UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS

A Geometric Avalanche Principle

" Documento Definitivo"

Doutoramento em Matemática

Especialidade de Geometria e Topologia

Luís Miguel Neves Pedro Machado Sampaio

Tese orientada por:

Professor Doutor Pedro Miguel Nunes da Rosa Dias Duarte

Documento especialmente elaborado para a obtenção do grau de doutor

2022

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS

A Geometric Avalanche Principle

Doutoramento em Matemática

Especialidade de Geometria e Topologia

Luís Miguel Neves Pedro Machado Sampaio

Tese orientada por:

Professor Doutor Pedro Miguel Nunes da Rosa Dias Duarte

Júri:

Presidente: Manuel Duque Pereira Monteiro Marques, Professor Catedrático, Faculdade de Ciências da Universidade de Lisboa

Vogais: Doutor Anders Karlsson, Professor Associado, *Section des Mathématiques* d*a Université de Genèv*e, Suíça;

Doutor Paulo César Rodrigues Pinto Varandas, Professor Associado, Instituto de Matemática da Universidade Federal da Bahia, Brasil;

Doutor Jorge Sebastião de Lemos Carvalhão Buescu, Professor Associado com Agregação, Faculdade de Ciências da Universidade de Lisboa;

Doutor Pedro Miguel Nunes da Rosa Dias Duarte, Professor Associado com Agregação, Faculdade de Ciências da Universidade de Lisboa, Orientador.

Documento especialmente elaborado para a obtenção do grau de doutor

Financiado pela Universidade de Lisboa ao abrigo do programa de bolsas de doutoramento BD2018

2022

”We are so unwise that we wander about in times that do not belong to us, and do not think of the only one that does; so vain that we dream of times that are not and blindly flee the only one that is.”
- Blaise Pascal, Pensées

Acknowledgments

Forevermore, anyone who glances at this thesis will not just see my name, they will also see my advisor's. From asking the correct questions, namely the ones that inspired this monograph, to the time he spent presenting me new topics and ideas, Pedro's influence has been a major factor in my mathematical work during the past few years.

On a personal level, Pedro's calm yet decisive nature that matches my own was also instrumental for the completion of the thesis. Throughout this endeavour he kept showing me a stance and belief that I had the ability to produce mathematics, whilst also requiring that I develop the necessary skills to do so. This reassurance was what kept me focused at the beginning stages of the PhD when the path was unclear and the doubts too many.

The breakthrough of this work came when I realized, and more importantly live by, the fact that Pedro's own PhD finished the year I was born; now it was my time to prove something and showcase my work. It was time to walk first and report i got lost later, not wait for a clear path to appear, mostly because it doesn't exist anymore. It is at this point I thank the first intervinient of the thesis - me.

Next I thank the LxDS group and all its members for the very nice seminars and discussions on dynamical systems. Despite my preference of working alone, the group was important to keep away the feeling of loneliness. For the proximity, I mention Ao Cai, who always keeps a big smile to show everyone, and Jamerson Bezerra for the occasional conversations.

I would also like to thank the great professors from the Department of Mathematics of the Faculty of Sciences of the University of Lisbon, which were big influences in my mathematical growth. In particular, mentioning the people that were part of this journey, I thank professors Maria Manuel Torres, Jean-Claude Zambrini, Maria João Pablo, José Francisco Rodrigues and Teresa Monteiro Fernandes for taking time of their days to guide me in the first year seminars and professors Hugo Tavares and Carlos André for the nice PhD courses. In all facets of life, I must also leave a special word to professor Carlota Gonçalves whose value goes far beyond the tangibles of academia, and hopefully will keep inspiring generations to come.

My gratitude extends as well to the department for supporting all my needs, namely room 6.2.39, where I am now writing this and have spent most of my days (and sometimes nights), during the past few years. Also thank the people that are usually forgotten, like Helena Afonso, whom would provide any missing commodities, or my old PhD colleagues who left a fan and a heater in the office.

Of course a word is also due to the University of Lisbon for providing me with a scholarship which allowed me to follow through with the project of the thesis. Despite the occasional bureaucratic problems, the *Núcleo de Formação ao Longo da Vida* was always very helpful.

Now comes the part I dedicate to my girlfriend Patricia who had to hear a myriad of soliloquies before any of my presentations. Thank you for understanding, or at least accepting, the time I spent looking at nothing thinking of something, the nights I barely slept or the times I forgot something because of mathematics.

Finally I thank all the close people who will most likely only read this part of the thesis, the jury who most likely will read everything but this part of the thesis and the kind souls that ever finds the will to read everything.

Abstract

In this thesis we obtain an abstract continuity theorem for the drift associated with a product of isometries in both Gromov hyperbolic spaces and symmetric spaces as well as the Lyapunov exponents for a product of linear operators over some Hilbert space. We obtain these results by following a recipe of having large deviations estimates and an avalanche principle; a result which allows us to take conclusion of global nature from local hypothesis.

As a main example, we apply the results to cocycles over Markov systems, where we prove the aforementioned large deviations estimates hold, thus providing a large class of examples. Upon presenting the linear setting we also mention the case of quasi-periodic linear cocycles. Whilst exploring Markov systems we also obtain a Fürstenberg type formula.

From the perspective of Gromov hyperbolic spaces, we prove their group of isometries is a topological group and how random products of isometries follow a multiplicative ergodic theorem for the drift, thus describing the behaviour of typical orbits.

Keywords: Large deviation estimates, multiplicative ergodic theory, abstract continuity theorem, avalanche principle

Resumo Alargado

Um tópico que surge naturalmente em diversas áreas da matemática e das suas aplicação é a compreensão de produtos aleatórios de operadores, escolhidos segundo alguma regra estocástica ou determinística. Realisticamente, não podemos esperar obter muita informação sobre o produto neste nível de generalidade, uma vez que tipicamente o comportamento assimptótico do produto deve divergir para alguma noção de ”infinito” do espaço. Para combater este problema, é costume introduzir alguma informação que descreve o comportamento assimptótico do produto. Um exemplo paradigmático de tal quantidade descritiva é o exponente de Lyapunov e a filtração de Oseledets que, através do teorema multiplicativo ergódico de Oseledets, descrevem completamente a dinâmica de um produto aleatório de matrizes.

Pelo teorema subaditivo ergódico de Kingman, a forma usual de garantir que tais quantidades descritivas existem é garantir que elas são subaditivas em relação ao produto. Isto acontece tanto com o maior exponente de Lyapunov como com o drift linear no caso de semicontracções em espaços métricos. Obtendo assim a existência destas quantidades, tentamos obter um teorema multiplicativo que explique o seu comportamento. e assim o comportamento do produto.

Uma vez percebida a dinâmica do produto, uma questão interessante é - o que acontece com estas quantidades, que podem ser vistas como funções, se perturbarmos subtilmente os elementos do produto, quer escolhendo elementos semelhantes, quer mudando a forma como são escolhidos? Chamamos a esta questão o problema da continuidade. Este é um problema bastante rico e complexo que iremos focar ao longo do resto do texto.

Outro problema interessante ao qual nos dedicaremos será perceber a taxa de convergência associada ao teorema de Kingman, uma vez que este nada nos diz sobre o tipo de convergência associado à existência das supracitadas quantidades descritivas. A forma natural de considerar esta taxa de convergência é utilizando a teoria de grandes desvios. Infelizmente, grandes desvios, tal como a continuidade, podem ser problemáticos, uma vez que em geral, dada a complexidade da regra que escolhe os operadores, são difíceis de obter.

Em 2001 Goldstein e Schlag, estabeleceram uma conexão entre os dois problemas provando que, para cóciclos de Schrödinger, estimativas de grandes desvios levam à continuidade. Além disso, esta continuidade pode ser quantificada pela força do desvio. Mais tarde, Duarte e Klein provam esta relação em geral para o exponente de Lyapunov de operadores lineares em dimensão finita. A esta conexão entre a propriedade estatistica de grandes desvios e a continuidade chamamos de teorema abstracto de continuidade.

Nesta tese iremos apresentar vários problemas. Em primeiro lugar iremos resolver é a existência de um teorema abstracto de continuidade para o drift em espaços hiperpólicos no sentido de Gromov. Por outras palavras, iremos considerar produtos de isometrias destes espaços e ver que o mesmo tipo de resultado obtido por Duarte e Klein para o exponente de Lyapunov também é válido para o drift. Em segundo lugar iremos abordar o tópico do drift em espaços simétricos. Estes espaços formam uma class de exemplos one a curvature pode tomar o valor zero para a qual temos várias ferramentas à nossa disposição. Umas dessas ferramentas das quais iremos fazer uso são as representações lineares. Assim sendo aproveitamos para apresentar as versões lineares dos resultados que iremos depois utilizar.

A parte principal do argumento do teorema abstracto de continuidade é a possibilidade de transportar controlos a uma certa escala finita em frente no tempo de forma a obter resultados válidos nas escalas seguintes. Na base desta possibilidade está um resultado pilar ao qual chamamos principio da avalanche. Dito isto, a novidade da tese é a obtenção de princípios da avalanche nos casos abordados anteriormente. O nome da tese, ”Geometric Avalanche Principle” surge da natureza geométrica do principio da avalanche e das suas aplicações ao longo do texto.

Dedicamos o capitulo 2 aos espaços hiperbólicos de Gromov. Estes espaços são ferramentas fundamentais no estudo da teoria geométrica de grupos, uma vez que apresentam uma forma sistemática de obter resultados de natureza algébrica a partir de argumentos geométricos para uma vasta classe de grupos. Grupos hiperbólicos, como são chamados, têm sido objecto de estudo deste os anos 80 e continuam a inspirar bastante investigação. Neste texto não estamos interessados no caso mais geral em que um grupo arbitrário, não necessariamente hiperbólico, actua por isometries nestes espaços. Assim sendo, este capitulo inclui uma breve apresentação dos conceitos essenciais de dinâmica e ergodicidade necessários. Em particular, provamos aqui um teorema ergódico que governa o comportamento do produto de isometrias.

No capitulo 3 apresentamos o teorema abstracto de continuidade do drift em espaços hiperbólicos. O resultado é obtido utilizado um argumento indutivo baseado na existência de grandes desvios e no principio da avalanche. Iremos também para além da continuidade, também podemos obter um critério abstracto de positividade do drift.

Seguimos então para o capitulo 4, onde iremos aplicar a teoria a sistemas de Markov. Assim provaremos que sistemas de Markov satisfazem grandes desvios para em seguida aplicar directamente o teorema abstracto de continuidade. Apresentamos também neste capitulo uma prova directa de continuidade para passeios aleatórios utilizando a fórmula de Fürstenberg, o que permite uma formulação mais simples do resultado neste caso. Por fim mostramos que para passeios aleatórios, se existir um número finito de operadores possíveis, com assumpções fracas o drift é analítico.

É no capitulo 5 que focamos o caso linear em dimensão infinita. O método será novamente o mesmo, procurar por estimativas de grandes desvios e provar que principio da avalanche é valido para daí obter um teorema abstracto de continuidade. Em seguida apresentamos um guião sobre como a vasta literatura sobre estimativas de grandes desvios de cociclos quasi-periódicos se adapta ao caso em dimensão infinita.

Terminamos a tese com o capitulo 6, onde descrevemos o teorema abstracto de continuidade para o drift em espaços simétricos. A ideia aqui é que os grupos de isometrias destes espaços admitem representações lineares, logo o resultado segue do resultado linear, fazendo assim um simples redução a um caso anterior.

Vários capítulos terminam com uma secção de notas bibliográficas sobre o que foi discutido ao longo do capitulo. Isto serve para mencionar trabalhos relevantes sobre o tema que não citámos anteriormente e fazer uma melhor contextualização da tese em relação que é conhecido.

Os ingredientes principais da tese são estimativas de grandes desvios, principio da avalanche e o teorema abstracto de continuidade. Estes irão aparecer várias vezes em variadas formas e por isso alertamos o leitor para uma possível fatiga na leitura da tese. Por outro lado isto mostra a força dos argumentos utilizados. Para reduzir este problema, nos capitulos 5 e 6 os argumentos resumem-se a reduções a casos anteriores ou feitos em alguma bibliografia. Apesar do baixo destaque dado a estes capitulos a sua importância para a motivação do trabalho da tese não pode ser desprezada, sendo assim fundamentais para completar a história deste trabalho.

Uma palavra também é necessária no que toca a pré-requisitos. A tese cobre uma vasto leque de tópicos e apresentá-los do zero não é o nosso objectivo. Assim sendo, assumimos que o leitor está familiarizado com tópicos de teoria ergódica e sistemas dinâmicos, algebra multilinear em espaços de Hilbert, geometria Riemanniana, grupos de Lie e espaços simétricos.

A última nota para o leitor trata o facto de que alguns teorema vêm seguidos de uma referência. Isto acontece quando ou não foi incluída uma prova para o resultado, ou a prova segue as linhas da referência citada sem alterações significativas.

Palavras-chave: estimativas de grandes desvios, teoria ergódica multiplicativa, teorema abstracto de continuidade, principio da avalanche

Contents

Chapter 1

Introduction

A topic that naturally appears in many areas of mathematics is that of a product of random operators, which are chosen according to some sensible rule. Realistically, we can't hope to obtain much information about the product itself as it may end up going to some notion of ”infinity”, so the typical solution has been to introduce descriptive information that encodes the asymptotic behaviour of the product. A paradigmatic example of such an encoding quantity is given by the Lyapunov exponents and Oseledets filtration which, through the Oseledets multiplicative ergodic theorem, fully govern the dynamics of random products of matrices.

Following Kingman (1968) ergodic theorem, the typical way to device the existence of such average limit quantities tracking the product has been to guarantee they are subadditive. This is the case with the top Lyapunov exponents in the linear case as well as the drift for semicontractions in metric spaces. Then we try to obtain some sort of multiplicative ergodic theorem that describes the behaviour of these quantities.

Once we are past understanding the dynamics of the product, an interesting question to consider is - what would happen to these limit quantities if we were to slightly perturb the elements of the product, either by choosing similar operators or changing the rule in which they were picked? We refer to this question as the continuity problem. This is a rich an complex problem which will be our focus throughout the text.

Despite its remarkable applications, Kingman's theorem doesn't tell us anything about the rate of convergence towards the limiting quantities. The natural way to study this rate of convergence is in measure, through large deviations estimate, which estimate the probability of being close to the limit quantity at a finite scale. Large deviations, just like continuity, pose a difficult problem as they can be hard to obtain since the rule responsible for picking the operators may be complicated to describe.

Goldstein and Schlag (2001) establish a connection between the two concepts by proving that, for Schrödinger cocycles, large deviation estimates lead to continuity. Moreover the modulus of continuity may be quantified by the strength of the deviation. Later, Duarte et al. (2016), prove this in general for linear operators of finite dimensional spaces. We call this connection the abstract continuity theorem.

In this thesis we aim at multiple problems. The main problem we shall tackle is the existence of an abstract continuity theorem for the drift in hyperbolic spaces. In other words, we will consider a random product of isometries of a Gromov hyperbolic space, consider its drift, and see if the same type of results hold. Then we concern ourselves with trying to look for other metric spaces where this type of results hold. A natural first place to look are symmetric spaces of noncompact type as these include cases of zero curvature whilst providing a myriad of tools to our study. In this direction we obtain a reduction of the problem to the heavily studied linear setting.

A pillar step in the argument associated with the abstract continuity is the ability to push local and finite scale controls forward in time to obtain results of global nature. At the core of this argument is the Avalanche Principle. With that, one of the novelties of the thesis is obtaining an avalanche principle for strongly hyperbolic spaces and extending the linear avalanche principle of Duarte and Klein to Hilbert spaces. To obtain these types of avalanche principles we will focus on geometric tools, thus the name of the thesis, ”Geometric Avalanche Principle”.

We dedicate chapter 2 to present Gromov hyperbolic spaces. These spaces appear as a fundamental tool in geometric group theory as they yield a systematic way to use geometry to obtain results of algebraic nature about a big class groups. Hyperbolic groups, as they are called, have been object of study since the late 1980’s and continue to inspire a lot of research. In this text we study the more general problem of a group, not necessarily hyperbolic, acting on a hyperbolic space. In this chapter we also briefly discuss dynamics and ergodicity, thus explaining what we meant by ”sensible” rule of choosing the operators. Of special importance here is the dynamics in hyperbolic spaces, which is governed by a multiplicative ergodic theorem which we introduce.

In chapter 3 we present the abstract continuity theorem for the drift on hyperbolic spaces. We obtain this result through an inductive argument based on the existence of large deviation estimates and a tool which quantifies the loss in passage from local to global analysis of the terms of a product. This tool allows us to transport quantities further in time at the cost of a quantified loss. The same argument is then used to obtain a criteria for the positivity of the drift.

We then proceed to chapter 4, where we obtain large deviation estimates for the drift in Markov systems, thus allowing us to obtain continuity. We also do a reinterpretation and simplified proof of the result for the case of random walks. We will use Markov systems as our main class of examples for the thesis.

Chapter 5 is dedicated to take a closer look at the infinite dimensional linear setting. The method will once again be the same, look for large deviation estimates and avalanche principle, and then proceed to obtain an abstract continuity result. As an application we succinctly present how to obtain large deviations estimates for the quasi-periodic setting.

We end with chapter 6, where we obtain and describe an abstract continuity theorem for the drift in the case of symmetric spaces. The idea here is that the isometry groups of symmetric spaces admit linear representations, so the result should follow from the linear one. Thus most of the chapter is based on representations so we can do this reduction.

Some chapters finish with a section on bibliographic notes on what was discussed. This serves to mention relevant works on the subject which we didn't find a way to cite previously without breaking the flow of the text.

As the reader could notice by now, the three main ingredients of the thesis are large deviation estimates, the avalanche principle and the abstract continuity theorem. These will appear repeatedly in multiple form so be ready for some possible exhaustion. We did our best to combat this problem by reducing some proofs to previous case and attempting to clean the presentation whenever possible. This makes chapters 5 and 6 a lot less developed from a point of view of examples and rigor, as there we explore what is known from previous chapters and other texts. On the other hand, we feel these chapters are essential to tell the story of the thesis.

Another word of caution we give the reader concerns the array of topics we discuss in the thesis. To fully present these subjects from the ground up would neither be feasible nor a goal of this work. Thus we assume the reader has some familiarity with ergodic theory (see Viana and Oliveira (2016), Viana (2014)), multilinear algebra in Hilbert spaces (see Temam (2012)), Riemannian geometry, Lie groups and symmetric spaces (see Helgason (2001)). We will recall these references during the text when they are relevant.

The last note to the reader is that in some theorems and propositions we present a reference to where the result may also be found. We do this when either we didn't include a proof, thus pointing the reader to a place where it can be found, or when we include a proof which boils down to a rewriting of the cited one.

We finish this introduction by mentioning that some of the material in the thesis can be found on ArXiV in Sampaio (2021) and Sampaio (2022) and is already submitted to peer review.

Chapter 2

Geometry and Dynamics on Gromov Hyperbolic Spaces

The theory of negatively curved spaces dates back to the first half of the 19th century, with the advent of hyperbolic geometry. Since then the theory immensely grew in both depth and applications. One of the central topics in the study of negatively curved spaces, dating back to Poincaré, is understanding the behaviour of isometric actions on such spaces, namely when the group responsible for the action is discrete. For example, such is the case with Fuchsian and Kleinian groups, which are discrete subgroups of isometries of $\mathbb{H}^2$ and $\mathbb{H}^3$, respectively.

During the twentieth century, new tools from differential and algebraic geometry allowed for further developments in the theory. Despite its importance, differential geometry can be too restrictive by requesting that our space has some differentiable structure. The realization of this fact led to an attempt at reconstructing most of the classical Riemannian theory from a purely metric point of view. One of the pioneers of this work was Alexandrov who reintroduced concepts like angle, length and curvature among others.

Alexandrov's notion of CAT(k) space (see Bridson and Haefliger (2013)), which in simplistic terms is a space whose curvature is bounded above by k, would come to prominence in the 1980's at the hand of Gromov, who used CAT(0) spaces to describe the global geometric properties of such spaces as well as the groups acting isometrically on them. It was known by Gromov's time that groups acting cocompactly on negatively curved manifolds had nice properties; for example, the word problem was solvable in them. Gromov thus attempted to identify directly such groups independently of the action, which led him to propose another class of negative curved spaces which are now coined Gromov hyperbolic spaces.

In this chapter we do a light presentation on a Gromov hyperbolic spaces; from definitions to basic properties to more intricate results we will need later. For a more thorough exposition on the topics check Väisälä (2005) and Das et al. (2017).

2.1 Generalities

Let X be a metric space, define the Gromov product in X as

$$\langle x\,,\,z\rangle_y := \frac{1}{2}\left(d(x,y) + d(z,y) - d(x,z)\right) \qquad \forall x,y,z \in X.$$

Definition 1 (Hyperbolic space)**.** We say that X is a Gromov hyperbolic space, or δ-hyperbolic space when we want to make δ evident, if for every x, y, z and w in X,

$$\langle x\,,\,z\rangle_w \geq \min\{\langle x\,,\,y\rangle_w\,,\,\langle y\,,\,z\rangle_w\} - \delta. \tag{2.1}$$

We call (2.1) the 4-point condition of hyperbolicity or Gromov's inequality.

We say that a group is word-hyperbolic if its Cayley graph is a Gromov hyperbolic space. Riemannian manifolds with pinched negative sectional curvature, or more generally CAT(-1) spaces, Cayley metrics on word-hyperbolic groups, Green metrics on word-hyperbolic groups and finitely punctured oriented surfaces are examples of Gromov hyperbolic spaces. A classical theorem (Theorem 3.3.7 in Das et al. (2017)) states that, in some sense, almost every finitely generated group is word-hyperbolic.

A metric space X is said to be geodesic if for every two points x and y in X, there exists an isometric embedding $\gamma : [0, d(x,y)] \to X$ connecting x to y. Throughout the text X will denote a separable, geodesic although not necessarily proper hyperbolic metric space. For geodesic spaces, Gromov hyperbolicity has more geometric flavour (Proposition 4.3.1 in Das et al. (2017)): X is δ-hyperbolic if there exists $\delta > 0$ such that for every triangle in X, any side is contained in a 3δ-neighbourhood of the other two, in other words, geometrically, triangles are thin.

In this thesis we will be interested in studying the behaviour of sequences approaching infinity in X, the natural way to deal with this convergence problem is to consider boundaries. Fortunately, hyperbolic spaces carry a natural boundary, called the Gromov boundary, which we now present.

We say that a sequence (x_n) in a hyperbolic space X with basepoint x_0 is a Gromov sequence if $\langle x_n\,,\,x_m\rangle_{x_0}$ tends to infinity as m and n tend to infinity. Two Gromov sequences (x_n) and (y_n) are equivalent, $(x_n) \sim (y_n)$, if $\langle x_n\,,\,y_n\rangle_{x_0}$ tends to infinity as n tends to infinity. Gromov's inequality implies that this is an equivalence relation. The Gromov boundary, denoted by ∂X, is the set of equivalence classes of Gromov sequences.

The Gromov product in X may be extended to its Gromov boundary: given $\xi, \eta \in \partial X$ and $y, z \in X$, let

$$\langle \xi \, , \, \eta \rangle_z := \inf \left\{ \liminf_{n,m \to \infty} \langle x_n \, , \, y_m \rangle_z \; : \; (x_n) \in \xi, \; (y_m) \in \eta \right\},$$
$$\langle x \, , \, \xi \rangle_z = \langle \xi \, , \, x \rangle_z := \inf \left\{ \liminf_{n \to \infty} \langle x_n \, , \, x \rangle_z \; : \; (x_n) \in \xi \right\}.$$

Denote by Bord$X$ the set $X \cup \partial X$. Given $1 < b \leq 2^{\frac{1}{\delta}}$ and $x \in X$ consider the symmetric map $\rho_{x,b}$: Bord$X$ $\times$ Bord$X \to \mathbb{R}$ given by

$$\rho_{x,b}(\xi \, , \, \eta) = b^{-\langle \xi \, , \, \eta \rangle_x}.$$

Definition 2 (Strongly hyperbolic space)**.** We say that a hyperbolic space X is a strongly hyperbolic space if there exists $1 < b < 2^{1/\delta}$ such that for every $x \in X$ the map $\rho_{x,b}$ satisfies the triangle inequality, in particular $\rho_{x,b}$ defines a metric in ∂X.

In general, $\rho_{x,b}$ doesn't have to be a metric in ∂X, however the Gromov boundary is still metrizable. Fix $x_0 \in X$ and denote by $\rho_b := \rho_{x_0,b}$. Using the Gromov inequality, for every $\xi, \eta, \zeta \in$ Bord$X$

$$\rho_b(\xi \, , \, \eta) \leq 2 \max\{\rho_b(\xi \, , \, \zeta), \rho_b(\zeta \, , \, \eta)\}.$$

By the classic Frink metrization theorem (see Frink (1937)), the map $\bar{D}_b$: Bord$X$ $\times$ Bord$X \to \mathbb{R}$ given by

$$\bar{D}_b(\xi \, , \, \eta) = \inf \sum_{i=0}^{n-1} \rho_b(\xi_i \, , \, \xi_{i+1})$$

where the infimum is taken over finite sequences of points ξ_i such that $\xi_0 = \xi$ and $\xi_n = \eta$, satisfies the triangle inequality. Moreover the following visual condition holds:

$$\rho_b(\xi \, , \, \eta)/4 \leq \bar{D}_b(\xi \, , \, \eta) \leq \rho_b(\xi \, , \, \eta) \text{ for every } \xi, \eta \in \partial X. \tag{2.2}$$

Using $\bar{D}_b$ we are now able to contruct a metric for the whole Bord$X$.

Proposition 1 (in Das et al. (2017))**.** *For every $\xi, \eta \in$ Bord$X$ let*

$$D_b(\xi, \eta) := \min\left\{(\log b) d(\xi \, , \, \eta) \, ; \, \bar{D}_b(\xi \, , \, \eta)\right\},$$

using the convention $d(\xi, \eta) = \infty$ if either ξ or η belong to ∂X and $\xi \neq \eta$. Then D_b is a metric in $BordX$ inducing in X the same topology as the metric d. Moreover, if X is complete then so is Bord$X$.

Proof. All conditions are immediate except for the triangle inequality. If $D_b(\xi, \eta) = (\log b)d(\xi, \eta)$ and $D_b(\eta, \zeta) = (\log b)d(\eta, \zeta)$, then

$$D_b(\xi, \zeta) \leq (\log b)d(\xi, \zeta) \leq (\log b)d(\xi, \eta) + (\log b)d(\eta, \zeta) \leq D_b(\xi, \eta) + D_b(\eta, \zeta),$$

and the same holds if $D_b(\xi, \eta) = \bar{D}_b(\xi, \eta)$ and $D_b(\eta, \zeta) = \bar{D}_b(\eta, \zeta)$.

So we focus on the case where the minimums are given by different expressions. Without loss of generality, assume $D_b(\xi, \eta) = (\log b)d(\xi, \eta)$ and $D_b(\eta, \zeta) = \bar{D}_b(\eta, \zeta)$, fix $\varepsilon > 0$ and take $\eta = y_0, y_1, ...y_n = \zeta$, a sequence such that

$$\sum_{i=0}^{n-1} b^{-\langle y_i\,,\,y_{i+1}\rangle_{x_0}} \leq \bar{D}_b(\eta, \zeta) + \varepsilon.$$

Consider also another sequence $x_i = y_i$ for $i > 0$ and $x_0 = \xi$. Now we use the immediate inequalities $b^{-t} \leq s \log b + b^{-(t+s)}$ for every $t, s > 0$ and $\langle x, y\rangle_{x_0} \leq \langle x, z\rangle_{x_0} + d(y, z)$ for every $x, y, z \in \mathrm{Bord}X$, to obtain

$$b^{-\langle \xi, y_1\rangle_{x_0}} \leq \log(b)d(\xi, \eta) + b^{-\langle \eta, y_1\rangle_{x_0}},$$

were we used $t = \langle \xi, y_1\rangle_{x_0}$ and $s = d(\xi, \eta)$. Hence

$$\begin{aligned} D_b(\xi, \zeta) \leq \bar{D}_b(\xi, \zeta) &\leq \sum_{i=0}^{n-1} b^{\langle x_i\,,\,x_{i-1}\rangle_{x_0}} \\ &= b^{-\langle \xi\,,\,y_1\rangle_{x_0}} + \sum_{i=1}^{n-1} b^{\langle y_i\,,\,y_{i-1}\rangle_{x_0}} \\ &\leq (\log b)d(\xi, \eta) + b^{-\langle \xi\,,\,y_1\rangle_{x_0}} + \sum_{i=1}^{n-1} b^{\langle y_i\,,\,y_{i-1}\rangle_{x_0}} \\ &\leq (\log b)d(\xi, \eta) + \bar{D}_b(\eta, \zeta)\varepsilon \\ &= D_b(\xi, \eta) + D_b(\eta, \zeta) + \varepsilon. \end{aligned}$$

Making ε as small as we want we obtain the intended result.

Let (x_n) be a converging sequence in X to some x with respect to d, then there exists an order $p \in \mathbb{N}$ after which $D_b(x_n, x) = (\log b)d(x_n, x)$, so (x_n) converges for the metric D_b. The reverse process implies a sequence converging in X with respect to D_b converges with respect to d. Hence the two metric yield the same converging sequences in X, so they induce the same topology.

Finally, suppose X is complete and let (x_n) be a Cauchy sequence. Then, for example by Ramsey's theorem, (x_n) is Cauchy for either $\bar{D}_b$ or d. If (x_n) is Cauchy with respect to d, then it converges since X is complete. If (x_n) is Cauchy with respect with $\bar{D}_b$, then (x_n) is a Gromov sequence, so it also converges. □

Example 1. Notice that Bord$X$ is bounded as $D_b \leq 1$. Typically one actually wants the boundary of a space to be compact. Unfortunately that is not always the case. Consider $X \subset \mathbb{R}^2$ a set given by countably many half-lines emanating from the origin in $\mathbb{R}^2$. Given $x, y \in X$, consider the distance

$$d(x,y) = \begin{cases} ||x-y|| & \text{, if } x \text{ and } y \text{ belong to the same half-line} \\ ||x|| + ||y|| & \text{, otherwise.} \end{cases}$$

This space is a tree, hence it is 0-hyperbolic. Notice however that the Gromov boundary of X is $\mathbb{N}$ with the discrete topology, which isn't compact. This is a consequence of the fact that locally compactness fails at the origin.

Proposition 2 (Proposition 3.4.18 in Das et al. (2017))**.** *The metric space* (BordX, D_b) *is compact if and only if X is proper.*

As a final note, the action of Isom(X), the group of isometries of X, on X extends to an action by homeomorphisms on Bord$X$ by taking every $(x_n) \in \xi \in \partial X$ to $(gx_n) \in g\xi \in \partial X$. Hence we equip Isom$(X)$ with the topology of uniform convergence tracking its behaviour in Bord$X$. With effect, given $1 < b \leq 2^{1/\delta}$, take

$$d_G(g_1, g_2) := \max\left\{ \sup_{\xi \in \text{Bord}X} D_b(g_1\xi, g_2\xi)\,;\, \sup_{\xi \in \text{Bord}X} D_b(g_1^{-1}\xi, g_2^{-1}\xi) \right\},$$

for every $g_1, g_2 \in$ Isom(X). We will prove that Isom(X) is a topological group when equipped with d_G. In particular, d_G is a distance.

2.2 Horofunction Compactification

Let X be a Gromov hyperbolic space. From the visual condition (2.2) $D_b \leq 1$, in particular, Bord$X$ is a bounded space when equipped with this metric. The main drawback of this construction is that in general Bord$X$ does not have to be compact. To remedy this problem we will construct another boundary of X and relate it to ∂X.

The horofunction compactification, which we present shortly, was also introduced by Gromov and can be considered for every metric space. With that said we will present the general construction and then explore the properties specific to hyperbolic spaces.

2.2.1 Construction

Let M be a separable metric space with basepoint x_0 (up to homeomorphism the construction will be independent of the choice of x_0). Consider the injective map

$$\rho : M \to C(M)$$

$$x \mapsto h_x(\cdot) = d(\cdot, x) - d(x, x_0).$$

where $C(M)$ stands for the set of continuous functions in M. Throughout the text we will also make use of the forms

$$h_x(z) = d(z, x_0) - 2\langle z\,,\, x\rangle_{x_0} \tag{2.3}$$
$$= \langle x\,,\, x_0\rangle_z - \langle x\,,\, z\rangle_{x_0}. \tag{2.4}$$

Notice that h_x are all 1-Lipschitz and satisfy $h_x(x_0) = 0$.

Endow the space $C(M) \subset \mathbb{R}^M$ with the product topology, that is the topology of pointwise convergence, which is equivalent to the compact-open topology. Then, using the triangle inequality one has

$$-d(z, x_0) \leq h_x(z) \leq d(z, x_0),$$

hence $\rho(M)$ may be identified with a subset of $\Pi_{z\in X}[-d(z, x_0), d(z, x_0)]$ which, by Tychonoff's theorem, is compact for the product topology. Therefore the closure $\overline{\rho(X)} =: M^h$ will be a compact set called the horofunction compactification of X. The elements in M^h are called horofunctions of M.

Proposition 3. *The horofunction compactification is compact, Hausdorff and second countable (hence metrizable).*

Proof. By hypothesis, M is a separable metric space, hence Hausdorff and second countable. Since $\mathbb{R}$ is also Hausdorff and second countable, so is $C(M)$ (for the compact-open topology). However for the subspace of $1-$Lipschitz functions normalized by taking the value 0 at x_0, the compact-open topology and the topology of pointwise convergence agree. □

If M is a proper space then the pointwise convergence coincides with the uniform convergence on compact sets from the usual construction of the horofunction compactification, in this case M^h contains M as an open and dense set. In the nonproper case, the image $\rho(M)$ may not be open in $\overline{\rho(M)}$, so we may not have a compactification in the usual sense.

Proposition 4. *Let* Isom(M) *be the group of isometries of* M*. Then the action of* Isom(M) *in* M *extends to an action by homeomorphisms on* M^h*, defined by*

$$g \cdot h(z) := h(g^{-1}z) - h(g^{-1}x_0), \tag{2.5}$$

for $g \in$ Isom(M) *and* $h \in M^h$.

Proof. We naturally transport the action on M to the action on $\rho(M)$ via $g \cdot h_x = h_{gx}$. Then

$$\begin{aligned} g \cdot h_x(z) &= h_{gx}(z) \\ &= d(z, gx) - d(gx, x_0) \\ &= d(g^{-1}z\, ,\, x) - d(x\, ,\, g^{-1}x_0) \\ &= h_x(g^{-1}z) - h_x(g^{-1}x_0), \end{aligned}$$

which we can transport to the whole of M^h. It is immediate that if $h_n \to h$ pointwisely, then $g \cdot h_n \to g \cdot h$ pointwisely, whence the action of Isom(M) on M^h is continuous. □

In this proof we set the notation $g \cdot h_x = h_{gx}$ which we will use henceforth without mention.

2.2.2 Horofunctions of a Gromov Hyperbolic Space

We will now explore the special properties of horofunctions in Gromov hyperbolic spaces. Let X once again be a Gromov hyperbolic space with basepoint x_0. We start by partitioning the horofunction compactification X^h in two: its finite part

$$X^h_F := \{h \in X^h \,:\, \text{inf}(h) > -\infty\}$$

and its infinite part

$$X^h_\infty := \{h \in X^h \,:\, \text{inf}(h) = -\infty\}$$

Both X^h_F and X^h_∞ are invariant for the action of Isom(X) on X^h. Clearly one has $\rho(X) \subset X^h_F$ and, in well behaved cases, one may actually get the equality. Another important remark to make concerns the fact that X^h_∞ need not be compact. Let us look again at example 1.

Example 2. Recall $X \subset \mathbb{R}^2$ is a set given by countably many half-lines emanating from the origin in $\mathbb{R}^2$. Given $x, y \in X$, consider the distance

$$d(x,y) = \begin{cases} ||x-y|| & \text{, if } x \text{ and } y \text{ belong to the same half-line} \\ ||x|| + ||y|| & \text{, otherwise.} \end{cases}$$

The sequence of horofunctions h_n given as a limit of sequences h_{x_k}, where x_k is the point at distance k from the origin in the n-th ordered half-line. Let y be a point in n-th half line, then $h_n(y) = -||y||$ whose infimum is $-\infty$, that is, $h_n \in X^h_\infty$. However that $h_n \to h_{x_0}$ which belongs to X^h_F, hence X^h_∞ is not compact.

Having access to two different boundaries of the space X we need to understand how the two interact with one another. For our purposes once we relate both boundaries we work with X^h when compactness is needed and ∂X if the problem requires a metric. We will start relating X^h and ∂X by proving that there exists a continuous, surjective $G-$equivariant map between X^h_∞ and ∂X. With that in mind, we will notice that horofunctions in X^h_∞ arise as limits of sequences of horofunctions (h_{x_n}) where (x_n) is Gromov.

Lemma 5 (Lemma 3.8 in Maher and Tiozzo (2018))**.** *Let X be a δ-hyperbolic space with basepoint x_0. Then for every horofunction $h \in X^h$ and points $x, y \in X$, the following inequality holds:*

$$\langle x\,,\, y\rangle_{x_0} \geq \min\{-h(x)\,,\, -h(y)\} - \delta,$$

moreover, for every $z \in X$,

$$\langle x\,,\, y\rangle_{x_0} \geq \min\{-h_x(z)\,,\, -h_y(z)\} - \delta.$$

Proof. Let $z \in X$. Using the triangle inequality one has

$$\begin{aligned}\langle x\,,\, z\rangle_{x_0} &= \frac{1}{2}(d(x_0,x) + d(x_0,z) - d(x,z)) \\ &\geq d(x_0,z) - d(x,z) \\ &= -h_z(x). \end{aligned} \tag{2.6}$$

Now, from the definition of hyperbolicity

$$\begin{aligned}\langle x\,,\, y\rangle_{x_0} &\geq \min\{\langle x\,,\, z\rangle_{x_0}\,,\, \langle z\,,\, y\rangle_{x_0}\} - \delta \\ &\geq \min\{-h_z(x)\,,\, -h_z(y)\} - \delta.\end{aligned}$$

The claim follows from the fact that every horofunction is the pointwise limit of functions of the form h_z.

The second inequality is analogous using $\langle x\,,\, z\rangle_{x_0} \geq -h_x(z)$. □

Lemma 6. *Let $h \in X^h_\infty$ be an horofunction and (x_n) a sequence such that $h_{x_n} \to h$ and (y_n) a sequence such that $h(y_n) \to -\infty$. Then the sequences (x_n) and (y_n) are Gromov.*

Proof. Using the first inequality in Lemma 5,

$$\lim_{n,m\to\infty} \langle y_n\,,\, y_m\rangle_{x_0} \geq \lim_{n,m\to\infty} \min\{-h(y_n)\,,\, -h(y_m)\} - \delta = +\infty,$$

hence (y_n) is Gromov.

Using Lemma 5 again, for every $z \in X$

$$\langle x_n \, , \, x_m \rangle_{x_0} \geq \min\{-h_{x_n}(z), -h_{x_m}(z)\} - \delta.$$

Taking the limit as m, n go towards infinity yields

$$\lim_{n,m\to\infty} \langle x_n \, , \, x_m \rangle_{x_0} \geq -\inf_{z\in X} h(z) - \delta = +\infty.$$

□

Proposition 7. *Let $h \in X^h_\infty$ be an horofunction. Let (x_n) and (y_n) be Gromov sequences such that $h_{x_n} \to h$ and $h(y_n) \to -\infty$, respectively. Then (x_n) and (y_n) are Gromov sequences and $(x_n) \sim (y_n)$, in particular all sequences (y_n) such that $h(y_n) \to -\infty$ converge to the same boundary point.*

Proof. All that is left to prove is that (x_n) and (y_n) are equivalent. Using Gromov's inequality together with Lemma 6 and (2.6)

$$\begin{aligned}\langle x_n \, , \, y_n \rangle_{x_0} &\geq \min\{\langle x_n \, , \, x_m \rangle_{x_0}, \langle x_m \, , \, y_n \rangle_{x_0}\} - \delta \\ &\geq \min\{\langle x_n \, , \, x_m \rangle_{x_0}, -h_{x_m}(y_n)\} - \delta.\end{aligned}$$

Taking the iterated limits towards infinity, we obtain

$$\begin{aligned}\lim_{n\to\infty} \langle x_n \, , \, y_n \rangle_{x_0} &\geq \lim_{n\to\infty} \liminf_{m\to\infty} \min\{\langle x_n \, , \, x_m \rangle_{x_0}, -h_{x_m}(y_n)\} - \delta \\ &= \min\left\{\lim_{n\to\infty} \liminf_{m\to\infty} \langle x_n \, , \, x_m \rangle_{x_0}, \lim_{n\to\infty} -h(y_n)\right\} - \delta \\ &= +\infty.\end{aligned}$$

□

Proposition 7 motivates the following definition:

Definition 3 (Local Minimum Map)**.** Define the local minimum map $\phi : X^h_\infty \to \partial X$ given by

$$\phi(h) = \lim_{n\to\infty} y_n = \xi,$$

where $(y_n) \in \xi$ is such that $h(y_n) \to -\infty$.

The local minimum map is Isom(X)-equivariant, continuous and surjective. Proofs for these properties of the local minimum map can be found in Maher and Tiozzo (2018), and we will include them as well for the sake of completeness.

Proposition 8 (Lemma 3.12, Proposition 3.14, Corollary 3.15 in Maher and Tiozzo (2018))**.** *The local minimum map $\phi : X^h_\infty \to \partial X$ is*

1. Isom$(X)-$*equivariant;*
2. *continuous;*
3. *surjective.*

Proof. **1.** Let $h \in X^h_\infty$, take (y_n) such that $h(y_n) \to -\infty$ as $n \to \infty$. Then, by definition of the action, one has

$$g \cdot h(gy_n) = h(y_n) - h(g^{-1}x_0) \to -\infty$$

hence $\phi(g \cdot h) = g\phi(h)$.

2. Let (h_n) be a sequence of horofunctions in X^h_∞ converging to some $h \in X^h_\infty$. Pick sequences (x_m) and $(y_{m,n})$ such that $h(x_m) \to -\infty$ and $h_n(y_{m,n}) \to -\infty$ as $m \to \infty$. We will now prove that $\lim_{n\to\infty}\langle \phi(h_n), \phi(h)\rangle_{x_0} = \infty$, which implies that $\lim_{n\to\infty} \phi(h_n) = \phi(h)$.

Let $N > 0$. Since (x_m) is Gromov and $h(x_m) \to -\infty$, there exists m_0 such that

$$h(x_{m_0}) \leq -N - 1$$

and for every $m, m' > m_0$

$$\langle x_m, x_{m'}\rangle_{x_0} \geq N + 1$$

Due to h being a pointwise limit, there exists $m_1 = m_1(N, n)$ such that for every $m \geq m_1 \geq m_0$

$$h_n(x_m) \leq -N$$

for n large enough. However, by definition of $y_{m,n}$, the same remains true for $h_n(y_{m,n})$. Hence by Lemma 5 we have

$$\langle x_{m_0}, y_{m,n}\rangle_{x_0} \geq \min\{-h_n(x_m), -h_n(y_{m,n})\} \geq N,$$

and using Gromov's inequality, for every $m, m' > m_1$

$$\langle x_{m'}, y_{m,n}\rangle_{x_0} \geq \min\{\langle x_{m_0}, y_{m,n}\rangle_{x_0}\ ,\ \langle x_m, x_{m'}\rangle_{x_0}\} - \delta \geq N - \delta.$$

Thus $\langle \phi(h_n), \phi(h)\rangle_{x_0} = \sup \liminf_{m,m'}\langle x_{m'}, y_{m,n}\rangle_{x_0} \geq N - \delta$, which concludes the proof of point 2.

3. Let $\xi \in \partial X$ and take $(x_n) \in \xi$. Due to compactness, h_{x_n} admits, up to a converging subsequence, a limit $h \in X^h$. Since $\inf h_{x_n} \leq h_{x_n}(x_n) = -d(x_0, x_n) \to -\infty$, h must in fact belong to X^h_∞. Finally, by Proposition 7, we can see that $\phi(h) = \xi$. □

The previous proposition points out that ϕ behaves like a quotient map. Consider in X^h the equivalence relation $h_1 \sim h_2$ if $\sup_{z\in X} |h_1(z) - h_2(z)|$ is finite. We dedicate the rest of this section to proving that the pre-images of the local minimum map are exactly the equivalence classes of $\sim$ in X^h_∞. As a consequence we will obtain that $X^h_\infty / \sim$ is homeomorphic to the Gromov boundary. This homeomorphism is a well known result for proper spaces (see Coornaert and Papadopoulos (2001)). In the nonproper cases, a mention of the result can be found in Maher and Tiozzo (2018), although a proof is not known to the author. We will not use this homeomorphism, although we believe it may help the reader with a better understanding on hyperbolic spaces.

Let us begin by associating each element ξ in ∂X with a function. Then we proceed to relate such functions with the limit horofunctions of h_{x_n}, where $(x_n) \in \xi$.

Definition 4 (Busemann Function)**.** Given $\xi \in \partial X$, we define the Busemann function associated with ξ as

$$B_\xi(z) = \langle \xi \, , \, x_0 \rangle_z - \langle \xi \, , \, z \rangle_{x_0},$$

for every $z \in X$.

In the following two lemmas we explore the continuity of the Gromov product. With effect we understand its behaviour upon considering Gromov sequences as arguments. This will also allow us to understand what we gain when we pass from Gromov hyperbolic to strongly hyperbolic spaces, as the latter are very well behaved at infinity.

Lemma 9 (Lemma 3.4.7 in Das et al. (2017))**.** *Let (x_n) and (y_n) be two Gromov sequences in a δ-hyperbolic space and fix $y, z \in X$. Then*

$$\limsup_{n\to\infty} \langle x_n \, , \, y \rangle_z \leq \liminf_{n\to\infty} \langle x_n \, , \, y \rangle_z + \delta$$
$$\limsup_{n,m\to\infty} \langle x_n \, , \, y_m \rangle_z \leq \liminf_{n,m\to\infty} \langle x_n \, , \, y_m \rangle_z + 2\delta$$

With the limits existing if the space is strongly hyperbolic.

Proof. Fix $n_1, n_2 \in \mathbb{N}$. By Gromov's inequality

$$\langle x_{n_1} \, , \, y \rangle_z \geq \min \{ \langle x_{n_1} \, , \, x_{n_2} \rangle_z \, , \, \langle x_{n_2} \, , \, y \rangle_z \} - \delta.$$

Taking the $\liminf$ over n_1 and the $\limsup$ over n_2 gives

$$\begin{aligned} \liminf_{n,m\to\infty} \langle x_n \, , \, y \rangle_z &\geq \min \left\{ \liminf_{n_1,n_2\to\infty} \langle x_{n_1} \, , \, x_{n_2} \rangle_z \, , \, \limsup_{n_2\to\infty} \langle x_{n_2} \, , \, y \rangle_z \right\} - \delta \\ &= \limsup_{n\to\infty} \langle x_n \, , \, y \rangle_z - \delta. \end{aligned}$$

Where the last equality comes from (x_n) being a Gromov sequence.

The second inequality is analogous using the following inequality which is immediate from iterating the 4-point condition of hyperbolicity:

$$\langle x\,,\,w\rangle_u \geq \min\{\langle x\,,\,y\rangle_u\,,\,\langle y\,,\,z\rangle_u\,,\,\langle z\,,\,w\rangle_u\} - 2\delta.$$

As for the statement regarding strongly hyperbolic spaces, notice we have the inequality

$$b^{-\langle x_{n_1}\,,\,y_{m_1}\rangle_z} \leq b^{-\langle x_{n_2}\,,\,y_{m_2}\rangle_z} + b^{-\langle x_{n_1}\,,\,x_{n_2}\rangle_z} + b^{-\langle y_{m_1}\,,\,y_{m_2}\rangle_z},$$

so taking the lim sup in n_1, m_1 and the lim inf in n_2, m_2 gives

$$b^{-\liminf_{m,n\to\infty}\langle x_n\,,\,y_m\rangle_z} \leq b^{-\limsup_{m,n\to\infty}\langle x_n\,,\,y_m\rangle_z}$$

and by continuity of the exponential the result follows. □

As a consequence, for strongly hyperbolic spaces we have the relation $\lim_{n,m\to\infty}\langle x_n\,,\,y_m\rangle_z = \langle \xi\,,\,\eta\rangle_z$, for every $(x_n) \in \xi$ and $(y_n) \in \eta$. In other words, in strongly hyperbolic spaces, the Gromov product is continuous. For regular hyperbolic spaces we incur in a loss upon doing the limits.

Lemma 10 (Lemma 3.4.10 in Das et al. (2017))**.** *Fix $\xi, \eta \in \partial X$ and $y, z \in X$. For all $(x_n) \in \xi$ and $(y_n) \in \eta$, we have*

1. $\langle \xi\,,\,y\rangle_z - \delta \leq \liminf_{n\to\infty}\langle x_n\,,\,y\rangle_z \leq \limsup_{n\to\infty}\langle x_n\,,\,y\rangle_z \leq \langle \xi\,,\,y\rangle_z + \delta$*;*
2. $\langle \xi\,,\,\eta\rangle_z - 2\delta \leq \liminf_{n,m\to\infty}\langle x_n\,,\,y_m\rangle_z \leq \limsup_{n,m\to\infty}\langle x_n\,,\,y_m\rangle_z \leq \langle \xi\,,\,\eta\rangle_z + 2\delta$.

Proof. The two leftmost inequalities are trivial and are included for symmetry We now prove 1. as the proof of 2. is analogue. Suppose that we are given two sequences (x_n^1), $(x_n^2) \in \xi$, let

$$x_n = \begin{cases} x^1_{n/2} & \text{, if n is even} \\ x^2_{(n+1)/2} & \text{, if n is odd.} \end{cases}$$

By the hyperbolicity condition, for every n,

$$\langle x_n\,,\,x_n^1\rangle_{x_0} \geq \min\left\{\langle x_n\,,\,x_n^2\rangle_{x_0}\,,\,\langle x_n^2\,,\,x_n^1\rangle_{x_0}\right\} - \delta,$$

which yields $(x_n) \in \xi$. Applying the previous lemma to x_n implies

$$\min_{i=1,2}\limsup_{n\to\infty}\langle x_n^i\,,\,y\rangle_z \leq \max_{i=1,2}\liminf_{n\to\infty}\langle x_n^i\,,\,y\rangle_z + \delta.$$

Now

$$\liminf_{n\to\infty}\langle x_n^2\,,\,y\rangle_z - \delta \leq \limsup_{n\to\infty}\langle x_n^2\,,\,y\rangle_z - \delta$$

$$\begin{aligned}
&\le \liminf_{n\to\infty} \langle x_n^1\,,\, y\rangle_z \\
&\le \limsup_{n\to\infty} \langle x_n^1\,,\, y\rangle_z \\
&\le \liminf_{n\to\infty} \langle x_n^2\,,\, y\rangle_z + \delta.
\end{aligned}$$

Taking the inf over all $(x_n^1) \in \xi$ one obtains the statement. □

In particular, these two lemmas tell us that in strongly hyperbolic spaces, the Gromov product is continuous. Finally we show that all horofunctions in X^h_∞ are close to a Busemann function, with the two concepts agreeing for strongly hyperbolic spaces.

Theorem 11. *Fix $\xi \in \partial X$ and $z \in X$. For all $(x_n) \in \xi$ we have*

$$B_\xi(z) - 2\delta \le \liminf_{n\to\infty} h_{x_n}(z) \le \limsup_{n\to\infty} h_{x_n}(z) \le B_\xi(z) + 2\delta$$

Proof. The proof is a direct application of the previous Lemma to (2.4) and the expression of Busemann functions. □

Corollary 12. *Let $\sim$ be the equivalence relation on X^h_∞ given by*

$$h_1 \sim h_2 \;\Leftrightarrow\; \sup_{z\in X} |h_1(z) - h_2(z)| < \infty$$

Then the local minimum map descends to an homeomorphism from $X^h_\infty/\sim$ onto ∂X. Moreover, if X is strongly hyperbolic, then $\sim$ is the equality relation.

Proof. Let us recall that we already have continuity and surjectivity. Using Proposition 7 and Theorem 11 we have that the image of two horofunctions $h_1,\ h_2 \in X^h_\infty$ under the local minimum is the same if and only if $||h_1 - h_2|| < 4\delta$; hence it descends to a continuous bijection between $X^h_\infty/\sim$ and ∂X. Finally, in X^h_∞ we consider the induced topology by X^h. Since X^h is compact, any closed set in X^h_∞ is compact, hence its image in ∂X is also compact, thus closed since ∂X is Hausdorff. This implies that ϕ is closed, hence the local minimum map descends to a closed map. Therefore $X^h_\infty/\sim$ and ∂X are homeomorphic. □

In our work, we often need to work with arbitrary sequences that escape towards the Gromov boundary. Thus a bad behaviour of ∂X could make our study a lot more difficult. With that in mind we will restrict our class of hyperbolic spaces to those with a good behaviour at the boundary

Definition 5 (Basic Assumption)**.** We say that a Gromov hyperbolic space satisfies the basic assumption (BA) if the local minimum map is a homeomorphism.

Notice that all strongly hyperbolic cases satisfy (BA), so it seems we have lost word-hyperbolic groups. This is not known however, since it remains as an open problem to know if every hyperbolic group admits a group of generators for which its Cayley graph satisfies (BA) (see Section 4 of Gilch and Ledrappier (2013)).

If X satisfies (BA) we set the notation h_ξ whenever $\phi(h_\xi) = \xi \in \partial X$, where ϕ stands for the local minimum map. Notice that this notation extends the previously used notation h_x for points in X, thus obtaining a consistent notation in Bord$X$.

We finish this subsection by proving a result that should be thought of as saying that the operator norm $||A||_{op}$ with $A \in \mathrm{SL}(2,\mathbb{R})$ and the norm $||A||_1 = \max\{||Ae_1||, ||Ae_2||\}$, where $\{e_1, e_2\}$ stands for the canonical basis, are equivalent.

Lemma 13. *Let X be a hyperbolic space satisfying (BA) and $\xi \neq \eta \in \partial X$. Take h_1 and h_2 such that $\phi(h_1) = \xi$ and $\phi(h_2) = \eta$, then for every $g \in \mathrm{Isom}(X)$ there exists a constant $K(\delta, \xi, \eta)$ depending on the hyperbolicity constant δ and the points ξ and η such that*

$$\max_{i=1,2} h_i(gx_0) \leq d(gx_0, x_0) \leq \max_{i=1,2} h_i(gx_0) + K(\delta, \xi, \eta).$$

Proof. For every $g \in G$

$$d(gx_0, x_0) = h_{y^i_m}(gx_0) + 2\langle y^i_m \, , \, gx_0\rangle_{x_0}.$$

Let (y^i_m) be Gromov sequences such that $h_{y^i_m} \to h_i$ for $i = 1, 2$ whence, using the Gromov inequality

$$\begin{aligned} d(gx_0, x_0) &= \max_{i=1,2} h_{y^i_m}(gx_0) + 2 \min_{i=1,2} \langle y^i_m \, , \, gx_0\rangle_{x_0} \\ &\leq \max_{i=1,2} h_{y^i_m}(gx_0) + 2\langle y^1_m \, , \, y^2_m\rangle_{x_0} + 2\delta \end{aligned}$$

By Lemma 6 and Proposition 7 we know (y^1_m) and (y^2_m) are not equivalent, so taking the inferior limit in m one obtains

$$\max_{i=1,2} h_i(gx_0) \leq d(gx_0, x_0) \leq \max_{i=1,2} h_i(gx_0) + K(\delta, \xi, \eta),$$

for a constant $K(\delta, \xi, \eta) = 2\langle \xi \, , \, \eta\rangle_{x_0} + 4\delta$. ∎

2.3 Group of Isometries

Some of the nice properties from hyperbolic spaces pass to their group of isometries $\mathrm{Isom}(X)$. In particular we will prove in this section that $\mathrm{Isom}(X)$ is a topological group. We start by noticing how $\mathrm{Isom}(X)$ behaves on the boundary. The results in this chapter can be found in Sampaio (2021).

Proposition 14. *Let X be a Gromov hyperbolic space $g \in \mathrm{Isom}(X)$ and $\xi, \eta \in \mathrm{Bord}X$, and take h_1 and h_2 such that $\phi(h_1) = \xi$ and $\phi(h_2) = \eta$ then*

$$\frac{1}{C(\delta)} b^{-\frac{1}{2}\left[h_1(g^{-1}x_0)+h_2(g^{-1}x_0)\right]} \leq \frac{\bar{D}_b(g\xi \,,\, g\eta)}{\bar{D}_b(\xi \,,\, \eta)} \leq C(\delta) b^{-\frac{1}{2}\left[h_1(g^{-1}x_0)+h_2(g^{-1}x_0)\right]},$$

where $C(\delta) = 4b^{6\delta}$. Moreover, $C(\delta) = 1$ if X is strongly hyperbolic.

Proof. We start by using the visual condition (2.2)

$$\frac{1}{4} \frac{\rho_b(g\xi \,,\, g\eta)}{\rho_b(\xi \,,\, \eta)} \leq \frac{D_b(g\xi \,,\, g\eta)}{D_b(\xi \,,\, \eta)} \leq 4 \frac{\rho_b(g\xi \,,\, g\eta)}{\rho_b(\xi \,,\, \eta)}.$$

Using the definition of Gromov product and some computations yields

$$\langle x \,,\, y \rangle_z - \langle x \,,\, y \rangle_{x_0} = \frac{1}{2}\big(\langle x \,,\, x_0 \rangle_z - \langle x \,,\, z \rangle_{x_0} + \langle y \,,\, x_0 \rangle_z - \langle y \,,\, z \rangle_{x_0}\big).$$

Take $(x_n) \in \xi$ and $(y_n) \in \eta$. Substituting in the equality above x by x_n, y by y_n and $z = g^{-1}x_0$ and taking limits, by Lemmas 9 and 10 we obtain

$$\frac{1}{2}\big(h_1(g^{-1}x_0) + h_2(g^{-1}x_0)\big) - 6\delta \leq \langle \xi \,,\, \eta \rangle_{g^{-1}x_0} - \langle \xi \,,\, \eta \rangle_{x_0} \leq \frac{1}{2}\big(h_1(g^{-1}x_0) + h_2(g^{-1}x_0)\big) + 6\delta.$$

Finally one has

$$\frac{\rho_b(g\xi \,,\, g\eta)}{\rho_b(\xi \,,\, \eta)} = b^{-\left(\langle g\xi \,,\, g\eta \rangle_{x_0} - \langle \xi \,,\, \eta \rangle_{x_0}\right)} = b^{-\left(\langle \xi \,,\, \eta \rangle_{g^{-1}x_0} - \langle \xi \,,\, \eta \rangle_{x_0}\right)},$$

and the result follows. □

Recall now that we set a metric in $\mathrm{Isom}(X)$ by the expression

$$d_G(g_1, g_2) := \max\left\{ \sup_{\xi \in \mathrm{Bord}X} D_b(g_1\xi, g_2\xi) \,;\, \sup_{\xi \in \mathrm{Bord}X} D_b(g_1^{-1}\xi, g_2^{-1}\xi) \right\}.$$

As we alluded to, we now prove the following theorem

Theorem 15. *Let X be a Gromov hyperbolic space and $\mathrm{Isom}(X)$ its group isometries, then d_G is a metric in $\mathrm{Isom}(X)$. Moreover $(\mathrm{Isom}(X), d_G)$ is a topological group.*

Proof. Since D_b is a metric, we are left with proving that $d_G(g_1, g_2) = 0$ implies that $g_1 = g_2$. Suppose $d_G(g_1, g_2) = 0$ and let $x \in X$. Notice g_1x and g_2x are both in X so

$$\bar{D}_b(g_1x, g_2x) \geq \rho_b(g_1x, g_2x)/4 > 0$$

hence $D_b(g_1x, g_2x) = \log(b)d(g_1x, g_2x)$. Therefore $d(g_1x, g_2x) = 0$ for every $x \in X$, implying $g_1 = g_2$.

All that remains is to see that the map $(g, g') \mapsto g^{-1}g'$ is continuous. This will follow from a series of inequalities. First, for every $(g, g'), (g_1, g_1') \in G \times G$,

$$d_G(g^{-1}g', g_1^{-1}g_1') \leq d_G(g^{-1}g', g^{-1}g_1') + d_G(g^{-1}g_1', g_1^{-1}g_1').$$

Clearly $d_G(g^{-1}g_1', g_1^{-1}g_1') \leq d_G(g^{-1}, g_1^{-1}) = d_G(g, g_1)$. Moreover, given $\xi \in \text{Bord}X$ we have

$$d(g^{-1}g'\xi, g^{-1}g_1'\xi) = d(g'\xi, g_1'\xi).$$

Next use Proposition 14 to obtain

$$\begin{aligned}
\bar{D}_b(g^{-1}g'\xi, g^{-1}g_1'\xi) &= \frac{\bar{D}_b(g^{-1}g'\xi, g^{-1}g_1'\xi)}{\bar{D}_b(g'\xi, g_1'\xi)}\bar{D}_b(g'\xi, g_1'\xi) \\
&\leq C(\delta)b^{-\frac{1}{2}(h_1(gx_0)+h_2(gx_0))}\bar{D}_b(g'\xi, g_1'\xi) \\
&\leq C(\delta)b^{d(gx_0,x_0)}\bar{D}_b(g'\xi, g_1'\xi),
\end{aligned}$$

for some horofunction $h_1, h_2 \in X^h$. Splitting into the two possible cases we have either

$$D_b(g^{-1}g'\xi, g^{-1}g_1'\xi) = (\log b)d(g^{-1}g'\xi, g^{-1}g_1'\xi) \leq \bar{D}_b(g^{-1}g'\xi, g^{-1}g_1'\xi)$$

or

$$D_b(g^{-1}g'\xi, g^{-1}g_1'\xi) = \bar{D}_b(g^{-1}g'\xi, g^{-1}g_1'\xi) \leq (\log b)d(g^{-1}g'\xi, g^{-1}g_1'\xi).$$

In either case, the previous controls yield

$$D_b(g^{-1}g'\xi, g^{-1}g_1'\xi) \leq C(\delta)b^{d(gx_0,x_0)}d_G(g', g_1').$$

Taking the supremum over ξ yields

$$d_G(g^{-1}g', g_1^{-1}g_1') \leq d_G(g, g_1) + C(\delta)b^{d(gx_0,x_0)}d_G(g', g_1'). \tag{2.7}$$

□

2.4 Generalities on Dynamics and Ergodicity

A dynamical system consists of a trio of notions - space, time and evolution. Typically a point in the *space* describes the state of a system at a particular *time*. The study of dynamical systems focuses on describing the *evolution* of points in space. Such problems arise naturally in other parts of mathematics as well as everyday applications. On top of that, these problems are typically hard, thus making dynamical systems a very active area of research in mathematics.

In some systems describing the state of the space is quite difficult, this happens for example if we try to model an ideal gas inside a container. In such cases, a full deterministic approach may not be reasonable, as the complexity of the problem is too vast, so we focus on the average behaviour of the system, thus studying the system as a whole. In such systems we describe dynamics with respect to some measure, and we call this area of mathematics Ergodic theory.

In this section we very briefly go over ergodic theory although we assume that the reader is familiar with some terms as well as the main theorems, such as Birkhoff's and Kingman's ergodic theorems. As a reference, the reader can find these topics in Viana and Oliveira (2016) or in the lecture notes of Sarig (2009). A lot of notations used in Chapter 3 are laid here.

Let (Ω, β, μ) be a standard probability space with measure μ and σ-algebra β. We say $T : \Omega \to \Omega$ is an ergodic transformation with respect to μ if it is measurable, preserves measure ($\mu(A) = \mu(T^{-1}A)$ for every measurable set A) and every measurable invariant set $A = T^{-1}A$ has measure 0 or 1. Typically we denote an ergodic transformation by (T, μ), where Ω and β are implicit in the measurability of T.

Given a topological group G, we say that G acts by semicontractions on M if for every $g \in G$, $d(gx, gy) \leq d(x, y)$. Notice that this includes actions by isometries. Given an ergodic transformation T, we say that a measurable map $a : \mathbb{N} \times \Omega \to G$ is a right multiplicative cocycle in G over T if $a(n+m, \omega) = a(n, \omega)a(m, T^n\omega)$. Given a Borel measurable $g : \Omega \to G$ consider its associated right multiplicative cocycle

$$a(n, \omega) = g^{(n)}(\omega) := g(\omega)g(T\omega)...g(T^{n-1}\omega),$$

for every $n \in \mathbb{N}$ and $\omega \in \Omega$. The cocycle a is thus comprised of the information (g, T, Ω, β), whenever it is clear we identify the cocycle with g. Fixed a basepoint $x_0 \in M$, we will refer to $g^{(n)}(\omega)x_0$ as the process.

Definition 6 (Integrable Cocycle)**.** Let x_0 be a basepoint in M. We say that a cocycle (g, T, Ω, β) is integrable if

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

One of the fundamental characteristics of an integrable cocycle is its drift

$$\ell(g) := \lim_{n\to\infty} \frac{1}{n} \int_\Omega d(g^{(n)}(\omega)x_0, x_0)d\mu(\omega) = \lim_{n\to\infty} \frac{1}{n} d(g^{(n)}(\omega)x_0, x_0),$$

where these limits exists by Kingman's ergodic theorem and the second is constant for almost every ω due to ergodicity, moreover none of these limits depend on the basepoint x_0.

In this thesis we are mostly interested in exploring continuity of the drift with respect to the cocycle g, this requires us to choose a space a cocycles and endow it with a topology. We devote this section to introducing this class of cocycles.

Let $G \subset \text{Isom}(X)$ be a subgroup of isometries of some Gromov hyperbolic space X. Let $1 \leq b \leq 2^{1/\delta}$. Given $g : \Omega \to G$ we denote by $g^{-1} : \Omega \to G$ the map that sends ω to $g(\omega)^{-1}$. Consider $S(\Omega, G)$ to be the space of integrable cocycles $g : \Omega \to G$ such that g^{-1} is also measurable and

$$d_\infty(g) := \sup_{\omega \in \Omega} b^{d(g(\omega)x_0, x_0)} < \infty$$

Define the following pseudometric

$$d_\infty(g_1, g_2) := \text{ess sup}_{\omega \in \Omega} d_G(g_1(\omega)\,,\, g_2(\omega)),$$

for every $g_1, g_2 \in S(\Omega, G)$. Analogously to the construction of L^∞, we define the equivalence relation

$$g_1 \sim g_2 \Leftrightarrow d_\infty(g_1, g_2) = 0$$

in $S(\Omega, G)$, so the set of equivalence classes $S^\infty(\Omega, G)$ becomes a metric space when equipped with d_∞.

We can now think of the drift as a map

$$\begin{aligned} \ell : S^\infty(\Omega, G) &\to \mathbb{R} \\ g &\mapsto \ell(g). \end{aligned}$$

The drift is given by the speed at which we move away from the basepoint x_0. This is exactly why we introduced boundaries in the previous section. The following theorem by Karlsson and Gouëzel makes this notion precise by guaranteeing the existence of a horofunction that tracks the process. In other words, there exists a specific direction that the trajectory follows towards the boundary.

Theorem 16 (Theorem 1.3 in Gouëzel and Karlsson (2020))**.** *Let (g, T, Ω, β) be a cocycle, for almost every ω, there exists an horofunction $h_\omega \in M^h$ such that*

$$\lim_{n\to\infty} \frac{1}{n} h_\omega(g^{(n)}(\omega)x_0) = -\ell(g).$$

Moreover, if M is separable and Ω is a standard probability space, one can choose the map $\omega \mapsto h_\omega$ to be Borel measurable.

The intuition behind the result is that we can think of $e^{h(z)}$ as the distance between a point associated to h and z. In other words the theorem states that the sequence $g^{(n)}(\omega)x_0$ approaches a certain point related to h at a linear rate given by the drift. Although the intuition is quite direct and this result is one of our stepping stones, its proof is quite intricate and long for our purposes.

Notice as well how general the Theorem is, as no hypothesis are placed in M. Less general versions of the result also exist as in Karlsson and Margulis (1999) or Karlsson and Ledrappier (2006).

2.5 Dynamics in Hyperbolic Spaces

Our goal in this section is to describe the typical behaviour of $g^{(n)}(\omega)x_0$ when X is a hyperbolic space and G is its group of isometries. We do that by adding information to Karlsson and Gouëzel's theorem:

Theorem 17 (Hyperbolic Multiplicative Ergodic Theorem)**.** *Let X be a separable geodesic Gromov hyperbolic space and (g, T, Ω, β) a cocycle with positive drift. For almost every ω in Ω there is a filtration of the horofunction boundary*

$$X^h_-(\omega) \subset X^h_+(\omega) = X^h,$$

such that:

1. *for every $h \in X^h_+(\omega) \backslash X^h_-(\omega)$*
$$\lim_{n\to\infty} \frac{1}{n} h(g^{(n)}(\omega)x_0) = \ell(g);$$
2. *for every $h \in X^h_-(\omega)$*
$$\lim_{n\to\infty} \frac{1}{n} h(g^{(n)}(\omega)x_0) = -\ell(g),$$
and given $h_1, h_2 \in X^h_-$, one has $\sup_{z\in X} |h_1(z) - h_2(z)| < \infty$.

Moreover the filtration is $G-$invariant, that is,

$$g(\omega) \cdot X^h_-(\omega) = X^h_-(T\omega)$$

and is measurable provided Ω is a standard probability space.

Remark 1. *In the case of (BA) and strongly hyperbolic spaces more can be said regarding X^h_-; in fact it consists of a single horofunction which is picked measurably, that is, the map $\omega \mapsto h^-_\omega \in X^h_-(\omega)$ is measurable. This fact will be very important later when we obtain large deviations for the Markov setting.*

In this version, the theorem describes the behaviour of all horofunctions. The idea here is that in negatively curved we are able to see every point at infinity from every other point at infinity as we can connect them by a quasi-geodesic (or a geodesic if the space is proper). Hence horofunctions in X^h_- see the process getting closer whilst horofunctions X^h_+ see the process going away, however since they are connected by a quasi-geodesic the rate is the same for both.

When we reach curvature zero curvature the property of connecting points at infinity by quasi-geodesics, typically called visibility, is lost (see Chapter 9 in Bridson and Haefliger (2013)), hence the result becomes more complicated allowing the limit of $h(g^{(n)}x_0)/n$ to take a continuum of values in the interval $[-\ell(g), \ell(g)]$. For proper CAT(0) spaces, the limit is given by $(2\sin(\theta/2) - 1)\ell(g)$, where θ is the angular distance between the point at infinity associated to the horofunction and the hitting point of the process, $\lim_{n\to\infty} g^{(n)}x_0$ (this is a consequence of Karlsson and Margulis (1999) ray approximation and Proposition 9.8 in Bridson and Haefliger (2013)).

Proof of Theorem 17. By Karlsson-Gouëzel's theorem, for almost every $\omega \in \Omega$ there is a measurably chosen horofunction such that

$$\lim_{n\to\infty} \frac{1}{n} h(g^{(n)}(\omega)x_0) = -\ell(\mu).$$

This choice being measurable is what makes the filtration measurable.

For such $\omega \in \Omega$, set

$$X^h_-(\omega) = \left\{ h \in X^h \;:\; \lim_{n\to\infty} \frac{1}{n} h(g^{(n)}(\omega)x_0) = -\ell(\mu) \right\}.$$

Let $h_\xi, h_\eta \in X^h_-(\omega)$ for some $\xi, \eta \in \partial X$. Then, by Proposition 14, for $1 < b \leq 2^{1/\delta}$ we have

$$\bar{D}_b(g^{-(n)}(\omega)\xi \,,\, g^{-(n)}(\omega)\eta) \geq \frac{1}{C(\delta)} b^{-\frac{1}{2}\left[h_\xi(g^{(n)}(\omega)x_0) + h_\eta(g^{(n)}(\omega)x_0)\right]} \bar{D}_b(\xi \,,\, \eta). \tag{2.8}$$

Using the fact ∂X is bounded and taking n large enough, we see that $\bar{D}_b(\xi \,,\, \eta)$ must be zero. Clearly, the same argument using (2.8) shows that there is no other equivalence class of horofunctions for which $\lim_{n\to\infty} \frac{1}{n} h(g^{-(n)}(\omega)x_0)$ takes a negative value.

Now

$$X^h_+(\omega) \backslash X^h_-(\omega) = \left\{ h \in X^h \;:\; \liminf_{n\to\infty} \frac{1}{n} h(g^{(n)}(\omega)x_0) \geq 0 \right\}$$

Let $h \in X^h_+(\omega) \backslash X^h_-(\omega)$ and $h_1 \in X^h_-(\omega)$, using (2.8) again together with the fact D_b is bounded from above by 1, we obtain that for every $n \in \mathbb{N}$

$$C \leq h(g^{(n)}(\omega)x_0) + h_1(g^{(n)}(\omega)x_0),$$

for some $C \in \mathbb{R}$. However, notice that $|h(\omega^n x_0)| \leq d(\omega^n x_0\,,\, x_0)$, whence

$$C - h_1(g^{(n)}(\omega)x_0) \leq h(g^{(n)}(\omega)x_0) \leq d(g^{(n)}(\omega)x_0\,,\, x_0)$$

for every $n \in \mathbb{N}$. Dividing both sides by n and taking limits one has

$$\ell(\mu) = \lim_{n\to\infty} -\frac{1}{n} h_1(g^{(n)}(\omega)x_0) \leq \lim_{n\to\infty} \frac{1}{n} h(g^{(n)}(\omega)x_0) \leq \ell(\mu),$$

which proves the statement.

For the G-invariance of X^h_-, first recall we are using the right action.

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} g(\omega) \cdot h(g^{(n)}\omega x_0) &= \lim_{n\to\infty} \frac{1}{n} \left(h(g(\omega)^{-1} g^{(n)}(\omega)x_0) - h(g(\omega)^{-1}x_0) \right) \\
&= \lim_{n\to\infty} \frac{n-1}{n} \frac{1}{n-1} \left(h(g^{n-1}(T\omega)x_0) - h(g(\omega)^{-1}x_0) \right) \\
&= \lim_{n\to\infty} \frac{1}{n-1} h\left(g^{(n-1)}(T\omega)x_0 \right),
\end{aligned}$$

for every $h \in X^h$, in particular, $g_0 \cdot X^h_-(\omega) = X^h_-(T\omega)$. □

Consider $S^\infty_+(\Omega, G)$ to be the subspace of $S^\infty(\Omega, G)$ consisting of the elements $g \in S^\infty(\Omega, G)$ with positive drift. By the previous theorem, for every $g \in S^\infty_+(\Omega, G)$ we can consider the almost everywhere defined map

$$\begin{aligned}
\xi_g : \Omega &\to \partial X \\
\omega &\mapsto \xi(g, \omega).
\end{aligned}$$

Denote by $S^1(\Omega, \partial X)$ be the space of bounded measurable maps from Ω to ∂X where we consider the metric

$$d_1(f_1, f_2) := \int_\Omega D_b(f_1(\omega), f_2(\omega)) d\mu(\omega),$$

for every $f_1, f_2 \in S^1(\Omega, \partial X)$. Since Ω is standard, ξ_g belongs to $S^1(\Omega, \partial X)$ due to the ending point of Theorem 17. Finally we define the map

$$\begin{aligned}
\xi : S^\infty_+(\Omega, G) &\to S^1(\Omega, \partial X) \\
g &\mapsto \xi_g.
\end{aligned}$$

Chapter 3

Abstract Continuity Theorem

Our goal in this chapter is to study the continuity of the drift and tracking point maps, ℓ and ξ, respectively. With effect we are interested in obtaining a general criteria for this continuity to occur. Duarte et al. (2016) explore the statistical property of large deviation estimates together with the avalanche principle to obtain this sought after continuity for the Lyapunov exponent. Just like the Lyapunov exponent, the drift is obtained as a limit of a subadditive process. We explore this similarity to transport the ideas to the metric setting.

With this in mind, we will start by proving an avalanche principle type result for hyperbolic spaces. Then we proceed to use large deviations with the avalanche principle in an inductive way to obtain our continuity result. In this chapter we stick to strongly hyperbolic spaces, as that is where our avalanche principle will be valid.

By the end of the chapter we also explore a positivity criteria for the drift using the avalanche principle and a similar inductive process as we have used before for continuity.

3.1 Statement of the Theorem

Let X stand for a strongly hyperbolic metric space with basepoint x_0. Define the finite scale drift of $g \in S^\infty(\Omega, G)$ at time $n \in \mathbb{N}$ as

$$\ell_n(g) := \frac{1}{n} \int_\Omega d(g^{(n)}(\omega)x_0, x_0) d\mu(\omega),$$

which clearly satisfies $\ell_n(g) \to \ell(g)$ as n goes to ∞.

Henceforth we fix $\mathcal{C} \subset S^\infty(\Omega, G)$ a class of cocycles equipped with some distance $d_\mathcal{C}$ such that $d_\mathcal{C}(g_1, g_2) \geq d_\infty(g_1, g_2)$. We do this as in some cases we need some extra restriction on the cocycles we consider, thus making it not enough to work with the entire $S^\infty(\Omega, G)$. In the same spirit we will denote by $\mathcal{C}_+$ the set $\mathcal{C} \cap S^\infty_+(\Omega, G)$.

Definition 7 (Large deviation estimates)**.** Fix $x_0 \in X$. A cocycle $g \in \mathcal{C}$ is said to satisfy a uniform large deviation estimates of exponential type if there are constants $r > 0$, C, $c > 0$ and for every $\varepsilon > 0$ there exists $\bar{n} = \bar{n}(\varepsilon)$ such that

$$\mu\left\{\omega \in \Omega \, : \, \left|\frac{1}{n} d(g_1^{(n)}(\omega)x_0, x_0) - \ell_n(g_1)\right| > \varepsilon\right\} < Ce^{-c\varepsilon^2 n}$$

for every $g_1 \in \mathcal{C}$ with $d_{\mathcal{C}}(g, g_1) < r$ and every $n \geq \bar{n}$.

Let $b > 1$, since $e^n = b^{n \log_b(e)}$ one can replace the exponential with any base in the large deviations. For computational reasons it will make sense to work with the same $1 < b \leq 2^{1/\delta}$ as in the metric D_b.

We are now ready to state the main theorem of the chapter.

Theorem 18. *Let X be a strongly hyperbolic metric space and (T, Ω, μ, β) be a measure preserving dynamical system. Given a class of cocycles $\mathcal{C}$, suppose every $g \in \mathcal{C}_+$ satisfies a uniform large deviation estimate, then*

1) The drift $\ell : \mathcal{C} \to \mathbb{R}$ is continuous;

2) The drift $\ell : \mathcal{C}_+ \to \mathbb{R}$ is locally Hölder continuous;

3) Moreover, $\xi : \mathcal{C}_+ \to S^1(\Omega, \partial X)$ is locally Hölder continuous.

The theorem yields not only continuity but also quantifies it. Exploring the proof one can obtain the explicit constants associated with the local Hölder continuity as well as the size of the neighbourhood where these hold.

3.2 Avalanche Principle

In its linear form the avalanche principle is a theorem that allows us to take conclusions of global nature on a product of linear mappings from local hypotheses on each map. Our goal in this section is to obtain a metric version of the result in hyperbolic spaces, where we are allowed to take conclusions of global nature based on the local nature of the objects at hand. The theorem itself will read as follows.

Theorem 19 (Avalanche Principle)**.** *Let X be a strongly hyperbolic space, $x_0, ..., x_n$ be a sequence of points in X and $\rho, \sigma > 0$ constants such that*

G) $d(x_{i-1}, x_i) \geq \rho, \quad i = 1, ..., n;$

A) $\langle x_{i-1}, x_{i+1}\rangle_{x_i} \leq \sigma, \quad i = 1, ..., n-1$*;*

P) $2\sigma < \rho - 2\delta$*;*

Then, for every $2 \leq k \leq n$*,*

1) $\langle x_0, x_k\rangle_{x_{k-1}} < \sigma + \frac{1}{\log b} b^{2\sigma-\rho+2\delta}$*,*

2) $d(x_0, x_k) > \rho + (k-1)(\rho - 2\sigma - 2\delta)$*,*

3) and the following inequality holds

$$\left| d(x_0, x_k) + \sum_{i=2}^{k-1} d(x_{i-1}, x_i) - \sum_{i=1}^{k-1} d(x_{i-1}, x_{i+1}) \right| \leq 2(k-1)\frac{1}{\log b} b^{2\sigma-\rho+2\delta}.$$

For CAT(-1) spaces, condition $P)$ may be replaced with $\sinh(\rho - \sigma) > 2\sinh(\rho/2)$, which is more general, specially for small values of ρ (see Oregón-Reyes (2020)). Our version applies to more general spaces and suffices for our applications.

Before we tackle the proof, let us make two remarks; first that the hypothesis imply

$$\langle x_{i-1}, x_{i+1}\rangle_{x_i} + \langle x_i, x_{i+2}\rangle_{x_{i+1}} \leq 2\sigma < \rho - 2\delta \leq d(x_i, x_{i+1}) - 2\delta, \tag{3.1}$$

secondly, that the left-hand side of the conclusion may be rewritten as

$$\left| d(x_0, x_k) - \sum_{i=1}^{k} d(x_{i-1}, x_i) + 2\sum_{i=1}^{k-1} \langle x_{i-1}, x_{i+1}\rangle_{x_i} \right|. \tag{3.2}$$

Proof of Theorem 19. We will base the proof in establishing two simple claims.

Claim 1:

$$|\langle x_0, x_k\rangle_{x_{k-1}} - \langle x_{k-2}, x_k\rangle_{x_{k-1}}| \leq \delta.$$

Let us use induction: The case $k = 2$ is trivial. For $k > 2$, notice that

$$\begin{aligned}
\langle x_0, x_{k-2}\rangle_{x_{k-1}} &= d(x_{k-1}, x_{k-2}) - \langle x_0, x_{k-1}\rangle_{x_{k-2}} \\
&\geq d(x_{k-1}, x_{k-2}) - \langle x_{k-3}, x_{k-1}\rangle_{x_{k-2}} - \delta && \text{by induction,} \\
&> \langle x_{k-2}, x_k\rangle_{x_{k-1}} + \delta && \text{by (3.1).}
\end{aligned}$$

Proceeding with the definition of hyperbolicity

$$\langle x_{k-2}, x_k\rangle_{x_{k-1}} \geq \min\{\langle x_0, x_{k-2}\rangle_{x_{k-1}}, \langle x_k, x_0\rangle_{x_{k-1}}\} - \delta,$$

where the minimum must be $\langle x_k, x_0\rangle_{x_{k-1}}$, otherwise we would get $\langle x_{k-2}, x_k\rangle_{x_{k-1}} > \langle x_{k-2}, x_k\rangle_{x_{k-1}}$. Whence

$$\langle x_0, x_k\rangle_{x_{k-1}} \leq \langle x_{k-2}, x_k\rangle_{x_{k-1}} + \delta.$$

Changing the roles of x_0, x_{k-2} we get the claim.

Claim 2: Our second claim is in fact item 1),

$$|\langle x_0, x_k\rangle_{x_{k-1}} - \langle x_{k-2}, x_k\rangle_{x_{k-1}}| \leq \frac{1}{\log b} b^{2\sigma-\rho+2\delta}.$$

Applying Lagrange's mean value theorem with $f(x) = b^{-x}$, followed by claim 1, the fact that X is strongly hyperbolic and the inequality $\langle x_0, x_{k-2}\rangle_{x_{k-1}} \geq d(x_{k-1}, x_{k-2}) - \langle x_{k-3}, x_{k-1}\rangle_{x_{k-1}} - \delta$ obtained in claim 1, yields

$$\begin{aligned}\left|\langle x_0, x_k\rangle_{x_{k-1}} - \langle x_{k-2}, x_k\rangle_{x_{k-1}}\right| &\leq \frac{1}{\log b} b^{\max\{\langle x_0,x_k\rangle_{x_{k-1}}, \langle x_{k-2},x_k\rangle_{x_{k-1}}\}} \left|b^{-\langle x_0,x_k\rangle_{x_{k-1}}} - b^{-\langle x_{k-2},x_k\rangle_{x_{k-1}}}\right| \\ &\leq \frac{1}{\log b} b^{\sigma+\delta} b^{-\langle x_0,x_{k-2}\rangle_{x_{k-1}}} \\ &\leq \frac{1}{\log b} b^{\sigma+\delta} b^{\langle x_{k-3},x_{k-1}\rangle_{x_{k-2}} - d(x_{k-1},x_{k-2})+\delta} \\ &\leq \frac{1}{\log b} b^{2\sigma-\rho+2\delta}.\end{aligned}$$

These claims were motivated by the relation

$$d(x_0, x_n) = d(x_0, x_{n-1}) + d(x_{n-1}, x_n) - 2\langle x_0, x_n\rangle_{x_{n-1}},$$

which makes it so by controlling $|\langle x_0, x_n\rangle_{x_{n-1}} - \langle x_{n-2}, x_n\rangle_{x_{n-1}}|$ by some quantity, then (3.2) can be controlled by $(n-1)$ times said quantity. □

Remark 2. *Notice that for hyperbolic spaces in general, claim 1 implies that*

$$\left|d(x_0, x_n) + \sum_{i=2}^{n-1} d(x_{i-1}, x_i) - \sum_{i=1}^{n-1} d(x_{i-1}, x_{i+1})\right| \leq 2(n-1)\delta.$$

Example 3. Let us look at the hyperbolic plane $\mathbb{H}^2$. The hyperbolic plane is strongly hyperbolic with $b = e$ and for every isometry g, $d(gx_0, x_0) = 2\log ||g||$. Consider $g_0, ..., g_{i-1} \in SL(2, \mathbb{R})$ isometries of $\mathbb{H}^2$. Finally take $x_0 = i$ and $x_j = g^{(j)}x_0$. Then the hypothesis read as follows

G) $d(x_{i-1}, x_i) \geq \rho \Leftrightarrow 2\log ||g_{i-1}|| \geq \rho \Leftrightarrow ||g_{i-1}||^2 \geq e^{\rho} := \mu$;

A) $\langle x_{i-1}, x_{i+1}\rangle_{x_i} \leq \sigma \Leftrightarrow \frac{||g_{i-1}g_i||}{||g_{i-1}|| \, ||g_i||} \geq e^{-\sigma} := \nu$;

P) $\mu^{-1} < e^{-2\delta}\nu^2$,

whilst the conclusion reads

$$\left|\log||g^{(n)}|| + \sum_{i=2}^{n-1} \log||g_{i-1}|| - \sum_{i=1}^{n-1} \log||g_{i-1}g_i||\right| \leq 2(n-1)e^{2\sigma-\rho+2\delta} = 2e^{2\delta}(n-1)\frac{1}{\mu\nu^2}.$$

Which is a restatement of the $\mathrm{SL}(2,\mathbb{R})$ version of the Avalanche principle, from Duarte and Klein (2017).

3.3 Continuity of the Drift

In this section we prove the first assertion of Theorem 18. This is done by following a specific route where we start by proving the continuity at a finite scale, then we transport the control to larger scales by an inductive step based on the Avalanche principle and the existence of large deviation estimates.

3.3.1 Finite Scale Continuity

Let us start by proving that at a finite scale the drift is continuous as well as understand this continuity rate, this is fulcral for the next step where we try to transport these controls forward in time.

Lemma 20. *Given $C > 0$, set $G_C = \{g \in G \,:\, d(gx_0, x_0) < C\}$. The map $G_C \to \mathbb{R}$ defined by $g \mapsto d(gx_0, x_0)$ is Lipschitz continuous.*

Proof. Let $g_1, g_2 \in G_C$. Notice that $|d(g_1x_0, x_0) - d(g_2x_0, x_0)| \leq d(g_1x_0, g_2x_0)$, so if $D_b(g_1x_0, g_2x_0) = (\log b)d(g_1x_0, g_2x_0)$ we are done, otherwise use the inequality $(\log b)x < b^{x/2}$

$$\begin{aligned} |d(g_1x_0, x_0) - d(g_2x_0, x_0)| &\leq d(g_1x_0, g_2x_0) \\ &\leq \frac{1}{\log b} b^{d(g_1x_0, g_2x_0)/2} \\ &\leq \frac{b^{d(g_1x_0, x_0)/2 + d(x_0, g_2x_0)/2}}{\log b} b^{d(g_1x_0, g_2x_0)/2 - d(g_1x_0, x_0)/2 - d(x_0, g_2x_0)/2} \\ &\leq \frac{C}{\log b} b^{-\langle g_1x_0\,,\, g_2x_0\rangle_{x_0}} \leq D_b(g_1x_0, g_2x_0) \leq D_G(g_1, g_2). \end{aligned}$$

We will use this trick multiple times throughout the text. ∎

Lemma 21. *Let $g \in S^\infty(\Omega, G)$, there exist $C = C(g) > 0$ and $r > 0$ such that if $g_1, g_2 \in S^\infty(\Omega, G)$ with $d_\infty(g_i, g) < r$ for $i = 1, 2$, then for every $n \in \mathbb{N}$ and μ-a.e $\omega \in \Omega$*

1. *$d_\infty(g_1) < C$;*
2. *$d_G(g_1^{(n)}(\omega), g_2^{(n)}(\omega)) \leq nC^{n-1} d_\infty(g_1\,,\, g_2)$.*

Proof. Point 1. is an immediate consequence of the continuity proven before. Denote by $T:\Omega\to\Omega$ the ergodic transformation at hand. From the proof that $\mathrm{Isom}(X)$ is a topological group, for every $\omega\in\Omega$, one has

$$\begin{aligned} d_G(g_1^{(n)}(\omega), g_2^{(n)}(\omega)) &\leq d_G(g_1(\omega), g_2(\omega)) + b^{d(g_1(\omega)x_0,x_0)} d_G(g_1^{(n-1)}(T\omega), g_2^{(n-1)}(T\omega)) \\ &\leq d_\infty(g_1,g_2) + C d_G(g_1^{(n-1)}(T\omega), g_2^{(n-1)}(T\omega)) \end{aligned}$$

so the claim follows upon taking the supremum. □

Proposition 22 (finite scale continuity)**.** *Let $g\in S^\infty(\Omega,G)$. For every $g_1,g_2\in S^\infty(\Omega,G)$ and for almost every $\omega\in\Omega$ there exists $C=C(g)>0$,*

$$\left|\frac{1}{n}d(g_1^{(n)}(\omega)x_0,x_0) - \frac{1}{n}d(g_2^{(n)}(\omega)x_0,x_0)\right| \leq \frac{C^n}{\log(b)}d_\infty(g_1,g_2) \leq \frac{b^{C_1 n}}{\log(b)}d_\infty(g_1\,,\,g_2).$$

where $C_1 := \log_b(C)$, in particular,

$$|\ell_n(g_1)-\ell_n(g_2)| < \frac{b^{C_1 n}}{\log(b)} d_\infty(g_1,g_2).$$

Recall that ℓ_n stands for the finite scale drift, hence the bottom inequality in the proposition follows from the upper one after integration on ω.

Proof. To soften notations, let us omit ω throughout the proof. explore this

$$\begin{aligned} \left|d(g_1^{(n)}x_0,x_0) - d(g_2^{(n)}x_0,x_0)\right| &\leq d(g_1^{(n)}x_0, g_2^{(n)}x_0) \\ &\leq \frac{1}{\log(b)} b^{d(g_1^{(n)}x_0,g_2^{(n)}x_0)/2} \\ &\leq \frac{b^{d(g_1^{(n)}x_0,x_0)/2+d(g_2^{(n)}x_0,x_0)/2}}{\log(b)} b^{-\langle g_1^{(n)}x_0\,,\,g_2^{(n)}x_0\rangle_{x_0}} \\ &\leq \frac{C^n}{\log b} d_G(g_1^{(n)}(\omega), g_2^{(n)}(\omega)) \\ &\leq n\frac{C^{2n}}{\log b} d_\infty(g_1,g_2) \end{aligned}$$

Which concludes the proof, by using C^2 as the C from the statement in the proposition. □

This proposition implies the continuity of the maps ℓ_n. Since the drift $\ell(g)$ may be given as $\inf_{n\geq 1}\ell_n(g)$, the upper semi-continuity of $\ell(g)$ follows from the following lemma.

Lemma 23. *Let M be a metric space and $f_n : M\to\mathbb{R}$ be a sequence of upper semi-continuous functions. Then, $f(x) = \inf_{n\geq 1} f_n(x)$, the pointwise infimum of these functions, is upper semi-continuous.*

Proof. Let $x \in M$ and take $\inf_{n \geq 1} f_n(x) = g(x) < r$, there must be $i \geq 1$ such that $f_i(x) < r$. Since $f_i(x)$ is upper semi-continuous, there must be a neighbourhood V of x such that for every $y \in V$ one has $f_i(y) < r$. Since $g_i(y) \leq f_i(y)$ for every y, we obtain $g(y) < r$ for every $y \in U$ thus proving the Lemma. □

Since ℓ is upper semi-continuous, it is continuous in the neighbourhood of the cocycles $g \in \mathcal{C}$ in which it is zero. With that said we focus cocycles in $\mathcal{C}_+$, where we obtain a stronger modulus of continuity.

3.3.2 Inductive Step

In this section we will understand how to pass the previously estabilished controls forward through an inductive step based on the large deviations estimates and the avalanche principle. Throughout the inductive process we will incur in some errors, we start with a lemma showing these errors are of the order of magnitude we want to control.

Lemma 24. *Let $g \in S^\infty(\Omega, G)$, if $n, n_0, n_1, r \in \mathbb{N}$ are such that $n_1 = n\, n_0 + r$ where $0 \leq r < n_0$, then*

$$-2\log_b(C)\frac{n_0}{n_1} + \ell_{(n+1)n_0}(g) \leq \ell_{n_1}(g) \leq \ell_{n\, n_0}(g) + 2\log_b(C)\frac{n_0}{n_1}.$$

Proof. Given $n_1 = n\, n_0 + r$ where $0 \leq r < n_0$ we have, for every ω, $g^{(n_1)}(\omega) = g^{(n\, n_0)}(\omega) g^{(r)}(T^{n\, n_0}\omega)$, whence

$$d(g^{(n_1)}(\omega)x_0\,,\, x_0) \leq d(g^{(n\, n_0)}(\omega)x_0\,,\, x_0) + d(g^{(r)}(T^{n\, n_0}\omega)x_0\,,\, x_0),$$

integrating both sides, one has

$$\ell_{n_1}(g) \leq \frac{n\, n_0}{n_1}\ell_{n\, n_0}(g) + \frac{r}{n_1}\ell_r(g).$$

which gives

$$\ell_{n_1}(g) \leq \ell_{n\, n_0}(g) + \frac{r}{n_1}\left[\ell^{(r)}(g) - \ell^{(n\, n_0)}(g)\right] \leq \ell_{n\, n_0}(g) + 2\log_b(C)\frac{r}{n_1}.$$

For the leftmost inequality write $n_1 = (n+1)\, n_0 + q$ where $q = n_0 - r$ and proceed similarly. □

The following proposition is the important step towards proving continuity of the drift. Its content is that if we obtain some control for time n_0, then we can transport it to time n_1 larger than n_0. To do this we break the orbit at time n_1 into smaller pieces of size n_0 which we then relate back with the larger piece of size n_1 by using the avalanche principle.

Proposition 25 (Inductive step)**.** *Let $g \in \mathcal{C}_+$ and $c, \bar{n}$ be the uniform large deviation parameters. Fix $\varepsilon = \ell(g)/100 > 0$ and denote $c_1 := \frac{c}{2}\varepsilon^2$. There are constants $C = C(g) > 0, r = r(g) > 0$, $\overline{n_0} = \overline{n_0}(g) \in \mathbb{N}$, such that for any $n_0 > \overline{n_0}$, if the inequalities*

$$\ell_{n_0}(g_1) - \ell_{2n_0}(g_1) < \eta_0$$
$$|\ell_{n_0}(g_1) - \ell_{n_0}(g)| < \theta_0$$

holds for any $g_1 \in \mathcal{C}$ such that $d(g_1, g) < r$ and if the positive numbers η_0, θ_0, satisfy

$$\theta_0 + 2\eta_0 < \ell(g) - 4\varepsilon,$$

then for every n_1 such that $|n_1 - e^{c_1 n_0}| < 1$ one has

$$|\ell_{n_1}(g_1) + \ell_{n_0}(g_1) - 2\ell_{2n_0}(g_1)| \leq C\frac{n_0}{n_1} \tag{3.3}$$

Furthermore,

$$\ell_{n_1}(g_1) - \ell_{2n_1}(g_1) < \eta_1 \tag{3.4}$$
$$|\ell_{n_1}(g_1) - \ell_{n_1}(g)| < \theta_1 \tag{3.5}$$

where

$$\theta_1 = \theta_0 + 4\eta_0 + C\frac{n_0}{n_1}$$
$$\eta_1 = C\frac{n_0}{n_1}.$$

From this point on in the text we will use the notation $a \lesssim b$ to convey that there exists a universal constant C such that $a \leq Cb$.

Proof. Throughout the proof C will stand for some constant which isn't a priori always the same. We start the proof with some assumptions, in particular, making r smaller if necessary, every g_1 with $d_\infty(g, g_1) < r$ satisfies large deviation estimates. We can also assume $\overline{n_0}$ to be large enough so that $|\ell_n(g) - \ell(g)| < \varepsilon$ for $n \geq \overline{n_0}$ which comes from the fact $\ell_n(g)$ converges to $\ell(g)$.

With that said, let g_1 be in the conditions above. Assume $n_1 = n\, n_0$ as otherwise we obtain an extra error of order n_0/n_1 which, by the previous lemma, is along the size of our control. Fix x_0 a basepoint in X and define, for every $0 \leq i \leq n-1$, the sequence of points

$$x_i(\omega) := g_1^{(n_0)}(\omega) g_1^{(n_0)}(T^{n_0}\omega) \dots g_1^{(n_0)}(T^{(i-1)\, n_0}\omega)\, x_0,$$

so that $x_n = g^{(n_1)}(\omega) x_0$ and for every $1 \leq i \leq n-1$,

$$d(x_i, x_{i-1}) = d(g_1^{(n_0)}(T^{(i-1)\, n_0}\omega)\, x_0\, ,\, x_0),$$

$$d(x_{i-1}, x_{i+1}) = d(g_1^{(n_0)}(T^{(i-1)\, n_0}\omega) g_1^{(n_0)}(T^{(i)\, n_0}\omega)\, x_0\,,\, x_0) = d(g_1^{(2n_0)}(T^{(i-1)\, n_0}\omega) x_0\,,\, x_0).$$

At this point we are going to use the large deviation estimates to verify the conditions of the avalanche principle are satisfied, with effect for every $m > \overline{n_0}$ there exists a set $\mathcal{B}_m$ whose measure does not exceed $e^{-c\varepsilon^2 m}$ such that for every $\omega \notin \mathcal{B}_m$

$$-\varepsilon \leq \frac{1}{m} d(g^{(m)} x_0, x_0) - \ell_m(g) \leq \varepsilon$$

in particular, if $\omega \notin \mathcal{B}_{n_0}$

$$\begin{aligned} \frac{1}{n_0} d(x_1, x_0) &= \frac{1}{n_0} d(g_1^{(n_0)}(\omega) x_0, x_0) \\ &\geq \ell_{n_0}(g_1) - \varepsilon \\ &> \ell_{n_0}(g) - \theta_0 - \varepsilon \\ &\geq \ell(g) - \theta_0 - \varepsilon, \end{aligned}$$

whence,

$$d(g_1^{(n_0)}(\omega) x_0, x_0) > n_0(\ell(g) - \theta_0 - \varepsilon) =: \rho_0.$$

Through the same process we obtain for every $\omega \notin \mathcal{B}_{2n_0}$

$$\frac{1}{2n_0} d(g_1^{(2n_0)}(\omega) x_0, x_0) \geq \ell^{(2n_0)}(g_1) - \varepsilon$$

as well as

$$\begin{aligned} \frac{1}{n_0} d(g_1^{(n_0)}(\omega) x_0, x_0) &\leq \ell_{n_0}(g_1) + \varepsilon \\ \frac{1}{n_0} d(g_1^{(n_0)}(T^{n_0}\omega) x_0, x_0) &\leq \ell_{n_0}(g_1) + \varepsilon, \end{aligned}$$

for every $\omega \notin \mathcal{B}_{n_0} \cup T^{-n_0}\mathcal{B}_{n_0}$. Hence, for every $\omega \notin \mathcal{B}_{2n_0} \cup \mathcal{B}_{n_0} \cup T^{-n_0}\mathcal{B}_{n_0}$

$$\begin{aligned} \langle x_0\,,\, x_2\rangle_{x_1} &= \left\langle x_0\,,\, g_1^{(2n_0)}(\omega) x_0 \right\rangle_{g_1^{(n_0)}(\omega) x_0} \\ &= \frac{1}{2}\Big(d(g_1^{(n_0)}(\omega) x_0, x_0) + d(g_1^{(n_0)}(T^{n_0}\omega) x_0, x_0) - d(g_1^{(2n_0)}(\omega) x_0, x_0) \Big) \\ &\leq n_0(\ell_{n_0} - \ell_{2n_0} + 2\varepsilon), \end{aligned}$$

in oher words,

$$\langle x_0\,,\, x_2\rangle_{x_1} < n_0(\eta_0 + 2\varepsilon) =: \sigma_0.$$

Similar computations yield the same controls for every $1 \leq i \leq n-1$, under appropriate assumptions. Moreover, by hypothesis, $2\sigma_0 - \rho_0 = n_0(\eta_0 + 3\varepsilon + \theta_0 - \ell(g)) \leq -\varepsilon n_0$ so choosing

n_0 large enough so that $-\varepsilon n_0 < -2\delta$, the AP applies outside the set $\mathcal{B}^*_{n_0} = \cup_{i=0}^{n-1} T^{in_0}\mathcal{B}_{n_0}$ where we obtain the control

$$\left| d(x_0, x_n) + \sum_{i=2}^{n-1} d(x_{i-1}, x_i) - \sum_{i=1}^{n-1} d(x_{i-1}, x_{i+1}) \right| \leq 2(n-1)\frac{1}{\log(b)} b^{2\sigma_0 - \rho_0 + 2\delta},$$

which translates to

$$\left| d(x_0, g^{(n_1)}(\omega)x_0) + \sum_{i=2}^{n-1} d(g_1^{(n_0)}(T^{(i-1)\,n_0}\omega)\,x_0\,,\,x_0) - \sum_{i=1}^{n-1} d(g_1^{(2n_0)}(T^{(i-1)\,n_0}\omega)x_0\,,\,x_0) \right| \lesssim n b^{-\varepsilon n_0}.$$

Dividing both sides by $n_1 = n\,n_0$, one now obtains

$$\Bigg| \frac{1}{n_1} d(x_0, g^{(n_1)}(\omega)x_0) + \frac{1}{n}\sum_{i=2}^{n-1} \frac{1}{n_0} d(g_1^{(n_0)}(T^{(i-1)\,n_0}\omega)\,x_0\,,\,x_0) \\ - \frac{2}{n}\sum_{i=1}^{n-1} \frac{1}{2n_0} d(g_1^{(2n_0)}(T^{(i-1)\,n_0}\omega)x_0\,,\,x_0) \Bigg| \lesssim b^{-\varepsilon n_0}.$$

Let $f(\omega)$ denote the bounded function on the left side. Notice that, for every $\omega \notin \mathcal{B}^*_{n_0}$, $|f(\omega)| \lesssim b^{-\varepsilon n_0}$, while in $\mathcal{B}^*_{n_0}$ the control $|f(\omega)| \leq C$ remains valid for some $C = C(g)$ since $g_1 \in \mathcal{C}$. On the other hand,

$$\int_\Omega f(\omega) d\mu(\omega) = \ell_{n_1}(g_1) + \frac{n-2}{n}\ell_{n_0}(g_1) - \frac{2(n-1)}{n}\ell_{2n_0}(g_1),$$

hence

$$\begin{aligned} \left| \ell_{n_1}(g_1) + \frac{n-2}{n}\ell_{n_0}(g_1) - \frac{2(n-1)}{n}\ell_{2n_0}(g_1) \right| &\leq \int_\Omega |f(\omega)| d\mu(\omega) \\ \int_{\Omega \setminus \mathcal{B}^*_{n_0}} |f(\omega)| d\mu(\omega) &+ \int_{\mathcal{B}^*_{n_0}} |f(\omega)| d\mu(\omega) \\ &\lesssim b^{-\varepsilon n_0} + C\mu(\mathcal{B}^*_{n_0}) \\ &\lesssim b^{-\varepsilon n_0} + C b^{-c_1 n_0} \\ &\lesssim b^{-c_1 n_0} < C\frac{n_0}{n_1} \end{aligned}$$

Having

$$\left| \ell_{n_1}(g_1) + \frac{n-2}{n}\ell_{n_0}(g_1) - \frac{2(n-1)}{n}\ell_{2n_0}(g_1) \right| < C\frac{n_0}{n_1}$$

one may write

$$\left| \ell_{n_1}(g_1) + \ell_{n_0}(g_1) - 2\ell_{2n_0}(g_1) - \frac{2}{n}\left[\ell_{n_0}(g_1) - \ell_{2n_0}(g_1)\right] \right| < C\frac{n_0}{n_1}$$

so that (3.3) holds:

$$|\ell_{n_1}(g_1) + \ell_{n_0}(g_1) - 2\ell_{2n_0}(g_1)| < C\frac{n_0}{n_1}.$$

The same process may be used to obtain (3.3) at times $2n_1$. Then by an immediate triangle inequality one obtains (3.4).

To prove (3.5) start by rewriting (3.3) as

$$|\ell_{n_1}(g_1) - \ell_{n_0}(g_1) + 2[\ell_{n_0}(g_1) - \ell_{2n_0}(g_1)]| < C\frac{n_0}{n_1}.$$

So

$$\begin{aligned}|\ell_{n_1}(g_1) - \ell_{n_1}(g)| &\leq |\ell_{n_1}(g_1) - \ell_{n_0}(g_1) + 2[\ell_{n_0}(g) - \ell_{2n_0}(g)]| \\ &\quad + |\ell_{n_1}(g) - \ell_{n_0}(g) + 2[\ell_{n_0}(g) - \ell_{2n_0}(g)]| \\ &\quad + 2|\ell_{n_0}(g_1) - \ell_{2n_0}(g_1)| + 2|\ell_{n_0}(g) - \ell_{2n_0}(g)| \\ &\quad + |\ell_{n_0}(g) - \ell_{n_0}(g_1)| \\ &< \theta_0 + 4\eta_0 + C\frac{n_0}{n_1} =: \theta_1\end{aligned}$$

□

3.3.3 Rate of Convergence

In this section we shall use the inductive step to understand exactly how pushing the controls though the natural numbers affects the convergence rate of the quantities at hand. In particular we obtain the rate convergence associated with the functions ℓ_n. These however will be too slow, hence we also look at $-\ell_n + 2\ell_{2n}$.

Lemma 26. *Let $\{x_n\}$ be a sequence converging to x such that for every $n \in \mathbb{N}$,*

$$|x_n - x_{2n}| < \frac{\log_b n}{n},$$

then, for every $n \in \mathbb{N}$

$$|x_n - x| \lesssim \frac{\log_b n}{n}.$$

Proof. Let $n \in \mathbb{N}$, then we can use a telescopic sum to write

$$\begin{aligned}|x_n - x| = \left|\sum_{i=0}^{\infty} x_{2^i n} - x_{2^{i+1}n}\right| &\leq \sum_{i=0}^{\infty} |x_{2^i n} - x_{2^{i+1}n}| \\ &\leq \sum_{i=0}^{\infty} \frac{\log_b(2^i n)}{2^i n} \lesssim \frac{\log_b n}{n},\end{aligned}$$

as the sum of the series is of order $\frac{\log_b n}{n}$. □

Proposition 27. *Let $g \in \mathcal{C}$. There are constants $r_1 > 0$, $\overline{n_0} \in \mathbb{N}$, $c_2 > 0$, $K < \infty$ all depending on g such that the following hold*

$$|\ell(g_1) - \ell_n(g_1)| < K \frac{\log_b n}{n}$$
$$|\ell(g_1) + \ell_n(g_1) - 2\ell_{2n}(g_1)| < b^{-c_2 n},$$

for every $n > \overline{n_0}$ and $g_1 \in S^\infty(\Omega, G)$ with $d_\infty(g, g_1) < r_1$.

Proof. Let us use the constants ε, c_1, C, r and $\overline{n_0}$ given in the inductive step. Consider the quantities $n_0^- = \overline{n_0}$, $n_0^+ = b^{c_1 \overline{n_0}}$ and set $\mathcal{N}_0 := [n_0^-, n_0^+]$. We shall also define $r_1 = \min\{r, b^{-3C_1\overline{n_0}}\}$. Then, by the finite scale continuity, for every $n_0 \in \mathcal{N}_0$, we have

$$|\ell_{2n_0}(g_1) - \ell_{2n_0}(g)| < b^{2C_1 n_0} d_\infty(g_1, g_2) \le b^{-C_1\overline{n_0}} \le \varepsilon,$$

choosing $\overline{n_0}$ large enough for the effect. Likewise

$$|\ell_{n_0}(g_1) - \ell_{n_0}(g)| < \varepsilon =: \theta_0,$$

Moreover

$$|\ell_{2n_0}(g) - \ell_{n_0}(g)| < |\ell_{2n_0}(g) - \ell(g)| + |\ell(g) - \ell_{n_0}(g)| < 2\varepsilon,$$

so that

$$|\ell_{2n_0}(g) - \ell_{n_0}(g)| < 2\varepsilon =: \eta_0,$$

and we have

$$\theta_0 + 2\eta_0 = 5\varepsilon < \ell(g) - 6\varepsilon.$$

Using the inductive process we now have $n_1^- = b^{c_1 n_0^-}$, $n_1^+ = b^{c_1 n_0^+}$ and define $\mathcal{N}_1 = [n_1^-, n_1^+]$. If $n_1 \in \mathcal{N}_1$ then $n_0 \lesssim \log_b(n_1)$. Now,

$$|\ell_{n_1}(g_1) + \ell_{n_0}(g_1) - 2\ell_{2n_0}(g_1)| < C\frac{n_0}{n_1} < K\frac{\log_b n_1}{n_1},$$

for some constant K. Moreover

$$\ell_{n_1}(g_1) - \ell_{2n_1}(g_1) < \eta_1$$
$$|\ell_{n_1}(g_1) - \ell_{n_1}(g)| < \theta_1$$

where

$$\theta_1 = \theta_0 + 4\eta_0 + C\frac{n_0}{n_1} < 13\varepsilon + K\frac{\log_b n_1}{n_1},$$
$$\eta_1 = C\frac{n_0}{n_1} < K\frac{\log_b n_1}{n_1}.$$

Furthermore,

$$\theta_1 + 2\eta_1 \leq 13\varepsilon + 3K\frac{\log_b n_1}{n_1} < 16\varepsilon < \ell(g) - 6\varepsilon.$$

Hence we can repeat the process, let $n_2^- = b^{c_1 n_1^-}, n_2^+ = b^{c_1 n_1^+}$, and define $\mathcal{N}_2 = [n_2^-, n_2^+]$, then, if $n_2 \in \mathcal{N}_2$, there exists $n_1 \in \mathcal{N}_1$ such that $n_1 \lesssim \log_b(n_2)$

$$|\ell_{n_2}(g_1) + \ell_{n_1}(g_1) - \ell_{2n_1}(g_1)| < C\frac{n_1}{n_2} < K\frac{\log_b n_2}{n_2}.$$

Moreover

$$\ell_{n_2}(g_1) - \ell_{2n_2}(g_1) < \eta_2$$
$$|\ell_{n_2}(g_1) - \ell_{n_2}(g)| < \theta_2$$

where

$$\theta_2 = \theta_1 + 4\eta_1 + C\frac{n_1}{n_2} < 13\varepsilon + 5K\frac{\log_b n_1}{n_1} + K\frac{\log_b n_2}{n_2},$$
$$\eta_2 = C\frac{n_1}{n_2} < K\frac{\log_b n_2}{n_2}$$

Inductively repeating the process we obtain intervals $\mathcal{N}_k$ whose union cover all natural numbers greater than n_1. Hence given $n > n_1$, there exists $k \geq 0$ such that $n \in \mathcal{N}_{k+1}$, so there is also $n_k \in \mathcal{N}_k$ so that

$$n = n_{k+1} = b^{c_1 n_k}.$$

Moreover

$$\ell_{n_{k+1}}(g_1) - \ell_{2n_{k+1}}(g_1) < \eta_{k+1} < K\frac{\log_b n_{k+1}}{n_{k+1}}$$

and

$$\begin{aligned}|\ell_{n_{k+1}}(g_1) - \ell_{n_{k+1}}(g_1)| &< \theta_{k+1}\\ &< \theta_k + 4\eta_{k+1} + C\frac{n_k}{n_{k+1}}\\ &< 13\varepsilon + 5K\sum_{i=1}^{k} K\frac{\log_b n_i}{n_i} + K\frac{\log_b n_{k+1}}{n_{k+1}},\end{aligned}$$

however, since n_k increase super-exponentially, the series $\sum_{i>0} \frac{\log_b n_i}{n_i}$ is convergent with sum of order $\frac{\log_b n_1}{n_1}$.

With that, for every $n \geq n_0$ we obtain

$$\ell_n(g_1) - \ell_{2n}(g_1) < K\frac{\log_b n}{n}$$

whence

$$|\ell_n(g_1) - \ell(g_1)| < K\frac{\log_b n}{n}.$$

Now,

$$\begin{aligned}\left|\ell_{n_{k+1}}(g_1) + \ell_{n_k}(g_1) - 2\ell_{2n_k}(g_1)\right| &< K\frac{\log_b n_{k+1}}{n_{k+1}} \\ &\leq Kc_1 n_k b^{-c_1 n_k} < b^{-\frac{c_1}{2}n}\end{aligned}$$

so

$$|\ell(g_1) + \ell_{n_k}(g_1) - 2\ell_{2n_k}(g_1)| < 2b^{-\frac{c_1}{2}n_k} < b^{-\frac{c_1}{3}n_k}$$

hence the result follows for $n > n_0$. □

3.3.4 Hölder Continuity of the Drift

We are finally ready to prove the first two items of Theorem 18

Proof. Consider $\overline{n_1} \in \mathbb{N}$, $r_1 > 0$, $r > 0$, c_2 as in proposition 27. Let $g \in \mathcal{C}_+$ with $\ell(g) > 0$ and take the function from $\mathcal{C}$ to $\mathbb{R}$

$$f_n := -\ell_n + 2\ell_{2n}$$

clearly $f_n(g) \to \ell(g)$, moreover an exponential rate of convergence holds for every $n \geq \overline{n_0}$,

$$|\ell(g_1) - f_n(g_1)| = |\ell(g_1) + \ell_n(g_1) - 2\ell_{2n}(g_1)| \leq b^{-c_2 n}.$$

Consider now $d_\infty(g_1, g_2) < \log(b)b^{-2(C_1+c_2)\overline{n_1}}$, and pick $n \geq \overline{n_1}$ such that

$$b^{-4(C_1+c_2)n} < d_\infty(g_1\ ,\ g_2) < b^{-2(C_1+c_2)n}.$$

Then for m equal to either n or $2n$ one has

$$|\ell_m(g_1) - \ell_m(g_2)| \leq \frac{b^{2C_1 n}}{\log b} d_{\mathcal{C}}(g_1, g_2) < b^{-2c_2 n}$$

thus

$$\begin{aligned}|f_n(g_1) - f_n(g_2)| &\leq |\ell_n(g_1) - \ell_n(g_2)| + 2|\ell_{2n}(g_1) - \ell_{2n}(g_2)| \\ &\leq 3b^{-2c_2 n} \leq b^{-c_2 n}.\end{aligned}$$

Finally one has

$$\begin{aligned}|\ell(g_1) - \ell(g_2)| &\leq |\ell(g_1) - f_n(g_1)| + |f_n(g_1) - f_n(g_2)| + |\ell(g_2) - f_n(g_2)| \\ &\leq 3b^{-c_2 n} \\ &\leq 3d_{\mathcal{C}}(g_1\ ,\ g_2)^\alpha,\end{aligned}$$

where $\alpha = \frac{c_2}{4(C_1+c_2)}$. □

3.3.5 Large Deviations Remark

Given $g \in \mathcal{C}_+$, by the rate of convergence, there exists a neighbourhood V of g in $\mathcal{C}$ and $\overline{n_1} \in \mathbb{N}$ such that the finite scale drifts ℓ_n converge uniformly to ℓ on V for every $n > \overline{n_1}$. Hence, for every $\varepsilon > 0$ there exists $\bar{n}(\varepsilon)$ such that for every $n \geq \bar{n}(\varepsilon)$ and $g_1 \in V$,

$$|\ell(g_1) - \ell(g)| < \varepsilon$$
$$|\ell_n(g_1) - \ell(g_1)| < \varepsilon.$$

Therefore large deviation estimates can be considered in a stronger manner

Definition 8 (Uniform large deviation estimates)**.** Given $g \in \mathcal{C}_+$ There exists a neighbourhood $V \subset \mathcal{C}$ of g and a constant $c > 0$ such that for every $\varepsilon > 0$, there exists $\bar{n}_0$ such that

$$\mu\left\{\omega \in \Omega \,:\, \left|\frac{1}{n} d(g_1^{(n)}(\omega)x_0\,,\, x_0) - \ell(g_1)\right| > \varepsilon\right\} < b^{-cn\varepsilon^2},$$

for every $g_1 \in V$ and $n \geq \bar{n}_0$.

3.4 Continuity of the tracking point

Proving the continuity of the tracking point is similar to proving the continuity of the drift although some of the hard work has already been done.

Let $g \in \mathcal{C}$, we start by considering the positional maps

$$p_g^{(n)} : \Omega \to X$$
$$\omega \mapsto g^{(n)}(\omega)x_0$$

and consider their limit in Bord$X$

$$p_g^{(\infty)}(\omega) := \lim p_g^{(n)}(\omega),$$

whose existence we shall discuss later in section 4.4.2. Notice that if $g \in \mathcal{C}_+$, then for almost every $\omega \in \Omega$

$$\xi_g(\omega) = p_g^{(\infty)}(\omega).$$

Given $g_1, g_2 \in \mathcal{C}$ we define the quantity

$$d_1(p_{g_1}, p_{g_2}) = \int_\Omega \bar{D}_b(g_1(\omega)x_0, g_2(\omega)x_0) d\mu(\omega). \tag{3.6}$$

The route to prove continuity of ξ is the same as the one done before for the drift ℓ. We check the finite scale continuity with respect to d_1 first and then we compute the rate of convergence. Since the space is strongly hyperbolic we then obtain

$$d_1(\xi_{g_1}(\omega), \xi_{g_2}(\omega)) = \lim_{n\to\infty} d_1(p_{g_1}^{(n)}, p_{g_2}^{(n)}).$$

3.4.1 Finite Scale Continuity

Proposition 28. *Let $g \in S^\infty_+(\Omega, G)$, there exist $c = c(g) > 0$, $r > 0$, $\varepsilon > 0$ and $C_2 = C_2(g, \varepsilon) < \infty$ such that for every $g_1, g_2 \in S^\infty(\Omega, G)$ with $d_\infty(g, g_i) < r$ if $n \geq \bar{n}(\varepsilon)$ and $d_\infty(g_1, g_2) < b^{-C_2 n}$, then for every ω outside a set of measure $\lesssim b^{-nc\varepsilon^2}$*

$$\overline{D}_b(p^{(n)}_{g_1}(\omega)\,,\, p^{(n)}_{g_2}(\omega)) \leq b^{-nc\varepsilon^2}.$$

Hence

$$d_1(p^{(n)}_{g_1}\,,\, p^{(n)}_{g_2}) \lesssim b^{-nc\varepsilon^2}.$$

Proof. Consider c to be the large deviation parameter. By the continuity of $\ell(g)$, take $0 < \gamma_1 < \ell(g) < \gamma_2$ close enough so that

$$\begin{aligned}\gamma_1 &< \inf\{\ell(g_*)\ :\ g_* \in S^\infty(\Omega, G) \text{ and } d_\infty(g, g_*) < r\}\\ &\leq \sup\{\ell(g_*)\ :\ g_* \in S^\infty(\Omega, G) \text{ and } d_\infty(g, g_*) < r\} < \gamma_2,\end{aligned}$$

as well as $\varepsilon > 0$ so that

$$c\varepsilon^2 \leq \gamma_1 \leq \ell(g_*) - \varepsilon \leq \ell(g_*) + \varepsilon \leq \gamma_2$$

for every $g_* \in S^\infty(\Omega, G)$ and $d_\infty(g, g_*) < r$.

For every $n \geq \bar{n}(\varepsilon)$, the deviation sets

$$\mathcal{B}_n(g_*) = \left\{\omega \in \Omega\ :\ \left|\frac{1}{n} d(g_*^{(n)}(\omega)x_0\,,\, x_0) - \ell(g_*)\right| > \varepsilon\right\}$$

have their measure bounded by $\lesssim b^{-nc\varepsilon^2}$.

Let $\omega \notin \mathcal{B}_n(g_1) \cup \mathcal{B}_n(g_2)$, then for $i = 1, 2$

$$\begin{aligned} d(g_i^{(n)}(\omega)x_0, x_0) &< n(\ell(g_i) + \varepsilon) < n\gamma_2,\\ d(g_i^{(n)}(\omega)x_0, x_0) &> n(\ell(g_i) - \varepsilon) > n\gamma_1.\end{aligned}$$

At this point, notice as in the proof of Proposition 22

$$\begin{aligned} d(g_1^{(n)}(\omega)x_0\,,\, g_2^{(n)}(\omega)x_0) &\leq \frac{1}{\log(b)} b^{d(g_1^{(n)}(\omega)x_0\,,\, g_2^{(n)}(\omega)x_0)/2}\\ &\leq \frac{b^{n\gamma_2}}{\log(b)} d_\infty(g_1, g_2).\end{aligned}$$

Finally, choosing $C_2 > \gamma_2$, provided $d_\infty(g_1, g_2) < b^{C_2 n}$,

$$\begin{aligned}\overline{D}_b(p^{(n)}_{g_1}(\omega)\,,\, p^{(n)}_{g_2}(\omega)) &\leq b^{-\langle g_1^{(n)}(\omega)x_0\,,\, g_2^{(n)}(\omega)x_0\rangle_{x_0}}\\ &\leq b^{\frac{1}{2}\left[d(g_1^{(n)}(\omega)x_0, g_2^{(n)}(\omega)x_0) - d(g_1^{(n)}(\omega)x_0, x_0) - d(g_2^{(n)}(\omega)x_0, x_0)\right]}\\ &\leq b^{-n\gamma_1} \leq b^{-nc\varepsilon^2}.\end{aligned}$$

□

3.4.2 Rate of Convergence

Proposition 29. *Let $g \in S^{\infty}_{+}(\Omega, G)$. There are constants $r > 0$, $\varepsilon > 0$ and $\overline{n_0} \in \mathbb{N}$, all depending on g, such that*

$$d_1(p^{(n)}_{g_1}\,,\, p^{(\infty)}_{g_1}) \lesssim b^{-nc\varepsilon^2}$$

for all $n \geq \overline{n_0}$ and for all $g_1 \in S^{\infty}(\Omega, G)$ with $d_\infty(g_1, g) < r$. In particular, $p^{(\infty)}_{g_1}$ is well defined.

Proof. Consider γ_1 and ε given as in the proof of the previous proposition and c the large deviation paramenter. As well as the deviation sets

$$\mathcal{B}_n(g_1) = \left\{ \omega \in \Omega \,:\, \left| \frac{1}{n} d(g_1^{(n)}(\omega)x_0\,,\, x_0) - \ell(g_1) \right| > \varepsilon \right\}$$

Recall the control, $d_\infty(g_1) < C$, for every $\omega \notin \mathcal{B}_n(g_1)$

$$\begin{aligned}\overline{D}_b(g_1^{(n)}(\omega)x_0\,,\, g_1^{(n+1)}(\omega)x_0) &\leq b^{\frac{1}{2}\left[d(g_1(T^n\omega)x_0,x_0) - d(g_1^{(n)}(\omega)x_0,x_0) - d(g_1^{(n+1)}(\omega)x_0,x_0)\right]} \\ &\leq \sqrt{C} b^{-n\gamma_1}\end{aligned}$$

Hence, for every $m > n$

$$\begin{aligned}\overline{D}_b(g_1^{(n)}(\omega)x_0\,,\, g_1^{(m)}(\omega)x_0) &\leq \sum_{i=n}^{m-1} D_b(g_1^{(i)}(\omega)x_0\,,\, g_1^{(i+1)}(\omega)x_0) \\ &\leq \sqrt{C} \sum_{i=n}^{m-1} b^{-i\gamma_1} \\ &\leq \frac{\sqrt{C}}{1 - b^{-\gamma_1}} b^{-n\gamma_1},\end{aligned}$$

hence $g_1^{(n)}(\omega)x_0$ is a Gromov sequence, in particular it converges to some point in ∂X. With this we obtain $\overline{D}_b(p^{(n)}_{g_1}\,,\, p^{(\infty)}_{g_1}) \lesssim b^{-nc\varepsilon^2}$. Integrating over ω yiels the result. □

The proof of item 3) in Theorem 18 is now analogue to that of item 1).

3.4.3 Large deviations remark

Given $g \in S^{\infty}_{+}(\Omega, G)$, by the rate of convergence, there exists a neighbourhood V of g in $S_{+}(\Omega, G)$ and $\overline{n_1} \in \mathbb{N}$ such that the finite scale drifts ℓ_n converge uniformly to ℓ on V. Hence, for every $\varepsilon > 0$ there exists $\bar{n}(\varepsilon)$ such that for every $n \geq \bar{n}(\varepsilon)$ and $g_1 \in V$,

$$\begin{aligned}|\ell(g_1) - \ell(g)| &< \varepsilon \\ |\ell_n(g_1) - \ell(g_1)| &< \varepsilon.\end{aligned}$$

Therefore large deviation estimates can be restated in a stronger manner

Definition 9 (Uniform large deviation estimates)**.** Given $g \in S^{\infty}_{+}(\Omega, G)$ There exists a neighbourhood $V \subset S^{\infty}(\Omega, G)$ of g and a constant $c > 0$ such that for every $\varepsilon > 0$, there exists $\bar{n}_0$ such that

$$\mu \left\{ \omega \in \Omega \, : \, \left| \frac{1}{n} d(g_1^{(n)}(\omega)x_0 \, , \, x_0) - \ell(g_1) \right| > \varepsilon \right\} < b^{-cn\varepsilon^2},$$

for every $g_1 \in V$ and $n \geq \bar{n}_0$.

3.5 Positivity Argument

In the previous sections we have seen that the avalanche principle can be used to transport continuity further in time. We now apply these techniques to transport positivity forward. Once again our secondary tool will be large deviation estimates, which for this purpose don't need them to be as strong. We say that a cocycle satisfies large deviation of polynomial type if there are constants $C, \sigma > 0$ such that for every $n \in \mathbb{N}$,

$$\int_{\Omega} \left| \frac{1}{n} d(g^{(n)}(\omega))x_0 \, , \, x_0) - \ell_n \right| d\mu(\omega) \leq CKn^{-\sigma}, \tag{3.7}$$

where $K := \sup_{n \in \mathbb{N}} \sup_{\omega \in \Omega} \frac{1}{n} d(g^{(n)}(\omega)x_0 \, , \, x_0)$.

Theorem 30 (Positivity criteria)**.** *Let G be the group of isometries of a δ-hyperbolic space X with basepoint x_0 and let (Ω, β, μ, T) be an ergodic measure transformation and $g : \Omega \to G$. Assume that there exists $q \in \mathbb{N}$ such that, for every $n > q$ the large estimates (3.7) hold for some given σ. Then, there exists $\overline{n_0} = \overline{n_0}(\sigma, \delta) \in \mathbb{N}$ such that if $n > n_0$ satisfies*

$$\begin{cases} \ell_{\overline{n_0}} > K\overline{n_0}^{-\sigma/4}, \\ \ell_{\overline{n_0}} - \ell_{2\overline{n_0}} < \frac{\ell_{\overline{n_0}}}{8}, \end{cases} \tag{3.8}$$

then $\ell > \ell_n/2$.

Notice that here we only request that our space is hyperbolic, this happens as we will use the remark to 19, which makes the conclusion of the avalanche principle

$$\left| d(x_0, x_n) + \sum_{i=2}^{n-1} d(x_{i-1}, x_i) - \sum_{i=1}^{n-1} d(x_{i-1}, x_{i+1}) \right| \leq 2(n-1)\delta.$$

Lemma 31 (Inductive Procedure)**.** *Let $n_0, k > 0$ such that, for every $n > n_0$ the polynomial large deviation estimates are satisfied. If*

$$\begin{cases} \ell_{n_0} > Kn_0^{-\sigma/4}, \\ \ell_{n_0} - \ell_{2n_0} < \frac{\ell_{n_0}}{8}, \end{cases} \tag{3.9}$$

and

$$\frac{6\kappa - 24}{8\kappa} K n_0^{1-\sigma/4} > 3\delta, \tag{3.10}$$

then defining $n_1 = n_0^{1+\tau}$ *where* $\tau = \frac{3}{8}\sigma$*, one has*

$$|\ell_{n_1} - 2\ell_{2n_0} + \ell_{n_0}| < C(\kappa) K \frac{n_0}{n_1}. \tag{3.11}$$

Moreover,

$$\begin{cases} \ell_{n_1} > K n_1^{-\sigma/4}, \\ \ell_{n_1} - \ell_{2n_1} < C(\kappa) K \frac{n_0}{n_1} < \frac{\ell_{n_1}}{8}. \end{cases} \tag{3.12}$$

Proof. Consider $q > 0$ and $r < n_0$ such that $n_1 = qn_0 + r$ and, for $0 \leq i \leq q$ set

$$x_i = g^{(n_0)}(\omega) g^{(n_0)}(T^{n_0}\omega) ... g^{(n_0)}(T^{n_0(i-1)}\omega) x_0.$$

Due to Lemma 24, we can once again assume that $r = 0$ as otherwise we incur in an error of the order of our control. We also consider $\sigma < 8/3$ so that $\tau < 1$. This is not a huge loss, as if we have large deviation estimates estimates for $\sigma > 8/3$, then they also hold for $\sigma < 8/3$. The aforementioned LDT estimate implies

$$\mu \left\{ \omega \in \Omega \, : \, |d(g^{(n_0)}(\omega) x_0, x_0) - n_0 \ell_{n_0}| > K\varepsilon n_0 \right\} \leq C \varepsilon^{-1} n_0^{-\sigma}.$$

Using the large deviation in this form with $K\varepsilon = \ell_{n_0}/\kappa$, one has

$$\min_{0 \leq i < q} d(x_i, x_{i+1}) = \min_{0 \leq i < q} d\left(g^{(n_0)}(T^{n_0 i}(\omega)) x_0 \, , \, x_0)\right) > n_0 \ell_{n_0} - \frac{1}{\kappa} n_0 \ell_{n_0} = \frac{\kappa - 1}{\kappa} n_0 \ell_{n_0},$$

on a set $\mathcal{B}_1 \subset \Omega$ such that

$$\mu(\Omega \backslash \mathcal{G}_1) < C \varepsilon^{-1} q \, n_0^{-\sigma} = C q \, n_0 \, n_0^{-1-3\sigma/4} \, n_0^{-\sigma/4} \varepsilon^{-1} \leq \kappa C n_1 \, n_0^{-1-3\sigma/4} \leq \kappa C n_0^{-\tau}.$$

In the same fashion, there is a set $\mathcal{B}_2 \subset \Omega$ such that $\mu(\Omega \backslash \mathcal{B}_2) < \kappa C n_0^{-\tau}$ on which

$$\begin{aligned} \max_{0 < i < q} \langle x_{i+1} \, , \, x_{i-1} \rangle_{x_i} &= \max_{0 < i < q} \frac{1}{2} \Bigg[d\left(g^{(n_0)}(T^{n_0 \, i}(\omega)) x_0 \, , \, x_0\right) + d\left(g^{(n_0)}(T^{n_0(i-1)}(\omega)) x_0 \, , \, x_0)\right) \\ &\qquad - d\left(g^{(n_0)}(T^{n_0 \, i}\omega) g^{(n_0)}(T^{n_0(i-1)}\omega) x_0 \, , \, x_0\right) \Bigg] \\ &\leq \frac{1}{2} \left[n_0(\ell_{n_0} + K\varepsilon) - n_0(\ell_{2n_0} - K\varepsilon) \right] \\ &= n_0 \left(\ell_{n_0} - \ell_{2n_0} + \frac{1}{\kappa} \ell_{n_0} \right) \end{aligned}$$

$$\leq \frac{1}{8} n_0 \ell_{n_0} + \frac{1}{\kappa} n_0 \ell_{n_0} = \frac{\kappa+8}{8\kappa} n_0 \ell_{n_0}.$$

Now, $2\frac{\kappa+8}{8\kappa} n_0 \ell_{n_0} < \frac{\kappa-1}{\kappa} n_0 \ell_{n_0} - 2\delta$ is an immediate consequence of condition (3.10), whence the Avalanche Principle applies,

$$\Bigg| d\left(g^{(n_1)}(\omega)x_0\,,\, x_0\right) + \sum_{i=2}^{q-1} d\left(g^{(n_0)}(T^{n_0 i}\omega)x_0\,,\, x_0\right) \\ - \sum_{i=1}^{q-1} d\left(g^{(n_0)}(T^{n_0(i+1)}\omega)g^{(n_0)}(T^{n_0 i}\omega)x_0\,,\, x_0\right)\Bigg| \leq 2q\delta.$$

Dividing both sides by n_1, one has

$$\Bigg|\frac{1}{n_1} d\left(g^{(n_1)}(\omega)x_0\,,\, x_0\right) + \frac{1}{q}\sum_{i=2}^{q-1} \frac{1}{n_0} d\left(g^{(n_0)}(T^{n_0 i}\omega)x_0\,,\, x_0\right) \\ - \frac{2}{q}\sum_{i=1}^{q-1} \frac{1}{2n_0} d\left(g^{(n_0)}(T^{n_0(i+1)}\omega)g^{(n_0)}(T^{n_0 i}\omega)x_0\,,\, x_0\right)\Bigg| \leq \frac{2\delta}{n_0}.$$

Call $f(\omega)$ the left-hand side of the inequality, integration over ω yields

$$\begin{aligned} \int_\Omega f(\omega)d\mu(\omega) &= \int_{\mathcal{G}_1\cap\mathcal{G}_2} f(\omega)d\mu(\omega) + \int_{\Omega\setminus\mathcal{G}_1\cap\mathcal{G}_2} f(\omega)d\mu(\omega) \\ &\leq \frac{2\delta}{n_0} + 2\kappa K n_0^{-\tau} \\ &\leq C_1 K \frac{n_0}{n_1}, \end{aligned}$$

as $n_0^{-\tau} = n_0/n_1$ and $\tau < 1$ ($C_1 = 2\delta + 2\kappa$)) . In other words

$$\left|\ell_{n_1} - 2\ell_{2n_0} + \ell_{n_0} - \frac{2}{q}\left[\ell_{n_0} - \ell_{2n_0}\right]\right| \leq C_1 K \frac{n_0}{n_1}$$

thus proving (3.11), as $\ell_{n_0} - \ell_{2n_0} < \ell_{n_0}/8 \leq K/8$.

For the upper inequality in (3.12), notice

$$\begin{aligned} \ell_{n_1} &\geq \ell_{n_0} - 2(\ell_{n_0} - \ell_{2n_0}) - C_1 K \frac{n_0}{n_1} \\ &\geq \ell_{n_0} - \frac{2\ell_{n_0}}{8} - C_1 K \frac{n_0}{n_1} \\ &\geq \frac{3K n_0^{-\sigma/4}}{4} - C_1 K \frac{n_0}{n_1} \\ &\geq K n_1^{-\sigma/4} \end{aligned}$$

Provided that n_0 is large enough.

For the first lower inequality, simply use the Avalanche Principle once again for the times $2n_0$ and subtract in (3.11); whilst the second is true provided n_0 is large enough. □

Proof of Theorem 30. Making n_0 large enough, the existence of κ in Lemma 31 is obvious and consider $C = C(\kappa) > 0$ as in the Lemma. Now given $n \geq \overline{n_0}$ in the conditions of Theorem 30, we will apply the previous Lemma inductively starting with $n_0 = n$ to obtain

$$|\ell_{n_{j+1}} - 2\ell_{2n_j} + \ell_{n_j}| < CK\frac{n_j}{n_{j+1}} \tag{3.13}$$

and

$$\begin{cases} \ell_{n_{j+1}} > Kn_{j+1}^{-\sigma/4} \\ \ell_{n_{j+1}} - \ell_{2n_{j+1}} < CK\frac{n_j}{n_{j+1}} < \frac{\ell_{n_{j+1}}}{8} \end{cases} . \tag{3.14}$$

where $n_{j+1} = n_j^{1+\tau}$ for $\tau = \frac{3}{8}\sigma$.

Whence

$$\begin{aligned} \ell_{n_{j+1}} &> \ell_{n_j} - 2(\ell_{n_j} - \ell_{2n_j}) - CK\frac{n_j}{n_{j+1}} \\ &> \ell_{n_j} - 3CK\frac{n_{j-1}}{n_j} \\ &> \ell_{n_0} - 2(\ell_{n_0} - \ell_{2n_0}) - 3CK\left[\sum_{i=1}^{j}\frac{n_i}{n_{i+1}}\right] \\ &> \frac{3}{4}\ell_{n_0} - 3CK\left[\sum_{i=1}^{j}\frac{n_i}{n_{i+1}}\right]. \end{aligned}$$

Since $\sum_{i=1}^{\infty}\frac{n_i}{n_{i+1}}$ is asymptotic to $n_0^{-\tau}$, the intended result follows since $\ell_{n_0} > Kn_0^{-\sigma/4}$. ☐

Having explored the proofs, something more can be said about how large $\overline{n_0}$ must be. First assume it is large enough so that there exists κ satisfying

$$\frac{6\kappa - 24}{8\kappa}K\overline{n_0}^{1-\sigma/4} > 3\delta,$$

$\overline{n_0}(\sigma)$ must satisfy

$$\frac{3}{4}\overline{n_0}^{-\sigma/4} - (2\delta - 2\kappa)\,\overline{n_0}^{-3\sigma/8} \geq \overline{n_0}^{-\sigma(1+3\sigma/8)/4}.$$

Remark 3. *In negatively curved Riemannian manifold with non-negative Ricci curvature, horofunctions are known to be plurisubharmonic functions (see Cheeger and Gromoll (1971)). Hence the argument used later in chapter 5 for quasi-periodic cocycles can be applied to this setting. We chose not to do it as it would require us to introduce additional notation from ergodict theory, harmonic analysis and potential theory which could disperse the thesis topics even further.*

3.6 Bibliographic Notes

The drift in the metric hyperbolic setting has a quite similar behaviour to the Lyapunov exponents in the linear cocycle setting. In fact, if we consider $SL(2,\mathbb{R})$ acting isometrically on the hyperbolic upper-half plane $\mathbb{H}^2$ the two concepts overlap. We explore this common playground to use techniques that were originally obtained for the study of the Lyapunov exponents in $SL(2,\mathbb{R})$, while working around the metric technicalities.

The type of abstract continuity and positivity results obtained in this chapter for hyperbolic spaces is new. The techniques employed however date back to Goldstein and Schlag (2001), where the first linear version of the avalanche principle for $\mathrm{SL}(2,\mathbb{R})$-cocycles appears. In their paper, Goldstein and Schlag (2001) use the avalanche principle to obtain continuity and positivity of the Lyapunov exponent of quasi-periodic Schrodinger cocycles. The continuity argument for general $\mathrm{SL}(d,\mathbb{R})$-cocycles as well as cocycles over non-invertible cocycles appeared later in Duarte et al. (2016). Note that the case of higher dimension Schrödinger cocycles in $\mathrm{SL}(d,\mathbb{R})$ can also be found in Schlag (2013). Han et al. (2020) deals once again with the problem of positivity of the Lyapunov exponent; here large deviations estimates are replaced with other quatifications of the convergence sets. The spirit of the argument however remains the same.

The continuity of the tracking point is related to the continuity of the Lyapunov filtration for $\mathrm{SL}(2)$-cocycles presented in Duarte et al. (2016), in fact for $\mathrm{SL}(2)$-cocycles the two concepts once again agree.

In this work we decided to use large deviations of exponential type. These can be relaxed to weaker type of deviations yielding weaker modulus of continuity.

Despite the similarities, there are several interesting natural actions of groups acting by isometries on hyperbolic spaces that escape the linear setting such as rank one semisimple Lie Groups acting on their symmetric spaces, Gromov hyperbolic groups acting on their Cayley graphs, mapping class groups on their curve complexes, the Cremona group acting on the Picard-Manin hyperbolic space among others (see Maher and Tiozzo (2018)). In some of these examples the fact we drop the usual properness condition is quite important.

A gain in our versions of the avalanche principle is that every constant is explicit and the result verses on chains of points rather than the operators themselves. A first version of the avalanche principle for CAT(-1) may be found in Oregón-Reyes (2020). The geometric applications presented by Oregón-Reyes exhibit the importance of working with chains of points.

Chapter 4

Markov Systems

In this chapter we apply the abstract continuity theorem to cocycles of isometries of hyperbolic space, where the base dynamics is governed by Markov systems. In view of the previous chapter all we must do is prove that exponential type large deviations hold. This path will only lead us to continuity in strong hyperbolic spaces.

In the second part of the chapter we assume (BA) holds, and prove continuity with respect to the measure in the case of random walks on the group. Random walks are simple examples of Markov systems where each successive element is picked in a random independently and identically distributed way. In other words, a measure on the group is enough to obtain a random walk, as the elements are picked with respect to this measure. We shall endow the space of probability measures with a topology to describe continuity in this setting.

4.1 Markov Systems

We now introduce Markov systems. For such systems the current configuration does not depend on the past nor the future but only on the present. We will begin by introducing the probabilistic language which we will use later, and then briefly present how to translate it into the dynamical language used previously through the Markov shift. Our presentation on the subject follows that of Duarte et al. (2016).

Definition 10 (Markov Kernel)**.** Let Γ be a metric space and let $\mathcal{F}$ be its Borel σ-algebra. A Markov kernel is a function $K : \Gamma \times \mathcal{F} \to [0, 1]$ such that

1. for every $\omega_0 \in \Gamma$, $E \mapsto K(\omega_0, E)$ is a probability measure on Γ;

2. The mapping $\omega_0 \mapsto K(\omega_0, \cdot)$ is continuous with respect to the weak-* topology in $\text{Prob}(\Gamma)$.

3. for every $E \in \mathcal{F}$, the function $\omega_0 \mapsto K(\omega_0, E)$ is $\mathcal{F}$-measurable.

A probability measure μ on $(\Gamma, \mathcal{F})$ is K-stationary if for every $E \in \mathcal{F}$,

$$\mu(E) = \int_\Gamma K(\omega_0, E)\mu(d\omega_0).$$

A set $E \in \mathcal{F}$ is said to be K-invariant when $K(\omega_0, E) = 1$ for all $\omega_0 \in E$ and $K(\omega_0, E) = 0$ for all $\omega_0 \in \Gamma \backslash E$. A K-stationary measure μ is called ergodic when there is no K-invariant set $E \in \mathcal{F}$ such that $0 < \mu(E) < 1$. Using the usual argument through Krein-Milman's theorem, ergodic measures are the extremal points in the convex set of K-stationary measures. A Markov system is a pair (K, μ), where K is a Markov kernel on $(\Gamma, \mathcal{F})$ and μ is a K-stationary probability measure.

In Duarte and Klein (2017), the considerations above are only done for compact Γ as this easily yields the existence of stationary measures, in this work however we will also need to work with non-compact spaces. Fortunately we will be able to find stationary measures for the non-compact cases that interest us.

We can define the iterated Markov kernels inductively, setting $K^1 = K$ and

$$K^{n+1}(\omega_0, E) = \int_\Gamma K^n(\omega_1, E)K(\omega_0, d\omega_1),$$

for $n > 1$.

Given (K, μ) a pair formed by a Markov Kernel a not necessarily stationary measure $\mu \in \text{Prob}(\Gamma)$, consider $\Omega = \Gamma^{\mathbb{N}}$ the space of sequences $\omega = (\omega_n)$ in Γ. The product space Ω is metrizable. Its Borel σ-algebra $\mathcal{B} = \mathcal{F}^{\mathbb{N}}$ is the product σ-algebra generated by the $\mathcal{B}$-cylinders, that is, generated by the sets

$$C(E_0, ..., E_m) := \{\omega \in \Omega \,:\, \omega_j \in E_j, \text{ for } 0 \leq j \leq m\},$$

where $E_0, ..., E_m \in \mathcal{F}$.

The set of $\mathcal{F}$-cylinders forms a semi-algebra on which

$$\mathbb{P}_\mu[C(E_0, ..., E_m)] := \int_{E_m} ... \int_{E_0} \mu(d\omega_0) \prod_{j=1}^m K(\omega_{j-1}, d\omega_j).$$

defines a pre-measure. By Carathéodory's extension theorem, it extends to a measure, still denoted $\mathbb{P}_\mu$ and often called the Kolmogorov extension of (K, μ), on $(\Omega, \mathcal{B})$.

Given a random variable $\zeta : \Omega \to \mathbb{R}$, its expected value with respect to μ in $(\Gamma, \mathcal{F})$ is

$$\mathbb{E}_\mu(\zeta) := \int_\Omega \zeta d\mathbb{P}_\mu.$$

If μ is δ_{ω_0} the Dirac measure at ω, then we soften the notation by setting $\mathbb{P}_{\omega_0} = \mathbb{P}_{\delta_{\omega_0}}$ and $\mathbb{E}_{\omega_0} = \mathbb{E}_{\delta_{\omega_0}}$.

By construction, the sequence of random variables $e_n : \Omega \to \Gamma$, given by $e_n(\omega) := \omega_n$ for $\omega = (\omega_n) \in \Omega$, is a Markov chain with initial distribution μ and transition kernel K, that is, for every $\omega \in \Gamma$ and $E \in \mathcal{F}$,

1. $\mathbb{P}_\mu[e_0 \in E] = \mu(E)$,

2. $\mathbb{P}_\mu[e_n \in E \,|\, e_{n-1} = \omega_{n-1}] = K(\omega_{n-1}, E)$.

Moreover the process $\{e_n\}_{n\in\mathbb{N}}$ is stationary with respect to $(\Omega, \mathcal{F}, \mathbb{P}_\mu)$ if and only if μ is K-stationary.

We now proceed to make sense of the dynamics behind Markov systems. Consider the shift map $T : \Omega \to \Omega$, $T(\omega_n) = (\omega_{n+1})$. The shift T is continuous and hence $\mathcal{B}$ measurable. It also preserves the measure $\mathbb{P}_\mu$. We call the triplet $(\Omega, \mathbb{P}_\mu, T)$ a Markov shift.

Definition 11 (Strongly Mixing)**.** Let B be a Banach space contained in $L^\infty(\Gamma)$. We say a Markov system (K, μ) is strongly mixing in B if there are constants $C > 0$ and $0 < \sigma < 1$ such that for every $f \in B$, all $x \in \Gamma$ and $n \in \mathbb{N}$,

$$\left| \int_\Gamma f(\omega_1) K^n(\omega_0, d\omega_1) - \int_\Gamma f(\omega_1)\mu(d\omega_1) \right| \le C\sigma^n ||f||_B.$$

Strongly mixing property of a Markov system is related Markov shift being mixing.

Definition 12. (Mixing transformation We say that a measure preserving transformation (Ω, T, μ) is mixing if for all measurable A and B in B

$$\lim_{n\to\infty} \mu(A \cap T^{-n}(B)) = \mu(A)\mu(B).$$

If we pick $A = B$ a T-invariant set in the definition of mixing, we obtain $\mu(A) = \mu(A)^2$, so A has either measure 0 or 1. In other words, mixing implies ergodic. The converse is not true. The two concepts of mixing now come together in the following proposition.

Proposition 32 (Proposition 5.1 in Duarte et al. (2016))**.** *If the Markov system (K, μ) is strongly mixing, then Markov shift $(\Omega, \mathbb{P}_\mu, T)$ is a mixing dynamical system.*

Suppose now that Σ is compact, any continuous $g \in C(\Sigma \times \Sigma, G) \subset S^\infty(\Sigma \times \Sigma, G)$, where $S^\infty(\Sigma \times \Sigma, G)$ is the subspace of $S^\infty(\Omega, G)$, consisting of cocycles which depend only on the first two variables, defines a cocycle in G over the Markov shift $(\Omega, \mathbb{P}_\mu, T)$, $a : \mathbb{N} \times \Omega \to G$ given by

$$a(n, \omega) = g^{(n)}(\omega) := g(\omega_0, \omega_1)g(\omega_1, \omega_2)...g(\omega_{n-1}, \omega_n).$$

From this point on we will also omit the reference to the ω's in $g^{(n)}(\omega)$ whenever there is no room for confusion, by simply writing $g^{(n)}$.

4.2 Large deviations for the drift in Markov systems

In this section we obtain that exponential type large deviations hold for the drift over Markov systems. Although the method used is based in Nagaev (1957), we will apply Duarte et al. (2016) recipe. In §4.2.1 we describe the recipe and ready the ingredients laid by Duarte and Klein whilst §4.2.2 is devoted to proving the large deviations. Many of the arguments displayed here are an adaptation of what was done in Sampaio (2021) for random walks, which will reappear in the next section.

Let us recall the reader once more that X stands for a δ-hyperbolic metric space with a basepoint x_0, G for its groups of isometries and b for a real number between 1 and $2^{1/\delta}$. We use Σ for a compact metric space where a Markov system (K, μ) lives.

Nagaev's method, also called the spectral method, is based on the contracting properties of suitable operators on Banach spaces. In order to ensure such properties hold, we need some hypothesis on the Markov system. We say that a cocycle $g \in S^{\infty}(\Sigma \times \Sigma, G)$ is irreducible with respect to (K, μ) if there is no map $h : \Sigma \to X^h$ such that

$$g(\omega_{n-1}, \omega_n)h(\omega_{n-1}) = h(\omega_n)$$

for $\mathbb{P}_{\mu}$-almost every ω_n. We now set $\mathcal{I}(K)$ to be the class of irreducible continuous cocycles with respect to (K, μ) in $C(\Sigma \times \Sigma, G) \subset S^{\infty}(\Sigma \times \Sigma, G)$. One can prove such a class is open in $S^{\infty}(\Sigma \times \Sigma, G)$.

Theorem 33. *Let Σ be a compact metric space, (K, μ) be a strongly mixing Markov system over Σ and $g \in \mathcal{I}(K)$ be a cocycle with positive drift. Then g satisfies uniform large deviations estimates.*

And as a corollary, by the abstract continuity theorem we finally obtain continuity of the drift in $\mathcal{I}(K)$, moreover this continuity is Hölder in a neighbourhood of cocycles with positive drift.

Theorem 34. *Let X be a strongly hyperbolic space, Σ be a compact metric space and (K, μ) be a strongly mixing Markov system over Σ. Then the drift $\ell : \mathcal{I}(K) \to \mathbb{R}$ is continuous, moreover it is locally Hölder continuous when restricted to the subset $\mathcal{I}(K)_+ \subset \mathcal{I}(K)$ formed by the cocycles with positive drift. Lastly, $\xi : \mathcal{I}(K)_+ \to S^1(\Sigma \times \Sigma, \partial X)$ is locally Hölder continuous.*

4.2.1 The Method

Consider a Markov system (K, μ) on a metric space Γ and let $\Omega = \Gamma^{\mathbb{N}}$. Given some Borel measurable observable $\zeta : \Gamma \to \mathbb{R}$, let $\hat{\zeta} : \Omega \to \mathbb{R}$ be the Borel measurable function $\hat{\zeta}(\omega) = \zeta(\omega_0)$. We call a sum process of $\zeta : \Gamma \to \mathbb{R}$ the sequence of random variables $\{S_n(\zeta)\}$ on $(\Omega, \mathcal{B})$,

$$S_n(\zeta)(\omega) := \sum_{i=0}^{n-1} \hat{\zeta} \circ T^j(\omega) = \sum_{i=0}^{n-1} \zeta(\omega_j).$$

An observed Markov system on Γ is a triple (K, μ, ζ) where (K, μ) is a Markov system on Γ and $\zeta : \Gamma \to \mathbb{R}$ is a Borel-measurable function.

Recall that $\mathbb{P}_{\omega_0}$ stands for the Kolmogorov extension of (K, δ_{ω_0}).

Definition 13 (Large deviations of exponential type)**.** We say that ζ satisfies large deviation of exponential type if there exist positive constants b, C, k, ε_0 and n_0 such that for all $n > n_0$, $0 < \varepsilon < \varepsilon_0$ and $\omega_0 \in \Gamma$,

$$\mathbb{P}_{\omega_0} \left\{ \omega \in \Omega \, : \, \left| \frac{1}{n} S_n(\zeta)(\omega) - \mathbb{E}_\mu(\zeta) \right| > \varepsilon \right\} \leq C b^{-k\varepsilon^2 n}.$$

We will obtain the large deviations in Theorem 34 by exploring the properties of contracting operators on suitable Banach spaces. Let us start by introducing the operators. Consider K a Markov kernel on a metric space Γ, the operator $Q_K : L^\infty(\Gamma) \to L^\infty(\Gamma)$, given by

$$(Q_K f)(\omega_0) = \int_\Gamma f(\omega_1) K(\omega_0, d\omega_1),$$

is called the Markov operator. The Markov operator allows us to characterize stationary measure in a more useful way, with effect, μ is K-stationary if and only if

$$\int Q_K f d\mu = \int f d\mu$$

for every $f \in L^1(\Gamma)$. Let now (K, μ, ζ) be an observed Markov system on a given metric space Γ, then we call the operator $Q_{K,\zeta} : L^\infty(\Gamma) \to L^\infty(\Gamma)$ given by

$$(Q_{K,\zeta} f)(\omega_0) := \int_\Gamma f(\omega_1) b^{\zeta(\omega_1)} K(\omega_0, d\omega_1),$$

and $b > 0$ the Laplace-Markov operator.

We will now follow closely Duarte et al. (2016) as we introduce a series of assumptions, eleven to be exact, which yield an abstract LDT. In the next section we make sense of this setting and prove the assumptions hold as to obtain the large deviations. The main difference between the two settings is the fact that we apply these results to not necessarily compact spaces.

Let $(\mathcal{M}, \text{dist})$ be a metric space of observed Markov systems (K, μ, ζ) on a given metric space Γ. Consider as well a scale of Banach algebras $(B_\alpha, ||\cdot||_\alpha)$ indexed in $\alpha \in [0,1]$, where each B_α is a space of bounded Borel measurable functions on Γ. We assume that there exists seminorms $v_\alpha : B_\alpha \to [0, +\infty)$ such that for every $0 \leq \alpha \leq 1$, A1) $||f||_\alpha = \upsilon_\alpha(f) + ||f||_\infty$,

A2) $B_0 = L^\infty(\Sigma)$ and $||\cdot||_0$ is equivalent to $||\cdot||_\infty$,

A3) B_α is a lattice, i.e., if $f \in B_\alpha$ then $\bar{f}, |f| \in B_\alpha$,

A4) B_α is a Banach algebra with unity $\mathbf{1} \in B_\alpha$ and $\upsilon_\alpha(\mathbf{1}) = 0$.

Assume also that for every $0 \le \alpha_0 < \alpha_1 < \alpha_2 \le 1$,

B1) $B_{\alpha_2} \subset B_{\alpha_1} \subset B_{\alpha_0}$,

B2) $\upsilon_{\alpha_0}(f) \le \upsilon_{\alpha_0}(f) \le \upsilon_{\alpha_0}(f)$, for every $f \in B_{\alpha_2}$,

B3) $\upsilon_{\alpha_1}(f) \le \upsilon_{\alpha_0}(f)^{\frac{\alpha_2-\alpha_1}{\alpha_2-\alpha_0}}\, \upsilon_{\alpha_2}(f)^{\frac{\alpha_1-\alpha_0}{\alpha_2-\alpha_0}}$, for every $f \in B_{\alpha_2}$.

The assuptions $A*)$ and $B*)$ exhaust our assuptions on the Banach algebras and will be the simple part of what is to come. Finally, for our assumptions on $\mathcal{M}$, assume there exists an interval $[\alpha_1, \alpha_0] \subset (0, 1]$ with $\alpha_1 < \alpha_0/2$ such that for every $\alpha \in [\alpha_1, \alpha_0]$ the following properties hold,

C1) $(K, \mu, -\zeta) \in \mathcal{M}$, whenever $(K, \mu, \zeta) \in \mathcal{M}$.

C2) The Markov operators $Q_K : B_\alpha \to B_\alpha$ are uniformly strongly mixing. That is, there exist $C > 0$ and $0 < \sigma < 1$ such that for every $(K, \mu, \zeta) \in \mathcal{M}$ and $f \in B_\alpha$,

$$\left|\left| Q_K^n f - \int_\Sigma f(\omega_0)\, d\mu(\omega_0) \right|\right|_\alpha \le C\sigma^n ||f||_\alpha.$$

C3) The operators $Q_{K,z\zeta}$ act continuously on the Banach algebra B_α uniformly in $(K, \mu, \zeta) \in \mathcal{M}$. With effect, we assume, there are positive constants c and M such that for $i = 0, 1, 2$, $|z| < c$ and $f \in B_\alpha$

$$Q_{K,z\zeta}(f\zeta^i) \in B_\alpha \text{ and } ||Q_{K,z\zeta}(f\zeta^i)|| \le M||f||_\alpha.$$

C4) Consider the family of maps $(K, \mu, \zeta) \to Q_{K,z\zeta}$ indexed in $|z| < c$, there exists $0 < \theta \le 1$ such that for every $|z| < b$, $f \in B_\alpha$ and $(K_1, \mu_1, \zeta_1), (K_2, \mu_2, \zeta_2) \in \mathcal{M}$,

$$||Q_{K_1,z\zeta_1} f - Q_{K_2,z\zeta_2} f||_\infty \le M||f||_\alpha \text{dist}\left((K_1, \mu_1, \zeta_1), (K_2, \mu_2, \zeta_2)\right)^\theta.$$

Under all these assumptions the following abstract LDT theorem holds:

Theorem 35 (in Duarte et al. (2016)). *Given $(K_0, \mu_0, \zeta_0) \in \mathcal{M}$ and $0 < s < \infty$ large enough (which can be made precise), there exists a neighbourhood V of $(K_0, \mu_0, \zeta_0) \in \mathcal{M}$, $C > 0$, $\varepsilon_0 > 0$ and $n_0 \in \mathbb{N}$, such that for every $(K, \mu, \zeta) \in V$, $0 < \varepsilon < \varepsilon_0$, $\omega_0 \in \Sigma$ and $n > n_0$*

$$\mathbb{P}_{\omega_0}\left[\left|\frac{1}{n}S_n(\zeta) - \mathbb{E}_\mu(\zeta)\right| \geq \varepsilon\right] \leq C b^{-\frac{\varepsilon^2}{s}n}.$$

which averaging over ω_0 with respect to μ yields

$$\mathbb{P}_{\mu}\left[\left|\frac{1}{n}S_n(\zeta) - \mathbb{E}_\mu(\zeta)\right| \geq \varepsilon\right] \leq C b^{-\frac{\varepsilon^2}{s}n}.$$

Remark 4. *By choosing a large s and $n \geq \bar{n}(\varepsilon)$ we can make $C = 1$, thus obtaining large deviations as in Definition 13.*

4.2.2 Obtaining the Large deviations

Let X be an Hyperbolic metric space, X^h, ∂X denote its horofunctions compactification and Gromov boundary, respectively. We denote by D_b the visual metric on ∂X, where $1 < b \leq 2^{1/\delta}$ is fixed. In this section we use Theorem 35 to obtain our large deviations for the drift. From this point on Σ stands for a compact metric space and $\Gamma = \Sigma \times \Sigma \times \partial X$

Verifying Conditions A*) and B*)

Given $0 \leq \alpha \leq 1$ and $f \in L^\infty(\Gamma)$, define

$$\upsilon_\alpha(f) := \sup_{\substack{(\omega_1,\omega_2)\in\Sigma\times\Sigma \\ \xi\neq\eta}} \frac{|f(\omega_1,\omega_2,\xi) - f(\omega_1,\omega_2,\eta)|}{D_b(\xi,\eta)^\alpha},$$

$$||f||_\alpha := ||f||_\infty + \upsilon_\alpha(f),$$

and set

$$\mathcal{H}_\alpha(\Gamma) := \{\, f \in L^\infty(\Gamma) \,:\, ||f||_\alpha < \infty \}\,.$$

the space of boundary Hölder continuous functions in Γ. We call $\upsilon_\alpha(f)$ the boundary Hölder exponent of f.

Proposition 36. *The family $\{\mathcal{H}_\alpha(\Gamma)\}$, for $0 \leq \alpha \leq 1$, consist of Banach algebras with norm $||f||_\alpha$ satisfying the conditions $A*)$ and $B*)$.*

Proof. It is a standard proof that $\mathcal{H}_\alpha(\Gamma)$ are Banach algebras. Now points $A1)$, $A3)$, $B1)$, $B2)$ are either clear or follow from some immediate computation. For point $A2)$ notice that for $\alpha = 0$ we have $||f||_\alpha \leq 2||f||_\infty$. Point $A4)$ follows from the immediate inequality

$$\upsilon_\alpha(fg) \leq ||f||_\infty \upsilon_\alpha(g) + ||g||_\infty \upsilon_\alpha(f).$$

For point $B3$), notice that given $\alpha_0, \alpha_2, s \in [0,1]$,

$$
\begin{aligned}
\upsilon_{s\alpha_0+(1-s)\alpha_2}(f) &= \sup_{\substack{(\omega_1,\omega_2)\in\Sigma\times\Sigma \\ \xi\neq\eta}} \frac{|f(\omega_1,\omega_2,\xi)-f(\omega_1,\omega_2,\eta)|^{s+(1-s)}}{D_b(\xi,\eta)^{s\alpha_0+(1-s)\alpha_2}} \\
&\leq \sup_{\substack{(\omega_1,\omega_2)\in\Sigma\times\Sigma \\ \xi\neq\eta}} \frac{|f(\omega_1,\omega_2,\xi)-f(\omega_1,\omega_2,\eta)|^{s}}{D_b(\xi,\eta)^{s\alpha_0}} \\
&\qquad \times \sup_{\substack{(\omega_1,\omega_2)\in\Sigma\times\Sigma \\ \xi\neq\eta}} \frac{|f(\omega_1,\omega_2,\xi)-f(\omega_1,\omega_2,\eta)|^{(1-s)}}{D_b(\xi,\eta)^{(1-s)\alpha_2}} \\
&= \upsilon_{\alpha_0}(f)^s \upsilon_{\alpha_2}(f)^{1-s},
\end{aligned}
$$

picking $s = \frac{\alpha_2-\alpha_1}{\alpha_2-\alpha_0}$ the result follows. □

Verifying Conditions C*)

Recall the space $S^\infty(\Sigma\times\Sigma, G)$ of bounded measurable cocycles $g : \Sigma\times\Sigma \to G$ introduced in section 1.3. Each cocycle $g \in S^\infty(\Sigma\times\Sigma,\, G)$ defines a Markov kernel on Γ given by

$$
\overline{K}_g(\omega_0,\omega_1,\xi) := \int_\Sigma \delta_{(\omega_1,\omega_2,g(\omega_1,\omega_2)^{-1}\xi)} K(\omega_1, d\omega_2),
$$

as well as an associated Markov operator $Q_g : L^\infty(\Gamma) \to L^\infty(\Gamma)$ with expression

$$
(\overline{Q}_g f)(\omega_0,\omega_1,\xi) := \int_\Sigma f(\omega_1,\, \omega_2,\, g(\omega_1,\omega_2)^{-1}\xi) K(\omega_1, d\omega_2).
$$

The reason for looking at the action of the inverse comes from the relation (2.5). For each $g \in C(\Sigma\times\Sigma,\, G)$ consider the measurable observable $\zeta_g : \Gamma \to \mathbb{R}$

$$
\zeta_g(\omega_0,\omega_1,\xi) := h_\xi(g(\omega_0,\omega_1)x_0). \tag{4.1}
$$

where h_ξ is the horofunction related to ξ through the local minimum map homeomorphism. Measurability of ζ_g follows from continuity. Notice that the set $\Omega \subset \Gamma^{\mathbb{N}}$ consisting of sequences $\kappa_n = (\omega_{n-1},\omega_n,\xi_n)$, where $\xi_n = (g(\omega_0,\omega_1)g(\omega_1,\omega_2)...g(\omega_{n-1},\omega_n))^{-1}\xi_0$ and notice that this is a set of full measure. The sum process in $\Sigma^{\mathbb{N}}$ is

$$
\begin{aligned}
(S_n\zeta)(\omega) &= \sum_{i=0}^{n-1} \zeta(\omega_i,\omega_{i+1},\xi_i) \\
&= \sum_{i=0}^{n-1} h_{\xi_i}(g(\omega_i,\omega_{i+1})x_0)
\end{aligned}
$$

$$= \sum_{i=0}^{n-1} (g(\omega_0, \omega_1) g(\omega_1, \omega_2) ... g(\omega_i, \omega_{i+1}))^{-1} \cdot h_{\xi_0}(g_i x_0)$$
$$= \sum_{i=0}^{n-1} h_{\xi_0}(g^{(i+1)}(\omega) x_0) - h_{\xi_0}(g^{(i)}(\omega) x_0)$$
$$= h_{\xi_0}(g^{(n)}(\omega) x_0).$$

These equalities are mostly a consequence of the property $g \cdot h_\xi = h_{g\xi}$ and (2.5). In what follows we will prove that provided $g \in C(\Sigma \times \Sigma,\, G)$ is irreducible with positive drift, then there exists a unique $\overline{K}_g$-stationary measure which we denote by μ_g. Finally we consider the space of observed Markov systems

$$\mathcal{M} := \{(\overline{K}_g, \mu_g, \pm\zeta_g) \,:\, g \in C(\Sigma \times \Sigma,\, G),\ \text{g is irreducible and } \ell(g) > 0\},$$

where μ_g is the K_g-stationary measure, with the metric

$$\text{dist}((\overline{K}_{g_1}, \mu_{g_1}, \zeta_{g_1}), (\overline{K}_{g_2}, \mu_{g_2}, \zeta_{g_2})) := d_\infty(g_1, g_2).$$

Due to the metric used, neighbourhoods in $\mathcal{M}$ are naturally identified with neighbourhoods in $C(\Sigma \times \Sigma,\, G)$. Our main goal for the remainder of this section is to prove the following proposition:

Proposition 37. *The space $\mathcal{M}$ satisfies the $C*)$ conditions.*

Notice that the $(\overline{Q}_g f)(\omega_0, \omega_1, \xi)$ does not depend on the variable ω_0. So we define $\mathcal{H}_\alpha(\Sigma \times \partial X)$ to be the space of functions f in $\mathcal{H}_\alpha(\Gamma)$ that do not depend on ω_0. Notice as well that $\mathcal{H}_\alpha(\Sigma \times \partial X)$ is still a family of Banach algebras satisfying $A*)$ and $B*)$. Our first goal is to prove that this space is invariant under the action of $\overline{Q}_g$:

Proposition 38. *The space $\mathcal{H}_\alpha(\Sigma \times \partial X)$ is invariant by the action of Q_g for α small enough.*

The proof of this proposition is based of the Lemmas 39, 40 and 41. First, given $g \in C(\Sigma \times \Sigma,\, G)$ and $0 < \alpha < 1$ define the average Hölder constant of g as

$$k_\alpha^n(g) := \sup_{\omega_0 \in \Sigma, \xi \neq \eta} \mathbb{E}_{\omega_0} \left[\left(\frac{D_b(g^{-(n)}\xi\,,\, g^{-(n)}\eta)}{D_b(\xi\,,\,\eta)} \right)^\alpha \right].$$

The relevance of $k_\alpha^n(g)$ becomes evident in the following lemma where we relate it with the contracting behaviour of the Markov operator of g.

Lemma 39. *Given $g \in S^\infty(\Sigma \times \Sigma,\, G)$, $f \in \mathcal{H}_\alpha(\Sigma \times \partial X)$ and $n \in \mathbb{N}$,*

$$v_\alpha(\overline{Q}_g^n f) \leq k_\alpha^n(g) v_\alpha(f).$$

Proof. Let $f \in \mathcal{H}(\Sigma \times \partial X)$ and $(\omega_0, \xi) \in \Sigma \times \partial X$, recall as well the random variables $e_n : \Omega \to \Sigma$ given by $e_n(\omega) = \omega_n$. Then notice

$$(\overline{Q}_g^n f)(\omega_0, \xi) = \mathbb{E}_{\omega_0}\left[f(e_n, g^{-(n)}\xi)\right].$$

Hence

$$\begin{aligned}
\upsilon_\alpha(\overline{Q}_g^n f) &\leq \sup_{\omega_0\in\Sigma, \xi\neq\eta\in\partial X} \frac{\mathbb{E}_{\omega_0}\left|f(e_n, g^{-(n)}\xi) - f(e_n, g^{-(n)}\eta)\right|}{D_b(\xi\,,\,\eta)} \\
&\leq \upsilon_\alpha(f) \sup_{\omega_0\in\Sigma, \xi\neq\eta} \mathbb{E}_{\omega_0}\left[\left(\frac{D_b(g^{-(n)}\xi\,,\,g^{-(n)}\eta)}{D_b(\xi\,,\,\eta)}\right)^\alpha\right] \leq \upsilon_\alpha(f)\, k_\alpha^n(g)
\end{aligned}$$

□

Lemma 40. *Given $g \in S^\infty(\Sigma \times \Sigma,\, G)$, the sequence $(k_\alpha^n(g))$ is sub-multiplicative, that is,*

$$k_\alpha^{n+m}(g) \leq k_\alpha^n(g) k_\alpha^m(g)$$

Proof. For every ξ, η in ∂X and $\omega_0 \in \Sigma$

$$\begin{aligned}
&\mathbb{E}_{\omega_0}\left[\left(\frac{D_b(g^{-(m+n)}\xi\,,\,g^{-(m+n)}\eta)}{D_b(\xi\,,\,\eta)}\right)^\alpha\right] \\
&\quad\leq \mathbb{E}_{\omega_0}\left[\left(\frac{D_b(g^{-(m)}\circ T^n g^{-(n)}\xi\,,\,g^{-(m)}\circ T^n g^{-(n)}\eta)}{D_b(g^{-(n)}\xi\,,\,g_1^{-(n)}\eta)}\right)^\alpha \left(\frac{D_b(g^{-(n)}\xi\,,\,g^{-(n)}\eta)}{D_b(\xi\,,\,\eta)}\right)^\alpha\right] \\
&\quad\leq \mathbb{E}_{\omega_0}\left[\left(\frac{D_b(g^{-(n)}\xi\,,\,g^{-(n)}\eta)}{D_b(\xi\,,\,\eta)}\right)^\alpha\right] \sup_{\xi'\neq\eta'\in\partial X} \mathbb{E}_{K^n(\omega_0,\cdot)}\left[\left(\frac{D_b(g^{-(m)}\xi'\,,\,g_2^{-(m)}\eta')}{D_b(\xi'\,,\,\eta')}\right)^\alpha\right].
\end{aligned}$$

Taking the supremum over ξ and η yields the result. □

Lemma 41. *Given $g \in S^\infty(\Sigma \times \Sigma,\, G)$ and $n \in \mathbb{N}$, for every $0 < \alpha < \frac{1}{n}$ there exists a constant $C = C(g)$, such that*

$$k_\alpha^n(g) \leq C(\delta) d_\infty(g).$$

Proof. Given $\omega_0 \in \Sigma$ and $\xi \neq \eta$ in ∂X, using Proposition 14,

$$\begin{aligned}
\mathbb{E}_{\omega_0}\left[\left(\frac{D_b(g^{-(n)}\xi\,,\,g^{-(n)}\eta)}{D_b(\xi\,,\,\eta)}\right)^\alpha\right] &\leq C(\delta)\mathbb{E}_{\omega_0}\left[b^{-\frac{\alpha}{2}(h_\xi(g^{(n)}x_0)+h_\eta(g^{(n)}x_0))}\right] \\
&\leq \mathbb{E}_{\omega_0}\left[b^{\alpha d(g^{(n)}x_0, x_0)}\right] \\
&\leq \mathbb{E}_{\omega_0}\left[b^{d(gx_0, x_0)}\right]
\end{aligned}$$

taking the supremum in ω_0 and $\xi \neq \eta$ we obtain the statement using Lemma 21. □

The following Lemma is where the necessity for the hyperbolic multiplicative ergodic theorem appears. The last part of the proof is analogous to that of Bougerol (1988) for the uniformity of the limit for Lyapunov exponents.

Lemma 42. *Let $g \in C(\Sigma \times \Sigma,\, G)$ positive drift,*

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}_{\omega_0}\left[h(g^{(n)}x_0)\right] = \ell(g) \tag{4.2}$$

uniformly on $(\omega_0, h) \in \Sigma \times X^h_\infty$.

We warn the reader that in the following proof we work with $\Gamma_1 = \Sigma \times \Sigma \times X^h$. We do this as we need compacity. With that in mind we are going to use the Markov Kernel in Γ_1 analogous to the one used in Γ, that is

$$\overline{K}_g(\omega_0, \omega_1, h) := \int_\Sigma \delta_{(\omega_1, \omega_2, g(\omega_1,\omega_2)^{-1}\cdot h)} K(\omega_1, d\omega_2),$$

which in turn gives rise to a Markov operator in the typical fashion. By compactness of Γ_1, there exists at least one $\overline{K}_g$-stationary measure μ. In what follows we drop the g in $\overline{K}_g$ and denote by $\mathbb{P}$ the Kolmogorov extension measure with respect to some $\overline{K}$-stationary measure in $\Omega = \Gamma_1^{\mathbb{N}}$.

The strategy of the proof is to first prove that the limit exists for every horofunction h and μ almost every ω. Then prove it is uniform on h and finally obtain its uniformity on ω_0. With that in mind we will prove four claims, are the Lemma should follow once those are done.

Claim 1: For every $h \in X^h_\infty$, $\lim_{n\to\infty} \frac{1}{n} h(g^{(n)}(\omega)x_0) = \ell(g)$ holds for $\mathbb{P}$ almost every ω.

Proof of Claim 1. Consider the observable $\zeta \colon \Gamma_1 \to \mathbb{R}$ defined by

$$\zeta(\omega_0, \omega_1, h) := h(g(\omega_0, \omega_1)x_0)$$

which is clearly continuous. Denote by $\text{Prob}_K(\Gamma_1)$ the space of K-stationary probability measures on Γ_1, which is non-empty by compactness of Γ_1. Just as before, consider the sum process $S_n\zeta$ generated by ζ along a K-Markov process on Γ_1 with initial state $(\omega_0, \omega_1, h_0) \in \Gamma$. This sum process can be realized as the process on $\Omega = \Sigma^{\mathbb{N}}$ defined by

$$(S_n\zeta)(\omega) := \sum_{j=0}^{n-1} \zeta(\omega_i, \omega_{i+1}, h_i) = h_0(g^{(n)}(\omega)\, x_0)$$

where $h_{i+1} = g(\omega_{i-1}, \omega_i)^{-1} \cdot h_i$ for every $i \geq 0$.

By Furstenberg-Kifer Theorems 1.1 and 1.4 in Furstenberg and Kifer (1983), letting

$$\beta := \sup\left\{ \int_\Gamma \zeta \, d\eta \;:\; \eta \in \text{Prob}_K(\Gamma_1) \right\}$$

then for $\mathbb{P}$-almost every $\omega \in \Omega$

$$\limsup_{n\to\infty} \frac{1}{n}\, h_0(g^{(n)}(\omega)x_0) = \limsup_{n\to\infty} \frac{1}{n}\,(S_n\zeta)(\omega) \leq \beta.$$

We claim now that $\int_{\Gamma_1} \zeta \, d\eta = \beta$ for every measure $\eta \in \mathrm{Prob}_K(\Gamma_1)$. Then changing ζ by $-\zeta$, the same argument implies that for $\mathbb{P}$-almost every ω

$$\lim_{n\to\infty} \frac{1}{n}\, h_0(g^{(n)}(\omega)x_0) = \beta.$$

By the Theorem 17 and its remark we must have $\beta = \ell(g)$. □

Claim 2: $\int_{\Gamma_1} \zeta \, d\eta = \beta$ for every measure $\eta \in \mathrm{Prob}_K(\Gamma_1)$.

Proof of Claim 2. If the claim were false there would be an ergodic measure $\eta \in \mathrm{Prob}_K(\Gamma_1)$ such that $\int_{\Gamma_1} \zeta \, d\eta = \beta_1 < \beta$. Consider the map

$$\begin{aligned} F : \Omega \times X^h &\to \Omega \times X^h \\ (\omega, h) &\mapsto (\sigma\omega, g(\omega_0, \omega_1)^{-1} \cdot h) \end{aligned}$$

which preserves the ergodic measure $\mathbb{P} \times \eta$. The observable ζ can be extended to $\bar{\zeta} \colon \Omega \times X^h \to \mathbb{R}$, $\bar{\zeta}(\omega, h) = \zeta(\omega_0, \omega_1, h)$. Moreover, with this notation, $(S_n\zeta)(\omega) = \sum_{j=0}^{n-1} \bar{\zeta}(F^j(\omega, h_0))$ is a Birkhoff sum. By Birkhoff's ergodic theorem, for η-almost every $h_0 \in X^h$ and $\mathbb{P}$-almost every $\omega \in \Omega$,

$$\lim_{n\to\infty} \frac{1}{n} h_0(g^{(n)}(\omega)x_0) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \bar{\zeta}(F^j(\omega, h_0)) = \beta_1$$

which together with Theorem 17 implies that $\beta_1 = -\ell(g)$ and $h_0 \in X^h_-(\omega)$. Next consider the family of sets

$$S_{\omega_0} := \left\{ h \in X^h : \; \mathbb{P}_{\omega_0}\{\omega \in \Omega : h \in X^h_-(\omega)\} = 1 \right\}.$$

The previous argument shows that $S_{\omega_0} \neq \emptyset$ for $\mathbb{P}$-almost every $\omega \in \Omega$. Again by the remark to Theorem 17 the set S_{ω_0} must be a single horofunction $S_{\omega_0} = \{s(\omega_0)\}$ and the function $s \colon \Sigma \to X_h$ is measurable. The invariance of X^h_- in Theorem 17 now implies that $g(\omega_0, \omega_1) \cdot s(\omega_0) = s(\omega_1)$, which proves that g is not irreducible. This contradiction implies that the claim is true. □

Claim 3: The convergence is uniform h.

Proof of Claim 3. Let us start by proving the uniformity in h, arguing by absurd, suppose there is a sequence of horofunctions $(h_n) \subset X^h_\infty$ converging to some h in X^h_∞ and $\varepsilon > 0$ such that

$$\lim_{n\to\infty} \frac{1}{n} \mathbb{E}_{\omega_0} \left[h_n(g^{(n)}x_0) \right] < \ell(g) - \varepsilon$$

Due to the compactness of X^h we can assume that h_n converges. Take $(y_m^n)_m \in X$ and $\xi_n \in \partial X$ two families of sequences such that $h_{y_m^n} \to h_n =: h_{\xi_n}$ and $y_m^n \to \xi_n$ as $m \to \infty$. Then

$$\begin{aligned}
\lim_{n\to\infty} h_n(g^{(n)}x_0) - d(g^{(n)}x_0\,,\, x_0) &= \lim_{n\to\infty}\lim_{m\to\infty} h_{y_m^n}(g^{(n)}x_0) - d(g^{(n)}x_0\,,\, x_0) \\
&= \lim_{n\to\infty}\lim_{m\to\infty} d(y_m^n\,,\, g^{(n)}x_0) - d(y_m^n\,,\, x_0) - d(g^{(n)}x_0\,,\, x_0) \\
&= \lim_{n\to\infty}\lim_{m\to\infty} -2\langle y_m^n, g^{(n)}x_0\rangle_{x_0} \\
&= \lim_{n\to\infty} -2\langle \xi_n, g^{(n)}x_0\rangle_{x_0},
\end{aligned}$$

where the last equality is a consequence of the continuity of the Gromov product in strongly hyperbolic spaces. Notice that the quantity $\langle \xi_n, g^{(n)}x_0\rangle_{x_0}$ goes to infinity if and only if both ξ_n and $g^{(n)}x_0$ converge to the same point in ∂X. If this were the case, by Proposition 4 in Sampaio (2021), $\lim_{n\to\infty} h(g^{(n)}x_0) = -\infty$, hence $h \in X^h_-(\omega)$. Therefore $\langle \xi_n, g^{(n)}x_0\rangle_{x_0}$ must $\mathbb{P}_{\omega_0}$ almost surely be finite as otherwise $h \in S_{\omega_0} = \emptyset$. Using dominated convergence theorem again,

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n}\mathbb{E}_{\omega_0}\left[h_n(g^{(n)}x_0)\right] &= \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_{\omega_0}\left[d(g^{(n)}x_0\,,\, x_0)\right] + \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_{\omega_0}\left[h_n(g^{(n)}x_0) - d(g^{(n)}x_0\,,\, x_0)\right] \\
&= \ell(g) + 0 = \ell(g),
\end{aligned}$$

which yields the claim. □

Claim 4: The convergence is uniform in ω_0.

Proof of Claim 4. Consider now, for $\omega_0 \in \Sigma$

$$q_n(\omega_0) = \sup\left\{\left|\frac{1}{n}\mathbb{E}_{\omega_0}\left[h(g^{(n)}x_0)\right] - \ell(g)\right| \,:\, h \in X^h\right\},$$

and notice the uniform bound $|q_n(\omega_0)| \leq \log_b(d_\infty(g)) + \ell(g)$. Due to the uniform limit in h proven above, using dominated convergence theorem

$$\lim_{n\to\infty}\int_\Sigma p_n(\sigma)d\mu(\sigma) = 0.$$

Let $\varepsilon > 0$. Consider $n > p$ to be specified later and take $a = a(p) := \sup_{\omega_0} q_p(\omega_0)$

$$\begin{aligned}
\left|\frac{1}{n}\mathbb{E}_{\omega_0}\left[h(g^{(n)}x_0)\right] - \ell(g)\right| &\leq \left|\frac{1}{n}\mathbb{E}_{\omega_0}\left[g^{-(p)}\cdot h(g^{(n-p)}x_0) + h(g^{(p)}x_0)\right] - \ell(g)\right| \\
&\leq \left|\frac{1}{n}\mathbb{E}_{\omega_0}\left[g^{-(p)}h(g^{(n-p)}x_0)\right] - \ell(g)\right| + \frac{1}{n}\mathbb{E}_{\omega_0}\left[h(g^{(p)}x_0)\right] \\
&\leq \left(\frac{n-p}{n}\right)\left|\mathbb{E}_{\omega_0}\left[\frac{1}{n-p}\mathbb{E}_{\omega_p}\left[g^{-(p)}h(g^{(n-p)}x_0)\right] - \ell(g)\right]\right| + \frac{p}{n}(\ell(g) + a),
\end{aligned}$$

from which

$$q_n(\omega_0) \le (Q^p q_{n-p})(\omega_0) + \frac{p}{n}(\ell(g) + a).$$

Now, taking p and n large enough, one has the following inequalities

$$\frac{p}{n}(\ell(g) - a)/n < \varepsilon/3,$$

as well as

$$\int_\Sigma q_{n-p}(\sigma)d\mu(\sigma) < \varepsilon/3,$$

moreover, by the strongly mixing condition

$$\sup_{\omega_0 \in \Sigma} \left| (Q^p q_{n-p})(\omega_0) - \int_\Sigma q_{n-p}(\sigma)d\mu(\sigma) \right| \le \varepsilon/3,$$

provided p is large enough and taking n large enough. Hence

$$q_n(\omega_0) \le \int_\Sigma q_{n-p}(\sigma)d\mu(\sigma) + 2\varepsilon/3 < \varepsilon.$$

□

In the following proposition we will use the relation, which is an immediate consequence of Proposition 14,

$$k_\alpha^n(g) \le \sup_{\omega \in \Sigma,\, \xi \in \partial X} \mathbb{E}_\omega \left[b^{-\alpha h_\xi(g^{(n)} x_0)} \right].$$

Proposition 43. *Given* $g_1 \in \mathcal{I}(K)$ *with positive drift, there exists a neighbourhood* V *of* g_1 *in* $C(\Sigma \times \Sigma, G)$ *and constants* $n_0 \in \mathbb{N}$, $0 < \alpha_1 < \alpha_0/2 < \alpha_0$, $C = C(g_1) > 0$, *and* $0 \le \sigma < 1$ *such that*

$$k_\alpha^n(g_2) \le C\sigma^n,$$

for all $g_2 \in V$, $n > n_0$, $\alpha \in [\alpha_0, \alpha_1]$ *and* $f \in \mathcal{H}_\alpha(\Sigma \times \partial X)$.

Proof. By Lemma 55,

$$\lim_{n\to\infty} \sup_{\xi \in \partial X} \left| \mathbb{E}_{\omega_0} \left[h_\xi(g^{(n)} x_0) \right] - \ell(g_1) \right| = 0.$$

In particular, there exists $n_0 \in \mathbb{N}$ such that $\mathbb{E}_{\omega_0} \left[h_\xi(g^{(n_0)} x_0) \right] \ge \frac{1}{\log b} > 0$ for every $\xi \in \partial X$.

Let $r > 0$ to be specified later and consider in $C(\Sigma \times \Sigma, G)$ the neighbourhood of g_1 given by the ball

$$V = B_r(g_1) := \{ g_2 \in C(\Sigma \times \Sigma, G) \,:\, d_\infty(g_1\,,\, g_2) < r \}.$$

Let $g_2 \in B_r(g_1)$, $\omega_0 \in \Sigma$, ξ in ∂X. Use the inequality

$$b^x < 1 + \log(b)x + \log(b)^2 \frac{x^2}{2} b^{|x|},$$

to obtain,

$$\begin{aligned}\mathbb{E}_{\omega_0}\left[b^{-\alpha h_\xi(g_1^{(n_0)}x_0)}\right] &\leq 1-\alpha\log(b)\mathbb{E}_{\omega_0}\left[h_\xi(g_1^{(n_o)}x_0)\right]\\ &\quad+\log(b)^2\frac{\alpha^2}{2}\mathbb{E}_{\omega_0}\left[(h_\xi(g_1^{(n_0)}x_0))^2 b^{|\alpha h_\xi(g_1^{(n_0)}x_0)|}\right]\\ &\leq 1-\alpha+\alpha^2 b^\alpha\left(\frac{\log(b)^2}{2}n_0^2\log_b(C)^2 C^{n_0\log b}\right),\end{aligned}$$

where C is a constant depending on g_1. Hence there exists α small enough so that the right-hand side becomes smaller than 1, which implies the existence of constants α_0 and $\alpha_1 < \alpha_0/2$ such that $k_\alpha^{n_0}(g_1) < \rho < 1$.

To extend this control to nearby cocycles let us introduce the following continuity type relation, using the mean value theorem and the argument around finite scale continuity in Proposition 22

$$\begin{aligned}\mathbb{E}_{\omega_0}\left|b^{-\alpha h_\xi(g_1^{(n_0)}x_0)}-b^{-\alpha h_\xi(g_2^{(n_0)}x_0)}\right| &\leq (\log b)\max_{i=1,2} b^{d(g_i^{(n_0)}x_0,x_0)}\left|h_\xi(g_1^{(n_0)}x_0)-h_\xi(g_2^{(n_0)}x_0)\right|\\ &\leq C^{n_0}d(g_1^{(n_0)}x_0,g_2^{(n_0)}x_0)\\ &\leq n_0 C^{2n_0}d_\infty(g_1,g_2)\end{aligned}$$

We can now choose r small enough to ensure there exists $\rho^*\in(\rho,1)$ such that

$$\mathbb{E}_{\omega_0}\left|b^{-\alpha h_\xi(g_1^{(n_0)}x_0)}-b^{-\alpha h_\xi(g_2^{(n_0)}x_0)}\right|\leq\rho^*-\rho.$$

Hence

$$\begin{aligned}\mathbb{E}_{\omega_0}\left[b^{-\alpha h_\xi(g_2^{(n_0)}x_0)}\right] &\leq \mathbb{E}_{\omega_0}\left[b^{-\alpha h_\xi(g_1^{(n_0)}x_0)}\right]+\left|\mathbb{E}_{\omega_0}\left[b^{-\alpha h_\xi(g_2^{(n_0)}x_0)}\right]-\mathbb{E}_{\omega_0}\left[b^{-\alpha h_\xi(g_1^{(n_0)}x_0)}\right]\right|\\ &\leq \mathbb{E}_{\omega_0}\left[b^{-\alpha h_\xi(g_1^{(n_0)}x_0)}\right]+\mathbb{E}_{\omega_0}\left|b^{-\alpha h_\xi(g_1^{(n_0)}x_0)}-b^{-\alpha h_\xi(g_2^{(n_0)}x_0)}\right|\\ &\leq \rho+(\rho^*-\rho)=\rho^*<1\end{aligned}$$

Due to the submultiplicativity, picking $\sigma=(\rho^*)^{\frac{1}{n_0}}$, for every $n\in\mathbb{N}$ there exists a constant $C>0$ such that

$$k_\alpha^n(g_2)\leq\mathbb{E}_{\omega_0}\left[b^{-\alpha h_\xi(g_2^{(n)}x_0)}\right]<C\sigma^n,$$

which completes the proof. □

The previous proposition now allows us to obtain the existence and uniqueness of the K_g stationary measures μ_g in a neighbourhood of g irreducible with positive drift.

Proposition 44. *Let $g\in S^\infty(\Sigma\times\Sigma,G)$ have positive drift. If for some $n\in\mathbb{N}$ and $\alpha<1$*

$$k_\alpha^n(g)^{1/n}<1,$$

then there exists a unique K_g stationary measure.

Proof. The proof is mostly taken from Sampaio (2021). The seminorms v_α are norms in the space $\mathcal{H}_\alpha(\Gamma)/\mathbb{C}\mathbf{1}$. Since $\overline{Q}_g^n\mathbf{1} = \mathbf{1}$, by hypothesis, $\overline{Q}_g^n$ acts in $\mathcal{H}_\alpha(\Gamma)/\mathbb{C}\mathbf{1}$ as a contraction. Using spectral theory (see chapter IX in Riesz and Nagy (2012) for example), there exists and invariant space H_0, isomorphic to $\mathcal{H}_\alpha(\Gamma)/\mathbb{C}\mathbf{1}$, such that $\mathcal{H}_\alpha(\Gamma) = H_0 \oplus \mathbb{C}\mathbf{1}$. Given $f \in \mathcal{H}_\alpha(\Gamma)$ we may write it as $c\mathbf{1} + h$ where $c \in \mathbb{C}$ and $h \in H_0$. With that in mind, define

$$\begin{aligned} \Lambda : \mathcal{H}_\alpha(\Gamma) &\to \mathbb{C} \\ c\mathbf{1} + h &\mapsto c. \end{aligned}$$

Now notice that $\overline{Q}_g$ is a positive operator, therefore so is Λ as

$$c\mathbf{1} = \lim_{n\to\infty} \left(c\mathbf{1} + \overline{Q}_g^n(h)\right) = \lim_{n\to\infty} \overline{Q}_g^n(f) \geq 0,$$

provided $f \geq 0$. Hence $c = \Lambda(f) \geq 0$. Positivity also implies continuity with respect to the uniform norm as

$$|\Lambda(\varphi)| \leq |\Lambda(||\varphi||_\infty \mathbf{1})|| = ||\varphi||_\infty.$$

Now since Γ is a metric space, the set of bounded Lipschitz functions in ∂X is dense in the space of bounded uniformly continuous functions $C_b(\Gamma)$. With effect, given $f \in C_b(\Gamma)$ one can take the functions

$$f_n(\xi) = \inf_{\eta\in\Gamma}\{f(\eta) - nD_b(\xi,\eta)\},$$

which are all bounded Lipschitz and uniformly converge to f. Since the space is bounded, the set of Lipschitz functions is contained in the space of Hölder functions, so $\mathcal{H}_\alpha(\Gamma)$ is dense in $C_b(\Gamma)$. Hence, Λ extends to a positive linear continuous functional $\hat{\Lambda} : C_b(\Gamma) \mapsto \mathbb{C}$.

Riesz-Kakutani-Markov for non-compact spaces (Theorem 1.3 in Sentilles (1972)) applies, so there exists a measure $\nu \in \text{Prob}(\Gamma)$ such that $\hat{\Lambda}(f) = \int_\Gamma f d\nu$ for every $f \in C_b(\Gamma)$. Finally, writing f once again as $c\mathbf{1} + h$ yields

$$\int_\Gamma \overline{Q}_g f d\nu = \hat{\Lambda}(\overline{Q}_g f) = c = \hat{\Lambda}(f) = \int_\Gamma f d\nu.$$

By yet another density argument, this holds for all $f \in L^1(\Gamma)$, therefore ν is K_g-stationary. This density of $C_b(\partial X)$ in $L^1(\Gamma)$ also justifies the uniqueness of the measure satisfying $\hat{\Lambda}(f) = \int_\Gamma f d\nu$. □

Henceforth $\mathcal{M}$ is welll defined and condition C1) is immediate. We now focus the remaining conditions.

Proposition 45 (Proposition 5.22 in Duarte et al. (2016))**.** *Given* $g_1 \in C(\Sigma \times \Sigma, G)$ *such that* $(\overline{K}_{g_1}, \mu_{g_1}, \zeta_{g_1}) \in \mathcal{M}$*, there exist a neighborhood* V *of* g_1 *in* $C(\Sigma \times \Sigma, G)$*, constants* $0 < \alpha_1 < \alpha_0/2 < \alpha_0 < 1$*,* $C > 0$ *and* $0 < \sigma < 1$ *such that for all* $g_2 \in V$*, and* $f \in \mathcal{H}_\alpha(\Sigma \times \partial X)$*,*

$$\left|\left|\overline{Q}^n_{g_2} f - \int_\Sigma f(\omega) d\mu_{g_2}(\omega)\right|\right|_\alpha \leq C\sigma^n ||f||_\alpha.$$

Proof. Take the neighbourhood V from the previous proposition, given $g_2 \in V$ and any K_{g_2} stationary measure μ_{g_2},

$$v_\alpha\left(\overline{Q}^n_{g_2} f - \int_\Sigma f(\omega) d\mu_{g_2}(\omega)\right) = v_\alpha(\overline{Q}^n_{g_2} f) \leq v_\alpha(f) k^n_\alpha(g_2) \leq C\sigma^n ||f||_\alpha. \tag{4.3}$$

So it remains to prove

$$\left|\left|\overline{Q}^n_{g_2} f - \int_\Sigma f(\omega) d\mu_{g_2}(\omega)\right|\right|_\infty \leq 2 \left|\left|\overline{Q}^n_{g_2} f\right|\right|_\infty \leq C\sigma^n ||f||_\alpha, \tag{4.4}$$

for possibly some other $C < \infty$ and $0 < \sigma < 1$.

For this purpose, consider as well as the operator $Q : L^\infty(\Sigma) \to L^\infty(\Sigma)$

$$(Qf)(\omega_1) := \int_\Sigma f(\omega_2)\, dK_{\omega_1}(\omega_2).$$

There is a natural projection $\pi : \Sigma \times \partial X \to \Sigma$ which induces a bounded linear embedding $\pi^* : L^\infty(\Sigma) \to \mathcal{H}_\alpha(\Sigma \times \partial X)$, $\pi^* f := f \circ \pi$. Notice that the range of this embedding is the subspace

$$\pi^* L^\infty(\Sigma) = \left\{ f \in \mathcal{H}_\alpha(\Sigma \times X^h) \, : \, v_\alpha(f) = 0 \right\}$$

and the following diagram commutes for every $n \in \mathbb{N}$

$$\begin{array}{ccc} L^\infty(\Sigma) & \xrightarrow{Q^n} & L^\infty(\Sigma) \\ {\scriptstyle \pi^*}\big\downarrow & & \big\downarrow{\scriptstyle \pi^*} \\ \mathcal{H}_\alpha(\Sigma \times \partial X) & \xrightarrow[\overline{Q}^n]{} & \mathcal{H}_\alpha(\Sigma \times \partial X). \end{array}$$

Given $f \in \mathcal{H}_\alpha(\Sigma \times X^h)$, by (4.3) the iterates $\overline{Q}^n f$ converge exponentially fast to the closed subspace $L^\infty(\Sigma) \equiv \pi^* L^\infty(\Sigma) \subseteq \mathcal{H}_\alpha(\Sigma \times X^h)$. On the other hand by assumption Q is strongly mixing on $L^\infty(\Sigma)$. Combining these two properties and the fact that Markov operators do not expand we get (4.4). □

The Laplace-Markov operator $Q_{g,z}$ of the observed Markov system (K_g, μ_g, ζ_g) is given by

$$(Q_{g,z} f)(\omega_0, \omega_1, \xi) = \int_\Sigma f(\omega_1, \omega_2, g(\omega_1, \omega_2)^{-1}\xi) b^{z h_\xi(g(\omega_1,\omega_2)x_0)}\, K(\omega_1, d\omega_2).$$

Lemma 46. *Given $g_1, g_2 \in S^{\infty}(\Sigma \times \Sigma, G)$ and $b > 0$, there is a constant $C_2 > 0$ such that for all $f \in \mathcal{H}_\alpha(\Sigma \times \partial X)$ and all $z \in \mathbb{C}$ such that Re $z \leq c$,*

$$||Q_{g_1,z}f - Q_{g_2,z}f||_\infty \leq C_2 d_\infty(g_1, g_2)^\alpha ||f||_\alpha.$$

Moreover, C_2 is bounded on a neighborhood of g_1.

Proof. Let $\xi \in \partial X$. Start by noticing that writting $z = x + yi$ with $x \leq c$

$$\begin{aligned} |b^{zh_\xi(g_1x_0)} - b^{zh_\xi(g_2x_0)}| &\leq \max_{i=1,2} b^{c\, d(g_ix_0,x_0)}\, |c\, h_\xi(g_1x_0) - c\, h_\xi(g_2x_0)| \\ &\leq c\, d_\infty(g_1, g_2) \max_{i=1,2} d_\infty(g_i)^c. \end{aligned}$$

Hence

$$\begin{aligned} |Q_{g_1,z}f - Q_{g_2,z}f| &\leq \mathbb{E}_{\omega_0}\left[|b^{zh_\xi(g_1x_0)} f(e_1, g_1^{-1}\xi) - b^{zh_\xi(g_2x_0)} f(e_1, g_2^{-1}\xi)|\right] \\ &\leq ||f||_\infty \mathbb{E}_{\omega_0}\left[|b^{zh_\xi(g_1x_0)} - b^{zh_\xi(g_2x_0)}|\right] + \max_{i=1,2} d_\infty(g_i)^c \mathbb{E}_{\omega_0}\left[|f(e_1, g_1^{-1}\xi) - f(e_1, g_2^{-1}\xi)|\right] \\ &\leq c\, d_\infty(g_1, g_2) \max_{i=1,2} d_\infty(g_i)^c ||f||_\infty + \max_{i=1,2} d_\infty(g_i)^c v_\alpha(f) \mathbb{E}_{\omega_0}\left[D_b(g_1^{-1}\xi, g_2^{-1}\xi)^\alpha\right] \\ &\leq C_2 ||f||_\alpha d_\infty(g_1, g_2)^\alpha. \end{aligned}$$

where $C_2 = \max\{c \max_{i=1,2} d_\infty(g_i)^c, \max_{i=1,2} d_\infty(g_i)^c\}$ which are bounded in a neighbourhood g_1. The last inequality is a consequence of $d_\infty(g_1, g_2) < d_\infty(g_1, g_2)^\alpha < 1$ and $D_b(g_1^{-1}\xi, g_2^{-1}\xi) \leq d_\infty(g_1, g_2)$. □

Proof of Proposition 37. Point $C1)$ is obvious. For $C3)$ recall from (4.1) that $||b^{\zeta_g}||_\infty = ||b^{h(gx_0)}||_\infty \leq d_\infty(g)$ which is finite, hence $\zeta_g \in \mathcal{H}_\alpha(\Sigma \times \partial X)$. Therefore $Q_{g,z\zeta}$ acts on $\mathcal{H}_\alpha(\Sigma \times \partial X)$ as the latter is a Banach algebra. Point $C2)$ is a consequence of Proposition 45 while $C4)$ follows from the previous Lemma. □

With this, we can now obtain the large deviations.

Proof of Theorem 34. Using Theorem 35, there exists V a neighbourhood of $g \in S^\infty(K)$ and constants $\varepsilon_0, C, k > 0$ such that for every $g_2 \in V$, $0 < \varepsilon < \varepsilon_0$, $h \in X_\infty^h$ and $n \in \mathbb{N}$

$$\mathbb{P}_\mu\left[\left|\frac{1}{n} h(g_2^{(n)}x_0) - \ell(g_2)\right| > \varepsilon\right] \leq C b^{-k\varepsilon^2 n}.$$

Using Lemma 5 in Sampaio (2021), one obtains that there exists an horofunction $h \in X_\infty^h$ such that

$$h(g_2^{(n)}x_0) \leq d(g_2^{(n)}x_0, x_0) \leq h(g_2^{(n)}x_0) + K(\delta).$$

where $K(\delta)$ is a constant depending on δ. Using this inequality we obtain the large deviations with a possible loss in the constant C. □

4.3 Random Walks Setting

In this section we slightly adjust the setting. Here we will be interested in random walks, in other words, we consider probability measures on the group and then tackle the same problems we have worked so far in this case. This change in setting requires that we reintroduce some concepts.

Let now $\text{Prob}(M)$ and $\text{Prob}_c(M)$ denote, respectively, the space of Borel probability measures and its subspace of with Borel probability measures with compact support in some metric space M. Let G be a topological group acting on a metric space M, we define the convolution

$$\begin{aligned} \star : \text{Prob}_c(G) \times \text{Prob}(M) &\to \text{Prob}(M) \\ (\mu, \nu) &\mapsto \mu \star \nu := \int_G g\nu d\mu(g), \end{aligned}$$

where $g\nu$ is the pushforward of ν under the action of g, in other words, the convolution is the average of the pushforwards with respect to μ. In particular, G acts on itself on the right, this allows for the definition of μ^n for every $\mu \in \text{Prob}_c(G)$ as the $n-th$ convolution of μ with itself. Let us stress the fact we are considering G acting on itself through the right action. As a side-note recall that if both μ and ν have compact support then so does $\mu \star \nu$.

Working in the degree of generality we intend to keep in this thesis, a problem could arise here. We need G to be second countable in order for the support of a measure in $\text{Prob}(G)$ to be well defined. Whence we restrict our attention to closed separable groups of isometries $G \subset \text{Isom}(X)$. With this in mind, throughout this section G always stands for a closed separable groups of isometries acting on a hyperbolic space X satisfying (BA).

Let $\mu \in \text{Prob}_c(G)$ then we will consider the product measure $\mu^{\mathbb{N}}$ which has compact support in $\Omega = G^{\mathbb{N}}$. Given $\omega = (g_0, g_1, ..., g_n, ...) \in \Omega$ we set the notation

$$\omega^n = g_0 g_1 ... g_{n-1},$$

as well as defining the Bernoulli shift $T : \Omega \to \Omega$ sending $\omega = (g_0, g_1, ...g_n, ...)$ to $T\omega = (g_1, ..., g_n, ...)$. We will denote by ω^{-n} the inverse of ω^n, that is, $(\omega^n)^{-1}$. Notice that T is an ergodic transformation with respect to $\mu^{\mathbb{N}}$.

Given $\mu \in \text{Prob}_c(G)$, by compacity we have

$$\int_G d(gx_0, x_0) d\mu(g) < \infty,$$

so we can define the drift

$$\ell(\mu) := \lim_{n\to\infty} \frac{1}{n} \int_G d(gx_0, x_0) d\mu^n(g) = \lim_{n\to\infty} \frac{1}{n} \int_G d(\omega^n x_0, x_0) d\mu^{\mathbb{N}}(\omega).$$

Notice that the measurability of the integrand functions follows from continuity. Moreover, due to ergodicity, by Kingman's Ergodic Theorem the limit is $\mu^{\mathbb{N}}$-almost surely equal to the limit of $d(\omega^n x_0, x_0)/n$.

Our goal this time is to understand the behaviour of $\ell(\mu)$ with respect to both small perturbations in μ. To understand what we mean by small perturbations in μ we shall introduce the Hölder Wasserstein distance. Let $L^\infty(G)$ stand for the space of Borel measurable functions $\varphi : G \to \mathbb{R}$ bounded in the sup norm. For every $0 < \alpha \leq 1$ and $\varphi \in L^\infty(G)$ define the α-Hölder constant

$$\upsilon_\alpha^G(\varphi) := \sup_{g, g' \in G,\, g \neq g'} \frac{|\varphi(g) - \varphi(g')|}{d_G(g, g')^\alpha}.$$

For every $\mu, \nu \in \mathrm{Prob}_c(G)$, we define the Wasserstein distance between them as

$$W_\alpha(\mu, \nu) = \sup_{\varphi \in L^\infty(G),\, \upsilon_\alpha^G(\varphi) \leq 1} \left| \int_G \varphi d\mu - \int_G \varphi d\nu \right|.$$

A detailed discussion on Wasserstein distances can be found in Villani (2009). One now feels tempted to consider the continuity of the drift with respect to the Wasserstein distance. This may be problematic as we could consider very close measures μ and μ_1 where the support of μ_1 contains elements that expand an arbitrarily large quantity which are picked with an equally small probability. Then the distance between μ and μ_1 would be small but the drift could be incommensurably different. To work around this problem, given $1 < b \leq 2^{1/\delta}$ for every $\lambda > 0$ set

$$G_\lambda := \left\{ g \in G \,:\, b^{d(gx_0, x_0)} < \lambda \right\}.$$

For Markov systems we introduced the concept of irreducibility. We now reintroduce it in the language of random walks.

Definition 14 (Irreducible measure)**.** We say that $\mu \in \mathrm{Prob}_c(G)$ is irreducible if there is no horofunction $h \in X^h$ such that $g \cdot h = h$ for μ-almost every g.

One can prove (see Proposition 5.3 in Duarte and Klein (2017)) that being irreducible is an open condition in $\mathrm{Prob}_c(G)$. Density seems a bit more delicate since in some cases, such as $X = \mathbb{R}$, there are no irreducible measures. However, due to the convexity of the space of probabilities, irreducible measures, if they exist, form a dense set. Thus the set of irreducible measures is either generic or empty.

Let once again M be a metric space on which G acts. Given $\mu \in \mathrm{Prob}_c(G)$, a measure $\nu \in \mathrm{Prob}_c(M)$ is μ-stationary if $\mu \star \nu = \nu$. The existence of stationary measures is extremely important to us. It is actually easy to see that when $M = X^h$ the existence of stationary measures follows from compactness. Our technique will actually allow for the existence of stationary measures in ∂X, which will require a finer treatment similar to what we have done in the previous section.

Theorem 47 (Furstenberg type formula)**.** *Let X be a hyperbolic space satisfying (BA) and G be a closed separable group of isometries of X. Let μ be an irreducible measure in* $\mathrm{Prob}_c(G)$*. Then for every μ-stationary measure $\nu \in$* $\mathrm{Prob}(\partial X)$

$$\ell(\mu) = \int_G \int_{\partial X} h_\xi(gx_0) d\nu(\xi) d\mu(g).$$

The Fürstenberg type formmula is a consequence of Claims 1 and 2 of the proof of Lemma 55. Notice that a μ stationary measure in ∂X yields a stationary measure in $X^h_\infty \subset X^h$ by pushing forward through the inverse of the local minimum map. Hence the result follows.

Notice that it also makes sense to define a Wasserstein distance in $\mathrm{Prob}(\partial X)$. Using this fact, show that every measure in a neighourhood of an irreducible measure with positive drift admits a unique stationary measure. Then we show that is stationary measure varies continuously to obtain the following theorem.

Theorem 48. *Let X be a hyperbolic space satisfying (BA) and G be a closed separable group of isometries of X. Given $\lambda > 0$, let $\mu \in$* $\mathrm{Prob}_c(G_\lambda)$ *be irreducible and $\ell(\mu) > 0$. Then there are constants $0 < \alpha \leq 1$, $C < \infty$ and $r > 0$ such that for every $\mu_1, \mu_2 \in$* $\mathrm{Prob}_c(G_\lambda)$ *if $W_\alpha(\mu_i, \mu) < r$, $i = 1, 2$, then*

$$|\ell(\mu_1) - \ell(\mu_2)| \leq C\, W_\alpha(\mu_1, \mu_2).$$

4.3.1 Wasserstein Distance and Convolution

Let now $G \subset \mathrm{Isom}(X)$ be a closed separable group. In this section we will explore the interplay between the Wasserstein distance and convolution.

Proposition 49. *Given $\lambda > 0$, let $\mu_1, \mu_2, \nu_1, \nu_2 \in$* $\mathrm{Prob}_c(G_\lambda)$*, for every $0 < \alpha \leq 1$*

$$W_\alpha(\mu_1 \star \mu_2, \nu_1 \star \nu_2) \leq W_\alpha(\mu_1, \nu_1) + C(\delta)^\alpha \lambda^\alpha W_\alpha(\mu_2, \nu_2).$$

Proof. Some parts of this proof will feel similar to the proof of Theorem 15. We also start with an inequality of the type

$$W_\alpha(\mu_1 \star \mu_2, \nu_1 \star \nu_2) \leq W_\alpha(\mu_1 \star \mu_2, \mu_1 \star \nu_2) + W_\alpha(\mu_1 \star \nu_2, \nu_1 \star \nu_2).$$

Let $\varphi \in L^\infty(G)$ with $\upsilon^G_\alpha(\varphi) \leq 1$, using the ideas from the previous proof,

$$\begin{aligned}\left| \int_G \varphi(g_1 g) d\mu_1(g_1) - \int_G \varphi(g_1 g') d\mu_1(g_1) \right| &\leq \int_G |\varphi(g_1 g) - \varphi(g_1 g')|\, d\mu_1(g_1) \\ &\leq \int_G d_G(g_1 g, g_1 g')^\alpha d\mu_1(g_1)\end{aligned}$$

$$\leq C(\delta)^\alpha \lambda^\alpha d_G(g,g')^\alpha.$$

In other words the map $g \mapsto \int_G \varphi(g_1 g) d\mu_1(g_1)$ is α-Hölder with constant $\leq C(\delta)^\alpha \lambda^\alpha d_G(g,g')^\alpha$. Hence

$$\left| \int_G \int_G \varphi(g_1 g_2) d\mu_1(g_1) d\mu_2(g_2) - \int_G \int_G \varphi(g_1 g_2) d\mu_1(g_1) d\nu_2(g_2) \right| \leq C(\delta)^\alpha \lambda^\alpha W_\alpha(\mu_2, \nu_2).$$

Transporting the inequalities from the proof of Theorem 15 once again, we obtain

$$|\varphi(gg_2) - \varphi(g'g_2)| \leq d_G(gg_2, g'g_2)^\alpha \leq d_G(g,g')^\alpha,$$

hence

$$\begin{aligned}
\left| \int_G \int_G \varphi(g_1 g_2) d\mu_1(g_1) d\nu_2(g_2) - \int_G \int_G \varphi(g_1 g_2) d\nu_1(g_1) d\nu_2(g_2) \right| &\leq \\
&\leq \int_G \left| \int_G \varphi(g_1 g_2) d\mu_1(g_1) - \int_G \varphi(g_1 g_2) d\nu_1(g_1) \right| d\nu_2(g_2) \\
&\leq W_\alpha(\mu_1, \nu_1),
\end{aligned}$$

taking the supremums over φ in the conditions above yields the result. ☐

Corollary 50. *Given $\lambda > 0$, let $\mu, \nu \in \mathrm{Prob}_c(G_\lambda)$. Then $\mu^n \in \mathrm{Prob}_c(G_{\lambda^n})$ and for every $0 < \alpha \leq 1$ and $n \in \mathbb{N}$*

$$W_\alpha(\mu^n, \nu^n) \leq W_\alpha(\mu, \nu) \sum_{i=0}^{n-1} C(\delta)^{i\alpha} \lambda^{i\alpha}.$$

Proof. For the first statement, notice that by the triangle inequality, for every $g_1, g_2 \in G_\lambda$

$$b^{d(g_1 g_2 x_0, x_0)} \leq b^{d(g_1 x_0, x_0) + d(g_2 x_0, x_0)} = b^{d(g_1 x_0, x_0)} b^{d(g_2 x_0, x_0)} \leq \lambda^2.$$

Direct applications of the previous Proposition yield the second statement as

$$\begin{aligned}
W_\alpha(\mu^n, \nu^n) &\leq W_\alpha(\mu^n, \mu \star \nu^{n-1}) + W_\alpha(\mu \star \nu^{n-1}, \nu^n) \\
&\leq W_\alpha(\mu, \nu) + C(\delta)^\alpha \lambda^\alpha W_\alpha(\mu^{n-1}, \nu^{n-1}) \\
&\leq W_\alpha(\mu, \nu) \sum_{i=0}^{n-1} C(\delta)^{i\alpha} \lambda^{i\alpha}.
\end{aligned}$$

☐

4.3.2 Existence and Uniqueness of the Stationary Measure

For every $f \in L^\infty(\partial X)$ and $0 < \alpha \le 1$ define

$$\upsilon_\alpha(f) := \sup_{\xi \neq \eta \in \partial X} \frac{|f(\xi) - f(\eta)|}{D_b(\xi, \eta)^\alpha},$$
$$||f||_\alpha := ||f||_\infty + \upsilon_\alpha(f).$$

Set

$$\mathcal{H}_\alpha(\partial X) := \{ \, f \in L^\infty(\partial X) \, : \, ||f||_\alpha < \infty \} \, .$$

the space of boundary Hölder continuous functions in ∂X. We call $\upsilon_\alpha(f)$ the Hölder constant of f. The space $\mathcal{H}_\alpha(\partial X)$ is Banach algebra with unity **1**.

Given $\mu \in \text{Prob}_c(G)$, define the Markov operator $Q_\mu : L^p(\partial X) \to L^p(\partial X)$ by

$$(Q_\mu f)(\xi) := \int_G f(g^{-1}\xi) d\mu(g),$$

for $1 \le p \le \infty$. A simple computation yields that for every $\nu \in \text{Prob}(\partial X)$ and $f \in L^1(\partial X)$

$$\begin{aligned} \int_{\partial X} (Q_\mu f)(\xi) d\nu(\xi) &= \int_{\partial X} \int_G f(g^{-1}\xi) d\mu(g) d\nu(\xi) \\ &= \int_{\partial X} f(\xi) d\mu \star \nu(\xi), \end{aligned}$$

which yields the following proposition.

Proposition 51. *Let $\mu \in \text{Prob}_c(G)$, then $\nu \in \text{Prob}(\partial X)$ is $\mu-$stationary if and only if for every $f \in L^1(\partial X)$*

$$\int_{\partial X} (Q_\mu f) d\nu = \int_{\partial X} f d\nu$$

We also have the following identity

$$\begin{aligned} (Q_\mu^n f)(\xi) &= \int_G f(g^{-1}\xi) d\mu^n(g) \\ &= \int_G \int_G f(g_{n-1}^{-1} g^{-1} \xi) d\mu^{n-1}(g) d\mu(g_{n-1}) \\ &= \int_G Q_{\mu^{n-1}} f(g_{n-1}^{-1}\xi) d\mu(g_{n-1}) \\ &= (Q_\mu(Q_{\mu^{n-1}} f))(\xi), \end{aligned}$$

in other words, for every $n \in \mathbb{N}$, $Q_{\mu^n} = Q_\mu^n$.

Given $\mu \in \text{Prob}_c(G)$ and $0 < \alpha < 1$ define the average Hölder constant of μ as

$$k_\alpha^n(\mu) := \sup_{\xi \neq \eta \in \partial X} \int_G \left(\frac{D_b(g^{-1}\xi \, , \, g^{-1}\eta)}{D_b(\xi \, , \, \eta)} \right)^\alpha d\mu^n(g).$$

Lemma 52. *For every* $f \in \mathcal{H}_\alpha(\partial X)$

$$\upsilon_\alpha(Q_{\mu^n} f) \leq k_\alpha^n(\mu)\upsilon_\alpha(f).$$

Proof. The proof is analogue to the proof of Lemma 39. □

In particular, the previous Lemma implies that the Markov operator restricts to a well defined operator in $Q_\mu : \mathcal{H}_\alpha(\partial X) \to \mathcal{H}_\alpha(\partial X)$. In the following Lemma we prove $k_\alpha^n(\mu)$ is submultiplicative, which emphasises the spectral character of the measurement k_n^α.

Lemma 53. *Let* $\mu \in \mathrm{Prob}_c(G)$*, for every* $m, n \in \mathbb{N}$

$$k_\alpha^{m+n}(\mu) \leq k_\alpha^m(\mu) k_\alpha^n(\mu).$$

Proof. The proof is analogue to the proof of Lemma 40. □

Proposition 54. *Let* $\mu \in \mathrm{Prob}_c(G)$*. If for some* $n \in \mathbb{N}$ *and* $0 < \alpha \leq 1$

$$k_\alpha^n(\mu)^{1/n} < 1,$$

then there exists a unique μ*-stationary measure* $\nu \in \mathrm{Prob}(\partial X)$*. Moreover, for every* $f \in \mathcal{H}_\alpha(\partial X)$*,*

$$\lim_{n\to\infty} Q_\mu^n(f) = \left(\int_{\partial X} f d\nu_\mu\right) \mathbf{1}.$$

Proof. For the second assertion, the seminorms υ_α are norms in the space $\mathcal{H}_\alpha(\partial X)/\mathbb{C}\mathbf{1}$. Since $Q_\mu^n \mathbf{1} = \mathbf{1}$, by hypothesis, Q_μ^n acts in $H_0 = \mathcal{H}_\alpha(\partial X)/\mathbb{C}\mathbf{1}$ as a contraction. Then $\lim Q_\mu^n(f)$ must be a constant function; by Proposition 51 it must be the constant function displayed in the statement. The remainder of the statement is equal to the Markov case. □

We will now focus on proving $k_\alpha^n(\mu_1) < 1$ in a neighbourhood of μ, provided μ is irreducible and $\ell(\mu) > 0$.

Lemma 55. *Let* $\mu \in \mathrm{Prob}_c(G)$ *be irreducible and* $\ell(\mu) > 0$*, then*

$$\lim_{n\to\infty} \frac{1}{n} \int_G h(gx_0) d\mu^n(g) = \ell(\mu)$$

uniformly on $h \in X_\infty^h$*.*

Proof. Is an immediate adaptation of the Markov version. □

Applying Proposition 14 one now has the inequality

$$k_\alpha^n(\mu) \leq C(\delta)^\alpha \sup_{h \in X^h} \int_G b^{-\alpha h(gx_0)} d\mu^n(g).$$

Since for every horofunction $h \in X^h$ and $g_1, g_2 \in G$

$$\begin{aligned} h(g_1 g_2 x_0) &= (h(g_1 g_2 x_0) - h(g_1 x_0)) + (h(g_1 x_0) - h(x_0)) \\ &= g_1^{-1} \cdot h(g_2 x_0) + h(g_1 x_0), \end{aligned}$$

one can easily verify that the process $\sup_{h \in X^h} \int_G b^{-\alpha h(gx_0)} d\mu^n(g)$ is submultiplicative. This fact is relevant to us since it allows us to pass from the spectral quantity $k_\alpha^n(\mu)$ to a more manageable one at the loss of a multiplicative constant. With that in mind we now take a closer look at the quantity on the right, in particular, next lemma tells us that $g \mapsto b^{-\alpha h(gx_0)}$ is Hölder continuous.

Lemma 56. *Given $\lambda > 0$, for every $g_1, g_2 \in G_\lambda$ one has*

$$\left| b^{-\alpha h(g_1 x_0)} - b^{-\alpha h(g_2 x_0)} \right| \leq \lambda^{2\alpha} d_G(g_1, g_2)^\alpha.$$

Proof. Start by noticing that $x \mapsto x^\alpha$ is α-Hölder with Hölder constant 1. Then we only need to control $|b^{-h(g_1 x_0)} - b^{-h(g_2 x_0)}|$, which we can do using the mean value theorem and the fact horofunctions are Lipschitz

$$\left| b^{-h(g_1 x_0)} - b^{-h(g_2 x_0)} \right| \leq \log(b)\lambda \left| h(g_1 x_0) - h(g_2 x_0) \right| \leq \lambda (\log b) d(g_1 x_0, g_2 x_0).$$

If $D_b(g_1^{-1} x_0, g_2^{-1} x_0) = \log(b) d(g_1^{-1} x_0, g_2^{-1} x_0)$ we are done, otherwise and immediate computation yields

$$\begin{aligned} (\log b) d(g_1 x_0, g_2 x_0) \leq b^{d(g_1 x_0, g_2 x_0)/2} &\leq b^{(d(g_1 x_0, x_0) + d(g_2 x_0, x_0))/2} b^{-\langle g_1 x_0 \, , \, g_2 x_0 \rangle_{x_0}} \\ &\leq \lambda D_b(g_1 x_0, g_2 x_0) \end{aligned}$$

so we are done. ∎

Remark 5. *In the course of this result we have proven that $g \mapsto h(gx_0)$ is Lipschitz.*

Proposition 57. *Given $\lambda > 0$ let $\mu \in$ Prob$_c(G_\lambda)$ be an irreducible measure with $\ell(\mu) > 0$. There are numbers $r > 0$, $0 < \alpha \leq 1$, $0 < k < 1$ and $n \in \mathbb{N}$ such that for every $\mu_1 \in$ Prob$_c(G_\lambda)$ satisfying $W_\alpha(\mu, \mu_1) < r$ one has $k_\alpha^n(\mu) \leq k$.*

Proof. We can mimic the proof of proposition to obtain the existence of $0 < \rho < 1$ and α small enough so that $\int_G b^{-\alpha h(gx_0)} d\mu^n(g) \leq \rho$. Fix such α and ρ for the remainder of the proof.

To extend this control to close measures notice that by the previous lemma $g \mapsto b^{-\alpha h(gx_0)}$ is α-Hölder with Hölder constant $\lambda^{2\alpha}$ for every $h \in X^h$. So taking μ, μ_1 with $W_\alpha(\mu, \mu_1) \leq r$, where r is at least smaller than 1 and chosen later,

$$\begin{aligned}\left|\int_G b^{-\alpha h(gx_0)}\mu^{n_0}(g) - \int_G b^{-\alpha h(gx_0)}\mu_1^{n_0}(g)\right| &\leq \lambda^{2\alpha} W_\alpha(\mu^{n_0}, \mu_1^{n_0}) \\ &\leq W_\alpha(\mu, \mu_1)\sum_{i=0}^{n-1} C(\delta)^{i\alpha}\lambda^{(i+2)\alpha}.\end{aligned}$$

So we can now choose r small enough to ensure there exists $\rho^* \in (\rho, 1)$

$$\left|\int_G b^{-\alpha h(g^{-1}x_0)}\mu^{n_0}(g) - \int_G b^{-\alpha h(g^{-1}x_0)}\mu_1^{n_0}(g)\right| \leq \rho^* - \rho.$$

Hence

$$\begin{aligned}\int_G b^{-\alpha h(gx_0)}\mu_1^{n_0}(g) &\leq \int_G b^{-\alpha h(gx_0)}\mu^{n_0}(g) + \left|\int_G b^{-\alpha h(gx_0)}\mu^{n_0}(g) - \int_G b^{-\alpha h(gx_0)}\mu_1^{n_0}(g)\right| \\ &\leq \rho^* < 1.\end{aligned}$$

Due to the submultiplicativity, picking $\sigma = (\rho^*)^{\frac{1}{n_0}}$, for every $n \in \mathbb{N}$ there exists a constant $C > 0$ such that

$$\sup_{h \in X^h} \int_G b^{-\alpha h(gx_0)} d\mu^n(g) < C\sigma^n.$$

Finally as observed before we now have

$$k_\alpha^n(\mu_1) \leq C(\delta)^\alpha \, C \, \sigma^n,$$

for every μ_1 with $W_\alpha(\mu, \mu_1) < \delta$. In particular, there exists $n \in \mathbb{N}$ for which this quantity is smaller than 1. □

4.3.3 Continuity

In the previous section we have proven that in a neighbourhood of every irreducible measure in G with positive drift all measures admit a unique stationary measure in ∂X. In the next Lemma we explore how the stationary measures behave under perturbations on the measure in G.

Lemma 58. *Given $\lambda > 0$, let $\mu_1, \mu_2 \in \mathrm{Prob}_c(G_\lambda)$ and $\nu_{\mu_1}, \nu_{\mu_2} \in \mathrm{Prob}(\partial X)$ their respective stationary measures. Suppose for some $0 < \alpha \leq 1$, $\max\{k_\alpha^n(\mu_1), k_\alpha^n(\mu_2)\} \leq k < 1$ (in particular, ν_{μ_1} and ν_{μ_2} exist), then for every $n \in \mathbb{N}$ and $f \in \mathcal{H}_\alpha(\partial X)$*

$$\left|\int_{\partial X} f d\nu_{\mu_1} - \int_{\partial X} f d\nu_{\mu_2}\right| \leq \frac{\upsilon_\alpha(f)}{1-k} W_\alpha(\mu_1, \mu_2).$$

Proof. The Markov operators satisfy

$$\begin{aligned}||Q_{\mu_1}f - Q_{\mu_2}f||_\infty &\leq \sup_{\xi\in\partial X}\left|\int_G f(g^{-1}\xi)d\mu_1(g) - \int_G f(g^{-1}\xi)d\mu_2(g)\right| \\ &\leq \upsilon_\alpha(f)W_\alpha(\mu_1,\mu_2).\end{aligned}$$

For the powers we get

$$\begin{aligned}\left|\left|Q^n_{\mu_1}f - Q^n_{\mu_2}f\right|\right|_\infty &\leq \sum_{i=0}^n ||Q^i_{\mu_2}(Q_{\mu_1}-Q_{\mu_2})(Q^{n-i-1}_{\mu_1}(f))||_\infty \\ &\leq \sum_{i=0}^n ||(Q_{\mu_1}-Q_{\mu_2})(Q^{n-i-1}_{\mu_1}(f))||_\infty \\ &\leq W_\alpha(\mu_1,\mu_2)\sum_{i=0}^n \upsilon_\alpha(Q^{n-i-1}_{\mu_1}(f)) \\ &\leq W_\alpha(\mu_1,\mu_2)\upsilon_\alpha(f)\sum_{i=0}^n k^{n-i-1} \\ &\leq \frac{\upsilon_\alpha(f)}{1-k}W_\alpha(\mu_1,\mu_2).\end{aligned}$$

Now $\lim_{n\to\infty} Q_{\mu_1}f = (\int_{\partial X} fd\nu_{\mu_1})\mathbf{1}$ and $\lim_{n\to\infty} Q_{\mu_2}f = (\int_{\partial X} fd\nu_{\mu_2})\mathbf{1}$ so

$$\left|\int_{\partial X} fd\nu_{\mu_1} - \int_{\partial X} fd\nu_{\mu_2}\right| \leq \sup_n \left|\left|Q^n_{\mu_1}f - Q^n_{\mu_2}f\right|\right|_\infty,$$

from which we obtain the result. □

Proof of Theorem 48. Let $\nu \in \text{Prob}(\partial X)$, for every $g, g' \in G_\lambda$, applying the mean value theorem with $x \mapsto b^{-\alpha x}$ as well as Lemma 56

$$\begin{aligned}\left|\int_{\partial X} h_\xi(gx_0)d\nu(\xi) - \int_{\partial X} h_\xi(g'x_0)d\nu(\xi)\right| &\leq \int_{\partial X} |h_\xi(gx_0) - h_\xi(g'x_0)|\, d\nu(\xi) \\ &\leq \frac{\max\{b^{d(gx_0,x_0)}, b^{d(g'x_0,x_0)}\}}{\alpha\log b}\int_{\partial X}|b^{-\alpha h_\xi(gx_0)} - b^{-\alpha h_\xi(g'x_0)}|d\nu(\xi) \\ &\leq \frac{\lambda}{\alpha\log b}\int_{\partial X}|b^{-\alpha h_\xi(gx_0)} - b^{-\alpha h_\xi(g'x_0)}|d\nu(\xi) \\ &\leq \frac{\lambda^3}{\alpha\log b}d_G(g,g')^\alpha,\end{aligned}$$

in particular, the map $g \mapsto \int_{\partial X} h(g^{-1}x_0)d\nu(g)$ is Hölder continuous.

Let $\mu \in \text{Prob}_c(G_\lambda)$ be irreducible with $\ell(\mu) > 0$. Then there exist $0 < \alpha \leq 1$ and a neighbourhood of μ in which all measures satisfy $k^n_\alpha(\mu_1) < 1$. Let μ_1, μ_2 be in a neighbourhood of μ and ν_{μ_1}, ν_{μ_2} their respective stationary measures. Using the Furstenberg type formula.

$$|\ell(\mu_1) - \ell(\mu_2)| \leq \left|\int_G\int_{\partial X} h_\xi(gx_0)d\nu_{\mu_1}(\xi)d\mu_1(g) - \int_G\int_{\partial X} h_\xi(gx_0)d\nu_{\mu_2}(\xi)d\mu_2(g)\right|$$

$$\leq \left| \int_G \int_{\partial X} h_\xi(gx_0) d\nu_{\mu_1}(\xi) d\mu_1(g) - \int_G \int_{\partial X} h_\xi(gx_0) d\nu_{\mu_1}(\xi) d\mu_2(g) \right|$$
$$+ \left| \int_G \int_{\partial X} h_\xi(gx_0) d\nu_{\mu_1}(\xi) d\mu_2(g) - \int_G \int_{\partial X} h_\xi(gx_0) d\nu_{\mu_2}(\xi) d\mu_2(g) \right|$$
$$\leq \frac{\lambda^3}{\alpha \log b} W_\alpha(\mu_1, \mu_2) + \frac{\lambda^3}{\alpha (\log b)(1-k)} W_\alpha(\mu_1, \mu_2).$$

□

4.4 Analyticity of the Drift

Gouëzel (2017), proved that the drift is analytic for random walks on word hyperbolic groups provided the measures have a fixed finite support. Despite solving the case of hyperbolic groups, Gouëzel left open the question of general groups acting on a hyperbolic space, which we now tackle. The argument is similar to Gouëzel's and dates back to Peres (1991). The method will boil down to applying the same ideas from the previous sections once again.

We are interested in considering measures supported in a fixed finite set of isometries, say $F = \{g_1, ..., g_d\}$. A probability measure μ in F is directly identified with probability vector $p = (p_1, p_2, ..., p_d)$ by

$$\mu_p = \sum_{i=1}^{d} p_i \delta_{g_i},$$

where δ_{g_i} stands for the Dirac measure at g_i. We will prove that the drift is analytic provided μ is irreducible and the drift $\ell(\mu)$ is positive. This section is a joint work with Jamerson Bezerra and the result of discussions between the two.

Theorem 59. *Let $F = \{g_1, g_2, ..., g_d\}$ be isometries of an hyperbolic space X satisfying (BA). Let $p = (p_1, p_2, ..., p_d)$ be a probability vector and $\mu_p = \sum_{i=1}^{d} p_i \delta_{g_i}$ a probability measure. If $\ell(\mu_p) > 0$, μ_p is irreducible and $p_i > 0$ for every $1 \leq i \leq d$, then ℓ is an analytic function in a neighbourhood of $(p_1, p_2, ..., p_d)$.*

Proof. The first step of the work is to consider our probability vector as living in $\mathbb{C}$ as this allows for a direct description of analyticity through holomorphic functions. So we consider $z = (z_1, z_2, ..., z_d)$ such that $\sum_{i=1}^{d} z_i = 1$, and take its Markov operator $Q_z : L^\infty(\partial X) \to L^\infty(\partial X)$

$$(Q_z f)(\xi) = \sum_{i=1}^{n} z_i f(g_i^{-1} \xi).$$

Notice that for each ξ, $(Q_z^n f)(\xi)$ are polynomials of degree n in the variables $(z_1, z_2, ..., z_d)$, hence holomorphic. Next we use the results from the previous section regarding the contractive behaviour of the Markov operator. Notice as well that since μ is irreducible and the drift is positive, as in Proposition 57, we have

$$k_\alpha^n(\mu_p) \leq C(\delta)^\alpha \, C \, \sigma^n. \tag{4.5}$$

where $C > 0$, $C(\delta)$ is as defined in 14 and $\sigma < 1$.

We will now see that provided f is Lipschitz, then $Q_z^n f$ converges uniformly for every z in

$$\Omega = \left\{ (z_1, z_2, ..., z_d) \in \mathbb{C}^n \, : \, \sum_{i=1}^{n} z_i = 1, \; \frac{|z_i|}{p_i} < \sigma_0 \text{ for } 1 \leq i \leq d \right\},$$

where $1 < \sigma_0 < \sigma^{-1}$ for the sigma given above. With effect, given $\xi, \eta \in \partial X$, we have

$$\begin{aligned} |(Q_z^n f)(\xi) - (Q_z^n f)(\eta)| &\leq \sum_{i=1}^{d} |z_i (Q_z^{n-1} f)(g_i^{-1}\xi) - z_i (Q_z^{n-1} f)(g_i^{-1}\eta)| \\ &\leq \sigma_0 \sum_{i=1}^{d} |p_i (Q_z^{n-1} f)(g_i^{-1}\xi) - p_i (Q_z^{n-1} f)(g_i^{-1}\eta)| \\ &= \sigma_0 \int_F |(Q_z^{n-1} f)(g_i^{-1}\xi) - (Q_z^{n-1} f)(g_i^{-1}\eta)| d\mu_p(g), \end{aligned}$$

hence by induction

$$|(Q_z^n f)(\xi) - (Q_z^n f)(\xi)| \leq \sigma_0^n \int_{F^n} |f(g^{-1}\xi) - f(g^{-1}\eta)| d\mu_p^n(g).$$

Assuming f is $L-$Lipschitz, using (4.5), we obtain

$$|(Q_z^n f)(\xi) - (Q_z^n f)(\xi)| \leq C L \sigma_0^n \sigma^n D_b(\xi, \eta). \tag{4.6}$$

Since $\sum_{i=1}^d z_i = 1$, applying the previous equation with $\eta = g_i^{-1}\xi$, we obtain

$$\begin{aligned} \left|(Q_z^{n+1} f)(\xi) - (Q_z^n f)(\xi)\right| &= \left| \sum_{i=1}^{d} z_i (Q_z^n f)(g_i^{-1}\xi) - (Q_z^n f)(\xi) \right| \\ &= \left| \sum_{i=1}^{d} z_i \left[(Q_z^n f)(g_i^{-1}\xi) - (Q_z^n f)(\xi) \right] \right| \\ &\leq d\sigma_0 C L (\sigma\sigma_0)^n, \end{aligned}$$

thus obtaining the intended uniform convergence as $\sigma\sigma_0 < 1$. We can project the first $d-1$ coordinates of Ω onto an open set of $\mathbb{C}^{d-1}$, which defines an analytic structure on Ω. The uniform convergence now implies that $\lim_{n\to\infty}(Q_z^n f)(\xi)$ is an analytic function of z, which by (4.6), doesn't depend on ξ.

Now we use $f_i(\xi) = h_\xi(g_i x_0)$. Notice that by the remark to Lemma 56 these are Lipschitz functions. So the limit of the functions

$$P_n(z) = \sum_{i=1}^{d} z_i (Q_z f_i)(\xi)$$

converges uniformly in Ω to an analytic function P. Now let $q = (q_1, q_2, ..., q_n) \in \Omega$ be a probability vector. Then for every f_i

$$(Q_q^n f_i)(\xi) = \int_{F^n} f_i(g^{-1}\xi) d\mu_q^n(g)$$

Hence

$$P_n(q) = \sum_{i=1}^{d} q_i (Q_q f_i)(\xi) = \int_{F^n} \left(\sum_{i=1}^{d} q_i h_{g^{-1}\xi}(g_i x_0) \right) d\mu_q^n(g),$$

where the rightmost side converges to the drift by the Fürstenberg type formula argument (claims 1 and 2 in Lemma 55). □

4.5 Bibliographic Notes

In this chapter we explored Markov systems in hyperbolic spaces. This has been done very scarcely in this degree of generality as most of the literature in the area sticks to random walks alone. In fact Goldsborough and Sisto (2021) made a proposal towards the study of problems regarding Markov chains in hyperbolic-like groups very recently. Thus the remainder of the references we present here will be based on the random walks case alone.

When it comes to large deviations, one may find the case of random walks in Sampaio (2021), which I decided to not include in the thesis as it intersects the more interesting case of Markov systems. Our work in this problem meets Boulanger et al. (2020), where a large deviations principle is proven in the setting of non-elementary measures in a countable group G. More recently Aoun and Sert (2022) obtained large deviation estimates as well as continuity in the proper case for cocompact actions of G in X. Our large deviations are weaker in the sense that the constants are local, although they apply more generally. The local nature of our results are a consequence of the methods applied. More precisely, we will use spectral techniques motivated by Duarte et al. (2016). Such methods have also been used in the case of hyperbolic groups in Björklund (2010), where a central limit theorem is presented.

Regarding continuity, it was known that if μ is non-elementary and supported on a finite set of some hyperbolic group, the drift is known to be analytic with respect to measures supported in the same set (see Gouëzel (2017); Gouëzel et al. (2018); Gilch and Ledrappier (2013)), here we expand this result to the case outside hyperbolic groups. As mentioned before, in the proper case our continuity results are also similar to the ones in Aoun and Sert (2022). Note however that our work allows Hölder continuity in compactly supported measures. As a consequence, we can consider more general measures and also allow for closeby isometries to be considered.

Although we define irreducible measures, the case where μ is non-elementary is also interesting to us. In Maher and Tiozzo (2018), it was proven that in this case the drift is positive; however, non-elementary measures are always irreducible by definition, in other words, our results apply in such situations.

The Fürstenberg type formula is another place where we borrow from Lyapunov exponents, namely from Furstenberg and Kifer (1983). Multiple metric versions appear in many papers, including the ones previously cited. Other references such as Carrasco et al. (2017) and Karlsson and Ledrappier (2011) are also relevant. We reprove the statement as these works don't relate back with irreducibility.

Chapter 5

On the Linear Setting

In this chapter we present a different setting where the operators are now linear, instead of isometries of some metric space, and we try to understand the average expanding factor of their product. It is well known, see Lian and Lu (2010), that the dynamics of product of linear operators are governed by the multiplicative ergodic theorem of Oseledets and the top Lyapunov will take the role of the drift.

Once again, questions regarding continuity of these limit quantities arises naturally. As is the theme of the thesis, we will present an avalanche principle and large deviations estimates to obtain an abstract continuity theorem; this time simply explaining how the tools for the finite dimensional case still exist in the infinite dimensional one of what changes from the previous cases. The finite dimensional version of these results can be found in Duarte et al. (2016). Here we will work with Hilbert spaces, which is why we mentioned Lian and Lu (2010). We recommend chapter 5 of Temam (2012) for a more thorough presentation of the prerequisites for this chapter, namely when it comes to exterior products and other multilinear algebra considerations regarding Hilbert spaces.

5.1 Generalities

Let H be a separable Hilbert space. We denote the natural norm in H by $|\cdot|$ and the operator norm in H by $||\cdot||$. Denote as well by $\mathcal{L}(H)$ the space of bounded linear operators from H to itself. Given a bounded linear operator we will denote by A^* its adjoint operator, which satisfies $\langle Au, v\rangle = \langle u, A^*v\rangle$ for every $u, v \in H$.

Definition 15. (Singular values) Given a bounded linear operator $A : H \to H$, we define it's $k-th$

singular value by the expression

$$s_k(A) = \sup_{\substack{F \subset H \\ \dim F = k}} \inf_{\substack{u \in F \\ |u|=1}} |Au|.$$

Trivially $s_1(A) = ||A||$, but we also have our disposal the relation $s_1(A)s_2(A) = ||\wedge^2 A||$, where $\wedge^2$ stands for the exterior power. In finite dimensions, we can assert that there are orthonormal vectors $u_1, ..., u_n \in H$ such that $|Au_i| = s_i(A)$. These vectors are the singular vectors and $s_i(A)$ are the singular values from the singular value decomposition. For the infinite dimensional case, singular value decomposition exists for example when A is compact. Nevertheless something may be said in the noncompact case. With effect, since $s_k(A)$ are decreasing, define

$$s_\infty(A) = \lim_{k \to \infty} s_k(A) = \inf_{k \geq 1} s_k(A).$$

Then we obtain the following theorem in the spirit of the singular value decomposition.

Theorem 60 (Chapter V, Proposition 1.3 in Temam (2012))**.** *Let H be a Hilbert space and B its unit ball. Let $A \in \mathcal{L}(H)$ and assume $s_k(A) > s_\infty(A)$ for some $k \in \mathbb{N}$. Then there exists a closed subspace space $E \subset H$ of dimension k such that $A(B)$ is included in an ellipsoid $\mathcal{E}$ which is a product of the ball centered at 0 with radius and $s_\infty(A)$ in $E^\perp$, and of the ellipsoid of E, whose axes are directed along the vectors Av_i, with length $s_i(A)$, where v_i are the orthogonal eigenvectors of A^*A spanning E.*

In particular, in the conditions of the theorem there exist two sets of orthonormal vectors $\{u_1, ...u_k\}$ and $\{v_1, ..., v_k\}$ such that $Au_i = s_i(A)v_i$ and $A^*v_i = s_i(A)v_i$. We call u_i's the singular direction of A. Notice for example that if $s_1 > s_2$ the theorem implies that there exists a unique direction u such that $Au = s_1(A)u$.

Theorem 61 (Avalanche Principle)**.** *Let $A_1, ..., A_n \in \mathcal{L}(H)$ be such that for every $1 \leq i \leq n$. There are universal constants $0 < C_1 < 1$, $C_2 > 0$ such that if for $0 \leq i \leq n-1$*

G) $\mathrm{gr}(A_i) := \frac{||A_i||}{s_2(A_i)} \geq a$;

A) $\frac{||A_i\, A_{i-1}||}{||A_i||\, ||A_{i-1}||} \geq b$;

P) $a^{-1} \leq C_1 b^2$.

Then, setting $A^{(n)} = A_n A_{n-1} ... A_1$

$$\left| \log ||A^{(n)}|| + \sum_{i=2}^{n-1} \log ||A_{i-1}|| - \sum_{i=1}^{n-1} \log ||A_i A_{i-1}|| \right| \leq C_2 \frac{n}{ab}.$$

Condition G) asserts that there exists a gap between the two first singular values, which implies some sort of contraction towards the most expanding direction, whereas condition A) requires a certain alignment between the most expanded direction of a map and the most expanding direction of the subsequent map. Finally condition P) quantitatively relates the two. Then we obtain a conclusion saying that the product satisfies a similar relationship of being aligned with its parcels in a quantifiable way.

Proof. See Duarte and Klein (2017). Having introduced the singular values and singular vectors, the proofs follow as in the finite dimensional case. □

Let (Ω, β, μ) be a standard space, given an ergodic transformation $T : \Omega \to \Omega$, we say that a measurable map $a : \mathbb{N} \times \Omega \to \mathcal{L}(H)$ is a multiplicative cocycle in $\mathcal{L}(H)$ over T if $a(n+m, \omega) = a(m, T^n\omega)a(n, \omega)$. To every Borel measurable $A : \Omega \to \mathcal{L}(H)$ we associate a right multiplicative cocycle

$$a(n, \omega) = A^{(n)}(\omega) := A(T^{n-1}\omega)...A(T\omega)A(\omega),$$

where $n \in \mathbb{N}$ and $\omega \in \Omega$. A cocycle is thus comprised of the information (A, T, Ω, β), whenever it is clear we denote it simply by A. Cocycles of this kind are also called linear cocycles.

Definition 16 (Integrable Cocycle)**.** We say that a linear cocycle (A, T, Ω, β) is integrable if

$$\int_\Omega \log^+ ||A(\omega)|| d\mu(\omega) < \infty,$$

where $\log^+(x) = \max\{\log(x), 0\}$.

One of the fundamental characteristics of an integrable cocycle is its top Lyapunov exponent

$$L_1(A) := \lim_{n\to\infty} \frac{1}{n} \int_\Omega \log ||A^{(n)}(\omega)|| d\mu(\omega) = \lim_{n\to\infty} \frac{1}{n} \log ||A^{(n)}(\omega)||,$$

where the first limit exists by Kingman's ergodic theorem whilst the second equality is true for almost every ω due to ergodicity. The Lyapunov exponent describes the exponential growth rate of the norm along orbits. By Kingman's ergodic theorem, the limits

$$\lambda_j(A) = \lim_{n\to\infty} \frac{1}{n} \log || \wedge^j A^{(n)}(\omega)|| = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{j} \log \left(s_i(A^{(n)}(\omega)) \right),$$

where $\wedge^j A$ denotes the j exterior power of A, also exist and are almost everywhere independent of ω. Then we may define the j-th Lyapunov exponent as

$$L_j(A) = \lambda_j(A) - \lambda_{j-1}(A) = \lim_{n\to\infty} \frac{1}{n} \log \left(s_j(A^{(n)}) \right).$$

Just as before we can now consider a class of cocycles where we tackle regularity problems. Consider $\mathcal{C}_m$ indexed on $m \in \mathbb{N}$ to be the space of integrable cocycles $A : \Omega \to \mathcal{L}(H)$ such that

$$|| \wedge_i A||_\infty := \sup_{\omega \in \Omega} || \wedge_i A(\omega)||$$

is finite, for $1 \leq i \leq m$. Equipped with some metric d such that $d(A, B) \geq ||A - B||_\infty$, for every $A, B \in \mathcal{C}_m$.

We can now think of the Lyapunov exponents as maps

$$\begin{aligned} L_k : \mathcal{C}_m &\to \mathbb{R} \\ A &\mapsto L_k(A). \end{aligned}$$

5.2 Abstract Continuity Theorem

Having the avalanche principle at our disposal an abstract continuity theorem for the top Lyapunov exponent can now be proven. Just as in the metric setting, in order to apply the avalanche principle, we need large deviations, however there is a small caveat this time, as the base dynamics provided by the ergodic transformation also needs to satisfy some type large deviation estimates.

Definition 17. An observable $\zeta : \Omega \to \mathbb{R}$ satisfies a base large deviation estimates if for every $\varepsilon > 0$ there exist $c > 0$ and n_0 depending on (ζ, ε) and C depending on ζ such that for every $n \geq n_0$

$$\mu \left\{ \omega \in \Omega \, : \, \left| \frac{1}{n} \sum_{j=0}^{n-1} \zeta(T^j \omega) - \int_\Omega \zeta d\mu \right| > \varepsilon \right\} < C e^{-cn}.$$

We will need these base large deviation estimates to prove that the conditions of the avalanche principle remain true through the inductive process as we will discuss later. Given a cocycle A define

$$L_i^n(A) = \frac{1}{n} \int_\Omega \log \left(s_i(A^{(n)}(\omega)) \right) d\mu(\omega).$$

Definition 18. We say that a cocycle $A \in \mathcal{C}_m$ satisfies uniform (fiber) large deviation estimates of exponential type if there are constants r, C, $c > 0$ and for every $\varepsilon > 0$ there exists $\bar{n} = \bar{n}(\varepsilon)$ such that

$$\mu \left\{ \omega \in \Omega \, : \, \left| \frac{1}{n} \log ||A_1^{(n)}(\omega)|| - L_1^n(A_1) \right| > \varepsilon \right\} < e^{-c\varepsilon^2 n}$$

for every $A_1 \in \mathcal{C}_m$ with $d(A, A_i) < r$ and every $n \geq \bar{n}$.

Theorem 62 (Abstract continuity theorem)**.** *Consider an ergodic system* (Ω, β, μ, T)*, a space of measurable cocycles* $\mathcal{C}_2$ *and a space of observables* Θ*. Assume the following*

1. *Θ is dense in $L^1(\Omega, \mu)$;*

2. *Every observable $\zeta \in \Theta$ satisfies a base large deviation estimate;*

3. *Every cocycle A in $\mathcal{C}_2$ such that $L_1 - L_2 > 0$ satisfies a uniform fiber large deviation estimates.*

Then the top Lyapunov exponent is continuous. Moreover it is locally Hölder continuous when restricted to the cocycles satisfying $L_1(A) > L_2(A)$.

Remark 6. *In the finite dimensional case we can work with determinants, which allows for a stronger result, namely that if $A \in \mathcal{C}_n$, where $\dim H = n$, such that $L_1 - L_2 > 0$ satisfies a fiber large deviation estimates, then every Lyapunov exponent is continuous. One can find this in Corollary 3.1 in Duarte et al. (2016).*

5.3 Large Deviations for Quasi-periodic Cocycles

Let $\mathbb{T}$ stand for the quotient group $\mathbb{R}/\mathbb{Z}$. Here $\mathbb{R}$ is the universal cover of $\mathbb{T}$, thus making sense of a topology, differentiable structure and Lebesgue measure μ in $\mathbb{T}$ (which agrees with the Haar measure). Let $\alpha \in \mathbb{T}$, we consider the map

$$\begin{aligned} T_\alpha &: \mathbb{T} \to \mathbb{T} \\ x &\mapsto x + \alpha \mod 1. \end{aligned}$$

Proposition 63. *The transformation T_α is ergodic with respect to μ if and only if α is irrational.*

Proof. Since $\mathbb{T}$ is compact, an alternative yet simple to verify it is equivalent, definition of ergodicity is that a system (T_α, μ) is ergodic if and only if all the functions $f \in L^2(\Omega, \mu)$ satisfying $f \circ T_\alpha = f$ almost everywhere are constant almost everywhere.

Thus let $f \in L^2(\Omega, \mu)$ and expand it in its Fourier series $f(x) = \sum_{k\in\mathbb{Z}} \hat{f}(k) e^{2\pi i k x}$ and $f \circ T_\alpha(x) = \sum_{k\in\mathbb{Z}} \hat{f}(k) e^{2\pi i k x} e^{2\pi i k \alpha}$. Now since $e^{2\pi i k x}$ are linearly independent and α_i are irrational, $f = f \circ T_\alpha$ if and only if $\hat{f}(k) = 0$ for every $n \neq 0$. With that we have proven (T_α, μ) is ergodic. □

For the remainder of the section we assume T_α is ergodic, in other words, α is irrational. However, in order to obtain the type of large deviations we want we must place an additional arithmetic assumption on α, namely that it can't be approximated by rationals with small denominators.

Definition 19. We say that $\alpha \in \mathbb{T}$ satisfies a Diophantine condition if

$$d(k \cdot \alpha, \mathbb{Z}) \geq \frac{\gamma}{|k|(\log |k|)^2}.$$

for some γ and every $k \in \mathbb{Z}\backslash\{0\}$.

The group, $\mathbb{R}/\mathbb{Z}$ can be identified with the interval $[0, 1]$ with the endpoints identified, in other words the circle $\mathbb{S} \subset \mathbb{C}$. Explicitly, $\mathbb{T} \to \mathbb{S}$ given by $x \mapsto e^{2\pi i x}$ is a homeomorphism. Using this identification, we say that a function f on $\mathbb{T}$ has holomorphic extension to an open domain $\mathbb{T} \subset \Omega \subset \mathbb{C}$, if there is an holomorphic function $\hat{f} : \Omega \to \mathbb{C}$, such that $\hat{f}(x) = f(x)$ for every $x \in \mathbb{T}$. Holomorphic extensions, when they exist, are unique by the interior uniqueness theorem.

Let now $r > 0$ and consider the annulus region of width r around $\mathbb{T}$ given by

$$\mathcal{A}_r := \{z \in \mathbb{C} \, : \, 1 - r \leq |z| \leq 1 + r\}.$$

Definition 20. Let H be a separable Hilbert space, $\mathcal{L}(H)$ the space of bounded linear operators from H to H with the operator norm and $\mathcal{L}(H)^*$ its dual. Now we set $\mathrm{GL}(H)$ as the set of invertible bounded linear operators in H. We define $C^\omega_r(\mathbb{T}, \mathrm{GL}(H))$ to be the set of all functions $A : \mathbb{T}^d \to \mathrm{GL}(H)$, such that for every $f \in \mathcal{L}(H)^*$, $f \circ A$ admits an holomorphic extension to $\mathcal{A}_r$.

In $C^\omega_r(\mathbb{T}, \mathrm{GL}(H))$ we define the distance

$$d(A, B) = \sup_{z \in \mathcal{A}_r} ||A(z) - B(z)||,$$

making $C^\omega_r(\mathbb{T}, \mathrm{GL}(H))$ a complete metric space, since it is a closed subspace of the space of functions $\mathbb{T} \to \mathrm{GL}(H)$ with the L^∞ norm.

Adapting the argument in Chapter 6 of Duarte et al. (2016) we obtain the follow large deviation estimates theorem:

Theorem 64. *Let $A \in C^\omega_r(\mathbb{T}, \mathrm{GL}(H))$ be a cocycle over an irrational rotation T_α, where α satisfies a Diophantine condition. Let $C < \infty$ be a constant such that $\log \sup_{z \in \mathcal{A}_r} |A(z)| < C$.*

For every small $\varepsilon > 0$ there is $c = c(C)$ and $\bar{n} = \bar{n}(\varepsilon, C) \in \mathbb{N}$ such that for every $n \geq \bar{n}$

$$\mu \left\{ x \in \mathbb{T} \, : \, \left| \frac{1}{n} \log ||A^{(n)}(x)|| - L^{(n)}(A) \right| > \varepsilon \right\} < e^{-c\varepsilon^2 n},$$

Notice that since deviation constants $c, \bar{n}$ only depend on the constant C, it is uniform. Thus the abstract continuity apply, whence:

Theorem 65. *Let $A \in C^\omega_r(\mathbb{T}, \mathrm{GL}(H))$ be a cocycle over an irrational rotation T_α, where α satisfies a Diophantine condition. Then the Lyapunov exponent $L : C^\omega_r(\mathbb{T}, \mathrm{GL}(H))$ is continuous.*

Moreover, if $A \in C^\omega_r(\mathbb{T}, \mathrm{GL}(H))$ has positive Lyapunov exponent, $L_1(A) > L_2(A)$, then the Lyapunov exponent is locally Hölder continuous around A.

5.3.1 The argument

In this subsection we explore the argument behind theorem 64. The argument will use mostly harmonic analysis, potential theory and number theory. No proofs will be included as these can be found in Duarte et al. (2016); this is mostly a simple exposition on the steps.

Definition 21. A function $u : \Omega \to [-\infty, \infty)$ is called subharmonic in the domain $\Omega \subset \mathbb{C}$ if for every $z \in \Omega$, u is upper semicontinuous at z and there exists $r_0(z)$ such that ht mean value property

$$u(z) \leq \int_0^1 u(z + re^{2\pi i\theta})d\theta,$$

for every $r \leq r_0(z)$.

A classical method for obtaining subharmonic functions is by considering $\log|f(z)|$ for some analytic function f. Moreover, the supremum of a collection of subharmonic functions is subharmonic, provided it is upper semicontinuous, thus given an analytic $A : \Omega \to \mathrm{GL}(H)$,

$$u(z) := \log||A(z)|| = \sup_{||v||=||w||=1} \log|\langle A(z)v\,,\,w\rangle|$$

is subharmonic in Ω. The same arguments yields that $u_A^{(n)}(z) = \frac{1}{n}\log||A^{(n)}(z)||$ are subharmonic. Moreover, one can easily prove that $u_A^{(n)}(z)$ are uniformly bounded in z, n and A.

At this point we note that for every $x \in \mathbb{T}$

$$\begin{aligned}\frac{1}{n}\log||A^{(n)}(x)|| - \frac{1}{n}\log||A^{(n)}(Tx)|| &= \frac{1}{n}\log\frac{||A(T^n x)^{-1}A^{(n)}(Tx)A(x)||}{||A^{(n)}(Tx)||}\\ &\leq \frac{1}{n}\log||A(T^n x)^{-1}||\,||A(x)||\\ &\leq \frac{2\sup_{z\in\mathcal{A}_r}\log||A(z)||}{n}\end{aligned}$$

and an analogous control holds for $\frac{1}{n}\log||A^{(n)}(Tx)|| - \frac{1}{n}\log||A^{(n)}(x)||$, so that we get the invariance principle

$$|u_A^{(n)}(x) - u_A^{(n)}(Tx)| \leq \frac{C}{n}.$$

Using then the triangle inequality one has

$$|u_A^{(n)}(x) - u_A^{(n)}(T^j x)| \leq \frac{Cj}{n}.$$

Thus picking R to be the integer part of $\log(n)$ for all $x \in \mathbb{T}$ we have

$$\left|u_A^{(n)}(x) - \frac{1}{R}\sum_{j=0}^{n-1} u_A^{(n)}(T^j x)\right| \leq \frac{CR}{n}.$$

Obtaining the invariance principle is why we require that our operators are invertible.

However, by Birkhoff's ergodic theorem for almost every $x \in \mathbb{T}$

$$\lim_{R\to\infty} \frac{1}{R}\sum_{j=0}^{R-1} u_A^{(n)}(T^j x) = \int_{\mathbb{T}} u_A^{(n)} = L^{(n)}(A).$$

This reduces the proof of large deviations to obtaining a quantitative version of Birkhoff's ergodic theorem, in other words we need to prove

$$\mu\left\{x \in \mathbb{T} \,:\, \left|\frac{1}{R}\sum_{j=0}^{R-1} u_A^{(n)}(T^j x) - \int_{\mathbb{T}} u_A^{(n)}\right| > \varepsilon\right\} \leq e^{-c\varepsilon^2 n},$$

holds uniformly in a neighbourhood of A. The first step in this direction is writing the observable u as its Fourier series $u(x) = \sum_{k\in\mathbb{Z}} \hat{u}(k)e^{2\pi ikx}$. Since $\hat{u}(0) = \int_{\mathbb{T}} u$ we have

$$u(x) - \int_{\mathbb{T}} u = \sum_{k\neq 0} \hat{u}(k)e^{2\pi ikx},$$

so with the Birkhoff averages we have

$$\frac{1}{R}\sum_{j=0}^{R-1} u_A^{(n)}(T^j x) - \int_{\mathbb{T}} u_A^{(n)} = \sum_{k\neq 0} \hat{u}(k)\frac{1}{R}\sum_{j=0}^{R-1} e^{2\pi ikx}e^{2\pi ikj\alpha}.$$

At this point we must control both the Fourier coefficients and the functions $e^{2\pi ikj\alpha}$. Controlling Fourier coefficients is where we use the subharmonic properties. It is known that Fourier coefficients decay sublinearly ($\hat{u}(k) \lesssim 1/|k|$) for continuously differentiable functions. Using Riesz representation theorem one can prove the same holds for subharmonic functions. The part $e^{2\pi ikj\alpha}$ is where the Diophantine condition is used. By quantifying how far $j\alpha$ is from being irrational we also control how small $|e^{2\pi ikj\alpha}|$ is.

5.4 Bibliographic Notes

We have now reached an important point in the thesis where the reader has tools for the drift, Lyapunov exponent and, risking repetition, virtually any subadditive quantity over a dynamical system. Albeit less relevant, we find that the drift yields a simple presentation of the topics, hence we were more thorough with its presentation. With that said, the theory developed for Markov systems also works for Lyapunov exponents, although one must be careful with substituting the horofunctions compactification with the unit ball equipped with the weak* topology.

As we have asserted before, the finite dimensional linear case was treated in Duarte et al. (2016). During our talks, Duarte P. asked me about the possibility of an abstract continuity theorem in infinite dimensions. That is the novelty of our work here, the operators live in some arbitrary Hilbert space. This may open new doors in the study of higher dimensional quasi-periodic Schrödinger cocycles (see Embree and Fillman (2019) for the setting).

Literature on one-dimensional quasi-periodic cocycles is quite vast, namely on the continuity problem. The study as expanded in multiple directions such as dropping the Diophantine condition (see Avila et al. (2014); Jitomirskaya and Marx (2012)), this however comes at the cost of not being able to quantify it in terms of Hölder continuity, which at times is important; reduce the regularity of the cocycles such as in Wang and Zhang (2015) or the noninvertible case allowing singularities in Avila et al. (2014).

The higher dimensional case (in $\mathbb{T}^d$) is more involved and complex (see Duarte and Klein (2017)). This is where the power of the abstract continuity theorem shows through. Our results in section 5.3 work for rational independent rotations of the torus with the appropriate Diophantine condition. Such treatment requires further analytic tools, for instance, pluri-subharmonic functions that we didn't want to touch upon in the thesis.

Chapter 6

Drift in Symmetric Spaces

In this chapter we showcase a class of Riemannian manifolds - symmetric spaces - where despite not having a direct avalanche principle, one can explore the algebraic properties associated with the group of isometries to represent it in a linear group thus using the linear avalanche principle and continuity theorems. Exploring these representations actually yields the correct framework as under mild assumptions the contracting behaviour fails for the horofunction compactification.

Symmetric spaces are also interesting as they come with quite large groups of isometries, typically larger than the space itself. Our considerations on symmetric spaces done in this chapter follow Helgason (2001) and Borel and Ji (2005).

6.1 Generalities

Let G be a Lie Group and $\mathfrak{g}$ its Lie algebra. We denote by C_g the conjugation by g. Its differential at e is the adjoint representation of G, $Ad_g \in \text{Aut}(\mathfrak{g})$. The map $Ad : G \to \text{Aut}(\mathfrak{g})$ given by $g \to Ad_g$ is differentiable, and its derivative is the adjoint representation of $\mathfrak{g}$, $ad : \mathfrak{g} \to \text{Der}(\mathfrak{g})$. It is well known that, $ad_x(y) = [x, y]$.

The symmetric bilinear form on $\mathfrak{g}$

$$B(x, y) = \text{tr}(ad_x \circ ad_y)$$

is called the Killing form. A Lie algebra is semisimple if it is non-abelian and its only ideals are $\{0\}$ and $\mathfrak{g}$. A classical result asserts that a finite dimension Lie algebra over a field of characteristic zero is semisimple if and only if its Killing form is non-degenerate. Finally, a Lie group is semisimple if its Lie algebra is.

Given $\mathfrak{g}$ a Lie algebra and B its Killing form. We say that θ is a Cartan involution if $B_\theta(x, y) = -B(x, \theta y)$ is positive definite. Any semisimple Lie algebra admits a Cartan involution, which is unique up to conjugation by automorphisms. For matrix groups, $x \to -x^*$ is a Cartan involution.

Since θ is an involution, it admits only the eigenvalues ± 1. We denote by $\mathfrak{k}$ and $\mathfrak{p}$ the eigenspaces corresponding to 1 and -1, respectively. This yields a $Ad(K)-$invariant decomposition of $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ where

$$[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k},\ [\mathfrak{k}, \mathfrak{p}] \subset \mathfrak{p},\ [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}, \tag{6.1}$$

$\mathfrak{k}$ is the Lie algebra of K and $\mathfrak{p}$ may be identified with the tangent space $T_{x_0}M$ of the symmetric space $M = G/K$. Moreover the restriction of B to $\mathfrak{k} \times \mathfrak{k}$ is negative definite. This decomposition is called the Cartan decomposition and it is orthogonal with respect to the Killing form. Moreover any subalgebra of $\mathfrak{p}$ is abelian and all maximal abelian subalgebras are conjugate by K. This yields a new Cartan decomposition of G, namely, if $A = \exp \mathfrak{a}$, $G = KAK$.

Example 4. Let us consider $G = \mathrm{SL}(d, \mathbb{R})$ and $K = \mathrm{SO}(d, \mathbb{R})$. The associated symmetric space X is the space of symmetric positive definite matrices. The lie algebra is the set $\mathfrak{sl}(d, \mathbb{R})$ of trace zero matrices. $\theta(A) = -A^T$ is the Cartan involution, yielding $\mathfrak{k} = \mathfrak{so}(d)$ as the set of antisymmetric matrices and $\mathfrak{p} = \mathfrak{sym}(d)$ as the set of symmetric matrices of trace zero. The Cartan decomposition is simply the known decomposition

$$x = \frac{1}{2}(x - x^T) + \frac{1}{2}(x + x^T).$$

The Cartan decomposition of $G = KAK$ is given by the singular value decomposition.

6.2 Gap Ratios

In previous iterations of the abstract continuity theorem we had either the distance between iterates or the gap ratio playing a fundamental role in the description of the problem. In either case, the requirement came from the contractive actions it would imply. In the case of symmetric spaces $M = G/K$, the generalized algebraic varieties will take the role of the contracted space. These appear naturally as quotients of G by its parabolic subgroups.

6.2.1 Roots

Consider $G = SL(d, \mathbb{C})$. One way to obtain the gap ratios is as eigenvalues of the map

$$\begin{aligned} Ad_g : \mathfrak{g} &\to \mathfrak{g} \\ A &\mapsto gAg^{-1}. \end{aligned}$$

We shall use this idea to obtain gap ratios in this more general setting.

Since the maximal subalgebras of $\mathfrak{p}$ are conjugate under K, let $\mathfrak{a}$ be one of them. A linear form $\lambda \in \mathfrak{a}^*$ is a root (or restricted root) if it is a nonzero form and the root space

$$\mathfrak{g}_\lambda = \{x \in \mathfrak{g} \,:\, ad_h(x) = [h, x] = \lambda(h)x, \, \forall h \in \mathfrak{a}\} \neq 0.$$

The set of roots is a root system in $\mathfrak{a}^*$ denoted $\phi = \phi(\mathfrak{g}, \mathfrak{a})$. The Weyl group $W = W(\mathfrak{g}, \mathfrak{a})$ may be identified with $N_K(\mathfrak{a})/Z_K(\mathfrak{a})$, where $N_K(\mathfrak{a})$ and $Z_K(\mathfrak{a})$ are, respectively, the normalizer and centralizer of $\mathfrak{a}$:

$$N_K(\mathfrak{a}) = \{h \in \mathfrak{a} \,:\, Ad_k(h) \in \mathfrak{a}\}$$
$$Z_K(\mathfrak{a}) = \{h \in \mathfrak{a} \,:\, Ad_k(h) = 0\}$$

Roots define hyperplanes $H_\alpha = \ker \alpha$. The connected components of $\mathfrak{a} - \cup_{\alpha\in\phi} H_\alpha$ are called the Weyl chambers. The action of the Weyl group can also be defined as the the group of reflections on H_α. This group acts simply transitively in the set of Weyl chambers of $\mathfrak{a}$. Hence we will fix a Weyl chamber, to be denoted by $\mathfrak{a}^+$. We define the set of positive roots as

$$\phi^+ = \phi^+(\mathfrak{g}, \mathfrak{a}) = \{\alpha \in \phi \,:\, \alpha > 0 \text{ on } \mathfrak{a}^+\}.$$

Analogously we can define ϕ^-, the set of negative roots. An element of ϕ^+ that cannot be written as a sum of two other positive roots is called simple. The set of simple roots is denoted by $\Delta = \Delta(\mathfrak{g}, \mathfrak{a})$. Let

$$\mathfrak{n} = \sum_{\alpha>0} \mathfrak{g}_\alpha, \qquad \mathfrak{n}^- = \sum_{\alpha<0} \mathfrak{g}_\alpha$$

By construction, these are nilpotent algebras exchanged by θ, normalized by $\mathfrak{a}$ such that

$$\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{z}(\mathfrak{a}) \oplus \mathfrak{n},$$

where $\mathfrak{z}(\mathfrak{a})$ is the centralizer of $\mathfrak{a}$. Notice as well that $\mathfrak{z}(\mathfrak{a}) = \mathfrak{k} \cap \mathfrak{z}(\mathfrak{a}) \oplus \mathfrak{a}$, we will denote by $\mathfrak{m}$ the space $\mathfrak{k} \cap \mathfrak{z}(\mathfrak{a})$.

The roots can also be transported to $A = \exp \mathfrak{a}$. Namely, given $a \in A$ and α a root. We define $a^\alpha = \exp \alpha(\log a)$. We will call these maps the roots of A. We also set a positive Weyl chamber A^+. The Cartan decomposition of $G = KA^+K$ yields a well defined map $\mu : G \to A^+$ called the Cartan projection. Due to the Cartan projection, we will define the roots on all G as $g^\alpha = \mu(g)^\alpha$. Root spaces can be defined directly through G as

$$\mathfrak{g}_\lambda = \{x \in \mathfrak{g} \,:\, Ad_a(x) = a^\lambda x, \, \forall a \in A\}$$

Example 5. Let us consider the lie algebra $\mathfrak{sl}(d, \mathbb{C})$, where we fix the maximal abelian algebra $\mathfrak{a}$, consisting of trace zero real diagonal matrices. The set of roots is given by $\{\alpha_{ij} \in \mathfrak{a}^*\}$, $i \neq j$, where

$$\alpha_{ij} : \mathfrak{a} \to \mathbb{R}$$
$$\text{diag}(a_1, ..., a_d) \mapsto a_i - a_j$$

and $\mathfrak{g}_{\alpha_{i,j}}$ consists of matrices whose entries besides the entry (i, j) are zero. The Weyl group is isomorphic to S_d, it acts by permuting coordinates on the diagonal. We fix the positive Weyl chamber

$$\mathfrak{a}^+ = \{\text{diag}(a_1, ..., a_d) \, : \, a_1 > a_2 > ... > a_d\}.$$

Then the set of positive roots is given by $\{\alpha_{i,j} \in \mathfrak{a}^* \, : \, i > j\}$, while the simple roots are of the form $\alpha_{i,i+1}$.

The roots in A are of the form $\text{diag}(a_1, ..., a_d)^{\alpha_{i,j}} = a_i/a_j$. Hence for a matrix in $g \in \text{SL}(d, \mathbb{C})$ one has $g^{\alpha_{i,j}} = s_i/s_j$ where s_i stands for the i-th singular value of g.

6.3 Representations

As we have stated before, the theory for these groups will boil down to the linear theory since they all admit a linear representation. In this section we broadly present the technicalities of this approach. Most of the details may be found for example in Benoist and Quint (2016).

6.3.1 Weights

Denote a representation $\rho : G \to \text{SL}(V)$ by (ρ, V). As an abuse of language, we will also call the tangent representation $\mathfrak{g} \to \mathfrak{sl}(V)$ by (ρ, V). Our representations are assumed to be faithful. For every character $\chi \in \mathfrak{a}^*$, set its weight space

$$V_\chi = \{v \in V : \forall a \in \mathfrak{a}, \, \rho(a)v = \chi(a)v\}.$$

Call the elements $\chi \in \mathfrak{a}^*$ such that $V_\chi \neq 0$ weights of ρ. The roots from the previous section are the weights for the adjoint representation.

A representation (ρ, V) is said to be irreducible if there is no proper nonzero subrepresentation $(\rho_{|W}, W)$, $W \subset V$. Suppose now that (ρ, V) is irreducible. In that case, if we endow the set of weights with a partial order given by

$$\chi_1 \leq \chi_2 \Leftrightarrow \chi_2 - \chi_1 \text{ is a linear combination of positive roots with positive coefficients}$$

then there is a largest weight, called the highest weight. Let χ be the highest weight and N the unipotent radical subgroup of G (whose tangent space at the identity is $\mathfrak{n}$), set $V_N := \{v \in V : \rho(N)v = v\}$ and denote by $V_{N,\chi}$ the intersection $V_N \cap V_\chi$ which we call the parabolic weight space of χ. We say a representation is proximal if $\dim V_{N,\chi} = 1$.

Recall that the Killing form defined an inner product on $\mathfrak{g}$ which is transported to $\mathfrak{g}^*$. For every simple root $\alpha \in \Delta$ we define its associated fundamental weight ϖ_α, by

$$2\frac{\langle \varpi_\alpha \,,\, \lambda\rangle}{\langle \lambda \,,\, \lambda\rangle} = \delta_{\alpha,\lambda}, \qquad \forall \lambda \in \Delta.$$

Theorem 66. *(in Tits (1971)) If G is a connected semisimple Lie group. For every simple root $\alpha \in \Delta$, there exists an irreducible proximal algebraic representation (ρ_α, V_α) whose highest weight χ_α is a multiple of ϖ_α, the fundamental weight associated with α.*

Remark 7. *Due to the defining expression of fundamental weights, one has that the roots for (ρ_α, V_α) are of the form χ_α, $\chi_\alpha - \alpha$ and $\chi_\alpha - \alpha - \sum_{\beta\in\Delta-\alpha} m_\beta \beta$, where $m_\beta \in \mathbb{N}$.*

6.3.2 Good Norms

We are finally ready to build the promised good norms which will relate to the avalanche principle via the presentations we have just constructed.

Lemma 67. *Let G be a connected semisimple Lie group and (V, ρ) be an irreducible representation of G with highest weight χ. Then there exists an hermitian scalar product ϕ on V such that*

- *ϕ is $K-$invariant;*
- *$\forall a \in A$, $\rho(a)$ is symmetric and diagonal;*
- *$\chi(\log \mu(g)) = \log |\rho(g)|$.*

Proof. Let $\mathfrak{g}_\mathbb{C}$ be the complexification of the Lie algebra of G and $\hat{\mathfrak{g}} = \mathfrak{k} + i\mathfrak{p}$ its compact real form. Denote by $\hat{G} \in \mathrm{SL}(V)$ a compact Lie group whose Lie algebra equals $\hat{\mathfrak{g}}$. Given an inner product $\langle\, ,\, \rangle$ on V, define the $\hat{G}$-invariant inner product

$$\phi(x, y) = \int_{\hat{G}} \langle kx, ky\rangle dk.$$

By construction of $\hat{G}$, this inner product is clearly K-invariant. Moreover $\rho(a)$ are symmetric as they are both real and hermitian. For every $a \in \mathfrak{a}^+$, the eigenvalues of $\rho(e^a)$ are the real numbers $e^{\chi'(x)}$ where χ' is some weight, since χ is the highest weight, one must have $\chi(a) = \log |\rho(e^a)|$, thus proving the third point.

□

6.4 Continuity of Lyapunov Exponents to the continuity of the Drift

Let $g \in G$ and x_0 be a base point in X, then $d(gx_0, x_0) = c|\log \mu(g)|$, for some constant $c > 0$. This makes the distance dependent solely on $\log \mu(g)$. However, $\log \mu(g)$ belongs to $\mathfrak{a}^+$, a subset of $\mathfrak{a}$ for which there exists a basis given by the representatives of the simple roots. That is, given the simple roots Δ, the elements $\{a_\alpha\} \in \mathfrak{a}$, such that $\alpha(\cdot) = \langle \cdot, a \rangle$, form a basis of $\mathfrak{a}$. Moreover, by definition of fundamental weight ϖ_α,

$$\log \mu(g) = \sum_{\alpha \in \Delta} \frac{\langle a_\alpha, a_\alpha \rangle}{2} \varpi_\alpha(\log \mu(g)).$$

which is a closed formula depending on $\varpi_\alpha(\log \mu(g))$. With that said, the continuity of the drift depends only on the continuity of $\lim_{n\to\infty} \varpi_\alpha(\frac{1}{n} \log \mu(g^{(n)}(\omega)))$ for every simple root α.

Using Theorem 66 recall that $\varpi_\alpha(\log \mu(g))$ equals to $\log |\rho_\alpha(g)|$ up to a multiplicative constant. Thus, given a cocycle $g : \Omega \to M$, the drift $\ell(g)$ is related to the Lyapunov exponents of $\rho_\alpha(g)$.

Definition 22. A space of measurable cocycles $\mathcal{C}$ is any class of measurable bounded $g : \Omega \to G$. We say that a map $g : \Omega \to G$ is bounded if $\sup_{\omega\in\Omega} |\rho_\alpha \circ g(\omega)|$ is finite for every simple root $\alpha \in \Delta$.

We will assume that $\mathcal{C}$ is equipped with a metric d, such that

$$d(g, h) > \sup_{\alpha\in\Delta} \sup_{\omega\in\Omega} |\rho_\alpha \circ g(\omega) - \rho_\alpha \circ h(\omega)|.$$

Since by assumption g is bounded, $||\rho_\alpha \circ g|| \in L^\infty(\mu)$. We then have $\log ||\rho_\alpha \circ g|| \in L^1(\mu)$, so by Kesten's theorem we can define the finite scale α-Lyapunov exponent

$$L_\alpha^{(n)}(g) = \int_\Omega \frac{1}{n} \log ||\rho_\alpha \circ A^{(n)}|| d\mu(x)$$

and the α-Lyapunov exponent $L_\alpha(g)$. In what follows, stronger conditions on $\frac{1}{n} \log |\rho_\alpha \circ A^{(n)}|$ are needed, namely we need large deviations.

Definition 23. An observable $\zeta : \Omega \to \mathbb{R}$ satisfies a base LDT if for every $\varepsilon > 0$ there exist $C > 0$, $0 < b \le 1$ and n_0 depending on (ζ, ε) such that for every $n \ge n_0$

$$\mu\left\{\omega \in \Omega \, : \, \left|\frac{1}{n}\sum_{j=0}^{n-1} \zeta(T^j\omega) - \int_\Omega \zeta d\mu\right| > \varepsilon\right\} < Ce^{-c(\varepsilon)n^b}.$$

We also need a large deviation for the cocycle itself.

Definition 24. A cocycle $g \in \mathcal{C}$ is said to satisfy an $\alpha-$LDT estimate if for every $\varepsilon > 0$ there exist $C > 0$, $0 < b \leq 1$ and n_0 depending on (ζ, ε) such that for every $n \geq n_0$

$$\mu \left\{ \omega \in \Omega \,:\, \left| \frac{1}{n} \log |\rho_\alpha \circ g^{(n)}(\omega)| - L_\alpha(g) \right| > \varepsilon \right\} < Ce^{-c(\varepsilon)n^b}.$$

In order to obtain continuity on the cocycle one needs a stronger conditions, namely, that the LDT estimates are uniform on a neighbourhood of the given one.

Definition 25. A cocycle $g \in \mathcal{C}$ is said to satisfy a uniform $\alpha-$LDT estimate if for every $\varepsilon > 0$ there exist $\delta = \delta(g, \varepsilon)$ and a $C > 0$, $0 < b \leq 1$ and n_0 depending on (ζ, ε) such that if $h \in \mathcal{C}$ with $d(g, h) < \delta$ and $n \geq n_0$ then

$$\mu \left\{ \omega \in \Omega \,:\, \left| \frac{1}{n} \log |\rho_\alpha \circ h^{(n)}(\omega)| - L_\alpha(h) \right| > \varepsilon \right\} < Ce^{-c(\varepsilon)n^b}.$$

We finish with the theorem in this setting.

Theorem 68 (Abstract Continuity Theorem)**.** *Consider an ergodic system (Ω, β, μ, T) a space of measurable cocycles $\mathcal{C}$ and a space of observables Θ. Assume the following*

1. *Θ is dense in $L^1(\Omega, \mu)$;*

2. *Every observable $\zeta \in \Theta$ satisfies a base LDT.*

3. *For every $\alpha \in \Delta$, every $g \in \mathcal{C}$ for which $g^\alpha > 0$ satisfies a uniform fiber $\alpha-$LDT estimate.*

Then, for every $\alpha \in \Delta$, $L_\alpha : \mathcal{C} \to [-\infty, \infty)$ are continuous functions of the cocycle over the distance d.

Moreover, if $g^\alpha > 0$ for every $\alpha \in \Delta$, then the drift is locally Hölder continuous.

Bibliography

Aoun, R. and Sert, C. (2022). Random walks on hyperbolic spaces: Concentration inequalities and probabilistic tits alternative. *Probability Theory and Related Fields*, pages 1–43.

Avila, A., Jitomirskaya, S., and Sadel, C. (2014). Complex one-frequency cocycles. *Journal of the European Mathematical Society*, 16(9):1915–1935.

Benoist, Y. and Quint, J.-F. (2016). *Random Walks on Reductive Groups*, volume 62. Springer.

Björklund, M. (2010). Central limit theorems for Gromov hyperbolic groups. *Journal of theoretical probability*, 23(3):871–887.

Borel, A. and Ji, L. (2005). Compactifications of symmetric and locally symmetric spaces. In *Lie theory*, pages 69–137. Springer.

Bougerol, P. (1988). Théorèmes limite pour les systèmes linéaires à coefficients markoviens. *Probability Theory and related fields*, 78(2):193–221.

Boulanger, A. R., Mathieu, P., Sert, C., and Sisto, A. (2020). Large deviations for random walks on hyperbolic spaces. *arXiv preprint arXiv:2008.02709*.

Bridson, M. R. and Haefliger, A. (2013). *Metric spaces of non-positive curvature*, volume 319. Springer Science & Business Media.

Carrasco, M., Lessa, P., and Paquette, E. (2017). A Furstenberg type formula for the speed of distance stationary sequences. *arXiv preprint arXiv:1710.00733*.

Cheeger, J. and Gromoll, D. (1971). The splitting theorem for manifolds of nonnegative ricci curvature. *Journal of Differential Geometry*, 6(1):119–128.

Coornaert, M. and Papadopoulos, A. (2001). Horofunctions and symbolic dynamics on Gromov hyperbolic groups. *Glasgow Mathematical Journal*, 43(3):425–456.

Das, T., Simmons, D., and Urbański, M. (2017). *Geometry and dynamics in Gromov hyperbolic metric spaces*, volume 218. American Mathematical Soc.

Duarte, P. and Klein, S. (2017). *Continuity of the Lyapunov exponents of linear cocycles*. Associação Instituto de Matemática Pura e Aplicada- IMPA.

Duarte, P., Klein, S., et al. (2016). *Lyapunov exponents of linear cocycles*, volume 3. Springer.

Embree, M. and Fillman, J. (2019). Spectra of discrete two-dimensional periodic schrödinger operators with small potentials. *Journal of Spectral Theory*, 9(3):1063–1087.

Frink, A. H. (1937). Distance functions and the metrization problem. *Bulletin of the American Mathematical Society*, 43(2):133–142.

Furstenberg, H. and Kifer, Y. (1983). Random matrix products and measures on projective spaces. *Israel Journal of Mathematics*, 46(1):12–32.

Gilch, L. and Ledrappier, F. (2013). Regularity of the drift and entropy of random walks on groups. *Publ. Mat. Urug.*, 14:147–158.

Goldsborough, A. and Sisto, A. (2021). Markov chains on hyperbolic-like groups and quasi-isometries. *arXiv preprint arXiv:2111.09837*.

Goldstein, M. and Schlag, W. (2001). Hölder continuity of the integrated density of states for quasi-periodic schrödinger equations and averages of shifts of subharmonic functions. *Annals of Mathematics*, pages 155–203.

Gouëzel, S. (2017). Analyticity of the entropy and the escape rate of random walks in hyperbolic groups. *Discrete Analysis*, 2017(7).

Gouëzel, S. and Karlsson, A. (2020). Subadditive and multiplicative ergodic theorems. *Journal of the European Mathematical Society*, 22(6):1893–1915.

Gouëzel, S., Mathéus, F., and Maucourant, F. (2018). Entropy and drift in word hyperbolic groups. *Inventiones mathematicae*, 211(3):1201–1255.

Han, R., Lemm, M., and Schlag, W. (2020). Effective multi-scale approach to the Schrödinger cocycle over a skew-shift base. *Ergodic Theory Dynam. Systems*, 40(10):2788–2853.

Helgason, S. (2001). *Differential geometry and symmetric spaces*, volume 341. American Mathematical Soc.

Jitomirskaya, S. and Marx, C. (2012). Analytic quasi-perodic cocycles with singularities and the Lyapunov exponent of extended harper's model. *Communications in mathematical physics*, 316(1):237–267.

Karlsson, A. and Ledrappier, F. (2006). On laws of large numbers for random walks. *The Annals of Probability*, 34(5):1693–1706.

Karlsson, A. and Ledrappier, F. (2011). Noncommutative ergodic theorems. *Geometry, rigidity, and group actions*, pages 396–418.

Karlsson, A. and Margulis, G. A. (1999). A multiplicative ergodic theorem and nonpositively curved spaces. *Communications in mathematical physics*, 208(1):107–123.

Kingman, J. F. (1968). The ergodic theory of subadditive stochastic processes. *Journal of the Royal Statistical Society: Series B (Methodological)*, 30(3):499–510.

Lian, Z. and Lu, K. (2010). *Lyapunov exponents and invariant manifolds for random dynamical systems in a Banach space*. American Mathematical Soc.

Maher, J. and Tiozzo, G. (2018). Random walks on weakly hyperbolic groups. *Journal für die reine und angewandte Mathematik*, 2018(742):187–239.

Nagaev, S. V. (1957). Some limit theorems for stationary markov chains. *Theory of Probability & Its Applications*, 2(4):378–406.

Oregón-Reyes, E. (2020). The avalanche principle and negative curvature. *Mathematische Zeitschrift*, 294(3):1397–1412.

Peres, Y. (1991). Analytic dependence of Lyapunov exponents on transition probabilities. In *Lyapunov exponents*, pages 64–80. Springer.

Riesz, F. and Nagy, B. S. (2012). *Functional analysis*. Courier Corporation.

Sampaio, L. M. (2021). Regularity of the drift for random walks in groups acting on Gromov hyperbolic spaces. *arXiv preprint arXiv:2109.13060*.

Sampaio, L. M. (2022). Continuity of the drift in groups acting on strongly hyperbolic spaces. *arXiv preprint arXiv:2204.08299*.

Sarig, O. (2009). Lecture notes on ergodic theory. *Lecture Notes, Penn. State University*.

Schlag, W. (2013). Regularity and convergence rates for the Lyapunov exponents of linear cocycles. *Journal of Modern Dynamics*, 7(4).

Sentilles, F. D. (1972). Bounded continuous functions on a completely regular space. *Transactions of the American Mathematical Society*, 168:311–336.

Temam, R. (2012). *Infinite-dimensional dynamical systems in mechanics and physics*, volume 68. Springer Science & Business Media.

Tits, J. (1971). Représentations linéaires irréductibles d'un groupe réductif sur un corps quelconque. *Journal für die reine und angewandte Mathematik*, 247:196–220.

Väisälä, J. (2005). Gromov hyperbolic spaces. *Expositiones Mathematicae*, 23(3):187–231.

Viana, M. (2014). *Lectures on Lyapunov exponents*, volume 145. Cambridge University Press.

Viana, M. and Oliveira, K. (2016). *Foundations of ergodic theory*. Number 151. Cambridge University Press.

Villani, C. (2009). *Optimal transport: old and new*, volume 338. Springer.

Wang, Y. and Zhang, Z. (2015). Uniform positivity and continuity of Lyapunov exponents for a class of c2 quasiperiodic schrödinger cocycles. *Journal of Functional Analysis*, 268(9):2525–2585.