly with p, as such, one usually denotes $g = \langle \cdot,\cdot \rangle$ and $g_p = \langle \cdot,\cdot \rangle_p$

An *isometry* between two Riemannian manifolds (M,g) and (N,h) is a diffeomorphism $f : M \to N$ such that $f^*h = g$, where f^* denotes the pullback. A smooth map $f : M \to N$ is called a *local isometry* at $p \in M$ if there is a neighbourhood $U \subset M$ such that $f|_U : U \to f(U)$ is an isometry.

Given a differentiable curve $c : [a,b] \to M$ on a Riemannian manifold we can calculate the *length* of c, denoted $l(c)$, as

$$l(c) = \int_a^b |c'(t)|dt.$$

The definition naturally extends to piecewise differentiable curves as summing the lengths of the portions on which c is smooth.

An important concept on the study of manifolds is how to differentiate vector fields. This doesn't pose any difficulty in $\mathbb{R}^n$ as tangent spaces are "connected by translations". Transporting this idea of "connecting" tangent spaces to a smooth manifold is not as easy but we shall now make it precise.

Definition 2.18. Given a smooth manifold M an *affine connection* on M is a $\mathbb{R}$-bilinear map

$$\begin{aligned} \nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) &\to \mathfrak{X}(M) \\ (X,Y) &\to \nabla_X Y \end{aligned}$$

satisfying C^∞-linearity on the first entry and the Leibniz rule on the second; that is,

1. $\nabla_{fX+gY}Z = f\nabla_X Z + g\nabla_Y Z$,
2. $\nabla_X(Y+Z) = \nabla_X Y + \nabla_X Z$,

3. $\nabla_X(fY) = f\nabla_X Y + X \cdot fY$,

for every $X, Y, Z \in \mathfrak{X}(M)$ and $f, g \in C^\infty(M)$.

Given $X, Y \in T_pM$ and $p \in M$ then $\nabla_X Y_p$ depends only on the values of Y along a curve to which X is tangent at p and X_p. Effectively, considering local coordinates $x : U \to \mathbb{R}^n$ where U is an open set in M and writing $X = \sum_{i=1}^n x_i \frac{\partial}{\partial x^i}$ and $Y = \sum_{i=1}^n y_i \frac{\partial}{\partial x^i}$ on U, we have

$$\nabla_X Y = \sum_{i=1}^n \left(X \cdot y_i + \sum_{j,k=1}^n \Gamma^i_{jk} x_j y_k \right) \frac{\partial}{\partial x^i},$$

where $\Gamma^i_{jk} : U \to \mathbb{R}$, known as the *Christoffel Symbols*, are determined by

$$\nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k} = \sum_{i=1}^n \Gamma^i_{jk} \frac{\partial}{\partial x^i}.$$

Locally one can uniquely determine an affine connection by a choice of Christoffel Symbols, however to define it globally such a choice must coincide whenever charts overlap. Although being an important tool, in this text we will avoid to calculate the Christoffel Symbols for big dimensions as they are n^3 which becomes a tedious computation quite quickly.

Another important concept which is related to affine connections and we also transport to manifolds is that of covariant derivative. A *vector field V along a curve* $c : I \to M$ is a map $V : I \to TM$ such that $V(t) \in T_{c(t)}M$.

Definition 2.19. Let M be a manifold equipped with some affine connection ∇. We define the *covariant derivative* of a vector field V along a curve $c : I \to M$ as the unique correspondence which associates with V another a vector field along c, denoted $\frac{DV}{dt}$ such that

1. $\dfrac{D(X+Y)}{dt} = \dfrac{DX}{dt} + \dfrac{DY}{dt}$,
2. $\dfrac{DfX}{dt} = \dfrac{df}{dt}X + f\dfrac{DY}{dt}$,
3. If $X(t) = Z(c(t))$, for every t in I, for some $Z \in \mathfrak{X}(M)$, then $\dfrac{DX}{dt} = \nabla_{c'(t)}Z$,

for every X, Y defined along c and $f \in C^\infty(M)$.

A vector field V defined along a curve is said to be *parallel* along a curve $c : I \to M$ if $\frac{DV}{dt}(t) = 0$ for all t in I. A curve γ is said to be a *geodesic of the affine connection* ∇ if $\frac{D\gamma'}{dt}(t) = 0$ for all t in I.

Note that, in local coordinates, both equations above define first and second order system of ODEs respectively, which by Picard-Lindelöf theorem have unique solutions for given initial conditions. The unique vector field $V : I \to TM$ parallel along a curve $c : I \to \mathbb{R}$ such that $V(0) = v$ is called the *parallel transport* of v along c.

Intuitively one looks at parallel vector fields as ones which do not change throughout the curve and at geodesics as curves with no acceleration. Such concepts become more clear when one considers Riemannian manifolds as these have a natural choice of connection with a stronger geometrical meaning.

An affine connection ∇ on a smooth manifold M is said to be *symmetric* if for all $X, Y \in \mathfrak{X}(M)$

$$\nabla_X Y - \nabla_Y X = [X, Y],$$

where $[\cdot,\cdot]$ denotes the Lie bracket. Locally symmetry is given by the Christoffel symbols being symmetric, that is $\Gamma^i_{jk} = \Gamma^i_{kj}$. If M has a Riemannian structure then ∇ is said to be *compatible with the metric* if for every $X,Y,Z \in \mathfrak{X}(M)$

$$X \cdot \langle Y,Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle.$$

The first immediate consequences of choosing a connection compatible with the metric is that the angle between parallel vector fields along a curve is constant, giving us a geometric interpretation to our intuition, and the fact that $|\gamma'(t)|$ is constant for any geodesic $\gamma : I \to \mathbb{R}$. In particular, one has that if $I = [a,b]$ then

$$l(\gamma) = \int_a^b |\gamma'(t)|dt = |\gamma'(t)|(b-a).$$

If $|\gamma'(t)| = 1$ for every $t \in I$, then γ is said to be *parametrized by arclength*.

Theorem 2.20 (Levi-Civita)**.** *Every Riemannian manifold (M,g) admits a unique symmetric affine connection ∇ compatible with the metric, denominated the Levi-Civita connection. Such connection may be obtained from the Kozsul Formula*

$$\begin{aligned} 2\langle \nabla_X Y, Z\rangle = {} & X \cdot \langle Y,Z\rangle + Y \cdot \langle X,Z\rangle - Z \cdot \langle X,Y\rangle \\ & - \langle [X,Z],Y\rangle - \langle [Y,Z],X\rangle + \langle [X,Y],Z\rangle. \end{aligned}$$

for $X,Y,Z \in \mathfrak{X}(M)$. Moreover, given local coordinates $x : U \to \mathbb{R}^n$, $U \subset M$ the Christoffel symbols for this connections are

$$\Gamma^i_{jk} = \frac{1}{2}\sum_{l=1}^n g^{il}\Big(\frac{\partial g_{kl}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^k} - \frac{\partial g_{jk}}{\partial x^k}\Big),$$

where $g_{ij} = g\big(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\big)$ and $(g^{ij}) = (g_{ij})^{-1}$.

Due to the nondegeneracy of the Riemannian metric, the tangent and cotangent bundle are isomorphic. This allows for an identification between vector fields and differential 1-forms. Namely given a function $f : M \to \mathbb{R}$ we define the *gradient* of f denoted $\mathrm{grad} f$ by $\mathrm{d}f(Y) = \langle \mathrm{grad} f, Y\rangle$. Define the *covariant Hessian* of f as the covariant 2-tensor

$$\begin{aligned} \mathrm{Hess}(f) : \mathfrak{X}(M) \times \mathfrak{X}(M) &\to \mathbb{R} \\ (X,Y) &\to \langle \nabla_X \mathrm{grad} f, Y\rangle \end{aligned}$$

Using compatibility of the metric we now have

$$X \cdot \langle \mathrm{grad} f, Y\rangle = \langle \nabla_X \mathrm{grad} f, Y\rangle + \langle \mathrm{grad} f, \nabla_X Y\rangle.$$

Notice that

$$\begin{aligned} X \cdot \langle \mathrm{grad} f, Y\rangle &= X \cdot df(Y) = X \cdot Y \cdot f \\ \langle \mathrm{grad} f, \nabla_X Y\rangle &= df(\nabla_X Y) = \nabla_X Y \cdot f. \end{aligned}$$

Using the equalities above we finally get

$$\nabla_X Y \cdot f = X \cdot Y \cdot f - \mathrm{Hess}(f)(X,Y).$$

This gives us a new way of expressing the Levi-Civita connection. Notice however that one usually does identifications of the tangent spaces which may lead to some sort of loss of information, that is, we may not be able to calculate the differentiation $\nabla_X Y$. This will become more evident later on.

Let $p \in M$ and consider $v \in T_pM$, as we remarked above the geodesic equation $\nabla_{c'(t)}c'(t) = 0$ is a second order system of ODEs. This system can be transformed into a first order system of ODEs in the tangent bundle TM as such there is $\mathscr{U}$ a neighbourhood of (p,v) in TM on which geodesics are uniquely determined by $\gamma(0,p,v) = p$ and $\gamma'(0,p,v) = v$. We can now define the exponential map as the correspondence

$$\begin{aligned} \exp : \mathscr{U} &\to M \\ (p,v) &\to \gamma(1,p,v). \end{aligned}$$

Often we focus on a single point obtaining the map

$$\begin{aligned} \exp_p : \mathscr{U} \cap (\{p\} \times T_pM) &\to M \\ v &\to \gamma(1,p,v) \end{aligned}$$

The important points about the exponential is that it is a local diffeomorphism from some open set $U \subset T_pM$ containing the origin onto some open set $V \subset M$ and the fact its differential is the identity map. On T_pM we can now consider $\varepsilon > 0$ small enough so that $\overline{B_\varepsilon(0)}$ is contained in U and define a *normal ball* of radius ε centred at p, $B_\varepsilon(p)$, as $\exp_p(B_\varepsilon(0))$. We are now in condition to state one of the famous facts of geodesics, the property of locally minimizing distances.

Theorem 2.21. *1. Let $(M, \langle\cdot,\cdot\rangle)$ be a Riemannian manifold and $p \in M$. For every q in some normal ball centred at p there is a unique geodesic $\gamma : I \to M$ joining p to q. Moreover, any other piecewise differentiable curve $c : J \to M$ joining p to q satisfies $l(c) \geq l(\gamma)$ with equality only being verified in case c is a reparametrization of γ.*

2. Conversely, given $p,q \in M$ if $\gamma : I \to M$ is a piecewise differentiable curve joining p to q such that for every other piecewise differentiable curve $c : J \to M$ verifies $l(c) \geq l(\gamma)$ then γ is a geodesic.

A Riemannian manifold is said to be *geodesically complete* if $\exp_p$ is defined on all T_pM for every p in M, in which case we can write any geodesic γ starting at p as $\gamma(t) = \exp_p(t\gamma'(0))$ for all values of $t \in \mathbb{R}$. On any connected Riemannian manifold we can define a distance function as

$$d(p,q) = \inf\{l(c) \mid c : I \to M \text{ is a piecewise differentiable curve joining p to q}\}$$

which makes (M,d) a metric space whose topology coincides with the original one. With this in mind we have our first theorem relating Riemannian manifolds to geodesic spaces.

Theorem 2.22 (Hopf-Rinow)**.** *Let M be a connected Riemannian manifold and $p \in M$ the following are equivalent:*

1. *M is geodesically complete.*
2. *M is complete as a metric space.*
3. *M is proper.*

Moreover,

1. *any of the above implies that for every $q \in M$ there is a geodesic γ joining p to q such that $l(\gamma) = d(p,q)$.*

2. *M is a geodesic metric space whose geodesics are Riemannian ones parametrized by arclength.*

Proof. All assertions except for the last one are part of the usual theorem. Let $\gamma\colon [0,l] \to M$ be a geodesic parametrized by arclength, then

$$d(\gamma(t),\gamma(s)) = \int_t^s |\gamma'(x)|dx = |s-t|$$

□

For the remainder of this section all Riemannian manifolds are supposed to be connected. They are also assumed to be complete unless more than one is presented in some statement, in which case we specify which must be complete. We shall now focus on the notion of curvature. The definition presented here is taken from [5], some texts prefer to present the symmetric but it is all a matter of convention.

Definition 2.23. Let $(M,\langle\cdot,\cdot\rangle)$ be a Riemannian manifold equipped with its Levi-Civita connection ∇. The *curvature* R of M is a correspondence that to every pair of vector fields $X,Y \in \mathfrak{X}(M)$ associates the map $R(X,Y) : \mathfrak{X}(M) \to \mathfrak{X}(M)$ given by

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

Perhaps more geometrically intuitive, and actually what Riemann introduced, is the notion of sectional curvature which can be viewed as the Gaussian curvature of some isometrically immersed two dimensional submanifold.

Definition 2.24. In the same conditions as above, let X and Y be two linearly independent vector fields, that is for every $p \in M$ $\{X_p,Y_p\}$ span some 2-dimensional subspace Π of T_pM. The sectional curvature of Π is given by

$$K(\Pi) = \frac{\langle R_p(X,Y)X_p,Y_p\rangle}{|X_p|^2|Y_p|^2 - \langle X_p,Y_p\rangle^2}$$

The model geodesic metric spaces M_k^n play an important role in Riemannian geometry, for starters they are Riemannian manifolds of dimension n and constant sectional curvature k. However their importance goes way beyond that, interfering on a topological level. Such is the content of the Killing-Hopf Theorem.

Theorem 2.25 (Killing-Hopf)**.** *Let M be a complete Riemannian manifold of dimension n with constant sectional curvature k, then the universal covering of M with the canonical metric is isometric to M_k^n.*

In order to relate curvature and sectional curvature any further we need the following technical result found in do Carmo's Riemannian geometry book [5].

Lemma 2.26. *Let M be a Riemannian manifold and $p \in M$. Define the map $R' : \mathfrak{X}(M)^3 \to \mathfrak{X}(M)$ by*

$$\langle R'_p(X,Y,W),Z_p\rangle = \langle X_p,W_p\rangle\langle Y_p,Z_p\rangle - \langle Y_p,W_p\rangle\langle X_p,Z_p\rangle$$

Then M has constant sectional curvature equal to K if and only if $R_p = KR'_p$ where R is the curvature of M.

An important point in the study of Riemannian geometry is how curvature influences the behaviour of geodesics or even the topology of our space. For that we will introduce the tool of Jacobi fields. Let $\gamma\colon [0,a] \to M$ be a geodesic in M parametrized by arclength. The expression

$$\frac{D^2 J}{dt}(t) + R(\gamma'(t), J(t))\gamma'(t) = 0$$

is called the *Jacobi equation* and any vector field J satisfying it is called a *Jacobi field*. The Jacobi equation is a second order ODE so that any Jacobi field is entirely determined by its initial conditions. Effectively, if $\gamma'(0) = v$ and $\frac{DJ}{dt}(0) = w \in T_v T_{\gamma(0)} M$, there is a curve $v :]-\varepsilon, \varepsilon[\to T_{\gamma(0)} M$ such that $v(0) = v$ and $v'(0) = w$. The Jacobi field is given by

$$J(t) = \left.\frac{\partial}{\partial s}\right|_{s=0} \exp_p(tv(s))$$

In essence, a Jacobi field is a variation of geodesics which allows us to study how geodesics deviate from one another based of curvature. We say that a Jacobi field is orthogonal to the geodesic if $\langle \gamma'(t), J(t)\rangle = 0$ for all t in the domain of γ.

Let $X \in \mathfrak{X}(M)$, by abuse of language we will still denote by X the vector $X_{\gamma(t)}$. Due to the lemma above, on Riemannian manifolds of constant curvature, for orthogonal Jacobi fields,

$$\begin{aligned}\langle R(\gamma'(t), J(t))\gamma'(t), X\rangle &= K\big(\langle \gamma'(t), \gamma'(t)\rangle \langle J(t), X\rangle - \langle \gamma'(t), J(t)\rangle\langle \gamma'(t), X\rangle\big)\\ &= K\langle J(t), X\rangle.\end{aligned}$$

Using nondegeneracy we can write the Jacobi equation as

$$\frac{D^2 J}{dt} + KJ = 0,$$

which is a linear system of ODEs whose solution is

$$J(t) = \begin{cases} \dfrac{\sin(t\sqrt{k})}{\sqrt{k}} W(t) & \text{if } k > 0,\\ tW(t) & \text{if k=0}\\ \dfrac{\sinh(t\sqrt{-k})}{\sqrt{-k}} W(t) & \text{if } k < 0,\end{cases}$$

where $W : I \to T_\gamma(t)M$ is a parallel vector field along $\gamma\colon I \to M$ such that $\langle W(t), \gamma(t)\rangle = 0$ and $|W(t)| = 1$ for all $t \in I$.

The following lemma is a simplification of the Rauch Comparison Theorem which applies whenever we compare with a space of constant sectional curvature. The spirit of the result remains the same; the rate at which geodesics spread apart may be described by sectional curvature. In positively curved spaces geodesics tend to converge whilst spreading apart in the negative case. Before the next result, recall D_k denoted the diameter of M_k^n.

Lemma 2.27 (Rauch)**.** *Let J be an orthogonal Jacobi vector field for some geodesic $\gamma : [0,l] \to M$ parametrized by arclength, denote by $K(t)$ the sectional curvature at $\gamma(t)$ for the 2-dimensional space spanned by $\{J(t), \gamma'(t)\}$, then, whenever $J(t) \neq 0$,*

$$|J|''(t) \geq -K(t)|J|(t).$$

Moreover, for $K(t) < k$ considering j_k the solution to the Cauchy problem $j_k''(t) = -kj_k(t)$ with initial conditions $j_k(0) = 0, j_k'(0) = 1$, if $J(0) = 0$ and $|\frac{DJ}{dt}|(0) = 1$, then for every $t < \min\{D_k, l\}$

$$|J|(t) \geq j_k(t).$$

Proof. [3] By compatibility with the metric one has

$$|J|'(t) = \frac{d}{dt}\Big\langle J(t), J(t)\Big\rangle^{\frac{1}{2}} = \Big\langle \frac{DJ}{dt}(t), J(t)\Big\rangle |J(t)|^{-1}.$$

Proceeding with the calculations

$$|J|''(t) = \Big\langle \frac{D^2J}{dt}(t), J(t)\Big\rangle |J(t)|^{-1} + \Big\langle \frac{DJ}{dt}(t), \frac{DJ}{dt}(t)\Big\rangle |J(t)|^{-1} - \Big\langle \frac{DJ}{dt}(t), J(t)\Big\rangle^2 |J(t)|^{-3}$$

which, upon applying the Jacobi equation and rearranging some terms, gives

$$|J|''(t) = -\Big\langle R(\gamma'(t), J(t))\gamma'(t), J(t)\Big\rangle |J(t)|^{-1} + \left(\left|\frac{DJ}{dt}(t)\right|^2 |J(t)|^2 - \Big\langle \frac{DJ}{dt}(t), J(t)\Big\rangle^2\right) |J(t)|^{-3}$$

By definition of sectional curvature and applying Cauchy-Schwarz inequality to the term in bracket we get the intended result

$$|J|''(t) \geq -K(t)|J|(t).$$

If j_k is in conditions above it can be written as

$$j_k(t) = \begin{cases} \dfrac{\sin(t\sqrt{k})}{\sqrt{k}} & \text{if } k > 0, \\ t & \text{if k=0}, \\ \dfrac{\sinh(t\sqrt{-k})}{\sqrt{-k}} & \text{if } k < 0. \end{cases}$$

Now assume that $K(t) < k$, $J(0) = 0$ and $|\frac{DJ}{dt}|(0) = 1$, taking into account the expression of j_k we can say that for $0 \leq t < D_k$

$$\begin{aligned} (|J|'(t)j_k(t) - |J|(t)j_k'(t))' =& (|J|''(t)j_k(t) - |J|(t)j_k''(t)) \\ \geq& (-K(t)|J|(t)j_k(t) + k|J|(t)j_k(t)) \\ \geq& 0. \end{aligned}$$

Thus $|J|'(t)j_k(t) - |J|(t)j_k'(t)$ is an increasing function being zero for $t = 0$, from which it is always non-negative for t in the above domain. As such, since $j_k(t) > 0$ for $0 \leq t < D_k$

$$\frac{|J|'(t)}{|J|(t)} \geq \frac{j_k'(t)}{j_k(t)}$$

which, by integration, yields

$$|J|(t) \geq j_k(t)c$$

for some constant $c \in \mathbb{R}$.

All that remains is to calculate c. By l'Hôpital's rule we have

$$\lim_{t\to 0}\frac{|J|(t)}{j_k(t)} = \frac{|J|'(0)}{j_k'(0)} = \lim_{t\to 0}\frac{|J|(t)}{t} = 1,$$

where the last equality comes from the fact the derivative of $\exp_p$ at zero is the identity map. □

Let $p,q \in M$, we say that q is *conjugate point* to p if it is a critical value of $\exp_p$. Equivalently there is a Jacobi field along a geodesic $\gamma\colon [0,l] \to M$ with $\gamma(0) = p$, $\gamma(l) = q$, $J(0) = J(l) = 0$, and $J(t) \neq 0$ for all $0 < t < l$.

Lemma 2.28. *Let M be a Riemannian manifold of non-positive sectional curvature. Then there are no conjugate points.*

Proof. Let $p \in M$, $\gamma\colon [0,l] \to M$ be a geodesic such that $\gamma(0) = p$ and J a Jacobi vector field satisfying $J(0) = 0$. Consider $f(t) = |J(t)|^2$, we have

$$\begin{aligned}
f'(t) &= 2\left\langle \frac{DJ}{dt}(t), J(t)\right\rangle \\
f''(t) &= 2\left|\frac{DJ}{dt}\right|^2 + 2\left\langle \frac{D^2J}{dt}(t), J(t)\right\rangle \\
&= 2\left|\frac{DJ}{dt}\right|^2 - 2\Big\langle R(\gamma'(t),J(t))\gamma'(t), J(t)\Big\rangle \\
&= 2\left|\frac{DJ}{dt}\right|^2 - K(t)\Big(|\gamma'(t)|^2|J(t)|^2 - \langle\gamma'(t),J(t)\rangle^2\Big) \\
&\geq 0.
\end{aligned}$$

This gives that f' is non decreasing. However, since $f'(0) = 0$ we have $f'(t) \geq 0$ for all $t \geq 0$. This means that f itself is non decreasing so $f(t) \geq f(0)$ for $t \geq 0$ and so there can be no conjugate points. □

Theorem 2.29. *Any local isometry $f : M \to N$, where (M,g) is a complete Riemannian manifold and (N,h) is a Riemannian manifold, is a covering map.*

Proof. Let $x \in N$, take $r = r(x) > 0$ small enough so that the exponential $\exp_x^N$ is a diffeomorphism mapping $B_r^{T_xN}(0)$ in T_xN onto $B_r^N(x) = \{y \in N | d_N(x,y) < r\}$, on which geodesics are uniquely determined by its endpoints.

Let $\gamma\colon [a,b] \to M$ be a geodesic in M

$$\begin{aligned}
l(f\circ\gamma) &= \int_a^b |f\circ\gamma(t)'|dt \\
&= \int_a^b |d_{\gamma(t)}f(\gamma'(t))|dt \\
&= \int_a^b |\gamma'(t))|dt = l(\gamma)
\end{aligned}$$

so f preserves lengths. Moreover, we can break γ into smaller curves which are still geodesics and get mapped to geodesics in N since f is smooth upon glueing all these parts together we still get a smooth curve which is a geodesic. In other words, f maps geodesics to geodesics.

Let $p \in f^{-1}(x)$, note that $f^{-1}(x)$ has to be discrete as f is a local diffeomorphism. Due to Hopf-Rinow theorem, since M is complete, we can say that $\exp_p^M(B_r^{T_pM}) = B_r^M(p)$. With this in mind, build the following commutative diagram

$$\begin{array}{ccc} B_r^{T_pM}(0) & \xrightarrow{\exp_p^M} & B_r^M(p) \\ {\scriptstyle d_pf}\big\downarrow & & \big\downarrow{\scriptstyle f} \\ B_r^{T_xN}(0) & \xrightarrow{\exp_x^N} & B_r^N(x). \end{array}$$

The diagram gives $f \circ \exp_p^M = \exp_x^N \circ d_p f$ where the second is a diffeomorphism, in particular injective, hence exp_p^M is a diffeomorphism. Therefore f restricted to $B_r^M(p)$ is a diffeomorphism.

We now claim that $f^{-1}(B_r^N(x)) = \{q \in M \mid d(q, f^{-1}(x)) < r\}$. If $d(f(q), x) < r$ we can choose a geodesic between x and $f(q)$ of length bounded by r, which lifts to a geodesic starting at some point p in $f^{-1}(x)$ and ending at q. Since f preserves distances which is given by geodesics, $d(q, f^{-1}(x)) < r$. For the other inclusion take q such that $d(q, f^{-1}(x)) < r$, then there is some p in $f^{-1}(x)$ such that, as f preserves lengths, $d(f(q), x) = d(q, p) < r$.

Let p_1, p_2 be two pre-images of x and suppose $B_r^M(p_1) \cap B_r^M(p_2) \neq \emptyset$. Take q in this intersection and the geodesics joining p_1 and p_2 to q. These would have to project into two different geodesics in $B_r^N(x)$ which is absurd. As such $f^{-1}(B_r^N(x))$ is the disjoint union $\cup_{p \in f^{-1}(x)} B_r^M(p)$. □

We finish the study of how curvature affects the behaviour of geodesics with the next two results. The first is a classical theorem relating topology and geometry based on the previous lemma, the second is an important immediate consequence of the first.

Theorem 2.30 (Cartan-Hadamard for Riemannian manifolds)**.** *Let (M, g) be a complete Riemannian manifold with non-positive sectional curvature and $p \in M$. Then $exp_p : T_pM \to M$ is a covering map. In particular, if M is simply connected, then exp_p is a diffeomorphism.*

Proof. Since M has non-positive curvature no two points are conjugate, so $\exp_p$ is an immersion, it is also surjective as M is complete. Consider in T_pM the Riemannian structure $h = exp_p * g$, making the exponential a local isometry. In (T_pM, h) lines through the origin are geodesics which means they are globally defined so that (T_pM, h) is complete. From the previous theorem if follows that the $\exp_p$ is a covering map. □

Corollary 2.31. *On any complete, simply connected Riemannian manifold M of non-positive curvature there is a unique geodesic, up to reparametrization, joining any two points.*

Proof. Let $\gamma_1 : [0,1] \to M$, $\gamma_2 : [0,1] \to M$ be two geodesics from p to q. That is, for some $v_1, v_2 \in T_pM$, $\gamma_1(t) = \exp_p(tv_1)$ and $\gamma_2(t) = \exp_p(tv_2)$. Clearly

$$\exp_p(v_1) = \gamma_1(1) = q = \gamma_2(1) = \exp_p(v_2).$$

Since the exponential is a diffeomorphism we get $v_1 = v_2$, which in turn implies $\gamma_1 = \gamma_2$. □

The strength of both Hopf-Rinow and Cartan-Hadamard theorems, albeit the second is a consequence of the first, is the ability to establish global properties from local ones. These theorems have counterparts in the theory of geodesic metric spaces although we will only have a need for the metric Cartan-Hadamard in a weaker version that we shall prove later.

Theorem 2.32 (Cartan-Hadamard for geodesic metric spaces)**.** *Let X be a complete, simply connected geodesic metric space. If X has its curvature bounded by $k \leq 0$ then X is CAT(k).*

Note that once again we are going from local to global. We are saying that being locally CAT(k) implies that the space is globally CAT(k). We will finish this section by relating the sectional curvature of a complete connected Riemannian manifold to its curvature as a geodesic metric space.

Lemma 2.33. *Fix some x in M_k^n and let $p \in M$. Suppose that for some positive $\varepsilon < D_k$ there exists a diffeomorphism φ from $B_\varepsilon(x)$ onto some open set $U \subset M$ such that $\varphi(x) = p$, for which*

1. *$\forall y \in B_\varepsilon(x)$ and $v \in T_yM$ we have $|d_y\varphi(v)| \geq |v|$,*
2. *$|d_y\varphi(v)| = |v|$ if v is tangent to the geodesic joining x to y,*

then

a. *$U = B_\varepsilon(p)$,*

b. *$\forall y,z \in B_{\varepsilon/2}(x)$, $d(\varphi(y), \varphi(z)) \geq d(y,z)$.*

Proof. [3] Let $c : [0,l] \to M$ be a piecewise differentiable curve totally contained in $B_\varepsilon(x)$ starting from x. Clearly $\varphi \circ c$ is a curve in M with initial point p. Moreover, due to 1. we have,

$$\begin{aligned} l(\varphi \circ c) &= \int_0^l |\varphi \circ c(t)'| dt \\ &= \int_a^b |d_{c(t)}\varphi(c'(t))| dt \\ &\geq \int_a^b |c'(t)| dt = l(c) \end{aligned}$$

with 2. implying that the equality holds if c is a geodesic. This proves that $U = B_\varepsilon(p)$.

Let y,z be in $B_{\varepsilon/2}(x)$. Then $d(y,z) \leq d(y,x) + d(z,x) < \varepsilon < D_k$. As such there is a unique geodesic between y and z in the metric sense which is also a Riemannian one. By the inequality given above we have the assertion $d(\varphi(y), \varphi(z)) \geq d(y,z)$. □

Lemma 2.34. *Let M be a complete Riemannian manifold of dimension n whose sectional curvature is bounded by k and p some point in M. There is V, a neighbourhood of p, and $\varepsilon > 0$ so small that for every $q \in V$ there exists a map $\varphi_q : B_\varepsilon(x) \to M$ satisfying the hypothesis of Lemma 2.33.*

Proof. [3] Given a small enough compact neighbourhood V of p, one can find $\varepsilon > 0$ such that $\exp_q : B_\varepsilon^{T_qM}(0) \to B_\varepsilon^M(q)$ is a diffeomorphism for every q in V. Take some linear isometry $\phi : T_xM_k^n \to T_qM$ and define $\varphi_q = \exp_q \circ \phi \circ {\exp_x}^{-1}$ so that the following diagram commutes

$$\begin{array}{ccc} B_\varepsilon^{T_xM_k^n}(0) & \xrightarrow{\phi} & B_\varepsilon^{T_qM}(0) \\ \downarrow{\scriptstyle \exp_x} & & \downarrow{\scriptstyle \exp_q} \\ B_\varepsilon^{M_k^n}(x) & \xrightarrow{\varphi_q} & B_\varepsilon^M(q). \end{array}$$

Let w be a vector tangent to the geodesic joining x to some y in $B_\varepsilon^{M_k^n}(x)$

$$|d_y\varphi_q(w)| = |d_{\phi({\exp_x}^{-1}(y))}\exp_q\left(d_{{\exp_x}^{-1}(y)}\phi\left(d_y\exp_x^{-1}(w)\right)\right)| = |w|.$$

Lastly, let $u,v \in T_xM_k^n$ be such that $|u| = |v| = 1$ and $\langle u,v \rangle = 0$. Construct, the following Jacobi fields

$$J_k(t) = \left.\frac{\partial}{\partial s}\right|_{s=0} \exp_x(t(u+sv))$$

$$J(t) = \frac{\partial}{\partial s}\bigg|_{s=0} \exp_q(t(\phi(u) + s\phi(v)))$$

where $s \in]-\delta, \delta[$ for some small δ. By construction, applying Rauch Lemma 2.27 and the expression of orthogonal Jacobi fields on M_k^n one has

$$|d_{\exp_x(tu)}\varphi_q(J_k(t))| = |J(t)| \geq j_k(t) = |J_k(t)|.$$

□

Theorem 2.35. *A complete connected Riemannian manifold is of curvature bounded by k as a geodesic metric space if its sectional curvature is bounded by k.*

Proof. Let M be a Riemannian manifold of sectional curvature bounded by k and $p \in M$. By the previous lemma we can find ε small enough so that for any $q \in B_{\varepsilon/4}(p)$ there is $\varphi_q : B_\varepsilon(x) \to M$ satisfying conditions from Lemma 2.33.

Let $a, b, c \in B = B_{\varepsilon/4}(p)$. Then we have

$$d(a,b) + d(a,c) + d(b,c) \leq \frac{3\varepsilon}{2} < 2D_k,$$

so there exist comparison triangles in M_k^n for the triangle Δabc. Take $\varphi_a : B_\varepsilon(x) \to B_\varepsilon(a)$. We can construct the comparison triangle so that $\bar{a} = \varphi_a^{-1}(a)$, $\bar{b} = \varphi_a^{-1}(b)$ and $\bar{c}$ is any point making $\bar{\Delta}\bar{a}\bar{b}\bar{c}$ a comparison triangle for Δabc. Let $i \in [ab]$, $j \in [ac]$ and consider $\bar{i} \in [\bar{a}\bar{b}]$ and $\bar{j} \in [\bar{a}\bar{b}]$ their respective comparison points. Take also $\bar{d} = \varphi_a^{-1}(c)$ and denote by $\bar{k}$ the point $\varphi_a^{-1}(j)$.

Notice

$$d_k(\bar{d}, \bar{b}) \ \leq \ d(\varphi_a(\bar{d}), \varphi_a(\bar{b})) = d(c,b) = d_k(\bar{c}, \bar{b}).$$

By the k-law of cosines $\angle \bar{d}\bar{a}\bar{b} < \angle \bar{c}\bar{a}\bar{b}$. Using the k-law of cosines again

$$d(i,j) \ \leq \ d(\varphi_a(i), \varphi_a(j)) \ = \ d_k(\bar{i}, \bar{k}) \ \leq \ d_k(\bar{i}, \bar{j}),$$

that is, B is CAT(k). Therefore M is of curvature bounded by k as a metric space. □

The converse to this theorem, although true, goes beyond our goals in this text. The equivalence is however quite important from an intuition standpoint as it gives us a more geometrical interpretation to sectional curvature since triangles are conceptually more natural than affine connections.

Theorem 2.36. *Let M be a complete, connected, simply connected Riemannian manifold of nonpositive sectional curvature, then M is a CAT(0) space.*

This follows with little effort from the corollary to Cartan-Hadamard on Riemannian manifolds and repeating the argument presented at the end of the previous proof.

2.4 Lie Groups

Lie Groups arise naturally as groups of isometries of Riemannian manifolds. Another important aspect of Lie Groups is the idea that the global object can be studied by some local linearisation, the Lie Algebra. In the present text we won't be interested in the Lie Algebras but rather on the whole group, both acting as isometries and as Riemannian manifolds themselves. A Lie Group G is a group with some smooth structure such that the two agree, that is, the group operation and inverse maps are smooth. In fact, one needs only to check that the map

$$\begin{aligned} q : G \times G &\to G \\ (g,h) &\to gh^{-1} \end{aligned}$$

is smooth.

On any Lie Group G the *translations from the left*

$$\begin{aligned} L_g : G &\to G \\ h &\to gh \end{aligned}$$

and *translations from the right* R_x given analogously are diffeomorphisms.

Example 2.37. Classical examples of Lie Groups include $(\mathbb{R},+)$ and $(\mathbb{S},\times)$ where we identify $\mathbb{S}$ with the complex numbers of modulus one whilst $\times$ is the usual operation. Another important class of examples are groups of matrices such as:

$$\begin{aligned} O(n) =& \{A \in M_{n\times n}(\mathbb{R}) \ : \ AA^T = I\}, \\ GL(n,\mathbb{R}) =& \{A \in M_{n\times n}(\mathbb{R}) \ : \ \det(A) \neq 0\}, \\ SL(n,\mathbb{R}) =& \{A \in M_{n\times n}(\mathbb{R}) \ : \ \det(A) = 1\}. \end{aligned}$$

A Riemannian structure on G for which L_x are isometries is said to be left invariant. Analogously we define right invariant if R_x are isometries and bi-invariant in the case we have both. Without loss of generality, suppose we are give some inner product $\langle \cdot,\cdot \rangle_e$ on T_eG, where e denotes the identity, then we can extend this inner product easily to a left invariant Riemannian structure by taking at $g \in G$

$$\langle u,v \rangle_g = \langle d_g L_{g^{-1}}(u), d_g L_{g^{-1}}(v) \rangle_e, \quad \forall u,v \in T_gG.$$

Let G be a group and M any set. We say that G *acts on* M if there is a mapping $\mu : G \times M \to M$, usually denoting $\mu(g,x) = g \cdot x$, such that $e \cdot p = p$ for all p in M and $g \cdot (h \cdot p) = (gh) \cdot p$ for all g,h in G and p in M.

Given a $p \in M$ define the *orbit of* p as $G \cdot p = \{g \cdot p \mid g \in G\}$. If the orbit $G \cdot p$ consist of only p then p is said to be a *fixed point* of the action. If for some $p \in M$ the orbit of p is the whole M then the action is said to be *transitive*. The *stabilizer* G_p is the group formed by the elements which fix p, that is $G_p = \{g \in G \mid g \cdot p = p\}$. The action is called *free* if every stabilizer contains only the identity. Finally an action is said to be *proper* if the map

$$\begin{aligned} G \times M &\to M \times M \\ (g,p) &\to (g \cdot p \ , \ p) \end{aligned}$$

is proper.

If G is a Lie Group and M is a smooth manifold we say that G acts smoothly on M if the action $\mu : G \times M \to M$ is smooth. From this point we work with such actions even if the definitions/concepts at hand are easily extended to topological or general groups.

The space of all orbits, called *orbit space* is denoted M/G. This notation is motivated from the fact that $p \sim q \Leftrightarrow q \in G \cdot q$ defines a equivalence relation for which orbits are equivalence classes.

Theorem 2.38. *Let M be a smooth manifold and G a Lie group acting freely and properly on M. Then M/G is a smooth manifold of dimension $\dim M - \dim G$. Moreover, the quotient map $\pi : M \to M/G$ is a submersion.*

The theorem above is already interesting when G is a discrete group (trivially a Lie Group). In this setting, the quotient map is a local diffeomorphism, moreover, it is in fact a covering map and G is the group of deck transformations of M. Recall that a diffeomorphism $f : M \to M$ is a *deck transformation* if $\pi \circ f = \pi$ and the group of all deck transformations is isomorphic to the fundamental group of M/G.

As an example consider $\mathbb{S}^n$ for $n \geq 2$ and the action given by $G = \{id, -id\}$. This action is clearly proper and free, thus $\mathbb{S}^n/G$ is a n-dimensional manifold, known as the real projective space $\mathbb{R}P^n$, whose fundamental group is $\mathbb{Z}_2$.

Suppose now that M is a Riemannian manifold and G some discrete subgroup of the group of isometries, $\mathrm{Isom}(X)$, acting properly and freely on M. Since π is a local diffeomorphism its differential at any point is in fact a linear isomorphism. Now define on M/G the Riemannian structure that makes π a local isometry, that is, given $p \in M/G$ choose $\bar{p}$ in the pre-image of p and define

$$\langle u, v \rangle_p = \langle (d_{\bar{p}}\pi)^{-1}(u), (d_{\bar{p}}\pi)^{-1}(v) \rangle_{\bar{p}}.$$

for all u, v in $T_p(M/G)$. From construction G is transitive on $\pi^{-1}(p)$, that is, given some other $\bar{q} \in \pi^{-1}(p)$ there is g in G such that $g(\bar{q}) = p$. Consequently, as G is the group of deck transformations, the definition above is independent of the $\bar{p}$ chosen.

We finish this section with a result on homogeneous spaces which we will need later. A *homogeneous space* is simply a manifold M with some Lie group G acting transitively on it.

Theorem 2.39. *Let G be a Lie Group and M some set such that G acts on M transitively. If for some $p \in M$ the stabilizer of p, G_p, is closed in G, then there is a unique smooth structure on M with respect to which the action is smooth. Moreover, M is diffeomorphic to the quotient G/G_p.*

2.5 Hyperbolic Plane Geometry

2.5.1 Upper-Half Plane Model

Throughout the text, the content of this section will be used as a privileged medium of showcasing examples. Hyperbolic geometry is a classical and heavily studied topic in mathematics, its advent inscribes itself in the thought revolution of the 19th century which spread through all areas of knowledge. Hyperbolic geometry appeared as a breakthrough from the shackles of Euclid's parallel postulate, even though the first semblances of hyperbolic geometry had appeared earlier due to attempts of proving the Euclid statement. Any attempt to prove the Euclid's axiom was in fact doomed to fail as it was later proved to be independent of the other axioms. The fundamental ideals behind hyperbolic geometry and the surface studies of Gauss were validated and came to their full form by the work of Riemann which we presented in a previous section. We choose to make a presentation based on what we've seen so far, however hyperbolic geometry can be introduced in independent fashion, we refer to [2] for a beautiful approach to the matter.

Consider the set $H = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$ which we identify with the invertible affine maps $h : \mathbb{R} \to \mathbb{R}$ given by $h(t) = yt + x$ and transport the operation coming from composition and the corresponding inverse, that is

$$\begin{aligned} m : H\times H &\to H & \qquad i : H &\to H \\ \big((x,y),(z,w)\big) &\to (yz+x\ ,\ yw) & (x,y) &\to \Big(-\frac{x}{y}\ ,\ \frac{1}{y}\Big). \end{aligned}$$

H is an open set of $\mathbb{R}^2$ so it inherits a differentiable structure for which both these maps are smooth, that is, H is a Lie Group.

The left translations $L_{(x,y)} : H \to H$ are given by $L_{(x,y)}(z,w) = (yz+x\ ,\ yw)$, thus, in the natural coordinates,

$$d_{(z_0,w_0)}L_{(x,y)} = \begin{bmatrix} \dfrac{\partial yz+x}{\partial z} & \dfrac{\partial yw}{\partial z} \\ \dfrac{\partial yz+x}{\partial w} & \dfrac{\partial yw}{\partial w} \end{bmatrix}_{(z_0,w_0)} = \begin{bmatrix} y & 0 \\ 0 & y \end{bmatrix} = yI$$

Consider the euclidean norm on $T_{(0,1)}H$, the left invariant metric on H can be given by

$$\begin{aligned} \langle u,v\rangle_{(x,y)} =& \langle d_{(x,y)}L_{(-\frac{x}{y},\frac{1}{y})}(u)\ ,\ d_{(x,y)}L_{(-\frac{x}{y},\frac{1}{y})}(v)\rangle_{(0,1)} \\ =& \Big\langle \frac{1}{y}\ u, \frac{1}{y}\ v\Big\rangle_{(0,1)} \\ =& \frac{1}{y^2}\langle u,v\rangle_{(0,1)}, \end{aligned}$$

in tensor notation, $g = \frac{1}{y^2}\big(dx\otimes dx + dy\otimes dy\big)$.

Using the Levi-Civita theorem to calculate the Christoffel symbols, $\Gamma^2_{11} = 1/y$, $\Gamma^1_{12} = \Gamma^1_{21} = \Gamma^2_{22} = -1/y$ and $\Gamma^1_{11} = \Gamma^1_{22} = \Gamma^2_{12} = \Gamma^2_{21} = 0$ and applying the symbols to the geodesic equations

$$\frac{d^2x^i}{dt^2} + \sum_{j,k=1}^{2} \Gamma^i_{jk}\frac{dx^j}{dt}\frac{dx^k}{dt} = 0, \quad i = 1,2,$$

we get the following system of ODEs

$$\begin{cases} \dfrac{d^2x}{dt^2} - \dfrac{2}{y}\dfrac{dx}{dt}\dfrac{dy}{dt} = 0, \\ \dfrac{d^2y}{dt^2} + \dfrac{1}{y}\Big[\Big(\dfrac{dx}{dt}\Big)^2 - \Big(\dfrac{dy}{dt}\Big)^2\Big] = 0. \end{cases}$$

Gladly the geodesics we are looking for are well known which will facilitate solving the system above. Start by exploring the case $dx/dt = 0$, for which we can rewrite the system as

$$\begin{cases} \dfrac{d^2x}{dt^2} = 0, \\ \dfrac{d^2y}{dt^2} - \dfrac{1}{y}\ \Big(\dfrac{dy}{dt}\Big)^2 = 0. \end{cases}$$

The solutions are of the form (a, be^t) where a and b are some constants. These solutions correspond to vertical lines. For the other case consider the expression

$$y\frac{dy}{dt}\Big(\frac{dx}{dt}\Big)^{-1} + x.$$

Differentiating the above and using the system,

$$\begin{aligned}
&\Big(\frac{dy}{dt}\Big)^2\Big(\frac{dx}{dt}\Big)^{-1}+y\,\frac{d^2y}{dt^2}\Big(\frac{dx}{dt}\Big)^{-1}-y\,\frac{dy}{dt}\frac{d^2x}{dt^2}\Big(\frac{dx}{dt}\Big)^{-2}+\frac{dx}{dt}\\
=&\Big(\frac{dy}{dt}\Big)^2\Big(\frac{dx}{dt}\Big)^{-1}-y\,\frac{1}{y}\Big[\Big(\frac{dx}{dt}\Big)^2-\Big(\frac{dy}{dt}\Big)^2\Big]\Big(\frac{dx}{dt}\Big)^{-1}+y\,\frac{dy}{dt}\Big[\frac{2}{y}\frac{dx}{dt}\frac{dy}{dt}\Big]\Big(\frac{dx}{dt}\Big)^{-2}+\frac{dx}{dt}\\
=&\Big(\frac{dy}{dt}\Big)^2\Big(\frac{dx}{dt}\Big)^{-1}-\frac{dx}{dt}+\Big(\frac{dy}{dt}\Big)^2\Big(\frac{dx}{dt}\Big)^{-1}+2\Big(\frac{dy}{dt}\Big)^2\Big(\frac{dx}{dt}\Big)^{-1}+\frac{dx}{dt}\\
=&0.
\end{aligned}$$

Therefore

$$y\,\frac{dy}{dt}+x\,\frac{dx}{dt}\;=\;c_1\,\frac{dx}{dt},$$

so the solutions must verify the circle equation $(x-c)^2+y^2=k^2$, for some constants c and k. These semicircles shall be parametrized by $(c+k\tanh(t),\frac{k}{\cosh(t)})$.

When working with hyperbolic geometry it is natural to work in a complex setting, namely we do a natural identification between H and $\mathbb{H}=\{x+iy\mid y>0\}$,

$$\begin{aligned}
f:\mathbb{H}&\to H\\
z&\to\Big(\frac{z+\bar z}{2},\frac{z-\bar z}{2i}\Big).
\end{aligned}$$

Doing the pullback $h=f^*g$ we obtain $h=-4\dfrac{dzd\bar z}{(z-\bar z)^2}$. We will only focus on the real part without further mention or changing the notation as that is where our tensor lies.

Let $a,b,c,d\in\mathbb{R}$, any complex map of the form,

$$f(z)=\frac{az+b}{cz+d},$$

where $ad-bc\neq 0$ is called a Möbius transformation; it is customary to extend the map to the Riemann sphere $\mathbb{C}\cup\{\infty\}$, by setting $x/0=\infty$. For the next arguments we could only ask for $ad-bc$ to be positive, however we will assume $ad-bc=1$ for simplicity. This is possible, as we can divide each entry in the expression above by $\sqrt{ad-bc}$. Notice

$$\begin{aligned}
2Im(f(z))=&f(z)-\overline{f(z)}\\
=&\frac{az+b}{cz+d}-\frac{a\bar z+b}{c\bar z+d}\\
=&\frac{z-\bar z}{|cz+d|^2}\\
=&2\frac{Im(z)}{|cz+d|^2}.
\end{aligned}$$

As such f maps $\mathbb{H}$ onto itself, moreover, it does so bijectively and allows us to extend the domain to $\partial\mathbb{H}=\mathbb{R}\cup\{\infty\}$ in the natural way. Finally notice

$$h(d_zf(u),d_zf(v))=-4\frac{f'(z)\overline{f'(z)}dz(u)d\bar z(v)}{(f(z)-\overline{f(z)})^2}$$

$$
\begin{aligned}
&= -4\frac{\dfrac{dz(u)d\bar{z}(v)}{|cz+d|^4}}{\left(\dfrac{z-\bar{z}}{|cz+d|^2}\right)^2} \\
&= -4\frac{dz(u)d\bar{z}(v)}{(z-\bar{z})^2} \\
&= h(u,v),
\end{aligned}
$$

that is, f is an isometry. In fact, all isometries preserving orientation are of this form, which leads to an identification between $\mathrm{Isom}_+(\mathbb{H})$ typically called Möbius transformations, $\mathrm{Möb}(\mathbb{H})$, and $PSL(2,\mathbb{R}) = SL(2,\mathbb{R})/\sim$ where $A \sim B \Leftrightarrow A = -B$. To obtain all isometries we just need to add some reflection on a geodesic.

Consider now points of the form ia and ib in $\mathbb{H}$, then the geodesic joining them is given by $\gamma(t) = ie^t$, $t \in [\log(a), \log(b)]$.

$$
\begin{aligned}
d(ia,ib) &= \int_{\log(a)}^{\log(b)} h(\gamma'(t),\gamma'(t))dt \\
&= \int_{\log(a)}^{\log(b)} -4\frac{ie^t(-ie^t)}{(ie^t+ie^t)^2}dt \\
&= \int_{\log(a)}^{\log(b)} 1dt \\
&= \log(b/a)
\end{aligned}
$$

To find the distance between two arbitrary points simply take an isometry that sends the geodesic containing them to the imaginary axis and use the above. Effectively if two points are in some vertical line $\mathrm{Re}(z) = a$ then use the transformation $f(z) = z - a$. If we consider a semi-circle with real end points $a < b$, then we use the map

$$f(z) = \frac{z-a}{z-b}$$

together with the easily verifiable expression, $|f(z) - f(w)| = |z-w||f'(z)|^{1/2}|f'(w)|^{1/2}$ to get

$$\cosh d(z,w) = 1 + \frac{|z-w|^2}{2\mathrm{Im}(z)\mathrm{Im}(w)}.$$

2.5.2 Poincaré Disk Model

Up until this point we've studied the upper half-plane model. We will now be interested in another model for the same geometry; or goal is to be able to visualize the entire space without any loss of information. To do so consider the complex open disk of radius 1, $\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$, and its usual boundary $\partial\mathbb{D} = \{z \in \mathbb{C} \mid |z| = 1\}$. One could do an independent study of hyperbolic geometry on the disk, however, it makes for a more concise presentation to consider the diffeomorphisms

$$
\begin{aligned}
i : \mathbb{D} &\to \mathbb{H} & \qquad i^{-1} : &\mathbb{H} \to \mathbb{D} \\
z &\to \frac{z+1}{iz-i} & z &\to \frac{iz+1}{iz-1}
\end{aligned}
$$

and transport all the information to $\mathbb{D}$ via the pullback i^*h, hence making these maps isometries. The pair $(\mathbb{D}, i^*h)$ is called the *Poincaré Disk Model*.

Geodesics in this model are the arcs of circles meeting $\partial\mathbb{D}$ orthogonally and its diameters. The isometries preserving orientation, Möb($\mathbb{D}$), are given by maps of the form

$$\begin{aligned} \varphi : \mathbb{D} &\to \mathbb{D} \\ z &\to \frac{\alpha z+\beta}{\bar{\beta}z+\bar{\alpha}} \end{aligned}$$

such that $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 - |\beta|^2 > 0$, which again we can normalize to be one. Finally the distance can be computed, for a point in the real axis

$$d(0,x) = \log\left(\frac{1+x}{1-x}\right)$$

and for generic points

$$d(z,w) = \log\left(\frac{|1-\bar{z}w|+|z-w|}{|1-\bar{z}w|-|z-w|}\right).$$

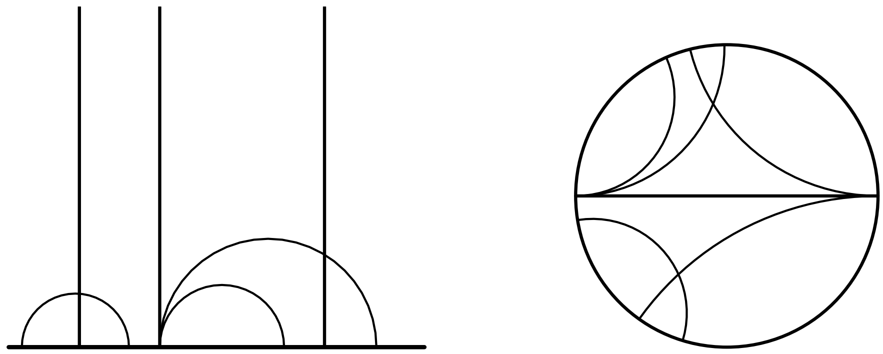

Figure 2.2: On the left we have the upper half plane and on the right we have the Poincaré Disk, both with their respective geodesics

Both the upper-half plane and the Poincaré disk models are simply connected spaces of curvature minus one, a third model of plane hyperbolic geometry was presented before, M_k^2. It can be proven that discrete subgroups of Möb($\mathbb{H}$), equivalently Möb($\mathbb{D}$), called *Fuchsian Groups*, act properly and freely by isometries, which, due to considerations from the previous section, gives us a way to look at Killing-Hopf theorem in the 2-dimensional case for constant negative curvature.

2.6 Symmetric Positive Definite Matrices

The space of symmetric positive definite matrices is a classical example in the study of symmetric spaces. All the results presented in this section are fairly known however we do a independent presentation using all the results obtained previously. We are specially interested in constructing the space above by studying its geodesics and curvature. With that in mind the section is divided in three subsections regarding each topic. The later two subsections rely heavily on calculations which unfortunately were unavoidable.

2.6.1 Construction of a Riemannian Structure

Denote by $Sym_+(n)$ the set of real symmetric positive definite $n \times n$ matrices. In a comprehensive way $Sym_+(n)$ consists of the real matrices such that $X = X^T$ and $\langle Xv, v\rangle$ is positive for all non-zero

$v \in M_{n\times 1}(\mathbb{R})$, where $\langle \cdot,\cdot \rangle$ denotes the euclidean product. Before starting with a proper construction for the Riemannian structure we present a first result in the direction of using Cartan-Hadamard results.

Theorem 2.40. *$Sym_+(n)$ is convex, hence simply connected.*

Proof. Let $X,Y \in Sym_+(n)$, $v \in \mathbb{R}^n \backslash \{0\}$ and $t \in [0,1]$, then

$$\begin{aligned}\big\langle (tX+(1-t)Y)v,v\big\rangle =&\big\langle tXv+(1-t)Yv,v\big\rangle \\ =&t\big\langle Xv,v\big\rangle +(1-t)\big\langle Yv,v\big\rangle \\ \geq&0.\end{aligned}$$

□

It is well known that $S \in Sym_+(n)$ if and only if it admits an orthonormal basis of eigenvectors with corresponding real positive eigenvalues. One usually writes this as $S = QDQ^T$ where $Q \in O(n)$ and $D \in Diag(n,\mathbb{R}^+)$. The following proposition will give us another way to express a matrix in $Sym_+(n)$.

Proposition 2.41. *A matrix S is in $Sym_+(n)$ if and only if it is of the form AA^T for some $A \in GL(n,\mathbb{R})$.*

Proof. Let $S \in Sym_+(n)$, then $S = QDQ^T$ in the conditions above. Consider the diagonal matrix $D^{1/2}$ whose entries are the square roots of the entries of D. We get that $S = QD^{1/2}(QD^{1/2})^T$. Conversely, a matrix of the form AA^T, with A invertible, is symmetric so it admits orthonormal basis of eigenvectors with corresponding real eigenvalues which remain to be proven positive. Now $QDQ^T = AA^T$, implies $D = Q^TA(Q^TA)^T$ which yields that all entries of D can't be negative. They can't be zero since A is invertible. □

Now let $A \in GL(n,\mathbb{R})$ and $S \in Sym_+(n)$, as $S = BB^T$ for some $B \in GL(n,\mathbb{R})$, we have that $ASA^T = AB(AB)^T$ which is still in $Sym_+(n)$. This tells us that the action

$$\begin{aligned}* : GL(n,\mathbb{R}) \times Sym_+(n) &\to Sym_+(n)\\ (A,S) &\to A * S := ASA^T\end{aligned}$$

is well defined.

Proposition 2.42. *There is a unique smooth structure on $Sym_+(n,\mathbb{R})$ that makes the action $*$ smooth.*

Proof. By theorem 2.39 we need only to show that the action is transitive and some stabilizer is closed. Let $S \in Sym_+(n)$, consider the invertible matrix $S^{1/2} = QD^{1/2}Q^T$. We have that $(S^{1/2})^{-1} * S = I$ so the action is transitive. The stabilizer of the identity is clearly given by the set of orthogonal matrices $O(n)$ which is closed in $GL(n,\mathbb{R})$, so we get our result. □

We could have obtained a smooth structure on $Sym_+(n)$ in a simpler way by noticing that it is an open set of $Sym(n)$; however, this action will now allow us to endow $Sym_+(n)$ with a Riemannian Structure and describe its isometries.

For the following computations start by noting that the map

$$\begin{aligned}\varphi_A : Sym_+(n) &\to Sym_+(n)\\ S &\to A * S\end{aligned}$$

for $A \in GL(n,\mathbb{R})$ is linear, meaning we can identify its differential at any point with itself,

$$\begin{aligned}\varphi_A : Sym(n) &\to Sym(n)\\ S &\to A * S\end{aligned}$$

having in mind $Sym(n)$ are identifications for the tangent spaces of $Sym_+(n)$.

Next consider the Frobenius inner product $\langle A,B\rangle_I = tr(A^T B) = tr(AB)$ on $Sym(n) \simeq T_I Sym_+(n)$. Let $A \in Sym_+(n)$, we have $A = A^{1/2} * I$, so we can transport the Frobenius inner product across the manifold by defining directly on $T_A Sym_+(n)$ the inner product

$$\begin{aligned}\langle X,Y\rangle_A &:= \langle A^{-1/2} * X, A^{-1/2} * Y\rangle_I =\\ &= \mathrm{tr}(A^{-1/2} X A^{-1/2} A^{-1/2} Y A^{-1/2}) =\\ &= \mathrm{tr}(A^{-1} X A^{-1} Y)\end{aligned}$$

Proposition 2.43. *The inner product $\langle\cdot,\cdot\rangle$ defined above gives a Riemannian structure for $Sym_+(n)$, moreover, φ_A and the map given by $i(S) = S^{-1}$ are isometries of $(Sym_+(n), \langle\cdot,\cdot\rangle)$.*

Proof. Note first that $\langle\cdot,\cdot\rangle_I$ is invariant by $O(n)$, for $M \in O(n)$, we have, for all $X,Y \in Sym(n)$

$$\begin{aligned}\langle M * X, M * Y\rangle_I &= \mathrm{tr}((MXM^T)^T MYM^T)\\ &= \mathrm{tr}(MXM^T MYM^T)\\ &= \mathrm{tr}(MXYM^T)\\ &= \mathrm{tr}(M^T MXY)\\ &= \mathrm{tr}(XY) =\\ &= \langle X,Y\rangle_I\end{aligned}$$

Suppose there is some $C \in GL(n,\mathbb{R})$, other than $A^{1/2}$ such that $C * I = A$, then $C^{-1}A^{\frac{1}{2}} \in O(n)$ and, since $\langle\cdot,\cdot\rangle_I$ is invariant by $O(n)$, we have that $\langle\cdot,\cdot\rangle$ is independent of the choice of transport, that is, it is well defined. Smoothness also follows directly as the trace is linear and $A^{-1}XA^{-1}Y$ depends smoothly on A.

It remains to show the statement about the isometries. For φ_A

$$\begin{aligned}{\varphi_A}^* \langle X,Y\rangle_B &= \langle A * X, A * Y\rangle_{A*B}\\ &= \langle AXA^T, AYA^T\rangle_{ABA^T}\\ &= \mathrm{tr}((ABA^T)^{-1} AXA^T (ABA^T)^{-1} AYA^T)\\ &= \mathrm{tr}((A^T)^{-1} B^{-1} A^{-1} AXA^T (A^T)^{-1} B^{-1} A^{-1} AYA^T)\\ &= \mathrm{tr}(B^{-1} X B^{-1} Y)\\ &= \langle X,Y\rangle_B\end{aligned}$$

The differential of i at a matrix $B \in Sym_+(n)$ is the map

$$\begin{aligned}d_B i : Sym(n) &\to Sym(n)\\ X &\to -B^{-1} X B^{-1}\end{aligned}$$

This follows from the fact that $I = id(A)i(A)$. Differentiating both sides at B we have, for every $X \in$

$T_BSym_+(n)$

$$0 = d_Bid(X)i(B) + id(B)d_Bi(X),$$

which can be rewritten as $d_Bi(X) = -i(B)d_Bid(X)i(B) = -B^{-1}XB^{-1}$.

Going back to the problem at hand, a straightforward computation gives

$$\begin{aligned} i^*\langle X,Y\rangle_B &= \langle d_Bi(X), d_Bi(Y)\rangle_{i(B)} \\ &= \langle B^{-1}XB^{-1}, B^{-1}YB^{-1}\rangle_{B^{-1}} \\ &= tr(BB^{-1}XB^{-1}BB^{-1}YB^{-1}) \\ &= tr(B^{-1}XB^{-1}Y) \\ &= \langle X,Y\rangle_B. \end{aligned}$$

□

Remark. This construction can be done whenever we have a transitive action of a Lie Group G with some compact stabilizer H. The homogeneous space G/H is a Riemannian Manifold with the action of G describing its isometries. Compactness of H ensures the existence of an invariant inner product so that the argument above can be used.

2.6.2 Connection, Geodesics and Metric

The first step into analysing the geometry of $Sym_+(n)$ is to find the Levi-Civita connection so we can then proceed to find the geodesics and geodesic distance. To do so, use the Koszul Formula; for $X,Y,Z \in \mathfrak{X}(Sym_+(n))$, and to ease notation write $X_A = X$, $Y_A = Y$ and $Z_A = Z$,

$$\begin{aligned} 2\langle \nabla_X Y, Z\rangle = &X\cdot\langle Y,Z\rangle + Y\cdot\langle X,Z\rangle - Z\cdot\langle X,Y\rangle \\ &- \langle [X,Z],Y\rangle - \langle [Y,Z],X\rangle + \langle [X,Y],Z\rangle. \end{aligned}$$

Denoting by $\cdot$ the derivation associated with a vector field,

$$\begin{aligned} X\cdot\langle Y,Z\rangle_A =&\mathrm{tr}(-A^{-1}XA^{-1}YA^{-1}Z \,+\, A^{-1}X\cdot YA^{-1}Z \\ &-A^{-1}YA^{-1}XA^{-1}Z \,+\, A^{-1}YA^{-1}X\cdot Z) \\ Y\cdot\langle X,Z\rangle_A =&\mathrm{tr}(-A^{-1}YA^{-1}XA^{-1}Z \,+\, A^{-1}Y\cdot XA^{-1}Z \\ &-A^{-1}XA^{-1}YA^{-1}Z \,+\, A^{-1}XA^{-1}Y\cdot Z) \\ -Z\cdot\langle X,Y\rangle_A =&\mathrm{tr}(A^{-1}ZA^{-1}XA^{-1}Y \,-\, A^{-1}Z\cdot XA^{-1}Y \\ &+A^{-1}XA^{-1}ZA^{-1}Y \,-\, A^{-1}XA^{-1}Z\cdot Y) \end{aligned}$$

and

$$\begin{aligned} -\langle [X,Z],Y\rangle_A &= \mathrm{tr}(-A^{-1}X\cdot ZA^{-1}Y \,+\, A^{-1}Z\cdot XA^{-1}Y) \\ -\langle [Y,Z],X\rangle_A &= \mathrm{tr}(-A^{-1}Y\cdot ZA^{-1}X \,+\, A^{-1}Z\cdot YA^{-1}X) \\ \langle [X,Y],Z\rangle_A &= \mathrm{tr}(A^{-1}X\cdot YA^{-1}Z \,-\, A^{-1}Y\cdot XA^{-1}Z). \end{aligned}$$

Putting these together we get

$$\begin{aligned}\langle \nabla_X Y, Z\rangle_A =& \operatorname{tr}(A^{-1}X\cdot YA^{-1}Z - \frac{1}{2}A^{-1}XA^{-1}YA^{-1}Z \\ & - \frac{1}{2}A^{-1}YA^{-1}XA^{-1}Z) \\ =& \operatorname{tr}\Big(A^{-1}\big(X\cdot Y - \frac{XA^{-1}Y + YA^{-1}X}{2}\big)A^{-1}Z\Big).\end{aligned}$$

Finally, the nondegeneracy of $\langle\cdot,\cdot\rangle$ gives

$$(\nabla_X Y)_A = X\cdot Y - \frac{XA^{-1}Y + YA^{-1}X}{2}.$$

To study the geodesics of our space start by considering the problem of joining I to some other matrix A in $Sym_+(n)$ by a geodesic which we claim to be $\gamma(t) = A^t$, $t\in[0,1]$. Note that, since A is symmetric positive definite, $\log(A)$ is well defined. Proceeding with the calculation one has

$$\begin{aligned}\nabla_{\gamma(t)}\gamma'(t) &= \gamma''(t) - \gamma'(t)\gamma(t)^{-1}\gamma'(t) \\ &= A^t\log(A)^2 - A^t\log(A)A^{-t}A^t\log(A) \\ &= A^t\log(A)^2 - A^t\log(A)^2 = 0.\end{aligned}$$

For the general case joining P to Q in $Sym_+(n)$ simply take the isometry that sends P to I, that is, $\varphi_{P^{-1/2}}$, note that it sends Q to $P^{-1/2}QP^{-1/2}$. As such, the curve $\gamma(t) = P^{1/2}(P^{-1/2}QP^{-1/2})^tP^{1/2}$, $t\in[0,1]$ is a geodesic joining P to Q. We shall prove later that there are no other geodesics.

As $\gamma(t)$ is a geodesic for the Levi-Civita connection we have that $|\gamma'(t)|$ is constant. This facilitates the calculations associated with the distance.

$$\begin{aligned}d(P,Q) &= \int_0^1 |\gamma'(t)|dt = \\ &= |\gamma'(0)| = \\ &= |\frac{d}{dt}\Big|_{t=0}(P^{1/2}(P^{-1/2}QP^{-1/2})^tP^{1/2})| \\ &= |(P^{1/2}\log(P^{-1/2}QP^{-1/2})P^{1/2})| \\ &= \operatorname{tr}^{1/2}(\log^2(P^{-1/2}QP^{-1/2})) \\ &= \left[\sum_{i=1}^n \log^2(\lambda_i(P^{-1}Q))\right]^{1/2}, \qquad (2.1)\end{aligned}$$

where $\lambda_i(P^{-1}Q)$ are the eigenvalues of $P^{-1}Q$. From these calculations we can also assert that geodesics parametrized by arclength joining I to A are given by

$$\gamma(t) = A^{t/\operatorname{tr}^{1/2}(\log^2(A))}.$$

It is customary to write the geodesics emanating from I as $\gamma(t) = e^{tS}$, where S is a symmetric matrix with $\operatorname{tr}(S^2) = 1$.

2.6.3 Curvature

The next step in the study of our space is giving an expression for the curvature and sectional curvature of $Sym_+(n)$. Let $X,Y,Z \in \mathfrak{X}(Sym_+(n))$ and $A \in Sym_+(n)$, remember

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

and again write $X_A = X$, $Y_A = Y$ and $Z_A = Z$,

$$\begin{aligned}
\nabla_Y \nabla_X Z =& \nabla_Y \Big(X \cdot Z - \frac{XA^{-1}Z + ZA^{-1}X}{2} \Big) \\
=& Y \cdot \Big(X \cdot Z - \frac{XA^{-1}Z + ZA^{-1}X}{2} \Big) \\
& - \frac{1}{2} YA^{-1} \Big(X \cdot Z - \frac{XA^{-1}Z + ZA^{-1}X}{2} \Big) \\
& - \frac{1}{2} \Big(X \cdot Z - \frac{XA^{-1}Z + ZA^{-1}X}{2} \Big) A^{-1} Y \\
=& Y \cdot X \cdot Z - \frac{Y \cdot (XA^{-1}Z) + Y \cdot (ZA^{-1}X) + YA^{-1}X \cdot Z + X \cdot ZA^{-1}Y}{2} \\
& + \frac{YA^{-1}XA^{-1}Z + YA^{-1}ZA^{-1}X + XA^{-1}ZA^{-1}Y + ZA^{-1}XA^{-1}Y}{4} \\
=& Y \cdot X \cdot Z - \frac{Y \cdot XA^{-1}Z - XA^{-1}YA^{-1}Z + XA^{-1}Y \cdot Z + Y \cdot ZA^{-1}X}{2} \\
& - \frac{-YA^{-1}ZA^{-1}X + YA^{-1}Z \cdot X + YA^{-1}X \cdot Z + X \cdot ZA^{-1}Y}{2} \\
& + \frac{YA^{-1}XA^{-1}Z + YA^{-1}ZA^{-1}X + XA^{-1}ZA^{-1}Y + ZA^{-1}XA^{-1}Y}{4}.
\end{aligned}$$

Analogously

$$\begin{aligned}
-\nabla_X \nabla_Y Z =& -X \cdot Y \cdot Z + \frac{X \cdot YA^{-1}Z - YA^{-1}XA^{-1}Z + YA^{-1}X \cdot Z + X \cdot ZA^{-1}Y}{2} \\
& + \frac{-XA^{-1}ZA^{-1}Y + XA^{-1}Z \cdot Y + XA^{-1}Y \cdot Z + Y \cdot ZA^{-1}X}{2} \\
& - \frac{XA^{-1}YA^{-1}Z + XA^{-1}ZA^{-1}Y + YA^{-1}ZA^{-1}X + ZA^{-1}YA^{-1}X}{4}.
\end{aligned}$$

Finally

$$\begin{aligned}
-\nabla_{[X,Y]} Z =& -[X,Y]Z + \frac{[X,Y]A^{-1}Z + ZA^{-1}[X,Y]}{2} = \\
=& -X \cdot Y \cdot Z + Y \cdot X \cdot Z + \frac{X \cdot YA^{-1}Z - Y \cdot XA^{-1}Z + ZA^{-1}X \cdot Y - ZA^{-1}Y \cdot X}{2},
\end{aligned}$$

and, after tedious calculations, we end up with a neat expression for the curvature,

$$2R(X,Y)Z = XA^{-1}YA^{-1}Z + ZA^{-1}YA^{-1}X - YA^{-1}XA^{-1}Z - ZA^{-1}XA^{-1}Y.$$

For X and Y, orthonormal vectors in $T_A Sym_+(n)$ spanning some 2-dimensional subspace Π,

$$K(\Pi) = \langle R(X,Y)X, Y \rangle$$

$$
\begin{aligned}
=&\frac{1}{2}\Big\langle (XA^{-1}YA^{-1}X+XA^{-1}YA^{-1}X-YA^{-1}XA^{-1}X-XA^{-1}XA^{-1}Y),Y\Big\rangle\\
=&\frac{1}{2}tr(A^{-1}XA^{-1}YA^{-1}XA^{-1}Y)+\frac{1}{2}tr(A^{-1}XA^{-1}YA^{-1}XA^{-1}Y)-\\
&-\frac{1}{2}tr(A^{-1}YA^{-1}XA^{-1}XA^{-1}Y)-\frac{1}{2}tr(A^{-1}XA^{-1}XA^{-1}YA^{-1}Y)\\
=&\mathrm{tr}\big((A^{-1/2}XA^{-1/2}A^{-1/2}YA^{-1/2})^2\big)-\mathrm{tr}\big((A^{-1/2}XA^{-1/2})^2(A^{-1/2}YA^{-1/2})^2\big)
\end{aligned}
$$

Simplifying the expression using the properties of the trace we finally arrive at

$$K(\Pi)=\mathrm{tr}\big((A^{-1}XA^{-1}Y)^2\big)-\mathrm{tr}\big((A^{-1}X)^2(A^{-1}Y)^2\big).$$

Theorem 2.44. *All sectional curvatures of $Sym_+(n)$ are nonpositive, hence it is a CAT(0) space.*

Proof. All that is left to show is

$$\mathrm{tr}\big((A^{-1}XA^{-1}Y)^2\big)-\mathrm{tr}\big((A^{-1}X)^2(A^{-1}Y)^2\big)\leq 0.$$

Start by rewriting the curvature formula obtaining

$$\mathrm{tr}\big((A^{-1/2}XA^{-1}YA^{-1/2})^2\big)-\mathrm{tr}\big((A^{-1/2}XA^{-1/2})^2(A^{-1/2}YA^{-1/2})^2\big)\leq 0.$$

Applying the change $P=A^{-1/2}XA^{-1/2}$ and $Q=A^{-1/2}YA^{-1/2}$ we have that both P and Q are symmetric matrices and we soften our notation to

$$\mathrm{tr}\big((PQ)^2\big)-\mathrm{tr}(P^2Q^2)\leq 0.$$

Consider the matrix $(PQ-QP)^2/2$, we can readily check that

$$\mathrm{tr}\Big(\frac{(PQ-QP)^2}{2}\Big)=\mathrm{tr}\big((PQ)^2\big)-\mathrm{tr}(P^2Q^2).$$

However $X=PQ-QP$ is antisymmetric so $tr(X^2)=tr(-X^TX)=-tr(X^TX)\leq 0$ which concludes our proof. □

The uniqueness of geodesics between any two points now also follows from the corollary to Cartan-Hadamard Theorem. This concludes our treatment of the space of symmetric positive definite matrices.

2.7 Horofunctions

Suppose X is a non-compact complete proper metric space. Our goal in this section is to give a compactification for X based on how geodesic rays behave towards "infinity". Such a compactification will have nice properties and geometric interpretation for CAT(0) spaces. Horofunctions can be introduced in various ways, however the core idea remains the same; geodesics displaying the same asymptotic behaviour are identified, that is, two geodesic rays $c,c':[0,+\infty[\to X$ are equivalent if $d(c(t),c'(t))\leq k$ for all $t\geq 0$.

The geometrical intuition given before is of major importance, we will however take a different approach to horofunctions so we can use powerful results to lessen the hardships along the road. To do so we will closely follow [11] and [3], focusing mainly on the first. Fix an "origin" x_0 in X and define, for every x in X the continuous map

$$\begin{aligned} D_x : X &\to \mathbb{R} \\ z &\to d(z,x) - d(z,x_0), \end{aligned}$$

and the map

$$\begin{aligned} i : X &\to C(X) \\ x &\to D_x \end{aligned}$$

where $C(X)$ denotes the space of continuous maps from X to $\mathbb{R}$.

Let us start by noticing that for every $x \in X$, $D_x(x_0) = 0$, that is, $i(X)$ is contained in the set of continuous functions such that $f(x_0) = 0$, which is homeomorphic to $C(X)/\mathbb{R}$, where functions differing by a constant are identified. This gives us a resemblance to the asymptotic geodesics.

Notice also that

$$\begin{aligned} d(D_x(y), D_x(z)) &= |d(y,x) - d(x_0,x) - d(z,x) + d(x_0,x)| \\ &= |d(y,x) - d(z,x)| \\ &\leq d(y,z), \end{aligned}$$

that is, D_x is 1-Lipschitz.

Definition 2.45. Let $\mathfrak{F} = \{f_\alpha : X \to Y\}_{\alpha \in A}$ be a family of continuous maps between metric spaces. Then $\mathfrak{F}$ is said to be:

1. *pointwise precompact* if for every x in X the set $\{f_\alpha(x) : \alpha \in A\}$ is compact;
2. *equicontinuous* if for every x in X and $\delta > 0$ there is some positive ε such that for all α in A, whenever $d_X(x,y) < \varepsilon$ one has $d_Y(f_\alpha(x), f_\alpha(y) < \delta$;
3. *uniformly equicontinuous* if for every $\delta > 0$ there is some positive ε such that for all α in A, for all $x, y \in X$ verifying $d_X(x,y) < \varepsilon$, one has $d_Y(f_\alpha(x), f_\alpha(y)) < \delta$;
4. *uniformly bounded* if there is some $x_0 \in X$ and M some constant such that for every $x \in X$ and $\alpha \in A$ one has $d_x(x_0, f_\alpha(x)) < M$.

Theorem 2.46 (Ascoli-Arzelà). *Let X be a compact metric space, Y a metric space and $\mathfrak{F} = \{f_\alpha : X \to Y\}_{\alpha \in A}$ a family of continuous functions. Then the following are equivalent:*

1. *$\{f_a : a \in A\}$ is precompact in $C(X,Y)$;*
2. *$\mathfrak{F}$ is pointwise precompact and equicontinuous;*
3. *$\mathfrak{F}$ is pointwise precompact and uniformly equicontinuous;*

As each D_x is 1-Lipschitz we have that $\{D_x : x \in X\}$ is equicontinuous and uniformly bounded on compact sets of X. Since X is proper, bounded closed sets are compact, whence by Ascoli-Arzelà every sequence $\{D_{x_n}\}$ has a subsequence which converges uniformly on compact sets. Consider $\hat{X} = \overline{i(X)}$ with the topology of uniform convergence on compact sets. We have seen so far that $\hat{X}$ is compact.

Proposition 2.47. *The space of continuous functions $C(X)$ with the topology of uniform convergence on compact sets is metrizable.*

Proof. Consider the following map

$$\begin{aligned} d: C(X) \times C(X) &\to \mathbb{R} \\ (f,g) &\to \sum_{i=1}^{+\infty} \frac{1}{2^i} \arctan\left(\max_{x \in B_i(x_0)} |f(x) - g(x)| \right). \end{aligned}$$

Since arctan is bounded it is well defined. The fact that d is positive definite and symmetric are trivial while the triangle inequality follows directly from

$$\arctan(x+y) \leq \arctan(x) + \arctan(y)$$

whenever x and y are positive numbers, whence d is a metric.

Let $\{f_\alpha\}_{\alpha \in A}$ be a net converging uniformly on compact sets to some f, that is, for every compact set K and $\delta > 0$, there is some $\beta \in A$ such that for every $\gamma \succ \beta$ we have, for every x in K, $|f_\alpha(x) - f(x)| < \delta$. Let $\varepsilon > 0$ and choose $\delta = \tan(\varepsilon)$ then we have,

$$\begin{aligned} d(f_\gamma, f) &= \sum_{i=1}^{+\infty} \frac{1}{2^i} \arctan\left(\max_{x \in B_i(x_0)} |f_\gamma(x) - f(x)| \right) \\ &\leq \sum_{i=1}^{+\infty} \frac{1}{2^i} \arctan\left(\tan(\varepsilon)\right) \\ &= \sum_{i=1}^{+\infty} \frac{\varepsilon}{2^i} \\ &= \varepsilon \end{aligned}$$

Therefore uniform convergence on compact sets implies convergence on the metric.

For the converse, let $\{f_n\}_{n \in \mathbb{N}}$ be a sequence converging to some f in $(C(X), d)$ and suppose there is a compact set $K \subset B_m(x_0)$ such that for some $x \in K$ one can take some positive ε making it so for every $p \in \mathbb{N}$ there is a q greater than p for which $d(f_q, f) > \varepsilon$. Then

$$\begin{aligned} d(f_q, f) &= \sum_{i=1}^{+\infty} \frac{1}{2^i} \arctan\left(\max_{x \in B_i(x_0)} |f_q(x) - f(x)| \right) \\ &\geq \frac{1}{2^m} \arctan(\varepsilon) \\ &> 0. \end{aligned}$$

Hence reaching an absurd, so $\{f_n\}_{n \in \mathbb{N}}$ converges to f uniformly on compact sets. □

Proposition 2.48. *Let X be a proper geodesic space, then $i: X \to i(X)$ is a homeomorphism.*

Proof. Start by noticing that given $z \in X$ one has

$$D_x(z) = d(x,z) - d(x,x_0) \geq -d(x_0,x) = D_x(x)$$

with equality only for $z = x$, that is, x is the only point minimizing D_x. Clearly this implies injectivity.

For continuity, let $x_n \xrightarrow[n\to+\infty]{} x$, then

$$\begin{aligned}|D_{x_n}(z) - D_x(z)| &= |d(x_n,z) - d(x_n,x_0) - d(x,z) + d(x,x_0)|\\ &\le |d(x_n,z) - d(x,z)| + |d(x_n,x_0) - d(x,x_0)|\\ &\le 2d(x_n,x)\end{aligned}$$

which goes to zero as n grows.

For continuity of the inverse, i^{-1}, let $D_{x_n} \xrightarrow[n\to+\infty]{} D_x$ uniformly on compact sets. Suppose there is some subsequence n_k for which $d(x_{n_k},x) \ge \varepsilon$. Take y_{n_k} on the geodesic segment $[xx_{n_k}]$ such that $d(x,y_{n_k}) = {}^{\varepsilon}/_{2}$. Then

$$\begin{aligned}D_{x_{n_k}}(y_{n_k}) &= d(x_{n_k},y_{n_k}) - d(x_{n_k},x_0)\\ &= d(x,x_{n_k}) - d(x,y_{n_k}) - d(x_{n_k},x_0)\\ &= D_{x_{n_k}}(x) - {}^{\varepsilon}/_{2}.\end{aligned}$$

Form construction, y_{n_k} belong to some compact set, so there is some converging subsequence y_{n_l}. Taking it to the limit we get $D_x(y) = D_x(x) - {}^{\varepsilon}/_{2} < D_x(x)$ which is absurd as x is the minimum of D_x. $\square$

Definition 2.49. The elements in $\hat{X} \setminus i(X)$ are called *horofunctions*. Given some horofunction D the level sets $D^{-1}(r)$ are called *horospheres* and $D^{-1}(]-\infty,r])$ are called closed *horoballs*.

Given a geodesic ray $\gamma\colon [0,+\infty[\to X$ with $\gamma(0) = x_0$, note that, by triangle inequality, for $t < s$ we have

$$\begin{aligned}d(\gamma(t),z) - d(\gamma(t),x_0) &\ge d(\gamma(s),z) - d(\gamma(s),\gamma(t)) - d(\gamma(t),x_0)\\ &= d(\gamma(s),z) - (s-t) - t\\ &= d(\gamma(s),z) - s\\ &= d(\gamma(s),z) - d(\gamma(s),x_0)\end{aligned}$$

and $d(\gamma(t),z) - d(\gamma(t),x_0) \ge -d(x_0,z)$. This discussion gives that the following is indeed well defined.

Definition 2.50. For any geodesic ray $\gamma\colon [0,+\infty[\to X$ such that $\gamma(0) = x_0$, define the *Busemann function*, $B_\gamma : X \to \mathbb{R}$, associated with the geodesic ray γ as

$$B_\gamma(z;x_0) = \lim_{t\to+\infty} d(\gamma(t),z) - d(\gamma(t),x_0).$$

The pointwise existence of this monotone limit implies uniform convergence on compact sets, as such $B_\gamma \in \hat{X}$. Moreover,

$$\begin{aligned}B_\gamma(\gamma(s);x_0) &= \lim_{t\to+\infty} d(\gamma(t),\gamma(s)) - d(\gamma(t),x_0)\\ &= \lim_{t\to+\infty} |t-s| - t\\ &= -s\end{aligned}$$

which tends towards $-\infty$ as $s \to +\infty$, but elements in $i(X)$ attain a minimum somewhere so B_γ is a horofunction.

Busemann functions are a classical tool for the study of geodesics in spaces with non-positive curvature.

Proposition 2.51. *Let X be a proper geodesic CAT(0) space, then every horofunction is a Busemann function.*

Proof. [11] Let D be some horofunction and write it as $\lim_{k\to+\infty} D_{x_k}$. Notice x_k must leave every compact set, otherwise there would be some converging subsequence $x_{k_l} \to x$ which, due to continuity, implies that $D_{x_{k_l}} \to D_x$ some element which is not a horofunction.

The proof of this result follows a very clear guideline, first we must construct a "good" geodesic ray and then we prove that D is the Busemann function given by that geodesic.

Let γ_n denote the geodesic ray starting at x_0 and going through x_n. Since x_k leaves any compact set and D_{x_k} converges to D on compact sets, given some positive ε and t, choose $N \in \mathbb{N}$ large enough so that whenever $d(z,x_0) \leq t$ one has $d(x_k,x_0) > t$ and $|D_{x_k}(z) - D(z)| < \varepsilon$. Finally take the sequence $y_k = \gamma_k(t)$.

Consider some $m,n > N$ and the geodesic triangle $\Delta x_m y_n x_0$. Construct in M_0^2 its comparison triangle $\bar{\Delta}\bar{x}_m\bar{y}_n\bar{x}_0$ on which we take $\bar{y}_m$ the comparison point of y_m. We are now allowed to work with simple euclidean geometry; drop the height from $\bar{y}_n$ onto some point $\bar{z}$ in $[\bar{x}_m\bar{x}_0]$ and take $\bar{w}$ on the line going through $\bar{x}_m$ and $\bar{x}_0$ such that $d_0(\bar{x}_0,\bar{w}) = t$ and $d_0(\bar{y}_m,\bar{w}) = 2t$ (figure 2.3).

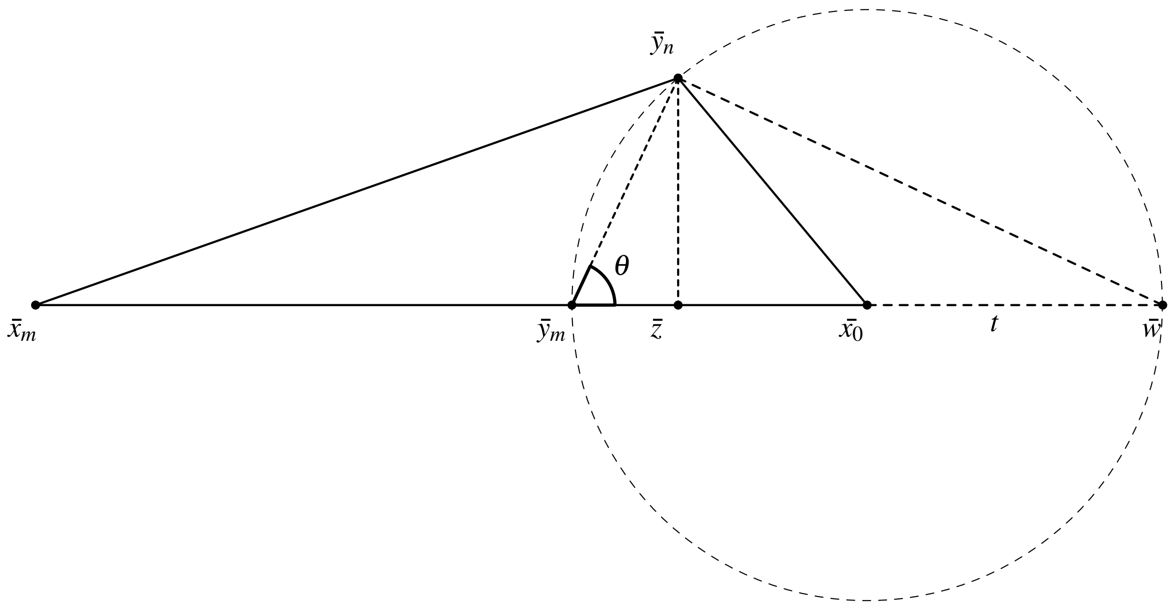

Figure 2.3: Geometrical construction to prove y_n is a Cauchy sequence.

Denoting $\theta = \angle\bar{y}_n\bar{y}_m\bar{x}_0$ and considering the right triangles we have

$$\begin{cases} \cos(\theta) = \dfrac{d_0(\bar{y}_m,\bar{z})}{d_0(\bar{y}_m,\bar{y}_n)} \\ \cos(\theta) = \dfrac{d_0(\bar{y}_m,\bar{y}_n)}{d_0(\bar{y}_m,\bar{w})} = \dfrac{d_0(\bar{y}_m,\bar{y}_n)}{2t} \end{cases}$$

The above yields $d(y_m,y_n) \leq d_0(\bar{y}_m,\bar{y}_n) = \sqrt{2td_0(\bar{y}_m,\bar{z})}$, so it remains to show that $d_0(\bar{y}_m,\bar{z})$ tends to zero. To do so, use CAT(0) and notice that $[\bar{x}_m\bar{y}_n]$ is the hypotenuse of $\bar{\Delta}\bar{x}_m\bar{y}_n\bar{z}$ so

$$d_0(\bar{y}_m,\bar{z}) = d_0(\bar{x}_m,\bar{z}) - d_0(\bar{x}_m,\bar{y}_m)$$

$$
\begin{aligned}
&\le d_0(\bar{x}_m,\bar{y}_n) - d_0(\bar{x}_m,\bar{y}_m)\\
&= d(x_m,y_n) - d(x_m,y_m)\\
&= d(x_m,y_n) - d(x_m,x_0) + d(x_m,x_0) - d(x_m,y_m)\\
&= d(x_m,y_n) - d(x_m,x_0) - t\\
&= d(x_m,y_n) - d(x_m,x_0) + d(x_n,x_0) - d(x_n,y_n)\\
&= D_{x_m}(y_n) - D_{x_n}(y_n)\\
&\le \sup_{d(y,x_0)\le t} |D_{x_m}(y) - D_{x_n}(y)|.
\end{aligned}
$$

Since D_{x_k} is a convergent sequence, it is a Cauchy sequence, so we can choose m,n large enough to make the above as small as we want. Therefore y_n is a Cauchy sequence, which, since X is complete, converges. Notice that our argument holds for a choice of t on compact subsets of $t \in [0,\infty[$ so define $\gamma(t) = \lim \gamma_n(t)$.

We must show γ is a geodesic. Let $t,s > 0$, given the convergence proven so far, there is p such that, for all $n > p$,

$$\max\{d(\gamma_n(t),\gamma(t)), d(\gamma_n(s),\gamma(s))\} < \frac{\varepsilon}{2}$$

Therefore

$$
\begin{aligned}
d(\gamma(t),\gamma(s)) =& d(\gamma(t),\gamma_n(t)) + d(\gamma_n(t),\gamma_n(s)) + d(\gamma_n(s),\gamma(s))\\
\le& \frac{\varepsilon}{2} + |s-t| + \frac{\varepsilon}{2}\\
=& |s-t| + \varepsilon.
\end{aligned}
$$

Repeating the calculations with $d(\gamma_n(t),\gamma_n(s))$, we obtain

$$|s-t| - \varepsilon \le d(\gamma(t),\gamma(s))$$

which, since ε is arbitrary, implies that γ is a geodesic.

It remains to show that $D(z) = B_\gamma(z;x_0)$, for that let $t_n = d(x_n,x_0)$ and consider

$$z_n = \begin{cases} \gamma(t_n) & \text{if n is odd} \\ x_n & \text{if n is even} \end{cases}$$

note that $D_{z_{2n}}(z)$ converges to $D(z)$ whilst $D_{z_{2n+1}}(z)$ goes towards $B_\gamma(z;x_0)$. To finish the proof we are going to prove $|D_{z_n} - D_{z_{n+1}}| \to 0$.

Let $\varepsilon > 0$, consider $r > \rho > 0$, and suppose $d(z,x_0) < \rho$ which later are chosen to be as large as needed for the argument, and take w_n the point on the segment $[z_n,x_0]$ at distance r from x_0.

$$
\begin{aligned}
|D_{z_n}(z) - D_{z_{n+1}}(z)| &= |d(z_n,z) - d(z_n,x_0) - d(z_{n+1},z) + d(z_{n+1},x_0)|\\
&= |d(z_n,z) - \big(d(z_n,w_n) + r\big) - d(z_{n+1},z) + \big(d(z_{n+1},w_{n+1}) + r\big)|\\
&= |d(z_n,z) - d(z_n,w_n) - d(z_{n+1},z) + d(z_{n+1},w_{n+1})|\\
&\le |d(z_n,z) - d(z_n,w_n) - d(w_n,z)|\\
&\qquad + |d(w_{n+1},z) + d(z_{n+1},w_{n+1}) - d(z_{n+1},z)|\\
&\qquad + |d(w_n,z) + d(w_{n+1},z)|.
\end{aligned}
$$

The last term tends to zero as $w_n \to \gamma(r)$. Construct $\bar{\Delta}\bar{z}_n\bar{x}_0\bar{z}$ in M_0^2 the comparison triangle for $\Delta z_n x_0 z$, $\bar{w}_n$ the comparison point of w_n and $\bar{z}^*$ the orthogonal projection of $\bar{z}$ onto the line going through $\bar{w}_n$ and x_0. Notice we are assuming that n is large enough so that $d(z_n, x_0) > r$ and $d_0(\bar{z}_n, \bar{z}^*) > r$ (see figure 2.4).

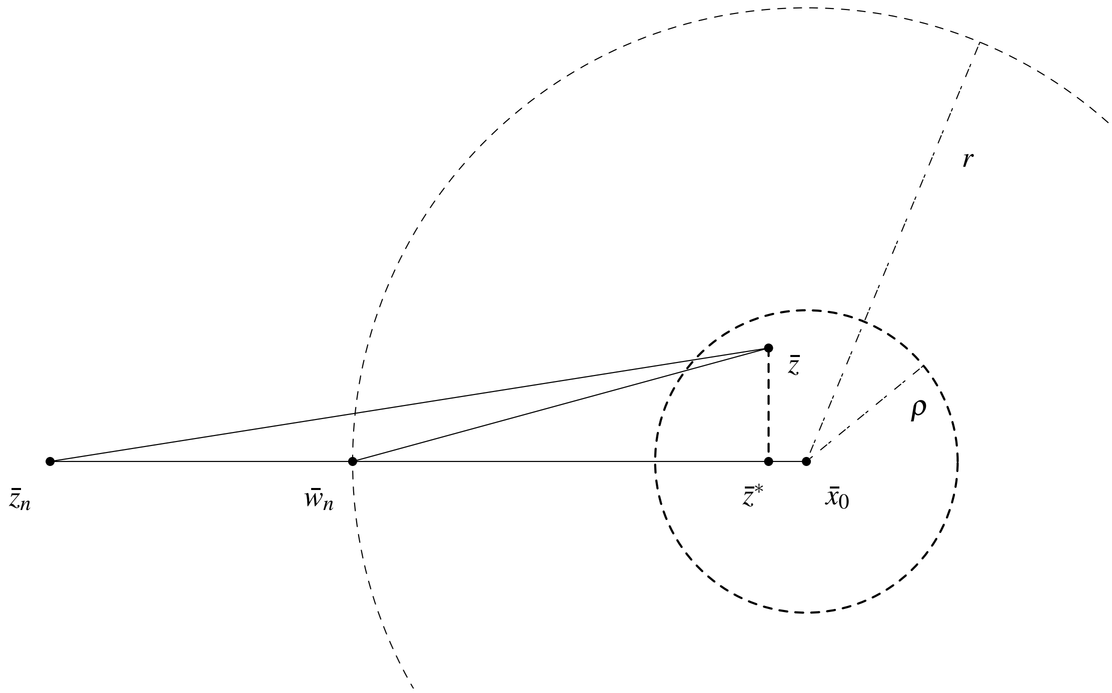

Figure 2.4: Geometrical construction to prove that every Horofunction is a Busemann function.

By the CAT(0) inequality, we have $d(w_n, z) \leq d_0(\bar{w}_n, \bar{z})$ which in turn implies that

$$\begin{aligned} d(w_n, z) + d(z_n, w_n) - d(z_n, z) &\leq d_0(\bar{w}_n, \bar{z}) + d_0(\bar{z}_n, \bar{w}_n) - d(\bar{z}_n, \bar{z}) \\ &= d_0(\bar{w}_n, \bar{z}) - d_0(\bar{w}_n, \bar{z}^*) + d_0(\bar{z}_n, \bar{z}^*) - d(\bar{z}_n, \bar{z}). \end{aligned}$$

The triangles $\bar{\Delta}\bar{z}_n\bar{z}^*\bar{z}$ and $\bar{\Delta}\bar{w}_n\bar{z}^*\bar{z}$ are right at $\bar{z}^*$, so, by Pythagoras theorem.

$$\begin{aligned} & d_0(\bar{w}_n, \bar{z})^2 = d_0(\bar{w}_n, \bar{z}^*)^2 + d_0(\bar{z}, \bar{z}^*)^2 \\ \Leftrightarrow\ & d_0(\bar{w}_n, \bar{z})^2 - d_0(\bar{w}_n, \bar{z}^*)^2 = d_0(\bar{z}, \bar{z}^*)^2 \\ \Leftrightarrow\ & \Big(d_0(\bar{w}_n, \bar{z}) - d_0(\bar{w}_n, \bar{z}^*)\Big)\Big(d_0(\bar{w}_n, \bar{z}) + d_0(\bar{w}_n, \bar{z}^*)\Big) = d_0(\bar{z}, \bar{z}^*)^2 \\ \Leftrightarrow\ & d_0(\bar{w}_n, \bar{z}) - d_0(\bar{w}_n, \bar{z}^*) = \frac{d_0(\bar{z}, \bar{z}^*)^2}{d_0(\bar{w}_n, \bar{z}) + d_0(\bar{w}_n, \bar{z}^*)} \leq \frac{\rho^2}{r}. \end{aligned}$$

Analogously $d_0(\bar{w}_n, \bar{z}) - d_0(\bar{w}_n, \bar{z}^*) \leq \frac{\rho^2}{r}$. Therefore

$$d(z_n, w_n) + d(w_n, z) - d(z_n, z) \leq \frac{2\rho^2}{r}.$$

The same inequality holds for $d(z_{n+1}, w_{n+1}) + d(w_{n+1}, z) - d(z_{n+1}, z)$. Choosing $r > 6\rho^2/\varepsilon$ we see that all the terms can be made smaller than $\varepsilon/3$, so we get our result. □

Consider an increasing sequence t_n, two geodesic rays γ_1, γ_2 giving rise to the same Busemann function and define

$$x_n = \begin{cases} \gamma_1(t_n) & \text{if n odd,} \\ \gamma_2(t_n) & \text{if n even.} \end{cases}$$

Both $D_{x_{2n}}$ and $D_{x_{2n+1}}$ converge to $B_{\gamma_1}(\cdot, x_0) = B_{\gamma_2}(\cdot, x_0)$ so $\lim D_{x_n}$ exists. The argument from previous proof shows that the segments $[x_0 x_n]$ converge to a single geodesic ray, so we must have $\gamma_1 = \gamma_2$, that is, Busemann functions are in fact uniquely determined by the geodesic ray, provided $\gamma_1(0) = \gamma_2(0) = x_0$.

Theorem 2.52. *Let X be a proper CAT(0) space and x_n a sequence such that $\lim d(x_n, x_0)/n \to s > 0$. Let D be a horofunction. Then $\frac{1}{n}D(x_n) \to -s$ if and only if $D = B_\gamma(\cdot, x_0)$ and $\frac{1}{n}d(x_n, \gamma(sn)) = 0$ for some geodesic ray γ starting at x_0.*

Proof. [11] Suppose we have the second. Since $B_\gamma(z, x_0)$ is the decreasing limit of $d(\gamma(sn), z) - sn$ we have $B_\gamma(z; x_0) \leq d(\gamma(sn), z) - sn$. However, we also have

$$\begin{aligned} d(\gamma(t), z) - t =& d(\gamma(t), z) - d(\gamma(t), x_0) \\ &\geq d(\gamma(t), z) - d(\gamma(t), z) - d(z, x_0) \\ &= -d(z, x_0) \end{aligned}$$

Thus

$$\frac{-d(x_n, x_0)}{n} \leq \frac{B_\gamma(x_n; x_0)}{n} \leq \frac{d(\gamma(sn), x_n) - sn}{n},$$

where both the left and right side tend to $-s$.

For the forward implication let γ be the geodesic ray defining $B_\gamma(\cdot; x_0) = D$ and $0 < n < t$ large enough for the arguments at hand. Consider also $\Delta x_n x_0 \gamma(st)$, $\bar{\Delta}\bar{x}_n \bar{x}_0 \bar{\gamma}(st)$ its comparison triangle in M_0^2 and $\bar{\gamma}(sn)$ the comparison point of $\bar{\gamma}(sn)$. Take w_t the point on the segment $[\bar{\gamma}(st)\bar{x}_0]$ such that $d_0(\bar{\gamma}(st), \bar{x}_n) = d_0(\bar{\gamma}(st), \bar{w}_t)$ and $\bar{z}$ the point in $[\bar{x}_0 \bar{\gamma}(st)]$ closest to $\bar{x}_n$ (see figure 2.5).

We will start by showing $\bar{z}$ is in fact an orthogonal projection onto the segment. Notice that if $\bar{z} = \bar{x}_0$, then $d_0(\bar{\gamma}(st), \bar{x}_n) \geq d_0(\bar{\gamma}(st), \bar{x}_0)$, so

$$\begin{aligned} B_\gamma(x_n; x_0) &= \lim_{t \to +\infty} d(\gamma(st), x_n) - d(\gamma(st), x_0) \\ &= \lim_{t \to +\infty} d_0(\bar{\gamma}(st), \bar{x}_n) - d_0(\bar{\gamma}(st), \bar{x}_0) \\ &\geq \lim_{t \to +\infty} d_0(\bar{\gamma}(st), \bar{x}_0) - d_0(\bar{\gamma}(st), \bar{x}_0) \\ &= 0, \end{aligned}$$

which is absurd by hypothesis. On the other hand if $\bar{z} = \bar{\gamma}(st)$ we can choose a bigger t so that $\bar{z}$ is the orthogonal projection of $\bar{x}_n$ onto the segment $[\bar{x}_0 \bar{\gamma}(st)]$.

Now $d_0(\bar{w}_t, \bar{\gamma}(sn)) = |d_0(\bar{\gamma}(st), \bar{w}_t) + d_0(\bar{\gamma}(sn), \bar{x}_0) - d_0(\bar{\gamma}(st), \bar{x}_0)| = |d(\gamma(st), w_t) + sn - st|$ which, as t tends to $+\infty$, converges to $D(x_n) + sn$, so, applying our hypothesis,

$$\lim \frac{d_0(\bar{w}_t, \overline{\gamma(sn)})}{n} = 0.$$

We have

$$\alpha_t = \angle \bar{x}_n \bar{\gamma}(st) \bar{z} = \sin^{-1}\Big(\frac{d_0(\bar{x}_n, \bar{z})}{d_0(\bar{\gamma}(st), \bar{x}_0)}\Big),$$

where the last tends to zero as t increases which means that so does $d_0(\bar{z}, \bar{w}_t) = d_0(\bar{x}_n, \bar{z})\tan(\alpha_t/2)$.

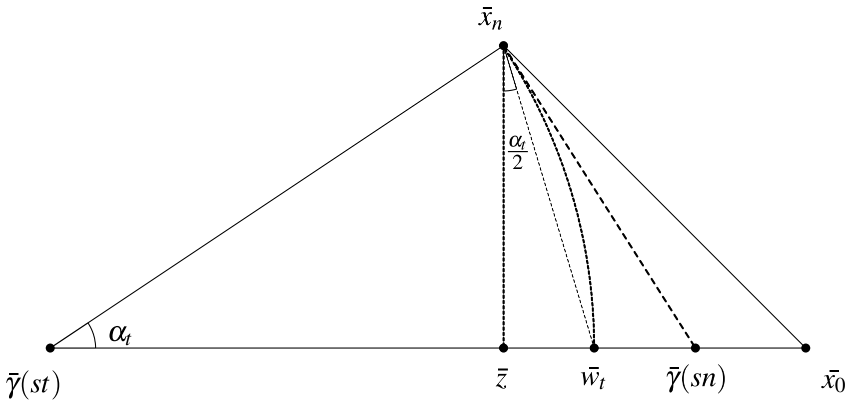

Figure 2.5: Geometrical construction for the proof of theorem 2.52

Consequently, there is some T_n for which whenever $t > T_n$

$$\lim_{n\to+\infty} \frac{d_0(\bar{z},\bar{w}_t)}{n} = 0.$$

Putting together what we have so far,

$$\lim_{n\to+\infty} \frac{1}{n} d_0(\bar{x}_0,\bar{z}) = \lim_{n\to+\infty} \frac{1}{n}\Big(d_0(\bar{x}_0,\bar{\gamma}(sn)) \pm d_0(\bar{w}_t,\bar{\gamma}(sn)) \pm d_0(\bar{z},\bar{w}_t)\Big) = s,$$

where the sign depend solely on the position of $\bar{\gamma}(st)$ relatively to $\bar{z}$ and $\bar{w}_t$.

By Pythagoras Theorem

$$\lim_{n\to+\infty} \frac{1}{n} d_0(\bar{x}_n,\bar{z}) = \lim_{n\to+\infty} \sqrt{\frac{d_0(\bar{x}_n,\bar{x}_0)^2 - d_0(\bar{z},\bar{x}_0)^2}{n^2}} = 0.$$

Finally for $t > T_n$, applying the CAT(0) property, we have

$$\begin{aligned} 0 \leq \frac{1}{n} d(x_n,\gamma(sn)) \leq & \frac{1}{n} d_0(\bar{x}_n,\bar{\gamma}(sn)) \\ = & \sqrt{\frac{d_0(\bar{x}_n,\bar{z})^2 + d_0(\bar{z},\bar{\gamma}(sn))^2}{n^2}} \xrightarrow[n\to+\infty]{} 0 \end{aligned}$$

□

Example 2.53. Let $\gamma : [0,+\infty[\to M_0^n$ be a geodesic ray, write it as $\gamma(t) = x_0 + tv$ where v is a vector in $\mathbb{S}^{n-1}$. Consider z the orthogonal projection of x onto the line $x_0 + tv$ with $t \in \mathbb{R}$ and $\{w_t\}_{t\geq 0}$ on the same line such that $d_0(\gamma(t),w_t) = d_0(\gamma(t),x)$. The following relation holds

$$\begin{aligned} 0 \leq \angle x\gamma(t)z = & \arctan\left(\frac{d(x,z)}{d(\gamma(t),z)}\right) \\ \leq & \arctan\left(\frac{d(x,z)}{d(\gamma(t),x_0) + d(x_0,z)}\right) \xrightarrow[t\to+\infty]{} 0. \end{aligned}$$

As such $d_0(w_t,z) = d(x,z)\tan(\frac{\angle x\gamma(y)z}{2})$, which tends to zero.

Therefore the associated Busemann function B_γ is given by

$$\begin{aligned}
B_\gamma(x;x_0) &= \lim_{t\to\infty} d(\gamma(t),x) - d(\gamma(t),x_0) \\
&= \lim_{t\to\infty} d(\gamma(t),w_t) - d(\gamma(t),x_0) \\
&= \lim_{t\to\infty} d(x_0,z) - d(w_t,z) \\
&= \pm\, d(x_0,z)
\end{aligned}$$

where the sign is given by whether the projection belongs to the geodesic ray or not, in a simpler notation, $B_\gamma(x;x_0) = \langle x_0 - x, v\rangle$. In this simple case the horospheres are precisely hyperplanes orthogonal to the line.

Example 2.54. Let us consider the Poincaré disk model $\mathbb{D}$ and suppose that our origin point x_0 is in fact the origin. Remember $c_{\mathbb{H}}(t) = ie^t$ is a geodesic for the upper half plane and transport it to the Poincaré disk via de isometries constructed in the hyperbolic geometry section, obtaining

$$c(t) = \frac{e^t - 1}{e^t + 1}.$$

Geometrically, complex multiplication is simply a rotation and some dilation or contraction, as such, given ξ in $\partial\mathbb{D}$, $\gamma(t) = \xi c(t)$, $t \in [0,+\infty[$ is the geodesic ray with origin in 0 and endpoint ξ. Then

$$\begin{aligned}
B_\gamma(z;0) &= \lim_{t\to+\infty} d(\gamma(t),z) - d(\gamma(t),0) \\
&= \lim_{t\to+\infty} d(\gamma(t),z) - t \\
&= \lim_{t\to+\infty} \log\Big(\frac{|1-\overline{\gamma(t)}z| + |z-\gamma(t)|}{|1-\overline{\gamma(t)}z| - |z-\gamma(t)|}\Big) - \log(e^t) \\
&= \lim_{t\to+\infty} \log\Big(\frac{1}{e^t}\frac{|1-\overline{\gamma(t)}z| + |z-\gamma(t)|}{|1-\overline{\gamma(t)}z| - |z-\gamma(t)|}\Big) \\
&= \lim_{t\to+\infty} \log\Big(\frac{1}{e^t}\frac{(|1-\overline{\gamma(t)}z| + |z-\gamma(t)|)^2}{|1-\overline{\gamma(t)}z|^2 - |z-\gamma(t)|^2}\Big) \\
&= \lim_{t\to+\infty} \log\Big(\frac{1}{e^t}\frac{(|1-\overline{\gamma(t)}z| + |z-\gamma(t)|)^2}{(1-\gamma(t)^2)(1-|z|^2)}\Big) \\
&= \lim_{t\to+\infty} \log\Big(\frac{1}{e^t}\frac{(|1-\overline{\gamma(t)}z| + |z-\gamma(t)|)^2}{4e^t(e^t+1)^{-2}(1-|z|^2)}\Big) \\
&= \log\Big(\frac{(|1-\bar{\xi}z| + |z-\xi|)^2}{4(1-|z|^2)}\Big) \\
&= \log\Big(\frac{|\xi - z|^2}{1-|z|^2}\Big).
\end{aligned}$$

Suppose we are now given some arbitrary x_0. Use the Möbius transformation

$$\begin{aligned}
f : \mathbb{D} &\to \mathbb{D} \\
z &\to \frac{z - x_0}{-\bar{x}_0 z + 1}
\end{aligned}$$

to transport x_0 to 0 and making any geodesic through x_0 of the form γ for some ξ. Making use of the

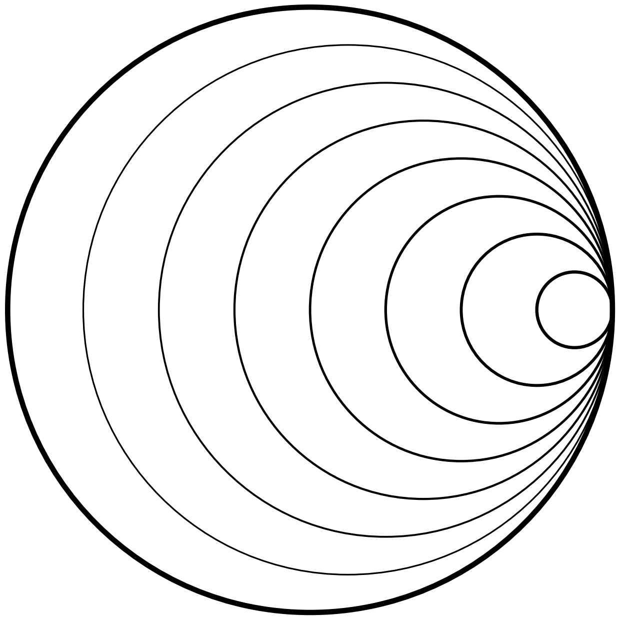

Figure 2.6: Horospheres on the Poincaré disk model

Busemann function we obtained in the case $x_0 = 0$ we have,

$$B_\gamma(z;x_0) = \log\Big(\frac{|\xi - z|^2}{1-|z|^2}\Big) - \log\Big(\frac{|\xi - x_0|^2}{1-|x_0|^2}\Big).$$

At the start of this section we gave emphasis to the geometrical ideas behind the horofunction compactification. Notice how in both examples the Busemann functions were determined by the asymptotic behaviour of the geodesics alone. For the euclidean space, asymptotic geodesic rays are exactly the parallel ones, whereas in the Poincaré disk the asymptotic behaviour of geodesics is decided by their endpoint alone. The last point we make is that since Busemann functions defined by asymptotic rays differ by a constant their associated horospheres are the same.

Next chapter we will shift our interest to the action of Isom(X). To go in that direction we finish with a result that may, on one hand, seem displaced, but on the other hand natural.

Proposition 2.55. *The action of Isom(X) on X extends continuously to an action of $\hat{X}$.*

Proof. The elements of $\hat{X}$ are either of the form D_x for some $x \in X$ or $\lim D_{x_n}$, for some $\{x_n\}_{n\geq 1} \subset X$. Start by defining $g \cdot D_x = D_{g(x)}$ for the first case. To extend it to the second, take $D = \lim D_{x_n}$ and extend the action by

$$\begin{aligned}
g \cdot D(z) =& \lim g \cdot D_{x_n}(z) \\
=& \lim D_{g(x_n)}(z) \\
=& \lim d(g(x_n), z) - d(g(x_n), x_0) \\
=& \lim d(x_n, g^{-1}(z)) - d(x_n, g^{-1}(x_0)) \\
=& \lim d(x_n, g^{-1}(z)) - d(x_n, x_0) + d(x_n, x_0) - d(x_n, g^{-1}(x_0)) \\
=& \lim D_{x_n}(g^{-1}(z)) - D_{x_n}(g^{-1}(x_0)) \\
=& D(g^{-1}(z)) - D(g^{-1}(x_0)).
\end{aligned}$$

Notice that the end result is independent of our choice of x_n, that is, the extension is well defined. Moreover, our construction was made exactly to make this extension continuous. $\square$

Chapter 3

The Noncommutative Ergodic Theorems

As we've asserted in the first chapter, one natural way to extend the average behaviour from the law of large numbers to the noncommutative case is by considering a random walk on a group. Another reason to do this extension is the fact every (locally compact Hausdorff) group acts by isometries on some metric space. The first to notice this geometric interpretation to the multiplicative ergodic theorem was Kaimanovich [6]. The main result we will focus on is an extension due to Anders Karlsson and Grigory Margulis [9] however we will start with a more elegant approach by Karlsson and François Ledrappier [7, 8].

So far we've denoted both probability and geodesic metric spaces by X as we were studying them independently. Now that we put them together we shall denote probability spaces by Ω and keep the X for geodesic spaces.

3.1 Karlsson-Ledrappier: The Proper Case

Keeping the notation from previous chapters, let (X,d) be a proper Polish space with a "marked point" point x_0 and G a topological group acting on X by isometries. Consider $(\Omega,\mathscr{B},\mu,T)$ some standard probability space , where $T:\Omega\to\Omega$ is ergodic with respect to μ and $g:\Omega\to G$ is a measurable map.

Consider the right random walk

$$Z_n(\omega):=g(\omega)g(T(\omega))\cdots g(T^{n-1}(\omega))$$

and the process $d(Z_n(\omega)\cdot x_0,x_0)$. The important fact to note is that this process is subadditive:

$$\begin{aligned} d(Z_{n+m}(\omega)\cdot x_0,x_0) \leq& d(Z_{n+m}(\omega)\cdot x_0,Z_n(\omega)\cdot x_0)+d(Z_n(\omega)\cdot x_0,x_0)\\ =& d(Z_m(T^n(\omega))\cdot x_0,x_0)+d(Z_n(\omega)\cdot x_0,x_0). \end{aligned}$$

Notice that the above is also true for semicontractions, which we will need next section. Assume $\int d(g(\omega)\cdot x_0,x_0)d\mu<\infty$, so the subadditive ergodic theorem implies that the *linear drift*

$$s=\lim_{n\to\infty}\frac{1}{n}d(Z_n(\omega)\cdot x_0,x_0)=\inf_{n\in\mathbb{N}}\frac{1}{n}\int_\Omega d(Z_n(\omega)\cdot x_0,x_0)d\mu(\omega)$$

exists for μ-a.e where the second equality is due to ergodicity. Whenever the assumption above is verified we say Z_n is *integrable*.

Before delving into the theorems it is important to assert the independence of the marked point. With that in mind, let x_1 be another point in X then

$$\begin{aligned} d(g(\omega)\cdot x_1, x_1) \leq& d(g(\omega)\cdot x_1, g(\omega)\cdot x_0) + d(g(\omega)\cdot x_0, x_0) + d(x_1, x_0) \\ =& 2d(x_0, x_1) + d(g(\omega)\cdot x_0, x_0). \end{aligned}$$

Hence $d(Z_n(\omega)x_1, x_1)$ verifies the integrability hypothesis. As for s, doing an analogous calculation

$$\lim_{n\to+\infty} \frac{d(Z_n(\omega)\cdot x_1, x_1)}{n} \leq \lim_{n\to+\infty} \frac{2d(x_0,x_1) + d(Z_n(\omega)\cdot x_0, x_0)}{n} = \lim_{n\to+\infty} \frac{d(Z_n(\omega)\cdot x_0, x_0)}{n}.$$

Since we can revert the inequality we have the equality.

Theorem 3.1 (Karlsson-Ledrappier)**.** *Let X be a proper metric space and Z_n an integrable right cocycle taking values in* $\mathrm{Isom}(X)$*. Then there is an almost everywhere defined mapping $\omega \to D_\omega = D$, where D is an horofunction, depending measurably on ω, such that*

$$\lim_{n\to\infty} -\frac{1}{n} D(Z_n(\omega)\cdot x_0) = s,$$

where s is defined above.

Proof. The proof of this theorem will be divided into multiple subsections with specific roles. On the first one we construct a dynamical system whose ergodic sums are $-D(Z_n(\omega)\cdot x_0)/n$, followed by creating a measure with some "good" properties and then completing the proof. During the proof there are two technicalities which we shall assume in the present manuscript, although giving reference to where one could find a discussion on the matter. In light of

$$\left|\frac{1}{n} D(Z_n(\omega)\cdot x_0)\right| \leq \frac{1}{n} d(Z_n(\omega)\cdot x_0, x_0)$$

we can focus the case $s > 0$ as the statement becomes trivial for $s = 0$.

Constructing a Dynamical System

Recall that proposition 2.55 tells us the action of G on X extends naturally to $\hat{X}$ by

$$g\cdot D(z) = D(g^{-1}\cdot z) - D(g^{-1}\cdot x_0).$$

For the elements in X we can see the action in both ways, as $g\cdot x = g(x)$ or $g\cdot D_x(z) = D_{g(x)}(z)$.

Define

$$\begin{aligned} F : \Omega \times \hat{X} &\to \mathbb{R} \\ (\omega, D) &\to -D(g(\omega)\cdot x_0). \end{aligned}$$

Since for every $x \in X$,

$$\begin{aligned} D_x(g(\omega)\cdot x_0) =& d(g(\omega)\cdot x_0, x) - d(x, x_0) \\ \geq& -d(x_0, g(\omega)\cdot x_0), \end{aligned}$$

with equality in the case $x = g(\omega)\cdot x_0$, we get

$$\max_{D\in\hat{X}} F(\omega,D) = d(x_0, g(\omega)\cdot x_0).$$

Consider also the skew-product

$$\begin{aligned} T_g : \Omega\times\hat{X} &\to \Omega\times\hat{X} \\ (\omega,D) &\to (T(\omega), g(\omega)^{-1}\cdot D) \end{aligned}$$

Using the fact $T_g^k(\omega,D) = (T^k(\omega), Z_k(\omega)^{-1}\cdot D)$ we can see

$$\begin{aligned} F_n(\omega,D) :&= \sum_{k=0}^{n-1} F(T_g^k(\omega,D)) \\ &= \sum_{k=0}^{n-1} F(T^k(\omega), Z_k(\omega)^{-1}\cdot D) \\ &= \sum_{k=0}^{n-1} -(Z_k(\omega)^{-1}\cdot D)(g(T^k(\omega))\cdot x_0) \\ &= \sum_{k=0}^{n-1} -D(Z_{k+1}(\omega)\cdot x_0) + D(Z_k(\omega)\cdot x_0) \\ &= -D(Z_n(\omega)\cdot x_0) + D(x_0) \\ &= -D(Z_n(\omega)\cdot x_0). \end{aligned}$$

From this equality and the discussion we've done for F we obtain $\max_{D\in\hat{X}} F_n(\omega,D) = d(x_0, Z_n(\omega)\cdot x_0)$, whence

$$s = \inf_{n\in\mathbb{N}} \frac{1}{n}\int_\Omega d(Z_n(\omega)\cdot x_0, x_0)d\mu(\omega) = \inf_{n\in\mathbb{N}} \frac{1}{n}\int_\Omega \max_{D\in\hat{X}} F_n(\omega,D).d\mu(\omega)$$

Constructing a Measure

Take $C(\hat{X})$ the Banach space of bounded continuous functions $f:\hat{X}\to\mathbb{R}$ with the supremum norm $||\cdot||_b$. Denote by $L^1(\Omega, C(\hat{X}))$ the space of measurable maps $f:\Omega\to C(\hat{X})$ for which

$$\int_\Omega \sup_{D\in\hat{X}} |f(\omega)(D)| d\mu(\omega) < \infty.$$

Finally, consider the space of probability measure ν on $\Omega\times\hat{X}$ such that $\nu(B\times\hat{X}) = \mu(B)$ for every measurable set B, in other words, ν projects onto μ. We shall denote this space by $\mathrm{Prob}_\mu(\Omega\times\hat{X})$.

For every $\nu\in\mathrm{Prob}_\mu(\Omega\times\hat{X})$ introduce the map

$$\begin{aligned} \nu : L^1(\Omega, C(\hat{X})) &\to \mathbb{R} \\ f &\to \int_{\Omega\times\hat{X}} f d\nu. \end{aligned}$$

We have that $f\to\nu(f)$ is linear and $|\nu(f)|\leq ||f||_b$. Therefore the dual of $\mathrm{Prob}_\mu(\Omega\times\hat{X})$ is contained in the dual of $L^1(\Omega, C(\hat{X}))$, on which we shall consider the weak topology coming from this duality, that is, the finest topology for which the linear functionals $f\to\nu(f)$ is continuous. The space $\mathrm{Prob}_\mu(\Omega\times\hat{X})$

is weakly sequentially compact, whence compact, as the given topology on $Prob_\mu(\Omega \times \hat{X})$ is metrizable. This is the first technicality we avoid; the details are mostly found in chapter 4 of [4].

Recall $Z_n(\omega) \cdot x_0 \in X \subset \hat{X}$. For each n, consider measures ν_n on $\hat{X}$ defined by the disintegration $\nu^n_\omega = \delta_{Z_n(\omega)\cdot x_0}$, that is,

a) ν^n_ω depends measurably on ω, meaning for every $A \in \mathscr{P}(\hat{X})$ the map $\omega \to \nu^n_\omega(A)$ is measurable;

b) ν^n_ω is a probability measure for μ-a.e. ω;

c) For every $A \in \mathscr{F} \otimes \mathscr{B}(\hat{X})$

$$\begin{aligned}\nu_n(A) &= \int_\Omega \int_{\hat{X}} 1_A(\omega, x) d\nu^n_\omega(x) d\mu(\omega) \\ &= \int_\Omega \delta_{Z_n(\omega)x_0}(A_\omega) d\mu(\omega)\end{aligned}$$

where $A_\omega = \{h \mid (\omega, h) \in A\}$, the usual section.

Let

$$\eta_n = \frac{1}{n}\sum_{i=0}^{n-1}(T^i_g)_*\nu_n$$

and take η to be a weak limit of these measures.

Completing the Proof

Let's start by putting together the dynamical system and the measure we constructed. We suppose η_n is already a converging sequence, otherwise just pass to the converging sublimit. First notice that being a weak limit η still projects onto μ; secondly

$$\begin{aligned}\eta(T_g^{-1}A) &= \lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}(T_g^{-i})_*\nu_n(T_g^{-1}A) \\ &= \lim_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{n}(T_g^{-i})_*\nu_n(A) \\ &= \lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}(T_g^{-i})_*\nu_n(A) - \frac{1}{n}\Big(\nu_n(A) - \nu_n(T_g^n(A))\Big) \\ &= \eta(A),\end{aligned}$$

so η is T_g invariant; thirdly, start by noticing,

$$\begin{aligned}\int_{\Omega\times\hat{X}} F d\eta_n &= \int_{\Omega\times\hat{X}} F d\Big(\frac{1}{n}\sum_{i=0}^{n-1}(T^i_g)_*\nu_n\Big) \\ &= \frac{1}{n}\int_{\Omega\times\hat{X}}\sum_{i=0}^{n-1} F\circ T^i_g d\nu_n \\ &= \frac{1}{n}\int_{\Omega\times\hat{X}} F_n d\nu_n.\end{aligned}$$

Being careful with the identification of elements in X with elements of $\hat{X}$,

$$\frac{1}{n}\int_{\Omega\times\hat{X}} F_n d\nu_n = \frac{1}{n}\int_{\Omega\times\hat{X}} -D(Z_n(\omega)\cdot x_0) d\nu_n(\omega, D)$$

$$
\begin{aligned}
&= \frac{1}{n}\int_{\Omega}\int_{\hat{X}} -D(Z_n(\omega)\cdot x_0)d\nu_{\omega}^{n}(D)d\mu(\omega) \\
&= \frac{1}{n}\int_{\Omega} -D_{(Z_n(\omega)\cdot x_0)}(Z_n(\omega)\cdot x_0)d\mu(\omega) \\
&= \frac{1}{n}\int_{\Omega} d(Z_n(\omega)\cdot x_0, x_0)d\mu(\omega) \\
&\geq s.
\end{aligned}
$$

On the other hand, for any $\theta \in \mathrm{Prob}_{\mu}^{T_g}(\Omega\times X)$

$$
\begin{aligned}
\int_{\Omega\times\hat{X}} F(\omega,D)d\theta(\omega,D) &= \int_{\Omega\times\hat{X}} \frac{1}{n}F_n(\omega,D)d\theta(\omega,D) \\
&\leq \int_{\Omega} \frac{1}{n}\max_{D\in\hat{X}} F_n(\omega,D)d\mu(\omega) \\
&= \frac{1}{n}\int_{\Omega} d(x_0, Z_n(\omega)\cdot x_0)d\mu(\omega).
\end{aligned}
$$

Taking the respective limit we obtain $\int_{\Omega\times\hat{X}} F d\eta = s$. Notice we can take the limits as we have $F \in L^1(\Omega, C(\hat{X}))$.

In fact, η might not be the measure we want yet as ergodicity may fail. To take care of this problem we appeal to subsection 1.2.2. Consider $\mathrm{Prob}_{\mu}^{s}(\Omega\times\hat{X})$ the subspace of $\mathrm{Prob}_{\mu}(\Omega\times\hat{X})$ given by the T_g invariant elements satisfying the equality above. We've proven this set is non empty as it contains η, it is also easily seen to be closed and convex. By Krein-Milman theorem it admits an extremal point η_0 which we've proven to be an extremal point of $\mathrm{Prob}^{T_g}(\Omega\times\hat{X})$ whence ergodic.

Finally we apply Birkhoff's Ergodic Theorem: there is a full measure set E in $\Omega\times\hat{X}$, with respect to η_0, such that for every $(\omega,D)\in E$

$$
\lim_{n\to+\infty} -\frac{1}{n}D(Z_n(\omega)\cdot x_0) = \lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1} F(T_g^i(\omega),D)) = \int_{\Omega\times X} F d\eta_0 = s.
$$

Due to the projection property, for μ-a.e. ω there is a set $E = E_{\omega}$ of horofunctions D satisfying the properties we were seeking.

This is where the second technicality appears. We could be satisfied with simply picking a pair in E and obtain a theorem that way; however, the statement is stronger than that, the choice of D_{ω} is done measurably. To do so we refer to the measurable choice theorem by Robert Aumann [1].

Theorem 3.2 (Measurable Choice Theorem)**.** *Let (Ω,μ) be a probability space, X a standard measurable space and E a measurable set in $\Omega\times X$ whose projection onto Ω has full measure. Then there is a measurable function $g:\Omega\to X$ such that $(\omega,g(\omega))\in E$ for almost every ω in Ω.*

□

As we've seen in Chapter 2, horofunctions are quite well behaved for CAT(0) spaces on which the previous theorem has a stronger geometrical meaning.

Theorem 3.3 (Karlsson - Margulis)**.** *Let $(\Omega,\mathscr{B},\mu,T)$ be a mpds, X a proper complete CAT(0) space and $g:\Omega\to Isom(X)$ a measurable map. If*

$$
\int_{\Omega} d(x_0, g(\omega)\cdot x_0)d\mu(\omega) < +\infty
$$

then, for μ-a.e. $\omega \in \Omega$, $\frac{1}{n}d(Z_n(\omega)\cdot x_0, x_0) \to s$ and if $s > 0$ there exists a geodesic ray emanating from x_0, γ_ω, such that

$$\lim_{n\to+\infty} \frac{1}{n} d(Z_n(\omega)\cdot x_0, \gamma_\omega(ns)) = 0.$$

Proof. Applying Karlsson-Ledrappier result one knows that, for almost every ω, there is an horofunction D for which

$$\lim_{n\to+\infty} -\frac{1}{n} D(Z_n(\omega)\cdot x_0) = s.$$

Since X is CAT(0) we know, by theorem 2.52, there is a unique geodesic γ for which

$$\lim_{n\to+\infty} \frac{1}{n} d(Z_n(\omega)\cdot x_0, \gamma_\omega(ns)) = 0.$$

□

3.2 Karlsson-Margulis: The Nonproper Case

In the previous section we presented Karlsson-Margulis theorem for proper spaces by using Karlsson-Ledrappier. In fact, the original statement included this class of spaces and was less restrictive on the action working with a semigroup of semicontractions rather then a group of isometries. On the other hand, this approach won't allow us to obtain an analogue to Karlsson-Ledrappier. The importance of considering the proper case (Karlsson-Ledrappier) is that it can be applied regardless of curvature, even if its geometrical meaning isn't as clear for such cases.

Theorem 3.4 (Karlsson-Margulis)**.** *Let S be a semigroup of semicontractions of some complete, uniformly convex, nonpositively curved in the sense of Busemann, metric space (X,d) with an origin point x_0. Just like before, let $(\Omega, \mathscr{B}, \mu, T)$ be an ergodic mpds and $g : \Omega \to S$ a measurable map. Define Z_n a right cocycle just like in the isometry case. If Z_n is integrable. then for μ-a.e. ω the following limit exists*

$$\lim_{n\to+\infty} \frac{1}{n} d(x_0, Z_n(\omega)\cdot x_0) = s.$$

Moreover, if $s > 0$, then for μ-a.e. ω there is a unique geodesic ray in X starting at x_0 such that

$$\lim_{n\to+\infty} \frac{1}{n} d(Z_n(\omega)\cdot x_0, \gamma_\omega(ns)) = 0.$$

Proof. [9] The first part of the proof is analogous to the proper case, $a(n,\omega) = d(x_0, Z_n(\omega)\cdot x_0)$ is subadditive, so applying subadditive ergodic theorem the limit

$$\lim_{n\to+\infty} \frac{1}{n} d(x_0, Z_n(\omega)\cdot x_0) = s. \tag{3.1}$$

exists for almost every ω.

Prep Work

We now tackle the part of the statement about $s > 0$. Let f be the function arising from lemma 2.12. Since $f(t)$ goes to zero as t does, pick $\{\varepsilon_i\}_{i\geq 0}$ small enough so that

$$f\left(\frac{2\varepsilon_i}{s-\varepsilon_i}\right) \leq 2^{-i}.$$

We shall refer to the quantity $2\varepsilon_i/(s-\varepsilon_i)$ as δ_i.

Define E as in proposition 1.16. From now on we will both assume ω is in E and satisfies the limit 3.1. From that proposition and the subadditive ergodic theorem, for every integer i there is K_i and infinitely many n_i such that, for all $K_i \leq k \leq n_i$,

$$a(n,\omega) - a(n-k, T^k\omega) \geq (s-\varepsilon_i)k$$

and

$$(s-\varepsilon_i)k \leq a(k,\omega) \leq (s+\varepsilon_i)k.$$

Pick a strictly increasing sequence $\{n_i\}_{i\geq 0}$ such that $n_i > n_{i+1}$ and the inequalities above hold. Then

$$a(n_i,\omega) - a(n_i-k, T^k(\omega)) + (s+\varepsilon_i)k \geq (s-\varepsilon_i)k + a(k,\omega),$$

that is,

$$a(k,\omega) + a(n_i-k, T^k(\omega)) \leq a(n_i,\omega) + 2\varepsilon_i k.$$

Hence,

$$\begin{aligned}
d(x_0, Z_k(\omega)\cdot x_0) + d(Z_k(\omega)\cdot x_0, Z_{n_i}(\omega)\cdot x_0) \leq& d(x_0, Z_k(\omega)\cdot x_0) + d(x_0, Z_{n_i-k}(T^k(\omega))\cdot x_0) \\
=& a(k,\omega) + a(n_i-k, T^k(\omega)) \\
\leq& a(n_i,\omega) + 2\varepsilon_i k \\
\leq& d(x_0, Z_n(\omega)\cdot x_0) + \frac{2\varepsilon_i}{s-\varepsilon_i} d(x_0, Z_k(\omega)\cdot x_0).
\end{aligned}$$

Choose now a sequence of geodesics emanating from x_0 passing through $Z_{n_i}(\omega)\cdot x_0$. Applying lemma 2.12 to the inequality above

$$d(\gamma_i(r_k), Z_k(\omega)\cdot x_0) \leq f(\delta_i) r_k,$$

with $r_k = d(x_0, Z_k(\omega))$.

Convergence of Geodesics

The idea is very similar to what we've done before for CAT(0) spaces: show that for every $t > 0$ the sequence $\{\gamma_i(t)\}_{i\geq 0}$ is Cauchy and then define the limit geodesic at t by the limit. Due to the construction of the n_i's

$$d(\gamma_{i+1}(r_{n_i}), \gamma_i(r_{n_i})) \ = \ d(\gamma_{i+1}(r_{n_i}), Z_{n_i}(\omega)\cdot x_0) \ \leq \ f(\delta_{n+1}) r_{n_i}$$

Using the definition of being nonpositively curved in the sense of Busemann, for $r_{n_i} > t$

$$\frac{1}{t} d(\gamma_{i+1}(t), \gamma_i(t)) \ \leq \ \frac{1}{r_{n_i}} d(\gamma_{i+1}(r_{n_i}), \gamma_i(r_{n_i})) \ \leq \ f(\delta_{i+1})$$

rearranging we obtain $d(\gamma_{i+1}(t), \gamma_i(t)) \leq f(\delta_{i+1})t$. Using triangle inequality

$$\begin{aligned}
d(\gamma_{i+m}(t), \gamma_i(t)) &\leq \sum_{j=1}^{m} d(\gamma_{i+j-1}(t), \gamma_{i+j}(t)) \\
&\leq \sum_{j=1}^{m} f(\delta_{i+j}) t
\end{aligned}$$

$$\leq \sum_{j=1}^{m} 2^{-(i+j)} t$$
$$\leq 2^{-i} t$$

Hence, $\{\gamma_i(t)\}_{i\geq 0}$ is Cauchy. The argument for γ being a geodesic is the same that was done for proposition 2.51. It now remains to complete the proof

Completing the Proof

Given any k we can choose i such that $K_i \leq k \leq n_i$.

$$\begin{aligned}
d(\gamma(sk), Z_k(\omega)\cdot x_0) \leq& d(\gamma(sk), \gamma_i(sk)) + d(\gamma_i(sk), \gamma_i(r_k)) + d(\gamma_i(r_k), Z_k(\omega)\cdot x_0)\\
\leq& 2^{-i} sk + |sk - r_k| + f(\delta_i) r_k\\
\leq& 2^{-i} sk + |sk - sk - \varepsilon_i k| + f(\delta_i)(s+\varepsilon_i)k\\
\leq& 2^{-i} sk + \varepsilon_i k + f(\delta_i)(s+\varepsilon_i)k\\
\leq& 2^{-i} sk + \varepsilon_i k + 2^{-i}(s+\varepsilon_i)k\\
\leq& 2^{-i+1} sk + 2\varepsilon_i k.
\end{aligned}$$

Therefore
$$\limsup_{n\to+\infty} \frac{1}{n} d(\gamma(ns), Z_n(\omega)\cdot x_0) \leq 0,$$
and the uniqueness of the geodesic follows from 2.10. □

3.3 Multiplicative Ergodic Theorem

The first space we shall apply Karlsson-Margulis theorem to is $Sym_+(n)$. Remember $GL(n,\mathbb{R})$ acts by isometries on the space of symmetric positive definite matrices. Our goal will be to obtain the multiplicative ergodic theorem stated in the first chapter.

Lemma 3.5. *Let $A \in GL(n,\mathbb{R})$ then*

$$\max\{\log^+ ||A||, \log^+ ||A^{-1}||\} \leq d(I, A*I) \leq 2\sqrt{n} \max\{\log^+ ||A||, \log^+ ||A^{-1}||\}$$

Proof. The matrix $(A*I)^T = A^T A$ is symmetric and positive definite, as such there is an orthonormal basis $v_1,\dots,v_n$ of eigenvectors with associated real eigenvalues $\lambda_1,\dots,\lambda_n$. Let $v\in\mathbb{R}^n$ and write it as $v = \sum_{i=1}^n a_i v_i$

$$\begin{aligned}
||Av||^2 =& \langle Av, Av\rangle\\
=& \langle A^T A v, v\rangle\\
=& \langle \sum_{i=1}^n \lambda_i a_i v_i, \sum_{i=1}^n a_i v_i\rangle\\
=& \sum_{i=1}^n \lambda_i a_i^2
\end{aligned}$$

Denoting by λ_{max} and λ_{min} the maximum and minimum of the singular values of A we have $||A|| = \lambda_{max}$ and $||A^{-1}||^{-1} = \lambda_{min}$.

Now, using the distance in $Sym_+(n)$ (2.1),

$$d(I, A*I) = \sqrt{\sum_{i=1}^{n} \log^2(\lambda_i^2)} = 2\sqrt{\sum_{i=1}^{n} \log^2(\lambda_i)}.$$

However

$$\max\{\log^+ ||A||, \log^+ ||A^{-1}||\} \leq 2\sqrt{\sum_{i=1}^{n} \log^2(\lambda_i)} \leq 2\sqrt{n}\max\{\log^+ ||A||, \log^+ ||A^{-1}||\}.$$

□

Before tackling the multiplicative ergodic theorem we need another little lemma.

Lemma 3.6. *Let $M \in M(d,\mathbb{R})$, then $||M|| = ||M^T||$.*

Proof. We have $||Mv||^2 = v^T M^T M v$ and $||M^T v||^2 = v^T M M^T v$ so $||M||^2$ and $||M^T||^2$ are both equal to the largest eigenvalue of $M^T M$ and MM^T respectively. Since a matrix has the same singular values as its transpose we get the equality. □

Theorem 3.7 (Multiplicative Ergodic Theorem)**.** *Let $(\Omega, \mathscr{B}, \mu, T)$ be a mdps, $A : \Omega \to GL(d,\mathbb{R})$ a measurable map such that $\log^+(||A^{\pm 1}||)$ is integrable. Consider $A_n(\omega) = A(T^{n-1}\omega)\cdots A(\omega)$ the left cocycle given by A. Then the limits*

$$\Lambda(\omega) := \lim_{n\to+\infty} (A_n(\omega)^T A_n(\omega))^{\frac{1}{2n}}$$

and

$$\lim_{n\to+\infty} \frac{1}{n} \log ||A_n(\omega)\Lambda^{-n}(\omega)|| \;=\; \lim_{n\to+\infty} \frac{1}{n} \log ||\Lambda^n(\omega) A_n^{-1}(\omega)|| = 0$$

exist for μ-a.e. ω.

Before proving the theorem let's see why this is the same as the version presented in Chapter 1 (1.9). The Lyapunov exponents $\chi_i(\omega)$ are the logarithms of the eigenvalues $\lambda_i(\omega)$ of $\Lambda(\omega)$ in increasing order ($\lambda_i(\omega) < \lambda_j(\omega)$ whenever $i < j$). For every i let $U_i(\omega)$ be the eigenspace associated with the eigenvalue λ_i. The filtration is given by $E_i(\omega) = \oplus_{j\leq i} U_j(\omega)$.

Let us explore why. Clearly $\chi_i(\omega)$ appear in increasing order as the logarithm is an increasing function, moreover, $E_i(\omega)$ form a flag. Now let $v \in E_i(\omega) \backslash E_{i-1}(\omega)$, then,

$$\begin{aligned}
\lim_{n\to+\infty} \frac{1}{n} \log ||A_n(\omega)v|| &= \lim \frac{1}{n} \log ||\lambda_i^n(\omega) A_n(\omega)\Lambda(\omega)^{-n} v|| \\
&= \lim \frac{1}{n} \Big(\log(\lambda_i(\omega)^n) + \log ||A_n(\omega)\Lambda(\omega)^{-n} v|| \Big) \\
&= \lim \frac{1}{n} \log(\lambda_i(\omega)^n) + \frac{1}{n} \log ||A_n(\omega)\Lambda(\omega)^{-n} v|| \\
&= \log(\lambda_i(\omega)) \\
&= \chi_i(\omega).
\end{aligned}$$

The above also implies that $A(\omega)E_i(\omega) = E_i(T(\omega))$ as

$$\lim_{n\to+\infty} \frac{1}{n} \log ||A_n(T(\omega))A(\omega)v|| = \lim_{n\to+\infty} \frac{1}{n} \log ||A_n(\omega)v||.$$

For the last point we need the property that $\log(\det(A)) = tr(\log(A))$. Then

$$\begin{aligned}
\frac{1}{n}\log\left(\det(A_n(\omega))\right) =& \frac{1}{n}\log\left(\det(A_n(\omega)^T A_n(\omega))^{\frac{1}{2}}\right) \\
=& \log\left(\det(A_n(\omega)^T A_n(\omega))^{\frac{1}{2n}}\right) \\
\to& \log\left(\det(\Lambda(\omega))\right) \\
=& \mathrm{tr}\left(\log(\Lambda(\omega))\right) \\
=& \sum_{i=1}^{k(\omega)} \chi_i(\omega)\left(\dim E_i(\omega) - \dim E_{i-1}(\omega)\right)
\end{aligned}$$

Proof. By lemma 3.5, $d(I, A(\omega)^T * I)$ is integrable so we are allowed to use Karlsson-Margulis theorem to the right cocycle $Z_n(\omega) = A_n(\omega)^T$. Doing so, for μ-a.e. ω there is

$$s(\omega) := \lim_{n\to+\infty} \frac{1}{n} d(I, Z_n(\omega) * I)$$

and, due to the geodesic approximation property, there is a symmetric matrix $S(\omega)$ such that

$$\lim_{n\to+\infty} \frac{1}{n} d(Z_n(\omega) * I, e^{s(\omega)nS(\omega)}) = 0$$

We now write $\Lambda(\omega)$ as the positive definite symmetric matrix $e^{s(\omega)S(\omega)/2}$, the above limit tells us the limit of $d(Z_n(\omega) * I, \Lambda(\omega)^{2n})/n$ is also zero. However, we also have the equalities

$$d(Z_n(\omega) * I, \Lambda(\omega)^{2n}) \;=\; d(Z_n(\omega) * I, \Lambda(\omega)^n * I) \;=\; d(I, Z_n(\omega)^{-n}\Lambda(\omega)^n * I) \;=\; d(\Lambda(\omega)^{-n} Z_n(\omega) * I, I),$$

which, using lemma 3.5, imply the limits

$$\lim_{n\to+\infty} \frac{1}{n}\log||\Lambda^{-n}(\omega)Z_n(\omega)|| \;=\; \lim_{n\to+\infty} \frac{1}{n}\log||Z_n^{-1}(\omega)\Lambda^n(\omega)|| = 0.$$

Using lemma 3.6, we obtain the limits from the theorem:

$$\lim_{n\to+\infty} \frac{1}{n}\log||A_n(\omega)\Lambda^{-n}(\omega)|| \;=\; \lim_{n\to+\infty} \frac{1}{n}\log||\Lambda^n(\omega)A_n^{-1}(\omega)|| = 0.$$

It remains to show $\Lambda(\omega)$ is the limit of $(Z_n(\omega)Z_n(\omega)^T)^{\frac{1}{2n}}$. To do so we shall consider two cases, effectively $\Lambda \neq I$ and $\Lambda = I$. For the first case consider γ_n, the sequence of geodesics joining I to $Z_n(\omega)$, as well as γ, the one joining I to Λ. We want to prove

$$\lim_{n\to+\infty} \gamma_n(t) = \gamma(t).$$

Let $t > 0$ and take $\bar{I}, \overline{Z_n * I}, \overline{\Lambda^n * I}$ forming a comparison triangle to the geodesic triangle given by $I, Z_n(\omega) * I$ and $\Lambda(\omega)^n * I$. Also take $\overline{\gamma_n(t)}$, $\overline{\gamma(t)}$ the comparison points for $\gamma_n(t)$ and $\gamma(t)$. Last points we will be taking are k_n and l_n the orthogonal projections of $\overline{\gamma_n(t)}$ and $\overline{Z_n * I}$ (see figure 3.1).

Notice

$$\frac{d(\overline{Z_n * I}, l_n)}{n} \leq \frac{d(\overline{Z_n * I}, \overline{\Lambda * I})}{n} \xrightarrow[n\to+\infty]{} 0$$

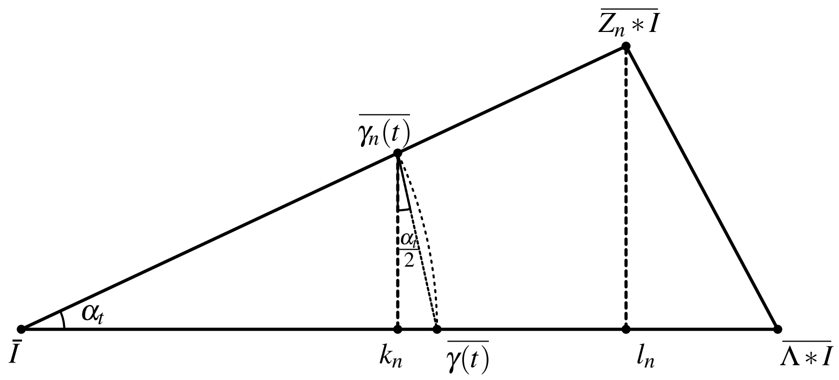

Figure 3.1: Geometrical construction used for the proof of the multiplicative ergodic theorem

which yields

$$\frac{d(\bar{I}, l_n)}{n} = \sqrt{\frac{d(\overline{Z_n * I}, \bar{I})^2}{n^2} - \frac{d(\overline{Z_n * I}, l_n)^2}{n^2}} \xrightarrow[n \to +\infty]{} s.$$

By triangle similarity,

$$\begin{aligned} d(\overline{\gamma_n(t)}, k_n) =& \frac{d(\bar{I}, k_n)}{d(\bar{I}, l_n)} d(\overline{Z_n * I}, l_n) \\ \leq& \frac{nt}{d(\bar{I}, l_n)} \frac{d(\overline{Z_n * I}, l_n)}{n} \xrightarrow[n \to +\infty]{} 0. \end{aligned}$$

Using the law of cosines we know $\alpha_t = \angle \overline{\gamma(t)} \bar{I} \overline{\gamma_n(t)} \to 0$. From this $d(k_n, \overline{\gamma(t)}) = \tan(\alpha_t/2) d(\overline{\gamma_n(t)}, k_n) \to 0$. Therefore

$$\begin{aligned} d(\gamma_n(t), \gamma(t)) \leq& d(\overline{\gamma_n(t)}, \overline{\gamma(t)}) \\ =& \sqrt{d(k_n, \overline{\gamma(t)})^2 + d(k_n, \overline{\gamma_n(t)})^2} \xrightarrow[n \to +\infty]{} 0, \end{aligned}$$

that is, $\gamma_n(t) \to \gamma(t)$.

In particular, we have

$$\frac{\log(Z_n(\omega) Z_n(\omega)^T)}{\mathrm{tr}^{1/2}\left(\log^2(Z_n(\omega) Z_n(\omega)^T)\right)} = \log\left(\gamma_n(1)\right) \xrightarrow[n \to +\infty]{} \log\left(\gamma(1)\right) = \frac{\log(\Lambda^2)}{\mathrm{tr}^{1/2}\left(\log^2(\Lambda^2)\right)}.$$

Let us first focus on relating the denominators starting with the equality $\mathrm{tr}^{1/2}\left(\log^2(Z_n(\omega) Z_n(\omega)^T)\right) = d(I, Z_n(\omega) * I)$ and proceeding with

$$\frac{d(I, \Lambda(\omega)^n * I)}{n} - \frac{d(Z_n(\omega) * I, \Lambda(\omega)^n * I)}{n} \leq \frac{d(I, Z_n(\omega) * I)}{n} \leq \frac{d(I, \Lambda(\omega)^n * I)}{n} + \frac{d(Z_n(\omega) * I, \Lambda^n * I)}{n}$$

from which, taking the limits, we obtain

$$\frac{\mathrm{tr}^{1/2}\left(\log^2(Z_n(\omega) Z_n(\omega)^T)\right)}{n} \xrightarrow[n \to +\infty]{} \mathrm{tr}^{1/2}\left(\log^2(\Lambda^2)\right).$$

Having taken care of the denominators, we have $\log(Z_n(\omega) Z_n(\omega)^T)/n \to \log(\Lambda(\omega)^2)$, equivalently, $\log(Z_n(\omega) Z_n(\omega)^T)/2n \to \log(\Lambda(\omega))$. Taking the exponential we obtain the wanted result.

It remains to tackle the case $\Lambda(\omega) = I$. Writing $Z_n(\omega)Z_n(\omega)^T = Q_n^T D_n Q_n$ where Q_n are orthogonal and D_n is a diagonal matrix comprised of the eigenvalues $\lambda_i(n)$. Notice we must have $d(Z_n(\omega) * I, \Lambda(\omega)^2)/n \to 0$, in other words, $\log\left(||(Z_n(\omega)Z_n(\omega)^T)^{\pm 1}||\right)/n \to 0$. Hence, for every $i \leq d$, $\lambda_i(n)^{1/n} \to 1$, that is, $D_n^{\frac{1}{2n}} \to I$. Finally

$$\begin{aligned} d((Z_n(\omega)Z_n(\omega)^T)^{\frac{1}{2n}}, I) &= d(Q_n^T D_n^{\frac{1}{2n}} Q_n, I) \\ &= d(Q_n^T D_n^{\frac{1}{2n}} Q_n, Q_n^T Q_n) \\ &= d(D_n^{\frac{1}{2n}}, I) \xrightarrow[n\to+\infty]{} 0, \end{aligned}$$

hence obtaining the the wanted result. □

3.4 Further Applications

Birkhoff's Ergodic Theorem

Let $(\Omega, \mathscr{B}, \mu, T)$ be an ergodic mpds. Consider $\mathbb{R}$ acting on itself by isometries through the map $g : \Omega \to \mathbb{R}$. By Karlsson-Ledrappier there is a horofunction D such that

$$\lim -\frac{1}{n} D\Big(\sum_{k=0}^{n-1} g(T^k(\omega))\Big) = s,$$

where $s = |\int_\Omega g(\omega) d\mu(\omega)|$. The only horofunctions of $\mathbb{R}$ are $D_1(x) = x$ and $D_{-1}(x) = -x$. Using D_1 if $\int_\Omega g(\omega) d\mu(\omega) < 0$ and D_{-1} if $\int_\Omega g(\omega) d\mu(\omega) > 0$ we obtain Birkhoff's Ergodic Theorem.

In this case one can actually find the geodesic to be $\gamma_\omega(t) = x_0 + \text{sgn}(\int_\Omega g d\mu) ts$ as

$$\lim \frac{1}{n} d\Big(x_0 + \sum_{k=0}^{n-1} g(T^k(\omega)), x_0 + \text{sgn}\Big(\int_\Omega g d\mu\Big) sn\Big) = \lim \frac{1}{n}\Big(\sum_{k=0}^{n-1} g(T^k(\omega)) - n\int_\Omega g(\omega) d\mu(\omega)\Big) = 0.$$

Cayley Graphs of Free Groups

Let p be a natural number greater than one, $\mathbb{F}_p$ be the free group on p generators, S the set of generators and S^{-1} the set of their inverses. We define the Cayley graph of $\mathbb{F}_p$ to be the graph whose set of vertices is $\mathbb{F}_p$ and two elements g and h are connected if and only if $g \in h(S \cup S^{-1})$, that is, they differ by an element on the right. Notice that the left action of $\mathbb{F}_p$ onto itself is an isometry. Moreover, the Cayley graph is a tree.

Let us consider the Bernoulli shift σ on the space of sequences over the space of $2p$ symbols, $\Omega = [2p]^{\mathbb{N}}$, which is ergodic. Take $T : \Omega \to [2p]$ the projection of the first coordinate. List the generators of and their inverses $S \cup S^{-1} = \{a_1, a_2, ..., a_{2p}\}$ and take the measurable map

$$\begin{aligned} g : \Omega &\to \mathbb{F}_p \\ x &\to a_{T(x)}. \end{aligned}$$

We will work with the cocycle $Z_n(x) = g(x)g(\sigma(x))\cdots g(\sigma^{n-1}(x))$. Fix the origin of the graph to be the identity element e. Notice that, denoting $k_i(x)$ the last element of $Z_n(x)$, we have

$$d(Z_n(x),e) = \sum_{i=0}^{n-1}(-1)^{\delta\left(k_i(x)^{-1},g(\sigma^i(x))\right)},$$

where δ denotes the Kronecker symbol. Since

$$\int_\Omega (-1)^{\delta(e,g(x))}d\mu = \int_{[0,1)} 1d\mu = 1,$$

the first part of Karlsson-Margulis asserts that

$$\frac{1}{n}d(Z_n(x),e) \to s.$$

We will want to show that $s > 0$ so we can use the second part of the statement, in fact we will try to calculate s. Start by noticing

$$\int_\Omega (-1)^{\delta\left(k_i(x)^{-1},g(\sigma^i(x))\right)}d\mu = \mu\{x \mid \sigma^i(x) \neq k_i(x)\} - \mu\{x \mid \sigma^i(x) = k_i(x)\}.$$

However, the two sets above are complementary so we have to look at

$$\begin{aligned}
1-2\mu\{x \mid \sigma^i(x) = k_i(x)\} =& 1 - 2\sum_{h\in S\cup S^{-1}} \mu\{x \mid \sigma^i(x) = h \wedge k_i(x) = h\} \\
\geq& 1 - 2\sum_{h\in S\cup S^{-1}} \mu\{x \mid \sigma^i(x) = h\}\mu\{x \mid k_i(x) = h\} \\
=& 1 - \frac{1}{p}\sum_{h\in S\cup S^{-1}} \mu\{x \mid k_i(x) = h\} \\
=& 1 - \frac{1}{p}(1-\mu\{x \mid k_i(x) = e\}).
\end{aligned}$$

Since the space has finite measure, pointwise almost everywhere convergence implies convergence in measure. Hence, for $\varepsilon < 1$, we have

$$\begin{aligned}
\mu\{\omega\in\Omega \mid k_i = e\} &= \mu\{\omega\in\Omega \mid \frac{1}{i}d(Z_i(\omega),e) = 0\} \\
&\leq \mu\{\omega\in\Omega \mid |\frac{1}{i}d(Z_i(\omega),e) - s| > s\varepsilon\} \xrightarrow[i\to+\infty]{} 0.
\end{aligned}$$

Therefore

$$s = \lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\int_\Omega (-1)^{\delta\left(k_i(x)^{-1},g(\sigma^i(x))\right)}d\mu = \frac{p-1}{p}.$$

The second statement of Karlsson-Margulis Theorem asserts that, for almost every x, there is a unique geodesic γ on $\mathbb{F}_p$ such that

$$\frac{1}{n}d\big(\gamma(\frac{p-1}{p}n), Z_n(x)\big) \xrightarrow[n\to+\infty]{} 0.$$

Intuitively we can look at the geodesic as the one that ultimately goes through the same vertices as $Z_n(x)$.

Bibliography

[1] Robert J Aumann. *Measurable Utility and the Measurable Choice Theorem*. Defense Technical Information Center, 1967.

[2] Alan F. Beardon. *The geometry of discrete groups*. Springer, 1983.

[3] Martin Bridson and André Haefliger. *Metric spaces of non-positive curvature*. Springer, 1999.

[4] Hans Crauel. *Random probability measures on Polish spaces*, volume 11. CRC press, 2003.

[5] Manfredo do Carmo. *Riemannian geometry*. Birkhauser, 1992.

[6] Vadim Kaimanovich. Lyapunov exponents, symmetric spaces, and a multiplicative ergodic theorem for semisimple lie groups. *Journal of Soviet Mathematics*, 47(2):2387–2398, 1989.

[7] Anders Karlsson. Ergodic theorems for noncommuting random products. 2008.

[8] Anders Karlsson and François Ledrappier. On laws of large numbers for random walks. *The Annals of Probability*, pages 1693–1706, 2006.

[9] Anders Karlsson and Gregory Margulis. A multiplicative ergodic theorem and nonpositively curved spaces. *Communications in mathematical physics*, 208(1):107–123, 1999.

[10] Walter Rudin. *Fourier analysis on groups*. Interscience, 1962.

[11] Omri Sarig. Lecture notes on ergodic theory. 2009.