Statistical Properties of Geometric Flows and Equidistribution

May 1, 2014

Introduction

The recent proof of the surface subgroup conjecture and the Ehrenpreis conjecture employed as their chief engine a quantitative understanding of the ergodic properties of the dynamical system generated by the geodesic flow: in particular their rate of mixing. This report aims to describe the techniques to gain quantitative estimates of these ergodic properties of these geometric flows, as well the techniques to apply them. The main example, and the heart of this report, will be the proof of the surface subgroup theorem, which says that any closed, orientable, irreducible hyperbolic 3 manifold admits an essential surface (one immersed in a way injective on fundamental groups).

Our analysis will be focused on homogeneous spaces of the form G/K for G a semisimple Lie group and K a maximal compact subgroup, known as symmetric spaces. Symmetric spaces are ubiquitous and a beautiful subject in their own right, and a brief introduction to their theory will be provided in section 2. Essentially, studying these offers an armageddon of tools for studying geometric flows: algebraic and geometric, both of which are particularly well behaved because of the rigid algebraic structure from semisimple Lie groups and highly symmetric geometric structure.

Spaces whose universal cover (with the pulled back metric) is a symmetric space are called *locally symmetric spaces*. There is a strong duality between those spaces whose universal cover is compact and is non-compact, which mirrors the duality between complex and compact semisimple Lie groups.

There are several other, mainly analytical, ways to prove quantitative results about the ergodic properties of geometric flows. For example, Dologpyat showed [16] that a wide class of hyperbolic flows are exponentially mixing. But we'll take a different approach, focusing on these symmetric spaces in which we can utilise representation theory to quantify mixing. However, a key advantage of this approach is that it automatically gives us control over the frame flow on hyperbolic manifolds using the algebraic relationship between the frame bundle and the isometry group.

After some basics and an outline of the theory of symmetric spaces, we'll focus on rank 1 locally symmetric spaces. We'll study closed locally symmetric spaces of the non-compact type. The ones of compact type are closed and all geodesics are closed and of the same length. Thus the geodesic flow is periodic and has no interesting ergodic properties. The main examples of the locally symmetric spaces we'll study are closed hyperbolic manifolds. Essentially, these are spaces where there is a unique isometry mapping any two geodesic segments of the same length to each other, and include hyperbolic manifolds. This corresponds to the real rank of the semisimple Lie group in question G, and indeed the geodesic flow corresponds to a particular subgroup $A \subset G$. We'll analyse the mixing properties by way of understanding the A action on a specific class of representations.

The techniques we use to analyse the representations are themselves useful. most importantly, we'll look at decompositions of semisimple Lie group, which give rigid control over the action of a group on itself by conjugation. These decompositions will be very useful for later parts of the report.

Once we've proven exponential mixing, we'll show that on hyperbolic manifolds, our bounds on the rate of mixing depending on short time solutions to the heat equation on our symmetric space (that is, the first eigenvalues of the Laplacian). This relates the mixing in a very strong way to the underlying geometry of the manifold, and we will cover a theorem of Cheeger.

We'll progress to a way of quantifying consequences of mixing via proving equidistribution of orbits of points in a larger class of symmetric spaces, affine symmetric spaces. These correspond to spaces G/H where H isn't necessarily compact and we actually have a pseudo-Riemannian structure. Equidistribution is a direct consequence of mixing, and in the sense that we'll prove, it is a way of quantifying mixing.

As such, we'll apply the equidistribution theorem to count the intersection of an orbit under a lattice with increasingly large sets, and in doing so, estimate the asymptotic behaviour. We'll use it to count integral points on a quadratic form, and such an estimate comprises the part of a relatively recent proof of Siegel's mass formula. We'll give a short explanation of this.

We are then in a position to tackle the proof of the surface subgroup theorem. The result of the theorem is impressive and the proof is constructive. Hyperbolic manifolds are locally symmetric spaces, and the results on exponential mixing will be readily applicable. A short introduction to hyperbolic geometry will be given. We'll give a short example illustrating the idea of using an 'abundance of objects' given by the mixing, in particular, gaining control over counting geodesics of a specific length. This example will demonstrate the key 'ergodic ideas' in the proof of the surface subgroup theorem.

The real challenge of the proof is to describe the symmetry given by equidistribution/mixing. That is, it won't be enough to show that there is an abundance of suitable objects, but we'll need to say with some precision where they are. We'll reap the full rewards of our algebraic approach, and utilise exponential mixing of the frame flow. The techniques used to do this will be be technical observations from hyperbolic geometry. Although we'll give sufficient preliminary information to the understand the motivation and language behind these, we will avoid details that obscure the main ideas of the proof.

We make the observation that explicit use of the quantitative rate of mixing has not appeared frequently in the literature, although less specific control over dynamical properties has been used. The applications we give should demonstrate some of the power of applying such quantitative techniques, which provide a new angle of attack for constructive proofs.

The chief source of background on symmetric spaces and Lie groups arises from Helgason's classic text [1]. The proof in Section 2 comes from Moore's paper [7]. Section 3, covering equidistribution, is based on the seminal paper of Eskin and McMcullen [6]. The proof of the surface subgroup theorem is found in [9].

I thank Prof. Vlad Markovic for setting this essay, which covered a large range of topics, as well as for useful comments and ideas.

Contents

Section 1

Basic concepts

1.1 Measures on groups and quotients

For details and proofs for this section, see [5, Chapter 8].

1.1.1 Haar measure

A fundamental tool in any analysis of topological groups (all of which we assume to be Hausdorff and second countable), we give some basic definitions and properties of Haar measure.

Definition 1.1.1. Let (X, μ) and (Y, ν) be measure spaces and $f : X \to Y$ a map. By $f_*\mu$, we mean the pushforward measure on Y given by $f_*\mu(A) = \mu(f^{-1}A)$

Theorem 1.1.1 (Haar measure)**.** *Let G be a locally compact group. Then there exists a unique, (up to scaling) right invariant, nontrivial Borel measure on G. That is a measure μ, right Haar measure, such that for all continuous functions f of compact support*

$$\int_G f(g)d\mu = \int_G f(gh)d\mu$$

for all $h \in G$.

Similarly, there exists a unique (up to scaling) left invariant measure, referred to as left Haar measure.

If μ is a right Haar measure, then for all $h \in G$, the measure $h_*\mu$ is also a right Haar measure (where h denotes the left translation by h on G). By the uniqueness of Haar measure, $h_*\mu = \Delta(h)\mu$. This observation gives rise to the following definition.

Definition 1.1.2. For $h \in G$, the function $\Delta_G : G \to \mathbb{R}^+$, $h \mapsto \Delta_G(h)$ where

$$h_*\mu = \Delta_G(h)\mu$$

is the modular function of G. G is unimodular if $\Delta \equiv 1$. It is clear from the definition that $\Delta_G : G \to \mathbb{R}^+$ is a homomorphism and indeed it is readily seen from the integral definition of left invariance that it is continuous

Remark. A left Haar measure is a right Haar measure if and only if G is unimodular. In this case, we refer to *the* Haar measure dg.

Example 1.1.1 (Unimodular groups)**.** The following examples of unimodular groups are important to the rest of this report.

1. Abelian groups. For any f, we have that $f(gh) = f(hg)$, so a left Haar measure is a right Haar measure.

2. Compact groups. Given a left Haar measure μ, the measure ν given by

$$\nu(A) = \int_G \chi_A(gk) d\mu(k)$$

is a left and right Haar measure, so G is unimodular. We usually normalise so that $dg(G) = 1$.

3. Semisimple Lie groups. Semisimple Lie groups admit a left-right invariant metric (via the Killing form). So the induced Riemannian measure is a left-right Haar measure.

4. Countable discrete groups (the Haar measure is just counting measure).

5. Groups admitting a lattice (see below).

1.1.2 Quotients

We begin with the following

Proposition 1.1.1. Let $H \subset G$ be a closed subgroup (not necessarily normal). G/H has a unique (up to scaling) measure invariant under left translation if and only if

$$\Delta_G|_H \equiv \Delta_H.$$

Remark. We refer to the measure $\widetilde{\mu}$ as the Haar measure on G/H (despite that G/H is not a group if H is not normal). We normalise so that

$$\int_{G/H} \int_H \phi(gh) d\mu|_H d\widetilde{\mu} = \int_G \phi d\mu$$

Definition 1.1.3. Let $\Gamma \subset G$ be a discrete subgroup. Then Γ is a lattice if the induced measure on G/Γ exists and is finite.

Proposition 1.1.2. If G admits a lattice, then G is unimodular.

1.2 Symmetric spaces and Lie Groups

This section will introduce the relevant concepts from Lie theory, the theory of symmetric spaces and demonstrate how the two can be used in combination to understand the geometry of these spaces. The goal will be to provide a context for the results of later chapters and to distil the relationship between the ergodic theory of geometric actions on spaces and the properties of the underlying groups.

1.2.1 Riemannian symmetric spaces

Let (M, g) denote a Riemannian manifold (which we will assume to be connected unless explicitly stated otherwise). Let G denote the isometry group of (M, g).

Definition 1.2.1. (M, g) is a symmetric space is a symmetric space if for each $p \in M$ there is an $s_p \in G$ such that $s_p^2 = \mathrm{id}$ and s_p is an isolated fixed point.

Remark. The isolated fixed point condition implies that $(ds_p)_p : T_pM \to T_pM = -\mathrm{id}$ and hence that s_p is locally the geodesic symmetry at p.

Each point has an isometric reflection through it. It's clear that if M is a symmetric space, then it must be geodesically complete, since any geodesic $\gamma : [0, t] \to M$ can be extended to all of $\mathbb{R}$ by setting $\gamma'(t + \epsilon) = s_t\gamma(t - \epsilon)$. The Hopf-Rinow theorem implies, in particular, that there is a unique, length minimising geodesic between any two points.

Remark. This implies that $G = \mathrm{Isom}^0(M)$ (the identity component of the isometry group of M) acts transitively on M, since we can take the geodesic reflections between the $\frac{1}{3}$-points of the unique, length minimising geodesic between any two points.

Theorem 1.2.1. *Let (M, g) be a Riemannian manifold. Then $G = \mathrm{Isom}^0(M)$ is a Lie group, such that its natural action is smooth and proper. Moreover, for any $p \in M$, $\mathrm{Stab}_G(p)$ is a compact subgroup of $O(\dim M)$.*

Proof. A full proof of the first fact can be found in [1, p. 169]. For an isometry k such that $k(p) = p$, we have that $k \circ \exp_p = \exp_p \circ dk_p$. In particular, if $dk_p = \mathrm{id}$, $k = \mathrm{id}$. Hence the map $K \to O(\dim M)$ given by $k \mapsto dk_p$ is an embedding of K into $O(n)$. Since K is closed, K must be compact.

Proposition 1.2.1. (i) Any Riemannian symmetric space is analytically diffeomorphic to $M = G/K_p$, via $gK \mapsto g \cdot p$.

(ii) $\sigma : g \mapsto s_p g_p^s$ is an involutive automorphism of G.

(iii) Let $\mathfrak{g}$ and $\mathfrak{k}$ denote the Lie algebras of G and K respectively. Then $d\sigma_e : \mathfrak{g} \to \mathfrak{g}$ gives rise to a decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. Moreover, $T_pM \cong \mathfrak{p}$ and the geodesic with tangent vector $X \in \mathfrak{p}$ is given by $\exp(tX)p$ and the parallel translates along the geodesic of $Y \in \mathfrak{p} = T_pM$ are given by $d\exp(tX))_p(Y)$.

Remark. It can be shown that a Riemannian symmetric space is in fact analytic (that is, the metric and the manifold are analytic) [1, p187]

1.2.2 Symmetric spaces as homogeneous spaces

Definition 1.2.2. A homogeneous $G-$space is a manifold M with a transitive, free Lie group action of a Lie group G. We denote by K_p the stabiliser of $p \in M$ under the G-action.

Proposition 1.2.2. K_p is a closed Lie subgroup of G.

Remark. In the case where M is a homogeneous G-space, the above enables us to write $M \cong G/K_p$ for some $p \in M$

The following notion of a symmetric space is a generalisation of the definition for Riemannian symmetric spaces given above.

Definition 1.2.3. A symmetric pair is a pair (G, K), where G is a connected analytic Lie group and K is an open subgroup of $\hat{K}$, where $\hat{K}$ is the fixed point set of some involutive analytic automorphism of G. It's a Riemannian symmetric pair if $\mathrm{Ad}(K) \subset \mathrm{Ad}(G)$ is compact.

The following proposition gives us a way to relate this to our previous definition of symmetric spaces, and gives us a glimpse into just how well behaved symmetric spaces are.

Proposition 1.2.3. Let (G, K) be a Riemannian symmetric pair, with involution σ. There is a G-invariant Riemannian metric g_0 on $M = G/K$ such that M is a Riemannian symmetric space. Moreover, the symmetry s_o at $o := eK$ satisfies

$$s_o \circ \pi = \pi \circ \sigma$$

where $\pi : G \to G/K$ is the canonical projection. Moreover, the Levi-Civita connection of the metric is independent of the choice of G-invariant metric.

Thus, combining the above proposition with *Proposition 1.2.1*, there is a correspondence between Riemannian symmetric spaces and Riemannian symmetric pairs (G, K). From now on, we will refer to symmetric spaces as $X = G/K$, fixing a base point o, and we avoid distinguishing between X as a coset space and X as a Riemannian manifold in its own right. We assume G is the identity component and thus connected.

For a symmetric space $X = G/K$ It is important to understand the action of G on X by isometries.

Let $X = G/K$. Denote by o the base point eK. Let $\pi : G \to X$ denote the the canonical submersion, $g \mapsto go$. Let s_o denote the symmetry at o and σ the corresponding involution of G. By *Proposition 1.2.3*, we make the observation that

$$d\pi_e \operatorname{Ad}(k)X = dk_o d\pi_e X$$

Identifying $\mathfrak{p}$ with the tangent space at o, we observe that for a vector $X \in T_oX$ the stabiliser subgroup (under isometries) of X is in K is isomorphic to the subgroup of K acting trivially with the adjoint action on $X \in \mathfrak{p}$. That is, the group with Lie algebra $\mathfrak{m} \subset \mathfrak{k}$ centralising $X \in \mathfrak{p}$.

1.2.3 Some examples

These examples are quintessential examples of symmetric spaces and will be used throughout.

1. Euclidean space $\mathbb{R}^n$ with the standard metric. Thus, $G = \mathbb{R}^n \times O(n)$. $K = O(n)$, and the symmetries are given by the reflections in perpendicular planes. The curvature is constant 0.

2. The sphere $S^n \subset \mathbb{R}^{n+1}$ with the metric induced by the embedding. $G^0 = SO(n+1)$ and $K = SO(n)$. The involution of G^0 corresponding to the symmetry at $N = (1, 0, \ldots, 0)$ is given by conjugation by the matrix $(-1)^{1+\delta_{1,j}}\delta_{ij}$.

3. The positive connected component of the hyperboloid in $\mathbb{R}^{n+1}$ given by the quadratic form
$$x_0^2 + \ldots + x_{n-1}^2 - x_n^2 = 1$$
with the metric induced by the embedding. The group $SO^0(n, 1)$, the connected component containing the identity of matrices with determinant 1 preserving the form, acts transitively. The stabiliser of the point $(1, 0, \ldots, 0)$ is $SO(n)$, and the involution is as above.

Definition 1.2.4. A locally symmetric space X is one that is locally isometric to a globally symmetric space.

Proposition 1.2.4. Let X be a complete locally symmetric space, and M its universal cover. Then M is a simply connected symmetric space with the pullback metric, and $X = \Gamma\backslash M$ for $\Gamma \subset \operatorname{Isom}(M)$.

Proof. For details, see [1, p187]. Essentially, we lift the metric to the universal cover. Then M is a simply connected Riemannian manifold with the pullback metric. The geodesic symmetries on X lift to global geodesic symmetries on M, and the metric on X is invariant under deck transformations, since it's the pullback of the one on M; hence $\pi_1(X)$ is realised as $\Gamma \subset \operatorname{Isom}(M)$.

Thus, a locally symmetric space X can be written as $\Gamma\backslash M$ for M a simply connected symmetric space. *Proposition 1.2.2* allows us the further expression for X

$$X = \Gamma\backslash G/K$$

1.2.4 Lie group theory of symmetric spaces

The classical textbook of Helgason [1, Chapters 1-5] provides much of the material for this section. For detailed proofs of propositions, the reader is advised to consult this. For details presented less densely, see [8]. We cover some basics of Lie theory in the context of symmetric spaces, with the aim being to understand ingredients of the classification, and to define the rank of a symmetric space. Restricting the rank of a symmetric space provides us with a key way to understand geometric flows on symmetric and locally symmetric spaces.

Throughout this section, let G be a Lie group and $\mathfrak{g}$ its Lie algebra.

We begin with some fundamental notions in Lie theory.

Definition 1.2.5. The Killing form on $\mathfrak{g}$ is the form $B(X,Y) = \operatorname{tr} \operatorname{ad}_X \operatorname{ad}_Y$.

Since the adjoint action of G is an action by automorphisms of the Lie algebra $\mathfrak{g}$, and by a standard fact about the Killing form [8] we have the following:

Proposition 1.2.5. The Killing form is a G-invariant inner product when G is semisimple and noncompact. If G is semisimple and compact, $-B$ is a G-invariant inner product.

Definition 1.2.6. Two Lie groups G, $\hat{G}$ are locally isomorphic if they have isomorphic Lie algebras.

Really, this implies that one is a covering group of the other. In the context of symmetric spaces, local isomorphism is the correct notion to look at.

Definition 1.2.7. An orthogonal symmetric Lie algebra is a pair $(\mathfrak{g}, s)$ where s is an involutive automorphism of $\mathfrak{g}$ and the Killing form of $\mathfrak{g}$ restricted to $\mathfrak{k}$, the fixed point set of s is negative definite. For G a Lie group and K a Lie subgroup, (G,K) is associated to $(\mathfrak{g}, s)$ if the Lie algebra of G is $\mathfrak{g}$ and the Lie algebra of K is $\mathfrak{k}$.

Proposition 1.2.6. Let $(\hat{G}, \hat{K})$ and (G,K) be a pair associated to an orthogonal symmetric Lie algebra $(\mathfrak{g}, s)$. Suppose K and $\hat{K}$ are connected, and $\hat{G}$ is simply connected. Then $(\hat{G}, \hat{K})$ is a Riemannian symmetric pair and G/K is a locally symmetric space for every G-invariant metric, and $\hat{G}/\hat{K}$ is the universal cover of G/K.

So, in order to understand simply connected symmetric spaces, it's enough to understand orthogonal symmetric Lie algebras.

Definition 1.2.8. A submanifold N of a Riemannian manifold (M, g) is totally geodesic if every geodesic on N is also a geodesic on M.

Definition 1.2.9. The rank of a symmetric space is the dimension of a maximal flat, totally geodesic submanifold

Recall that for a globally symmetric space $M = G/K$, with base point for the G-action o, and decomposition of $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, we have a natural identification of T_oM with $\mathfrak{p}$, with geodesics with tangent at o given by $X \in \mathfrak{p}$ corresponding to $\exp(tX)o$. The following is a computation, and details are in [1, p190].

Proposition 1.2.7. Let $\mathfrak{s} \subset \mathfrak{p}$ be a subspace. Then if

$$X, Y, Z \in \mathfrak{s} \Rightarrow [X,[Y,Z]] \in \mathfrak{s} \quad (*)$$

$\exp_o(\mathfrak{s})$ is a totally geodesic submanifold of M, where $\exp_o$ denote the exponential map for the metric on M at o. Conversely, if S is a totally geodesic submanifold with $o \in S$, $T_oS \subset \mathfrak{p}$ satisfies $(*)$. Moreover, $\exp_o(\mathfrak{s})$ is flat if and only if $\mathfrak{s}$ is abelian.

As a result, the rank of a symmetric space is the dimension of a maximal abelian sub algebra of $\mathfrak{p}$. After an explanation of the classification of symmetric spaces, we shall provide a geometric description for the rank.

1.2.5 Decompositions of Lie groups

We take a brief foray into decompositions of semisimple Lie groups, which are incredibly useful in their own right. Some of the tools are also important for the classification. Let G be a connected Lie group and $\mathfrak{g}$ its Lie algebra.

Definition 1.2.10. An involution θ of $\mathfrak{g}$ is a Cartan involution if

$$-B(X, \theta(Y))$$

is a positive definite bilinear form on $\mathfrak{g}$

Proposition 1.2.8. Any real semisimple Lie algebra has a Cartan involution, and any two Cartan involutions are conjugate

By construction, the Cartan involution is self adjoint with respect to the Killing form, and hence gives rise to an eigenspace decomposition of $\mathfrak{g}$

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$$

where $\mathfrak{k}$ is the $+1$ eigenspace and $\mathfrak{p}$ is the -1 eigenspace. Such a decomposition is known as a *Cartan decomposition.*

Proposition 1.2.9. Suppose G is non compact. Let K be the subgroup of G with Lie algebra $\mathfrak{k}$. Then K is maximally compact in G.

Let $\mathfrak{a}$ be a maximal abelian subalgebra of $\mathfrak{p}$. ad_X for $X \in \mathfrak{a}$ is self-adjoint with respect to the Killing form and hence diagonalisable. Since $\mathfrak{a}$ is abelian, the operators ad_X for $X \in \mathfrak{a}$ are simultaneously diagonalisable.

Definition 1.2.11. By a restricted root system relative to $\mathfrak{a}$, we mean a minimal collection Σ of $\alpha \in \mathfrak{a}^*$ such that

$$\mathfrak{g} = \sum_{\alpha \in \Sigma} \mathfrak{g}_\alpha$$

where

$$[X, Y] = \alpha(X)Y \text{ for all } Y \in \mathfrak{g}_\alpha$$

Now the sets α^0 given by $X \in \mathfrak{a}$ such that $\alpha(X) = 0$ form a finite collection of hyperplanes in $\mathfrak{a}$, and disconnect $\mathfrak{a}$ into connected components $\mathcal{C}_i$, the Weyl chambers.

Indeed there is a rich combinatorial structure to the restricted root system and the Weyl chambers, but it shall not be of significant relevance here.

Each Weyl chamber $\mathcal{C}_i$ determines a notion of positivity on Σ.

Definition 1.2.12. For a fixed Weyl chamber $\mathcal{C}_i$, we refer to a positive system of roots as

$$\Sigma^+ = \{\alpha \in \Sigma : \alpha(X) > 0 \text{ for all } X \in \mathcal{C}_i\}$$

Given a choice of positive roots, we denote the algebra given by the sum of the positive roots spaces by $\mathfrak{n}$. That is,

Definition 1.2.13.

$$\mathfrak{n} = \sum_{\alpha \in \Sigma^+} \mathfrak{g}_\alpha$$

Remark. By K, A and N, we mean the subgroups of G with Lie algebras $\mathfrak{k}$, $\mathfrak{a}$ and $\mathfrak{n}$ respectively. This notation, unless otherwise specified, is used in each chapter.

The following proposition will be used repeatedly in the report, and is fundamental to understanding the structure of semisimple Lie groups

Proposition 1.2.10 (Iwasawa decomposition)**.** There is a direct sum decomposition

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$$

As a consequence, the map

$$K \times A \times N \to G,\ (k, a, n) \mapsto kan$$

is a diffeomorphism.

The proof is standard and can be found in [He, p 451]

Definition 1.2.14. A is referred to as a maximal $\mathbb{R}$-split torus.

If G is compact and A has dim 1, then A is isomorphic to S^1. If G is non compact and A has dim 1, then G is isomorphic to $\mathbb{R}$.

Remark. All Cartan involutions are conjugate, and hence all such decompositions are conjugate.

Definition 1.2.15. The dimension of $\mathfrak{a}$ is the *rank* of G.

From *Proposition 1.2.7*, note that for a symmetric space G/K with G semisimple, its rank agrees with the rank of G as a Lie group.

1.2.6 Duality and Classification

There is an explicit classification for simply connected globally symmetric spaces, which boils down to the classification of complex semisimple Lie algebras. This will provide a framework for restricting our attention in the next section, where we aim to prove properties of geometric flows on rank 1 locally symmetric spaces. Understanding the classification (even without a full understanding of the proof) gives an insight into precisely where the results of the next section are applicable, as well as being a beautiful result in its own right.

Fix an orthogonal symmetric Lie algebra $(\mathfrak{g}, s)$ as defined above. s gives rise to a decomposition $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$. By $\mathfrak{g}^*$, we denote the subset $\mathfrak{k} + i\mathfrak{p}$ of $\mathfrak{g} \otimes \mathbb{C}$ with the inherited bracket. By s^*, we mean the map

$$s^* : T + iX \mapsto T - iX \text{ for } T \in \mathfrak{k},\ X \in \mathfrak{p}$$

Definition 1.2.16. $(\mathfrak{g}, s)$ is of the compact type if $\mathfrak{g}$ is compact and semisimple. If $\mathfrak{g}$ is noncompact, and $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is a Cartan decomposition, then $(\mathfrak{g}, s)$ is of noncompact type. If $\mathfrak{p}$ is an abelian ideal of $\mathfrak{g}$, then $(\mathfrak{g}, s)$ is of Euclidean type.

Proposition 1.2.11. Every orthogonal symmetric algebra with trivial centre is isomorphic to a direct sum of ideals of the compact type, non-compact type and the Euclidean type

We define the type of a Riemannian symmetric space as the type of the corresponding orthogonal symmetric Lie algebra.

As a consequence, we have the following on the level of Riemannian symmetric spaces.

Theorem 1.2.2. *Let M be a simply connected Riemannian symmetric space.*

$$M = M_0 \times M_- \times M_+$$

where M_0 is of the Euclidean type, M_- is of the non-compact type and M_+ is of the compact type

A characterisation of the types of Riemannian symmetric spaces is given by sectional curvature:

Theorem 1.2.3. *A Riemannian symmetric space of non-compact type has nonpositive sectional curvature. One of compact type has nonnegative sectional curvature. Spaces of Euclidean type are flat.*

The following is the key statement in the duality theorem

Theorem 1.2.4. *Let $(\mathfrak{g}, s)$ be orthogonal symmetric. Then so is $(\mathfrak{g}^*, s^*)$. If $(\mathfrak{g}, s)$ is of the non-compact type, then $(\mathfrak{g}^*, s^*)$ is of the compact type (and vice versa). Moreover, if $(\mathfrak{g}_1, s_1)$ is isomorphic to $(\mathfrak{g}_2, s_2)$, then $(\mathfrak{g}_1^*, s_1^*)$ is isomorphic to $(\mathfrak{g}_2^*, s_2^*)$*

Symmetric spaces of the Euclidean type are those locally isometric to Euclidean spaces - and are largely uninteresting in terms of their ergodic properties. The theorem above means that to classify simply connected spaces, it's enough to classify those of the compact type. Note that simply connected compact Lie groups are canonically Riemannian symmetric spaces with G-invariant metric given by translates of the Killing form followed by integration over G, and symmetry at the identity given by $g \mapsto g^{-1}$.

Proposition 1.2.12. Every orthogonal symmetric Lie algebra of the compact type corresponds to a compact semisimple Lie group

Compact semisimple Lie groups have been completely classified. Thus, by duality, so have Riemannian symmetric spaces.

Finally, we include a geometric interpretation of the rank, as described in *Section 1.2.4*.

Proposition 1.2.13. Let M be a symmetric space of the non compact type or the compact type. Let G denote $\mathrm{Isom}^0(M)$. Let A and A' be maximal flat totally geodesic sub-manifolds of M. Then A and A' are closed (as subsets). Suppose $q \in A$ and $q' \in A'$. There is an isometry $h \in G$ with $h(A) = A'$ and $h(q) = q'$. Finally, for $X \in T_qM$, there is an isometry g such that $dg_x(X) \in T_qA$.

1.2.7 Examples, continued

Consider $\mathbb{H}^n$ and S^n as described in previous sections. These are simply connected Riemannian symmetric spaces, corresponding to $SO(n,1)$ and $SO(n+1)$ respectively. The corresponding Lie algebras are given by

$$\mathfrak{so}(n,1) = \begin{pmatrix} 0 & y \\ y^t & X \end{pmatrix}$$

where X is an $n \times n$ skew symmetric matrix and $y \in \mathbb{R}$ and $\mathfrak{so}(n+1)$ is the skew symmetric matrices. Recall that the involution was given by conjugation by the matrix $A_{ij} = (-1)^{1+\delta_{1j}}\delta_{ij}$. This gives rise to the decomposition of $\mathfrak{so}(n+1)$ into

$$\mathfrak{k} = \{\begin{pmatrix} 0 & 0 \\ 0 & X \end{pmatrix} : X \text{ a skew symmetric } n \times n \text{ matrix}\}$$

and

$$\mathfrak{p} = \{\begin{pmatrix} 0 & Y \\ -Y^t & 0 \end{pmatrix} : Y \text{ a } 1 \times n \text{ matrix}\}$$

Remark. Considering the case $Y_1 = \begin{pmatrix} 1 & \dots & 0 \end{pmatrix}$ and $Y_2 = \begin{pmatrix} 0 & 1 & \dots & 0 \end{pmatrix}$ shows that any maximal abelian subspace has dimension 1. So $SO(n,1)$ has rank 1.

Now the map

$$\begin{pmatrix} 0 & iY \\ -iY^t & X \end{pmatrix} \mapsto \begin{pmatrix} 0 & Y \\ Y^t & X \end{pmatrix}$$

is clearly an isomorphism of $\mathfrak{so}(n+1)$ onto $\mathfrak{so}(n,1)$, and the conjugation by A gives the involution

Hence we have exhibited the duality between $\mathbb{H}^n$ and S^n.

It's clear even geometrically that S^n (and therefore $\mathbb{H}^n$) have rank 1: for $x \in A$, where A is any totally geodesic submanifold, x has a neighbourhood embedded isometrically in some sphere. So the rank is 1. This argument shows in fact that if M has negative or positive sectional curvatures, then the rank is 1.

We now finish off the Iwasawa decomposition of $SO(n,1)$. We have that $K = SO(n) \subset SO(n,1)$, and $A = \exp tX$ for $X \in \mathfrak{p}$ given by $Y = (1,\dots,0)^t$ under the isomorphism. Observe that the collection of matrices of the form

$$\begin{pmatrix} 0 & z & -z \\ -z^t & 0 & 0 \\ z^t & 0 & 0 \end{pmatrix}$$

for $z \in \mathbb{R}^{n-1}$ are ad_X eigenvectors of eigenvalue $+2$, for X as shown with $Y = Y_1$. Indeed, this is an abelian subalgebra of the correct dimension, and we denote the corresponding subgroup by N. Thus the Iwasawa decomposition is

$$SO(n,1) = SO(n) \times A \times N$$

1.3 Geometric flows

With this combination of algebraic properties and geometric ones, we shall seek to study the dynamical properties of geometric flows on locally symmetric spaces, which can be written as $\Gamma\backslash G/K$. Of most interest to us is the frame flow and the geodesic flow. This section will begin by outlining these flows in generality, and describing the measures on the spaces which they act. Finally, we seek to interpret these flows in terms of subgroups of the group G defining our locally symmetric spaces.

1.3.1 Preliminaries

We first recall some basic definitions and results about geodesics. Let (M, g) be a Riemannian manifold of dimension n. We make the additional assumption that M is geodesically complete. Note that homogeneous spaces or closed manifolds are all geodesically complete, so the definitions given will apply for them.

Definition 1.3.1. The unit tangent bundle

$$T^1M = \bigcup_x \{v \in T_xM \mid ||v||_g = 1\}$$

Since M is geodesically complete, we are able to make the following definition

Definition 1.3.2. Let $v \in T_xM$ and let $\gamma_v : (-\infty, \infty) \to M$ be the unique geodesic which has $\gamma(0) = x$ and $\dot{\gamma}(0) = v$. The action of geodesic flow $(g_t)_{t\in\mathbb{R}}$ on v is given

$$g_t(v) = \dot{\gamma}_v(t)$$

That is, the geodesic flow at time t gives the parallel transport of a tangent vector v at time t along the unique geodesic having v as a tangent at time 0.

Since geodesics are parallel, the geodesic flow acts on T^1M. We consider a similar space, the frame bundle $\mathcal{F}(M)$ and induce a similar action.

Definition 1.3.3. The frame bundle $\mathcal{F}(M)$ is the fibre bundle, with fibre F_x at a point $x \in M$

$$F_x = \{ \text{ all ordered bases of } T_xM \ \}$$

We induce the topology from the natural $GL(TM)$ action, and this action makes $\mathcal{F}(M)$ into a principal $GL(n, \mathbb{R})$ bundle.

1.3.2 The Liouville measure

The goal of this section is to construct natural measures on T^1M and $\mathcal{F}(M)$ with respect to which the geodesic flow is measure preserving.

Recall that on any oriented Riemannian manifold, we have the Riemannian measure given on Borel subsets by integration of the volume form. We denote this measure by $d\mu_g$.

For S_x the fibre of T^1M at x, let ν_x denote the Riemannian measure on the fibre $S_x \subset (\mathbb{R}^n, g_x)$ induced by the restriction of the metric.

Definition 1.3.4. The Liouville measure Λ on T^1M is given locally by the product measure

$$d\Lambda = d\mu_g d\nu_x$$

Proposition 1.3.1. The Liouville measure is invariant under the geodesic flow

Proof. The metric gives a pairing $TM \cong T^*M$. The geodesic flow on the cotangent bundle is a Hamiltonian flow. Since the canonical symplectic form is closed, by Cartan's magic formula, we have that Hamiltonian flows are flows by symplectomorphisms. In particular, they are volume preserving.

Locally, a frame $A \in \mathcal{F}(M)$ is denoted by

$$A = (p, u_1, u_2, \dots u_n)$$

where $u_i \in TM$. We have a similar definition for the Liouville measure on $\mathcal{F}(M)$ is given locally by

Definition 1.3.5. The Liouville measure Λ on $\mathcal{F}(M)$ is defined locally by the product

$$d\Lambda = d\mu_g d\nu_x d\nu_x \dots d\nu_x$$

where ν_x is as defined above.

We define the frame flow $(h_t)_{t\in\mathbb{R}}$ on a geodesically complete manifold by

Definition 1.3.6. The action of the frame flow h_t is given by the relation locally by

$$h_t(p, u_1, \dots, u_n) = (g_t(p, u_1), v_2, \dots, v_n)$$

where v_i is the parallel transport of u_i along the geodesic at p determined by u_1.

It is clear that the frame flow preserves the Liouville measure Λ on $\mathcal{F}(M)$, since the inner product at $g_t(p)$ is given by the parallel transport of the inner product at p, and thus $v_i d\nu_y = u_i d\nu_x$

1.4 Ergodic Theory

We summarise some basic definitions and results from ergodic theory. Throughout this section, (X, Ω, μ) will denote a measure space and $L^p(X)$ will denote $L^p(X, \mu)$ unless otherwise specified.

Definition 1.4.1. For $T : X \to X$ measurable, T is measure preserving if $\mu(T^{-1}B) = \mu(B)$ for all measurable sets B.

Definition 1.4.2. A measure preserving transformation T is ergodic if $T^{-1}(A) = A \cup N$ for N a null set implies that

$$\mu(A) = 0 \text{ or } \mu(X \setminus A) = 0$$

We now generalise to the actions of a second countable locally compact group G on a compact Hausdorff topological space X with a Borel probability measure μ.

Definition 1.4.3. An action of G by homeomorphisms is continuous if the induced map $G \times X \to X$ is. μ is invariant if $g_*\mu = \mu$ for all $g \in G$

Remark. Although it is not clear that in general such a measure μ exists, we will study mainly homogeneous spaces which possess induced invariant measures.

Definition 1.4.4. The action of G is ergodic if any measurable set such that $g^{-1}A = A \cup N$ where N is null for all $g \in G$ implies that

$$\mu(A) = 0 \text{ or } \mu(X \setminus A) = 0$$

Proposition 1.4.1. An action of G is ergodic with respect to an invariant measure μ if and only if the only μ measurable complex valued functions that are G-invariant are constant almost everywhere

Definition 1.4.5. A sequence $g_n \in G$ tends to ∞, and we write $g_n \to \infty$ if any compact subset of G contains only finitely many elements of G.

The following definition is fundamental to the examples we study.

Definition 1.4.6. The G action is mixing if

$$\mu(A \cap g_n^{-1} B) \to \mu(A)\mu(B)$$

as $n \to \infty$ for all sequences $g_n \to \infty$ as $n \to \infty$.

Remark. Mixing clearly implies ergodicity: for an invariant set A, the definition implies that $\mu(A) = \mu(A)^2$

Example 1.4.1 (Geometric flows)**.** Let (M, g) be a closed Riemannian manifold, T^1M its unit tangent bundle (equipped with the Liouville measure Λ) and $G = (g_t)_{t\in\mathbb{R}}$ the geodesic flow. Then Λ is G-invariant. The action being mixing says that for any set A, the *proportion* of a set B that is in A at time t of the flow is approximately constant and equal to the measure of A. This gives a very clear visual depiction of what it means for a group action to be mixing.

If a group G acts on X and μ is an invariant measure for the action, then we have an induced unitary representation on $L^2(X, \mu)$ via

$$f \mapsto f \circ g$$

Thus the statement of ergodicity is equivalent to saying that the only nontrivial subrepresentation is the trivial representation. Mixing is rewritten as

$$(\psi \circ g_n, \phi) = \int_X \psi \circ g_n, \overline{\phi} d\mu \to \left(\int_X \psi d\mu\right) \left(\overline{\int_X \phi d\mu}\right)$$

as $n \to \infty$ for all sequences $g_n \to \infty$ as $n \to \infty$, since simple functions are dense in $L^2(X, \mu)$.

Section 2

Rank 1 locally symmetric spaces

2.1 Basic properties of rank 1 locally symmetric spaces

We include some preliminary facts to make precise which spaces our results apply to.

A rank 1 locally symmetric space of the non-compact type is $X = \Gamma\backslash G/K$ where

(i) $\tilde{X} = G/K$ is the universal cover

(ii) G is a connected semisimple Lie group (the identity component of the isometry group of M)

(iii) K is a maximal compact subgroup

(iv) The dimension of a maximal torus in G is 1, or equivalently, the maximum totally geodesic submanifold of $\tilde{M}$ has dimension one.

Proposition 2.1.1. Let M be a rank 1 Riemannian symmetric space. Then M is of the Euclidean type, non-compact type or compact type

Proof. If not, then it can be written as a product, by *Theorem 1.2.2*. Clearly, the product of two maximal flat totally geodesic sub-manifolds will be a maximal flat totally geodesic sub-manifold of dimension greater than 1

The cases $M = S^1$ and $M = \mathbb{R}$ are uninteresting. The following explains why we shall only consider symmetric spaces of the noncompact type.

Proposition 2.1.2. Let M be a compact symmetric space of the compact type. Then the geodesic flow on M is periodic.

Proof. Write $M = G/K$. A maximal torus in G is homeomorphic to S^1. By *Proposition 1.2.13*, all geodesics are isometric. Thus all are closed and of the same length.

Theorem 2.1.1. *Let M be a Riemannian symmetric space of rank 1. Then M is a two point homogeneous space. That is, for any $p, q, p', q' \in M$ with $d(p,q) = d(p',q')$, there is an isometry f with $f(p) = p'$ and $f(q) = q'$.*

This follows from *Proposition 1.2.13*, (see [1, 211, 355] for details), along with the observation that rank 1 symmetric spaces of the Euclidean type are $\mathbb{R}$ and S^1.

We use the following.

Proposition 2.1.3. Let (M, g) be a Riemannian manifold that is a two point homogeneous space. Then $G = \mathrm{Isom}^0(M, g)$ acts transitively on the unit tangent bundle.

Using this, we get a a convenient description for the unit tangent bundle of a symmetric space

Proposition 2.1.4. Let $X = \Gamma\backslash G/K$ be a locally symmetric space of rank 1. Then

$$T^1X = \Gamma\backslash G/M$$

where M is the centraliser K of a maximal split torus in as in the Iwasawa decomposition.

Proof. Let $\tilde{X} = G/K$ the universal cover of X, and let π denote the projection. Then

$$\pi^*T^1X = T^1G/K$$

Now G acts transitively on the unit tangent bundle. Hence $T^1G/K = G$. Furthermore, the stabiliser of $v \in T_oG/K$ corresponds to a vector in $\mathfrak{p}$ where $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. Since we're in rank 1, the stabiliser of a vector X is the stabiliser of a maximal split torus under isometries - which corresponds to is the stabiliser in K of $X \in \mathfrak{p}$ under the Ad action, and thus corresponds to M.

Let $\gamma \in \pi_1(T^1X, x)$ be a loop. Then γ corresponds to an isometric deck transformation $g_\gamma \in$ of G/K. Now each fibre is simply connected, so the induced dg_γ corresponds to the appropriate deck transformation for the loop in T^1M.

2.2 Classification of rank 1 locally symmetric spaces

The classification of simple Lie groups, gives rise to a classification of simply connected rank 1, irreducible symmetric spaces of non-compact type (see *Section 1.2.6* for details). They are as follows. $\tilde{X} = G/K$ with K maximal compact in G and

- $G = SO(n, 1)$, the group of real matrices of determinant 1 fixing the form

$$f(x) = x_1^2 + \ldots + x_n^2 - x_{n+1}^2$$

  on $\mathbb{R}^n$. Locally symmetric spaces for this G are hyperbolic manifolds

- $G = SU(n, 1)$, the group of complex matrices of determinant 1 fixing the Hermitian form

$$f(z) = z_1\bar{z}_1 + \ldots z_n\bar{z}_n - z_{n+1}\bar{z}_{n+1}$$

  Locally symmetric spaces of this form are complex hyperbolic manifold

- $G = Sp(n, 1)$. The matrices satisfying, for $J_{n+1} = \begin{pmatrix} 0 & I_{n+1} \\ -I_{n+1} & 0 \end{pmatrix}$ and $K = \begin{pmatrix} -I_n & & & \\ 0 & I & & \\ 0 & 0 & -I_n & \\ 0 & 0 & 0 & I \end{pmatrix}$,

$$\psi J_{n+1}\psi^{-1} = J_{n+1} \text{ and } \psi^t K\psi = K$$

  Locally symmetric spaces corresponding to these are the quaternionic hyperbolic manifolds.

- F_4^{-20}. Corresponds to 'octonionic hyperbolic space'. This does not occur in any infinite family and does not feature extensively in the literature.

2.3 Outline of results

This section is devoted to understanding the proof and use of the following theorem. The proof follows arguments in [7].

Theorem 2.3.1. *Let X be a closed, rank 1 locally symmetric space. Then the geodesic flow g_t is exponentially mixing on T^1X, with respect to the Liouville measure. That is, for smooth functions $\psi, \phi \in C^\infty(T^1X)$, we have the following*

$$\left| \int_{T^1X} (\psi \circ g_t)\overline{\phi} d\Lambda - \left(\int \psi d\Lambda\right) \overline{\left(\int \phi d\Lambda\right)} \right| \leq K(\psi, \phi)(e^{-q(X)t})$$

We shall in the process also prove the following:

Theorem 2.3.2. *Let M be a closed hyperbolic manifold. Then the frame flow is exponentially mixing on $\mathcal{F}(X)$.*

The idea of the proof is as follows. We write $X = \Gamma \backslash G/K$. G/K is a Riemannian symmetric space, and the geodesic flow on X corresponds to the image under the projection of the geodesic flow on $\tilde{X} = G/K$. The geodesic flow is represented by a 1-parameter group generated by $Y \subset \mathfrak{p} \cong T_o\tilde{X}$. Thus the stabiliser of $Y \in T_o^1 X$ corresponds to the centraliser in K of the 1-parameter subgroup generated by Y. Thus we have the decomposition for $T^1\tilde{X} = G/M$, and since $\tilde{X}$ and X are locally isometric, with X closed, we have the decomposition

$$T^1X = \Gamma \backslash G/M$$

The geodesic flow is thus given by $\Gamma gM \mapsto \Gamma g g_t M$. Since the Liouville measure is invariant under isometries it corresponds to the Haar measure on the coset space. Since M is compact, we can normalise so that

$$\int_{\Gamma \backslash G/M} \int_M \psi(\Gamma g m) dm dg = \int_{\Gamma \backslash G} \psi(g) dg$$

Thus, we study the action of $A = (g_t)_{t \in \mathbb{R}}$ by studying $L^2(\Gamma \backslash G)$ which is naturally a unitary representation of G. In the case where K is actually transitive on frames, as in the case for hyperbolic manifolds (where the K is $SO(n)$), we see that $\Gamma \backslash G$ is in fact the frame bundle $\mathcal{F}(X)$. We show the following, by studying the representation theory G on $L^2(\Gamma \backslash G)$

$$\int_{\Gamma \backslash G} \psi \circ g_t \overline{\phi} dg \leq K(\psi, \phi) e^{-q(X)|t|}$$

2.4 Proof of the main theorem

We first recall *Proposition 2.0.5*: $T^1X = \Gamma \backslash G/M$. Note that in the Iwasawa decomposition, we can choose A to be the 1-parameter subgroup generated by the geodesic flow, and M the corresponding centraliser of this. Hence, we have that the geodesic flow acts on $T^1X = \Gamma \backslash G/M$ via

$$g_t(\Gamma g M) = \Gamma g g_t M$$

Example.

Consider a closed, oriented, hyperbolic 3 manifold $\mathbf{M}$. Then $\mathbf{M} = \Gamma \backslash \mathbb{P}SL(2, \mathbb{C})/SO(3)$. We have that A is the group generated by

$$\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$$

for $t \in \mathbb{R}$, and $M = \begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix}$. So

$$T^1\mathbf{M} = \Gamma \backslash \mathbb{P}SL(2, \mathbb{C})/M$$

Work in the universal cover $\mathbb{H}^3$. WLOG, our base point is the origin in the ball model. Fixing a unit tangent vector, it's clear that the isometries fixing it correspond to rotations of the boundary. So choosing a particular basis for $\mathbb{R}^3$, the stabiliser of $M \cong S^1$ is as shown. Furthermore, geodesics through the origin are straight lines, parametrized by A as shown. Thus the same holds in the projection to $T^1\mathbf{M}$.

2.4.1 Basic representation theory

Let G be any semisimple Lie group. We cover basic concepts in representation theory in the context of proving the theorem.

Definition 2.4.1. Let H be a complex Hilbert space and $\mathcal{B}(H)$ denote bounded linear operators on H. A representation of G is a homomorphism

$$\pi : G \to \mathcal{B}(H)$$

such that for all $v \in H$, the map

$$\pi_v : G \to H,\ g \mapsto \pi(g)v$$

is continuous.

A representation is unitary if $\pi(g)$ is unitary for all $g \in G$.

Definition 2.4.2. Let π be a representation of G on a Hilbert space H. Then the matrix coefficients are:

$$\pi_{u,v}(g) = \langle \pi(g)u, v \rangle$$

Notice that for fixed u, v, the matrix coefficients $\pi_{u,v} \in C(G)$, the space of continuous functions $f : G \to \mathbb{C}$.

Proposition 2.4.1. The action of G on $\Gamma\backslash G$ induces a unitary representation of G on $L^2(\Gamma\backslash G)$, via

$$\rho(g) : f(\Gamma h) \mapsto f(\Gamma h g^{-1})$$

This is clear since G acts smoothly and preserves the measure.

We write $L^2(\Gamma\backslash G) \ominus 1$ for functions orthogonal to the constants, and we will refer to this representation by (ρ, H).

Remark. Let $A = (g_t)$ denote the geodesic flow. Our theorem, in the language of matrix coefficients, is

$$|\rho_{\psi,\phi}(g_t)| < K(\psi, \phi)e^{-q(X)t}$$

Remark. Where it is clear, we will simply write gv for the action of g on a representation vector $v \in V$.

We now make the following definition, which shall be made use of later

Definition 2.4.3. A unitary representation π is *tempered* if for all $u, v \in H$ and $\epsilon > 0$ we have that

$$\pi_{u,v} \in L^{2+\epsilon}(G)$$

and the following concept is fundamental

Definition 2.4.4. A representation (π, H) of G is irreducible if the only proper, closed $\pi(G)$-invariant subspaces of H are $\{0\}$ and H.

In the representation theory of semisimple groups, an important role is played by the parabolic subgroups. These are constructed as follows.

2.4.2 Parabolic subgroups

Let σ be a Cartan involution of $\mathfrak{g}$. Then

$$\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$$

Let $\mathfrak{a}$ be a maximal abelian subspace of $\mathfrak{p}$. We're restricting to rank 1 - but these are all self adjoint with respect to Killing form. So we can simultaneously diagonalise. Choose a root system for the eigenspaces. Pick a Weyl chamber, and a selection of positive root spaces corresponding to it. Then

Definition 2.4.5. A Borel subgroup of G is the group B with Lie algebra $\mathfrak{b}$ given by

$$\mathfrak{b} = \mathfrak{m} + \mathfrak{a} + \mathfrak{n}$$

where $\mathfrak{m} \subset \mathfrak{k}$ centralises $\mathfrak{a}$.

A parabolic subgroup P of G is one that contains a Borel subgroup.

From the above, we can see that any parabolic subgroup can be written is $P = QAN$

Remark. Since $\mathfrak{k} + \mathfrak{b} = \mathfrak{g}$, we have that

$$G = KB = KMAN$$

2.4.3 Induced representations and series

In the case of discrete groups, for a subgroup H of G and a unitary representation (V, σ) of H, we set

$$\operatorname{Ind}_H^G V = \{f : G \to V \text{ (continuous)} \mid f(xh) = h^{-1}f(x)\}$$

Note that this definition means that f is constant on cosets xH, with the G action given by

$$gf(x) = f(g^{-1}x)$$

But in our more general setting, where G is a semisimple Lie group, we want $\operatorname{Ind}_H^G V$ to be a Hilbert space. A natural inner product to suggest would be to define (on continuous functions of compact support over H)

$$(\psi, \phi) = \int_{G/H} (\psi, \phi)_V d(xH)$$

which makes sense since V is a unitary representation. Now the induced representation is only unitary if the measure $d(xH)$ is G-invariant. In general, from the preliminaries, such a measure will only exist if H is unimodular (since G is). So we need a normalisation factor that ensures G acts unitarily and a restriction on the representation space that enables this.

Restrict to the case where $P = MAN$ is a parabolic subgroup.

First set

$$\rho = \frac{1}{2} \sum_{\lambda \in \Sigma^+} m_\lambda \lambda$$

where Σ^+ is a choice of positive roots and m_λ is their multiplicity.

Since A is abelian, all its complex representations are 1 dimensional. For $\nu \in \mathfrak{a}_{\mathbb{C}}^*$, we have that e^ν is a representation of A. Hence, given a finite dimensional representation σ of M, we have the representation

$$V_{\sigma,\nu} = \sigma \otimes e^\nu \otimes \mathbf{1}$$

Definition 2.4.6. Let G be a semisimple Lie group, K its maximal compact subgroup. Let $P = MAN$ be a parabolic subgroup such that $G = KMAN$, and $V_{\sigma,\nu}$ the representation as above. Let

$$C_{\rho,\nu,\sigma} = \left\{ F \in C(G,V) \,|\, F(xman) = e^{-(\rho+\nu)\log a}\sigma(m^{-1})F(x) \right\}$$

And let

$$(F,G) = \int_K (F(k),G(k))\,dk$$

Let G act linearly on this space by

$$gF(x) = F(g^{-1}x)$$

We define the induced representation $(V_{P,\sigma,\nu}, U_{P,\sigma,\nu})$ as the completion of this space with respect to the inner product given, and the induced operators of G.

We label these parabolically induced representations according to σ and ν:

Definition 2.4.7. We say

(a) $U_{P,\sigma,\nu}$ is an *principal series representation* if P is a Borel subgroup

(b) If ν is imaginary, $U_{P,\sigma,\nu}$ is a unitary principal series representation

(c) If ν is imaginary, $\sigma = \mathbf{1}$, and $|\nu| < 1$, then we have a spherical series representation

Spherical complementary series will be of use to us later.

Definition 2.4.8. Let (π, V) be a complex representation of G on a Hilbert space V. For K a subgroup of G and $v \in V$, we say that v is K-finite if

$$\dim\langle \pi(k)v | k \in K\rangle < \infty$$

We can now define

Definition 2.4.9. The spherical complementary series is the restriction of a spherical series representation to the completion of its space of K-finite vectors

If a representation has a K finite vector, we can find a K-fixed vector by integration over K (since K is compact).

2.4.4 Representations induced from the Borel subgroup

Let us return to the problem at hand, where G is a semisimple Lie group of rank 1. That is, $G = KAN = KMAN$ where $A = (a_t)_{t\in\mathbb{R}} = \exp(tX)$ choosing a basis of $\mathfrak{a}$. We have seen that the action of A corresponds to the geodesic flow.

Fix an induced representation $(W,U) = (W_{\sigma,\nu}, U_{\sigma,\nu})$. The space W is the completion of a subspace

$$C_{\sigma,\lambda} \subset C(G, V_\sigma)$$

possessing a the invariance

$$f(xman) = e^{-(\rho+\nu)\log a}\sigma(m^{-1})f(x)$$

and we use the norm

$$\int_K ||f(v)||_\sigma^2 dk$$

We now wish to understand the matrix coefficients

$$\rho_{f,g}(a_t)$$

However, the action of A is rather difficult to isolate from the representation given. Our strategy will be to break up the action and find a unitarily equivalent representation where the A action makes more sense.

Since $G = KB$, we have the following

Proposition 2.4.2. A function $f \in W$ is uniquely determined by its values on K

Hence W can be viewed as a subspace of $L^2(K, V_\sigma)$.

The A action is given by

$$a.f(x) = f(a^{-1}x) = f(x'm'a'n') = e^{-\nu \log a'}\sigma(m)f(x')$$

We will denote the spatial $x \mapsto x'$ by S, so

$$a : f \mapsto \psi(a)f(S(x))$$

We require a factor $\psi(a)$, which has non-unit modulus. This is because the A action isn't measure preserving. If we had $K \cong \Omega$ and $dk = f(\omega)d\omega$, but with the measure $d\omega$ invariant under the induced A-action, and we made a unitary transformation $L^2(K, V_\sigma) \to L^2(\Omega, V_\sigma)$ with the representation of G defined by to be the equivalent one intertwined by the transformation, we would have

$$af(\omega) = \psi'(a)f(a\omega)$$

but with $|\psi'| = 1$. Our goal is to find such an Ω, where the A-action is also understandable.

2.4.5 Important example: hyperbolic 3 manifolds

Let $G = \mathbb{P}SL(2, C)$, which is locally isomorphic to $SL(2, \mathbb{C})$, so has the same algebra $\mathfrak{sl}(2, \mathbb{C})$. The notation will be as in the basic concepts section.

- $\mathfrak{sl}$ are the traceless matrices
- $X \mapsto -X^*$ is a Cartan involution
- $\mathfrak{p}$ is the space of matrices with $X^* = X$, so Hermitian, traceless matrices
- A maximal abelian real subalgebra of $\mathfrak{p}$ is given by $X = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
- Let $Y = \begin{pmatrix} 0 & z \\ 0 & 0 \end{pmatrix}$. This is a 2 dimensional subspace of $\mathfrak{g}$ and $XY - YX = 2Y$. So we have found $\mathfrak{n}$.
- Putting this together, this gives that for $SL(2, \mathbb{C})$

$$K = SU(2),\ A = \left\{ \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} : t \in \mathbb{R} \right\},\ N = \left\{ \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} : z \in \mathbb{C} \right\}$$

  and $M = S^1 \subset SU(2)$. So for $\mathbb{P}SL(2, \mathbb{C})$ we have the induced maps

We're considering the representation $L^2(K, W_\sigma)$. Let's study the $A = (a_t)$ action on this via

$$a_t \circ f(k) = f(a_{-t}k)$$

Now $a_{-t}k = k_t m' a' n'$, so by the Iwasawa decomposition we have that

$$a_t F(k) = F(k_t m' a' n') = e^{-(2+\nu)t}\sigma(m)F(k')$$

We first analyse the action of $a_t : k \mapsto k_t$. For simplicity, we work in $SL(2, \mathbb{C})$. Then $K = SU(2)$ and $M = S^1$.

- $$SU(2) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} : |a|^2 + |b|^2 = 1 \right\} \text{ and } M = \left\{ \begin{pmatrix} a & 0 \\ 0 & \bar{a} \end{pmatrix} \right\}$$
- Multiplication by $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ gives a map from one copy of S^1 ($\bar{b} = 0$) to the other.

- Let's compute $a_t k = k_t a_s n'$:

$$a_t \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} = \begin{pmatrix} u_t & v_t \\ -\bar{v}_t & \bar{u}_t \end{pmatrix} \begin{pmatrix} e^s & ze^s \\ 0 & e^{-s} \end{pmatrix}$$

  Giving

$$\begin{pmatrix} e^t a & e^t b \\ -e^{-t}\bar{b} & e^{-t}\bar{a} \end{pmatrix} = \begin{pmatrix} e^s u_t & ze^s u_t + e^{-s} v_t \\ -e^s \bar{v}_t & -ze^s \bar{v}_t + e^s \bar{u}_t \end{pmatrix}$$

- Since $|u_t| \leq 1$, $s \to \pm\infty$ as $t \to \pm\infty$. The only fixed points are those represented by

$$e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } e_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

- Thus $a_t k \to e_1$ as $t \to \infty$ and $a_t k \to e_2$ as $t \to -\infty$.
- We write $K/M = A \times L$, a cross section of the orbits where $a_t(s,l) = (t+s,l)$. Thus

$$dk = f(t,l)dtdl$$

- We want to understand the asymptotic behaviour of $f(t,l)$ as $|t| \to \infty$
- Let $P = SU(2) \cap \{a \neq 0\}$, and $\bar{P} = SU(2) \cap \{b \neq 0\}$. Then $e_1 \in P$ and $e_2 \in \bar{P}$, and P and $\bar{P}$ have full measure in K/M So

$$K/M = P/M \cup \bar{P}/M = \bar{N}/MAN \cup N/MAN$$

  where $\bar{N} = \left\{ \begin{pmatrix} 1 & 0 \\ z & 1 \end{pmatrix} : z \in \mathbb{C} \right\}$
- Thus, we look at N to study the asymptotic behaviour as $t \to -\infty$. This is because $N \cong \mathfrak{n}$ (via the exponential map), and we understand the action a_t on $\mathfrak{n}$
- $N/MAN = \bar{P}/M \equiv \mathbb{C} = \exp \begin{pmatrix} 0 & z \\ 0 & 0 \end{pmatrix}$. Since $aNa^{-1} = aN$ modulo MAN, with a slight abuse of notation,

$$a_t \exp(z) = \exp(\mathrm{Ad}_{a_t} z)) = \exp(2tz)$$

- As $t \to -\infty$, $a_t n \to 0$ for $n \in \mathfrak{n}$.
- Let $\mathbf{x} = 1$ and $\mathbf{y} = i$ under the identification $\mathbb{C} \cong \mathbb{R}^2$.
- Now $a_t \mathbf{x} = e^{2t}\mathbf{x}$ and $a_t \mathbf{y} = e^{2t}\mathbf{y}$. We have coordinates therefore on N given by t, θ, $\mathbf{n} = e^{2t}\cos\theta \mathbf{x} + e^{2t}\sin\theta \mathbf{y}$.
- The Jacobian for the change of coordinates for x_i, y_i the euclidean ones is given by

$$J = \begin{pmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} 2e^{2t}\cos\theta & -e^{2t}\sin\theta \\ 2e^{2t}\sin\theta & e^{2t}\cos\theta \end{pmatrix}$$

- Hence $dn = dxdy = 2e^{4t}dtd\theta$ for $t < 0$.
- Now suppose $dk = \psi(n)dn$. It's clear that $\psi(n) \to 0$ as $n \to \infty$.
- So $dk = 2\psi(n)e^{4t}dtd\theta$. So as $t \to -\infty$, $f(t,l) \leq ce^{4t}$.
- We can repeat the computation for $\bar{N}$. The subset of K/M corresponding to the circle $|z| = 1$ in $\mathfrak{n}$ transforms to another compact set in $\bar{\mathfrak{n}}$, so the constant c we retrieve in the estimate for f is different. Nonetheless, we still recover the asymptotic estimate $f(t,l) \leq Ke^{-4|t|}$. This estimate is the heart of the proof of exponential mixing.

We now aim to generalise the analysis to any rank 1 semisimple Lie group.

2.4.6 The Weyl group and Bruhat cells

Recall the Cartan decomposition

$$\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$$

Now K acts on $\mathfrak{p}$ with the Ad action.

Proposition 2.4.3. Let $x \in \mathfrak{p}$ and $\mathfrak{a}$ be maximal abelian in $\mathfrak{p}$. Then the $\mathrm{Ad}(K)$ orbit of x intersects $\mathfrak{a}$ orthogonally.

In our case, each 1-d subspace is maximal abelian, and so $\mathrm{Ad}(K)$ orbits are spheres or odd dimensional projective spaces. We define the Weyl group of G by

$$W(G) = \mathrm{Stab}_K(\mathfrak{a}) / \mathrm{Stab}_K(x)$$

Since K acts isometrically, clearly, in our case $W(G) = \mathbb{Z}_2$ generated by the reflection perpendicular to x. It is easy to see that this is well defined in rank 1, but it's well defined in general since all such $\mathfrak{a}$ are conjugate. The Weyl group is a transformation group on the Weyl chambers.

Proposition 2.4.4 (Bruhat decomposition)**.** Let G be a rank 1 semisimple Lie group of noncompact type. Then

$$G = B \cup BwB$$

where $w \in K$ represents the non-identity element of $W(G)$

Now recall (from the Iwasawa decomposition) that

$$\mathfrak{g} = \sum_{\lambda \in \Sigma^+} \mathfrak{g}_\lambda + \mathfrak{g}_0 + \sum_{-\lambda \in \Sigma^+}$$

Let $\bar{\mathfrak{n}} = \sum_{-\lambda \in \Sigma^+} \mathfrak{g}_\lambda$ and $\bar{N} = \exp \mathfrak{n}$. Clearly, $wNw = \bar{N}$. So from the Bruhat decomposition, we have that

$$NwB/B = BwNw/\bar{N}/B = \bar{N}$$

and $G/B = N \cup \bar{N}$.

Now the dimension of B is strictly less than that of G (since G is non compact), so we have that $G/B = \bar{N}$ up to a closed set of measure 0.

We will use the following

Proposition 2.4.5. The natural map $K/M \to G/B = KB/MAN$ is a diffeomorphism.

The purpose for considering the Bruhat decomposition is to generalise the previous analysis from $SL(2, \mathbb{C})$. We can look at any integral on K/M as an integral on N or on $\bar{N}$, and hence compute the asymptotic behaviour. Observe, each choice of N contains exactly 1 fixed point for the Ad_A action - so there are exactly two in this case.

2.4.7 Changeover to $A \times L$

We repeat our study of $\mathbb{P}SL(2, \mathbb{C})$ for the general case In $K/M = G/B$ we have two Bruhat cells N and $\bar{N}$ corresponding to a choice of positive roots (there are only two Weyl chambers). Additionally, N and $\bar{N}$ are isomorphic to their Lie algebras through the exponential map. The A action on K/M corresponds to conjugation by A on G/B and thus to the Ad_A action on $\mathfrak{g}$ and similarly for N and $\bar{N}$. Using $\mathrm{Ad}_{\exp X} = \exp(\mathrm{ad}_X)$, we see that on $\mathfrak{n} = \mathfrak{n}_1 + \mathfrak{n}_2$ that the action of A is given by $\exp(\alpha(X))$ on $\mathfrak{n}_1$ and $\exp(2\alpha(X))$ on $\mathfrak{n}_2$, and similarly on the opposite Bruhat cell. It's clear therefore that the A action has two fixed points on K/M and that the flow is towards one as $t \to \infty$ and the other as $t \to -\infty$.

If we take a cross section of the orbits (minus the fixed points) for the A-action L (necessarily compact), and consider $K/M = A \times L$, then we have that the measure

$$dk = f(t, l) dt d\mu_L$$

where dt is Haar measure on A and $d\mu_L$ is some (necessarily finite) measure on L. We can bound the asymptotic behaviour of $f(t,l)$ in t by estimating how fast a point approaches each fixed point.

The following lemma is the meat of why we are interested in induced representations, and indeed it will be the reason for exponential mixing.

Lemma 2.4.1.

$$f(t,l) \leq ce^{-2\rho(X)|t|}$$

for some constant K

Proof. Since $\mathfrak{n} \cong N = K/M := A \times L = \mathbb{R} \times L = \bar{N} \cong \bar{\mathfrak{n}}$ up to an open dense set, we analyse the asymptotic behaviour under the A action by comparing the measures dk, and $dt \otimes dl$ to dn, the Lebesgue measure on the vector space $\mathfrak{n}$, and $d\bar{n}$ on $\bar{\mathfrak{n}}$. Fixing $L \subset \mathfrak{n}$, the corresponding subset L' in $\bar{\mathfrak{n}}$ will also be compact, so we have

$$dk = \psi(n)dn = f(t,l)dtdl = f(t,l')dtdl'$$

and $f(t,l') = g(l)f(t,l)$, where g is some bounded function.

We include details of the estimate as $t \to -\infty$. The other estimate will follow using the opposite Bruhat cell.

$N \cong \mathfrak{n}$ via the exponential map. We can take as a cross section of the orbits the space corresponding to the unit sphere in $\mathfrak{n}$. If we take a basis of unit eigenvectors for the Ad_X action on $\mathfrak{n}$ given by e_i and the induced coordinates x_i, we get coordinates on the unit sphere given by $\lambda_i e_i$ where $\lambda_n = \sqrt{1 - \sum_i |\lambda i|^2}$. Letting t denote the coordinate induced by the a_t action, we get the relation

$$x_i = \lambda_i e^{\rho_i(X)t}$$

where ρ_i is the root corresponding to the eigenvector e_i. We compute the Jacobian determinant

$$|J| = |\det \begin{pmatrix} \rho_1(X)e^{\rho_1(X)t} & e^{\rho_1(X)t} & \dots & 0 \\ \vdots & & \ddots & \\ \rho_n(X)e^{\rho_n(X)t} & \frac{\lambda_1}{\lambda_n}e^{\rho_n(X)t} & \dots & \frac{\lambda_{n-1}}{\lambda_n}e^{\rho_n(X)t} \end{pmatrix}| = e^{2\rho(X)t}|\det \begin{pmatrix} \rho_1(X) & 1 & \dots & 0 \\ \vdots & & \ddots & \\ \rho_n(X) & \frac{\lambda_1}{\lambda_n} & \dots & \frac{\lambda_{n-1}}{\lambda_n} \end{pmatrix}|$$

So we see that $dn = dx_1 \dots dx_n \leq Ke^{2\rho(X)t}dtd\mu$ as $t \to -\infty$, where μ is Lebesgue measure on the unit sphere.

Suppose that $dk = \psi(n)dn$. Then $\psi(n) \to 0$ as $n \to \infty$. So we have that $dk = K\psi(n)e^{2\rho(X)t}dtd\mu$. As $n \to 0$ as $t \to -\infty$, we have that $f(t,l) \leq Ke^{2\rho(X)t}$ as $t \to -\infty$.

The image of the unit sphere in $\mathfrak{n}$ in $\bar{\mathfrak{n}}$ will again be compact, we can perform a similar computation for $\bar{N}$ to get the corresponding estimate as $t \to \infty$.

□

So now, we have understood the A action on the induced representation, using the example of $\mathbb{P}SL(2,\mathbb{C})$ as guidance.

Lemma 2.4.2. *Suppose u and v are continuous functions on K, orthogonal to constants. Then for a representation (ρ, V) induced from the Borel subgroup $B = MAN$*

$$|\rho_{u,v}(a_t)| = \left| \int_K \langle a_t.u(k), v(k) \rangle dk \right| \leq Ke^{-2p|t|}$$

Proof. First observe that from our construction of induced representations, for any integrable function f in the representation space

$$\int_K f(k)dk = \int_M \int_{K/M} f(km)dmdk = \int_M \sigma(m)f(k)dkdm = c_\rho f(k)$$

So our computation on K/M suffices to understand ρ.

We realise the representation space as functions on $A \times L$, instead of functions on K. Since $dk = f(t,l)dtdl$, making a unitary transformation from $L^2(K,W) \to L^2(A \times L, W)$

$$v(k) \mapsto f(t,l)^{\frac{1}{2}}v(t,l)$$

Write $\tilde{v}(t,l)$ for the image of v.

Now the induced A action is measure preserving. Writing $k = (s,l)$ in the new coordinates,

$$a_t v(k) = \psi(a_t)v(k_t) = \tilde{\psi}(a_t)\tilde{v}(t+s,l)$$

Since a_t is measure preserving now, $|\tilde{\psi}| = 1$.

$$|\rho_{u,v}(a_t)| \leq c_\rho \int_{\mathbb{R}\times L} |f(t+s,l)^{\frac{1}{2}} f(s,l)^{\frac{1}{2}} \langle u(t+s,l), v(s,l)\rangle| dsdl$$

From the computations above and the Cauchy-Schwarz inequality (using that continuous functions on compact spaces are bounded)

$$|\rho_{u,v}(a_t)| \leq c_\rho K e^{-2\rho(X)|t|} \int_{\mathbb{R}} e^{-2\rho(X)|s|} ds \leq K' e^{-2\rho(X)|t|}$$

□

Thus, we have proven the following

Theorem 2.4.1. *Let G be a rank 1 semisimple Lie group and $A = \exp(tX)$ its maximal torus. Let (π, V) be a subrepresentation of a representation induced from a unitary representation of the Borel subgroup. Then for $u, v \in V$ which are continuous restricted to K, we have exponential decay of correlations. That is, the matrix coefficients satisfy*

$$|\pi_{u,v}(a_t)| \leq K(u,v)e^{-2\rho(X)|t|}$$

where ρ is half the sum of the restricted roots relative to X.

Remark. Observe that

$$\pi^{\otimes k}_{u^{\otimes k}, v^{\otimes k}}(a_t) = (\pi_{u,v}(a_t))^k$$

and so, if $\pi^{\otimes k}$ is a subrepresentation of a representation induced from a unitary representation of the Borel subgroup, then

$$|\pi_{u,v}(a_t)| \leq K(u,v)e^{-2\frac{\rho(X)}{k}|t|}$$

We now reduce the study of the matrix coefficients under the A action of the representation $L^2(\Gamma\backslash G) \ominus 1$ to its study on these induced representations.

2.4.8 Tempered Representations

Theorem 2.4.2. *A unitary representation of G connected, semisimple, with finite centre H is tempered if and only if H is a subrepresentation of $\mathrm{Ind}_B^G \pi$, where B is the minimal parabolic subgroup of G.*

Theorem 2.4.3. *Let π_0 be a tempered representation of G, connected, semisimple and finite centre. Then π_0 is a subrepresentation of a representation induced from a unitary finite dimensional representation of $P = M_p A_p N_p$ in the way described above*

P is a parabolic subgroup: any subgroup containing the Borel subgroup is known as a parabolic subgroup.

The result is based on a Mackey type formula to see that our representation occurs as a subrepresentation of Ind_B^G. By the Bruhat decomposition, $P\backslash G/B$ is finite, so we can express the restriction to P of a representation induced from B in terms of those induced from the restrictions, and then use transitivity of induction to see that the original representation is indeed what we want it to be.

2.4.9 Sobolev vectors for representations

In the section above, we showed that if π is induced from a representation of the Borel subgroup, then for u and v continuous functions on K, we have exponential decay of correlation coefficients. Now, we want to find a way of understanding which vectors of a particular representation, when viewed in the induced picture, are represented by continuous functions on K. For the representation we're interested in, $L^2(\Gamma\backslash G) \ominus 1$, we need to find a condition on these functions that mean that the corresponding functions in the induced representation on K are continuous. Indeed, we have no a priori way to know *how* the representation $L^2(\Gamma\backslash G) \ominus 1$ sits in $\mathrm{Ind}_B^G V$.

We first need the notion of smooth vectors for a representation:

Definition 2.4.10. Let π be a representation of G on a Hilbert space H. Then $v \in H$ is a C^k-vector for π if

$$f_\pi : G \to H,\ g \mapsto \pi(g)h$$

is a C^k map.

Although a general notion, in our situation, where the representation is essentially a function space of G, we have the following

Proposition 2.4.6. If $f \in L^2(\Gamma\backslash G)$ is a C^k of compact support function on the manifold $\Gamma\backslash G$, then f is a C^k vector for the representation

Proof. Denote by $p : G \to \Gamma\backslash G$ the canonical projection. Easy to see that $df \circ dp$ is the point wise derivative. Compactness ensures L^2 convergence. For details, see [3, 4.4.1]

For a representation of G V, denote by V^k the C^k vectors of G.

We will use the following standard theorem:

Theorem 2.4.4 (Garding's theorem)**.** *Let H be a unitary representation of G. Then smooth vectors are dense in H.* For a proof, see [3, Proposition 4.4.1.1] For a representation of $K \subset G$, we define

Definition 2.4.11. $v \in L^2(\Gamma\backslash G)$ is a Sobolev vector of index α for K if

$$\Delta_K^{\frac{s}{2}} \in L^2(\Gamma\backslash G)$$

We write W_π^s for the Sobolev vectors of the representation π, and define the Sobolev norm $\|.\|_s$ by $\|f\|_s = \|\Delta_K^{\frac{s}{2}}\|_2$ where $d\pi$ is the map induced on the universal enveloping and Δ_K is the Laplacian of K (see below for the construction). We define the operator $\Delta_K^{\frac{s}{2}} f$ by taking the inverse Fourier transform of order $\frac{s}{2}$ of $\Delta_K f$, where it exists.

It is clear from the definitions that the spaces W_π^s are the completions of V^∞ with respect to the Sobolev norms. The theory of Sobolev spaces is well understood, and unnecessary to develop here. We use the following (standard) fact.

Theorem 2.4.5. *Consider the representation of K given by $V = L^2(K, H)$. If $v \in V$ is a Sobolev vector of index α greater than $(\dim K)/2$, then v is a weakly continuous function on K. In particular, it is bounded, and $\max \|f\|$ is determined by its Sobolev norms*

See [7, Section 2] for details.

Remark. We will only ever consider smooth functions, and use them to approximate indicator functions. The Sobolev norms of a function f with $\|df\|$ large are greater. This is of relevance in the case of indicator functions as the sizes of the neighbourhoods we consider decrease. Nonetheless, it is possible to estimate and bound the Sobolev norms and hence $\max f$.

2.4.10 The Laplace operator on M

As a result of the above, we are reduced to showing that the representation on $L^2(\Gamma\backslash G) \ominus 1^{\otimes k}$ is tempered for some k. The following theorems will not be proved, but are essential for completing the proof.

Theorem 2.4.6. *Let G have no simple factor covered by $SO(n, 1)$ or $SU(n, 1)$. Then for any representation π of G, $\pi^{\otimes k}$ is tempered. Moreover, if $G = SU(n, 1)$ or $G = SO(n, 1)$, $\pi^{\otimes k}$ is tempered if and only if there exists an s_0 such π does have a subrepresentation any spherical complementary series for $s > s_0$.*

Remark. It's an easy computation that the spherical series have matrix coefficients not in $L^{\frac{2}{1-s}}$, which coupled with the alternative definition of tempered representations given above, shows that such representations require a large k before $\pi^{\otimes k}$ is tempered.

So, to show that some power of $L^2(\Gamma\backslash G) \ominus 1$ is tempered, we need to detect the presence of spherical complementary series in the cases $G = SO(n, 1)$ or $G = SU(n, 1)$.

We state first a simple lemma of importance in representation theory introducing which will motivate the use of the Laplacian

Lemma 2.4.3 (Schur's lemma)**.** *Let (π, V) be an irreducible representation of G on a Hilbert space V, and T any G-equivariant operator with a spectral decomposition. Then T is scalar.*

Proof. By the spectral decomposition, each eigenspace of T is a Hilbert space, on which G acts. Hence each eigenspace is a subrepresentation, and hence only one can occur.

In order to use this, we will state (without proof) the following

Theorem 2.4.7. *Let G be a semisimple Lie group. Then its series representations, constructed above are irreducible* See [31, 5.5.2]

Universal Enveloping Algebra

Let $\mathfrak{g}$ be a Lie algebra. The universal enveloping algebra $U(\mathfrak{g})$ is defined by

$$\left(\bigoplus_k \underbrace{\mathfrak{g} \otimes \ldots \otimes \mathfrak{g}}_{k \text{ times}}\right) \Big/ \sim$$

where $\sim$ is the equivalence relation $a \otimes b - b \otimes a = [a, b]$.

The usefulness of this algebra comes from the following

Proposition 2.4.7. $U(\mathfrak{g})$ is an associative algebra, with bracket $[a, b] = a \otimes b - b \otimes a$. Furthermore, any lie algebra homomorphism $\mathfrak{g} \to \mathfrak{E}$ extends uniquely to an algebra homomorphism $U(g) \to \mathfrak{E}$.

We have a distinguished second order element given by the following process.

- Choose a basis $X_1, \dots, X_n$ for $\mathfrak{g}$, and denote the dual basis by $\xi_1, \dots, \xi_n$. Then we have

$$\Xi = \sum_{i=1}^{n} X_i \otimes \xi_i \in \mathfrak{g} \otimes \mathfrak{g}^* \underbrace{\equiv \mathfrak{g} \otimes \mathfrak{g}}_{\text{via Killing form}} \hookrightarrow U(\mathfrak{g})$$

- Since the Killing form is G- invariant, Ξ is in the centre of $U(\mathfrak{g})$

Definition 2.4.12. Let (ρ, H) be a representation of a semisimple group G, and H^∞ its dense subspace of C^∞ vectors. Then for we define the representation of $\mathfrak{g}$ via

$$d\rho : \mathfrak{g} \to B(H),\, X \mapsto \phi,, \text{ where } \phi(v) = \left.\frac{d}{dt}\right|_{t=0} \exp(tX)v$$

It can be checked that this is indeed a Lie algebra homomorphism. In particular, it induces a representation on smooth vectors (also called $d\rho$) of $U(\mathfrak{g})$.

Definition 2.4.13. The Laplace operator Δ_G for the representation, defined on smooth vectors of H of G is $\Delta_G = d\rho(\Xi)$

We observe that, by the construction, Δ_G is G-equivariant (where it is defined).

In the case that G is a Lie group acting on a manifold M, with the natural representation on $L^2(M)$, we have that the corresponding action of $U(\mathfrak{g})$ is by differential operators on smooth vectors. [1, p 473].

Suppose further that G acts on M by isometries. Then Δ_G is the unique (up to scalar) differential operator on smooth functions that commutes with isometries: in particular, it's the Laplace-Beltrami operator on $C^\infty(M)$.

Proposition 2.4.8. Suppose that (ρ, H) is an irreducible representation of G. Then Δ_G is a scalar on its smooth vectors

Proof. Δ_G is G-equivariant and has a spectral decomposition on C^∞ vectors, which form a dense subspace by Garding's theorem.

We will apply this proposition and the previous proposition to complete the proof.

2.4.11 Finishing the proof of the main theorem

We only need to consider the cases $G = SO(n, 1)$ or $G = SU(n, 1)$, by *Theorem 2.4.6.*

Proposition 2.4.9. Suppose $L^2(\Gamma\backslash G)$ contains spherical complementary series H_s in its direct integral decomposition, corresponding to s. Then H_s is irreducible, and has a K-invariant function (see the remark after the definition of a spherical complementary series), i.e. a function f on $X = \Gamma\backslash G/K$.

Since H_s is irreducible, the Laplacian Δ_G acts by a scalar on smooth vectors, by *Proposition 8.* We denote the scalar value of Δ_G on H_s by $F_G(s)$

Theorem 2.4.8. *For $G = SO(n, 1)$ and $H = SU(n, 1)$, we have that*

$$F_G(s) = F_H(s) = \frac{1 - s^2}{4}$$

Proof. These representations can be directly realised as functions where the Laplacian is directly computable, and the eigenvalue so. See [13, chapter 1] and [7, Section 4]

Since smooth vectors are are dense, we can take our K-fixed vector f to be smooth. Thus, presence of spherical complementary series corresponding to s lead to eigenvalues corresponding to $F_G(s)$ (or $F_H(s)$) on M.

Remark. As such, we have spherical series bounded above 0 if and only if the Laplacian has the spectral gap.

Let us summarise what we have shown thus far. We first showed that for subrepresentations of those induced from the Borel subgroup, the action of a_t is exponentially mixing on those functions bounded on K. We quoted a standard Sobolev space fact, which gave a sufficient condition on the Sobolev norm of a function to be bounded on K, and we made the observation that in the representation we're interested in $L^2(\Gamma\backslash G)\ominus 1$, smooth functions on the unit tangent bundle satisfied this requirement. We gave a brief description of the characterisation of tempered representations as those induced from the Borel subgroup. We stated that for our symmetric spaces, for $G \neq SO(n,1), SU(n,1)$ that for any representations, a finite tensor power is tempered. For complex hyperbolic space and real hyperbolic space, a condition on the presence of spherical complementary series in the direct integral decomposition was stated. Finally, we have explained how spherical series in this case led to eigenvectors of the Laplacian corresponding to s. So, once we have shown that closed hyperbolic manifold and closed complex hyperbolic manifolds have the spectral gap, we will have proven the following theorem.

Theorem 2.4.9. *Let M be a rank 1 locally symmetric space. For smooth functions $f, g \in C^\infty T^1(M)$. Let g_t denote the geodesic flow. Then*

$$\left| \int_T^1 (M)(g_{t*}\phi) \cdot \overline{\psi} d\Lambda - \int \phi\overline{\psi} d\Lambda \right| \leq K(\phi,\psi)e^{-q_t}$$

where q

(i) depends on the first eigenvalue of the Laplacian on M in the case that M is hyperbolic or complex hyperbolic. From the remark following Theorem 2.3.6*, an explicit dependence can be computed.*

(ii) is universal for quaternionic and octonionic space.

2.4.12 Spectral Gap and Geometry

We first recall some standard theory of the Laplacian:

Proposition 2.4.10. Let (M, g) be a closed Riemannian manifold. Then

(i) Δ_g is formally self adjoint on $C^\infty(M)$

(ii) Δ_g has a self-adjoint extension to $L^2(M)$, and the domain of the extension is $W^2(M)$, (the Sobolev space of index 2)

(iii) The natural embedding $W^2(M) \hookrightarrow L^2(M)$ is compact, and hence the self-adjoint extension has compact resolvent

(iv) The spectrum of Δ_g is contained in $\mathbb{R} \geq 0$

and we apply the following theorem to show that the Laplacian on M does indeed have the spectral gap.

Theorem 2.4.10. *Let A be a self-adjoint (unbounded) operator on a Hilbert space with compact resolvent. Then the the spectrum of A is discrete, real and has no finite accumulation point.*

This completes the proof of *Theorem 2.3.9.* We include a further analysis to determine how the first eigenvalue of the Laplacian depends on the underlying geometry.

Definition 2.4.14. Let M be a closed Riemannian manifold. Define the Cheeger constant $h(M)$ by

$$h(M) = \inf_A \frac{s(\partial A)}{\mu(A)}$$

where μ is normalised Riemannian measure on M and where s is the induced Riemannian measure on an $(n-1)$ dimensional submanifold ∂A, where the infimum is taken over $\mu(A) \leq \frac{1}{2}$.

Remark. A first eigenvalue of the Laplacian of λ_1 corresponds to behaviour of order $e^{-\lambda_1 t}$ of the heat equation on M. [2]. Thus the following inequality is unsurprising.

Theorem 2.4.11 (Cheeger's theorem)**.** *Let (M, g) be a closed Riemannian manifold, and let $\lambda_1(M)$ denote the first positive eigenvalue of Δ_g, the Laplacian on M. Then*

$$\lambda_1(M) \geq \frac{h^2(M)}{4}$$

Proof. The proof is essentially an application of the following formula, which itself is just an application of Fubini's theorem. We follow [10, Section 1]

Lemma 2.4.4 (Coarea formula)**.** *Let u be a positive smooth function on M, and denote by ∇u the gradient of u (i.e. defined by $g(\nabla u, X) = du(X)$). Then for any positive, measurable ϕ*

$$\int_M \phi |\nabla u| d\mu_M = \int_0^\infty \int_{x \in u^{-1}(t)} \phi(x) d(s(u^{-1}(t)))dt$$

where we assume Sard's theorem to ensure the integral makes sense.

In particular,

$$\int_M |\nabla u| d\mu = \int_0^\infty \int_{x \in u^{-1}t} d(s(u^{-1}(t)))dt \geq h \int_0^\infty \min\{\mu(u \geq s), (1 - \mu(u \geq s))\} ds$$

For a smooth function f on M, set choose a median m and set $f^+ = \max(f - m, 0)$ and f^- similarly. Then

$$\int |f - m|^2 d\mu = \int_0^\infty \mu((f^+)^2 \geq t)dt + \int_0^\infty \mu((f^-)^2 \geq t)dt$$

So, applying the above and Cauchy-Schwarz (and using that $\nabla f^2 = f \nabla f$)

$$h \int |f - m|^2 d\mu \leq \int |\nabla (f+)^2| + |\nabla (f^-)^2| d\mu \leq 2 \left(\int |f - m|^2 d\mu \right)^{\frac{1}{2}} \left(\int |\nabla f|^2 d\mu \right)^{\frac{1}{2}}$$

Hence for any median m of f, we have

$$\frac{h^2}{4} \int |f - m|^2 d\mu \leq \int |\nabla f|^2 d\mu$$

By spectral theory/eigenfunction expansion, we have that (for f orthogonal to constants)

$$\lambda_1 \int f^2 d\mu \leq \int |\nabla f|^2 d\mu$$

Now since M is compact and $\int f^2 \leq \int |f - \mu|^2 d\mu$ for $\int f d\mu = 0$, choosing $f = f_1$ achieving this inequality gives us that

$$\frac{h^2}{4} \int f^2 d\mu \leq \frac{h^2}{4} \int |f - m|^2 d\mu \leq \int |\nabla f|^2 d\mu$$

and

$$\frac{h^2}{4} \leq \lambda_1(M)$$

□

2.4.13 The Frame flow

Finally, we show that the frame flow is exponentially mixing for hyperbolic manifolds.

Proposition 2.4.11. $G = SO(n,1)$ acts transitively on $\mathcal{F}(\mathbb{H}^n)$. Hence, for any hyperbolic manifold M, we have that $M = \Gamma\backslash G$. The action of the geodesic flow is given by translations.

Proof. We can compute that K, the stabiliser of a point, is isomorphic to $SO(n)$ via the isotropy representation. Thus, G acts transitively on $\mathcal{F}(\mathbb{H}^n)$. Since M is locally isometric to $\mathbb{H}^n$,

$$M = \Gamma\backslash G$$

Finally, that the geodesic flow is given by translations is easily seen from this construction.

If M is closed, then it has spectral gap, and thus we have exponential decay of matrix coefficients for C^∞ vectors for the representation $L^2(\Gamma\backslash G)$. Hence the frame flow is exponentially mixing for smooth functions on $\Gamma\backslash G = \mathcal{F}(M)$.

Section 3

Mixing, equidistribution and applications

In this section, we turn to a wider class of symmetric spaces, affine symmetric spaces, and we explore some consequences of mixing. We follow [6], and aim to quantify equidistribution in these spaces. We cover some basic applications to number theory; with our goal being, in particular understanding the contribution of ergodic theory of affine symmetric spaces to the proof of Siegel's mass formula.

3.1 Preliminaries

Let us restrict our attention to G being a connected, semisimple Lie group, with finite center.

Definition 3.1.1. Let H be the fixed point set of some involution. Then the homogeneous space G/H is an affine symmetric space

Remark. Since we are assuming G connected, symmetric spaces as described above are included in this definition. When H is non-compact, the resulting symmetric bilinear form constructed in section 1 need no longer be positive definite; that is, affine symmetric spaces include pseudo-Riemannian symmetric spaces.

Recall from section 1 that G is a locally compact, semisimple Lie group, and is therefore unimodular. Let $\Gamma \subset G$ be a discrete (in particular, unimodular) subgroup.

Proposition 3.1.1. We have a unique measure ν on $\Gamma\backslash G$ which satisfies

$$\int f dg = \int_{\Gamma\backslash G} \int_{\Gamma} f(\gamma g) d\gamma d\nu$$

Definition 3.1.2. Γ is a lattice if ν is a finite measure. It is a weakly irreducible lattice if Γ has dense projection to any G/G', for G' any nontrivial normal, non-compact subgroup.

Definition 3.1.3. Let $(g_n)_{n\in\mathbb{N}}$ be a sequence in G. If any compact subset of G contains only finitely many elements of (g_n) then we say that $g_n \to \infty$. If every compact subset of G/H (viewed as a union of elements of G) contains only finitely many elements of (g_n), we say that $g_n \to \infty$ over H.

For any function F on G, , we say

$$F(g) \to a \text{ as } g \to \infty \text{ if } F(g_n) \to a \text{ for every sequence } g_n \to \infty$$

Remark. For the rest of this section, we write $m(A)$ we mean the measure of a set A with the appropriate measure, when it is unambiguous.

The following is the main theorem of this section.

Theorem 3.1.1. *Let G/H be an affine symmetric space, and let Γ be a weakly irreducible lattice in G such that $\Gamma \cap H$ is also a lattice in H. Let $X = \Gamma\backslash G$ and*

$$Y = (\Gamma \cap H)\backslash H$$

Then for any $f \in L^2(X)$

$$\frac{1}{m(Y)} \int_{Yg} f(h)dh \to \frac{1}{m(X)} \int_X f(x)d\nu$$

as $g \to \infty$ over H

The statement above is a reformulation of *Defintion 1.1.2*, equidistribution. To develop an intuition for this, we give a specific example.

Example 3.1.1. Let $G = \mathbb{P}SL_2(\mathbb{C})$. Then G acts transitively on the space of geodesics in $\mathbb{H}^3$. The stabiliser of any geodesic must commute with the reflection in the hyperbolic plane orthogonal to that geodesic, and is thus the fixed point set of conjugation by that element. So G/H is the space of geodesics in $\mathbb{H}^3$. For a hyperbolic 3 manifold $M = \Gamma\mathbb{H}^3$, $\Gamma\backslash G/H$ is the space of geodesics in H. The theorem says that for a function on $T^1M = \Gamma\backslash G$, its average along closed geodesics of the same length converges to its overall average, as the distance between translates in the universal cover tends to infinity. That is, geodesics are equidistributed on closed hyperbolic manifolds.

Theorem 4.1.1 is a consequence of mixing (see below). We examine during the proof whether it is possible, in the case that G/H is a rank 1 *symmetric* space, to gain estimates on the rate of convergence of the above, under the action of $G \supset A = (g_t)_{t\in\mathbb{R}}$, a maximal torus.

We will then apply these results to gain quantitative estimates on

$$|\Gamma[H] \cap B_n|$$

for a wide class of sets $B_n \subset G/H$, $m(B_n) \to \infty$ and we will use these estimates to understand ergodic aspects of the proof of Siegel's mass formula.

3.2 Mixing

A key ingredient of the proof will be the following theorem

Theorem 3.2.1. *The action of G on $\Gamma\backslash G/H$ is mixing. That is, for ψ and $\phi \in L^2(\Gamma\backslash G/H)$, we have*

$$(\phi \circ g, \psi)_{L^2} \to (\phi, \psi)_{L^2}$$

as $g \to \infty$

The structure of the proof is similar to the previous section, and comes from proving decay of the matrix coefficients

$$\rho_{\psi,\phi}(g)$$

for the representation induced on $\mathrm{L}^2(\Gamma\backslash G/H)$ by $\psi \mapsto \psi \circ g$ We require the finite centre assumption, so that we may assume that up to a finite cover, $G = G_0 \times K$, which is an ingredient for the proof of mixing.

Remark. We showed that the action for G of rank 1 of $(g_t)_{t\in\mathbb{R}}$ is exponentially mixing for smooth functions on $\Gamma\backslash G/K$. As a result, we will be able to carry forward the estimates gained from this to the main theorems of this section.

3.3 Mixing implies equidistribution

The proof will utilise, as before, decompositions of the Lie group G as a result of the involution fixing H. However, in this particular case, the involution σ is not a Cartan involution. Nevertheless, we have the following decompositions, along with the Iwasawa decomposition from earlier.

We have

Proposition 3.3.1. Let G, H be as above. Then there exists a decomposition of Lie groups

$$G = HAK$$

where A is abelian and K is maximal compact.

and

Proposition 3.3.2. Let G be any semisimple Lie group, and H the fixed point set of some involution. Then the natural map

$$HMAN \to G$$

is a submersion at the identity. In particular, it is an open mapping in some neighbourhood of the identity.

See [6] for details and proofs.

Example 3.3.1. *The corresponding decompositions for $G = SL(n, \mathbb{R})$.*

The Iwasawa decomposition of $SL(n, \mathbb{R})$ has $K = SO(n, \mathbb{R})$, A the group of positive definite diagonal matrices with determinant 1 and N the upper triangular matrices in G.

Proof. Let $X \in SL(n, \mathbb{R})$ be a matrix. Fix a basis $e_1, \ldots, e_n$ of $\mathbb{R}^n$ and an inner product. Using the Gram-Schmidt orthogonalisation process, we can find a basis of orthogonal vectors $f_1, \ldots, f_n$, such that the change of basis matrix is upper triangular. Then the diagonal matrix with diagonal entry $A = 1/||f_i||$. Thus we have that

$$ABX \in SO(n)$$

proving the decomposition.

If we let H be fixed point subgroup of the involution given by conjugation by

$$\begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

$H = SL(n-1, \mathbb{R}) \subset SL(n, \mathbb{R})$ and that

$$G = SL(n-1, \mathbb{R}) \times A \times SO(n, \mathbb{R})$$

as an analytic manifold.

3.3.1 Proof of Equidistribution

We are now in a position to state the wavefront lemma

Let $G = HAK$ using the polar decomposition above. We have:

Lemma 3.3.1 (Wavefront lemma)**.** *For any open neighbourhood of the identity in G, there exists an open set $V \subset G$ such that*

$$HVg \subset HgU$$

for all $g \in AK$

The next section will be devoted to proving the wavefront lemma. For now, we assume it.

Proof of Theorem 4.1.1

Recall the statement of the theorem: let $X = \Gamma\backslash G$ of finite volume and $Y = \Gamma \cap H\backslash H$.

$$\frac{1}{m(Y)} \int_{Yg} f(h)dh \to \frac{1}{m(X)} \int_X f(x)d\nu$$

Let

$$\mathbb{E}_X(f) = \frac{1}{m(X)} \int_X f(x)d\nu$$

We prove the theorem for continuous functions of compact support (in particular, uniformly continuous functions), a dense subspace of $L^2(X)$. Let $\alpha \in C_0(X)$. By the uniform continuity of α, for all g and ϵ, there exists a neighbourhood of the identity such that

$$|\alpha(gu) - \alpha(g)| < \epsilon$$

(For example, pick a invariant metric on G and a δ such that $d(x, y) < \delta \Rightarrow |\alpha(x)-\alpha(y)| < \epsilon$)

Fix a sequence $(g_n)_{n\in N} \to \infty$ over H

By the wavefront lemma, we have a neighbourhood of the identity V such that

$$HVg \subset HgU \text{ for all } g \in AK$$

Notice that since $G = HAK$, we can assume that our sequence is contained in AK. Hence, we have that

$$YVg_n \subset Yg_nU$$

The chief idea of the proof is to estimate the integral over our thickened Y in two ways. First we use the mixing of the G-action to show that the integral

$$\frac{1}{m(YV)} \int_{YVg_n} \alpha(g)dg \to \mathbb{E}_X(\alpha)$$

Then we apply the following lemma

Lemma 3.3.2. *Let $X = \Gamma\backslash G$ and let ν denote the Haar measure on X as defined above. Let Y be an H-orbit, and W a subset of G. Then the Haar measure of X restricted to YW, say ν_x can be written as*

$$\int_W p(w)\mu_{Hw}dg$$

where the integral is with respect to the Haar measure on G and μ_{Hw} is the Haar measure on Hw.

Thus we retrieve the following estimate

$$\inf_{u\in U} \frac{1}{m(Y)} \int_{Yg_nu} \alpha(g)dg \leq \frac{1}{YV} \int_{YVg_n} \alpha(g)dg \leq \sup_{u\in U} \frac{1}{m(Y)} \int_{Yg_nu} \alpha(g)dg$$

Note that

$$\int_{Yg_n} \alpha(hu^{-1})dh = \frac{1}{m(Y)} \int_{Yg_nu} \alpha(g)dg$$

By our choice of U, the each bound on the inequality above varies at most by at most ϵ from $\int \frac{1}{m(Y)} \int_{Yg_n} \alpha(g)dg$, and hence we have the result.

3.3.2 Explicit estimates in rank 1

In the proof above, restrict to the case where K is maximal compact, and $G = (g_t)_{t\in\mathbb{R}}$ is the geodesic flow.

We choose neighbourhoods U_t such that $|\alpha(gu) - \alpha(g)| < \delta(t)$ and V_t such that

$$HVg_s \subset Hg_sU$$

for all $s \in \mathbb{R}$.

We assume without loss of generality that $\int_X f(x)d\nu = 0$, that is $\mathbb{E}_X(\alpha) = 0$. Now suppose that

$$\left|\int_{YV_tg_s} \alpha(g)dg\right| \leq Ke^{-q|s|} \qquad (*)$$

Then we have that

$$\left|\frac{1}{m(Y)}\int_{Yg_s} \alpha(g)dg - \int_{YV_tg_s} \alpha(g)dg\right| \leq 2\delta(t)$$

So

$$\left|\frac{1}{m(Y)}\int_{Yg_s} \alpha(g)dg\right| \leq 2\delta(t) + Ke^{-qs}$$

In general this doesn't tell us much, so we need to look at $(*)$ a little more closely.

$$\int_{YV_tg_s} \alpha(g)dg = (g_{s*}\chi(YV_t), \alpha)$$

where $\chi(YV)$ denotes the indicator function of YV and $(\cdot, \cdot)$ is the usual L^2 inner product. We proved exponential mixing for smooth functions. We could approximate characteristic functions using cut-off functions. However, the constant bounding the exponential decay term depended on the Sobolev norm of the function, which in turn depends on the norm of its derivatives. So as we take smaller neighbourhoods, we lose control over the actual size of the matrix coefficient, despite being in control of its asymptotic behaviour.

In the special case that $G = SL(2, \mathbb{R})$, a theorem of Ratner's, for decay of correlations in $L^2(\Gamma\backslash G/K)$ gives us a slightly stronger result, that we can apply. Her proof is a technical analysis of differential equations satisfied by matrix coefficients, and study of their asymptotic behaviour. We include only the statement, and details are in [R].

Theorem 3.3.1 (Ratner). *Let $S = \Gamma\backslash G/K$ be a Riemann surface of finite volume. Then for $\phi, \psi \in L^2(S) \ominus 1$*

$$(g_{t*}\phi, \psi) \leq K(S, \epsilon)||\psi||_{L^2}\|\phi\|_{L^2}e^{-(q(M)-\epsilon)|t|}$$

(where $q(M) > 0$ and ϵ small)

Write

$$D_\alpha(U) = \sup_{g\in G, u\in U} |\alpha(gu) - \alpha(gu)|$$

Applying Ratner's theorem and using the above definition, we get:

$$\left|\frac{1}{m(Y)}\int_{Yg_s} \alpha(g)dg\right| \leq 2D_\alpha(U_t) + K(M, \epsilon)||\alpha||m(YV_t)e^{-(q(M)-\epsilon)|s|}$$

Choose a collection of neighbourhoods U_s such that $D_\alpha(U_s) < e^{-(q(M)-\epsilon)|s|}$. Plugging in, we get the following

$$\left|\frac{1}{m(Y)}\int_{Yg_s} \alpha(g)dg\right| \leq (K(M, \epsilon)\|\alpha\|m(YV_s) + 1)\, e^{-(q(M)-\epsilon)|s|} \leq (K(M, \epsilon)\|\alpha\|m(\Gamma\backslash G) + 1)\, e^{-(q(M)-\epsilon)|s|}$$

We have shown that tangent circles on genus $g \geq 2$ closed Riemann surfaces along a geodesic are exponentially equidistributed.

3.4 The Wavefront Lemma

Let G be a semisimple Lie group, with Cartan decomposition $G = HAK$. Given a neighbourhood U of the identity, we need to construct a neighbourhood of the identity V such that

$$HVg \subset HgU$$

for all $g \in AK$

The decompositions detailed in the appendix give us tighter control over the structure of the group G. We make recourse to our usual trick: study the adjoint action on A on the vector group N.

Lemma 3.4.1. *Let $\mathcal{C}$ be a Weyl chamber, and let N be the corresponding subgroup. Then there exist arbitrarily small neighbourhoods of the identity in N such that $a^{-1}Ua \subset U$ for all $a \in \exp(\overline{\mathcal{C}})$*

Proof. Just the same as in the previous chapter: we apply the identity $\operatorname{Ad}\exp = \exp\operatorname{ad}$ to prove the result for $\mathfrak{n}$, which is isomorphic to N.

Example 3.4.1. Consider $G = SL(n, \mathbb{R})$. Recall that N was the space of upper triangular matrices, and A was the diagonal matrices with positive entries. *Lemma 2.4.2* really corresponds to the following observations:

(a) The Weyl chamber corresponding to N is

$$\mathcal{C} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1 > x_2 \ldots > x_n\}$$

(b) The resulting exponential map on $\overline{\mathcal{C}}$ gives matrices with decreasing entries

(c) Clearly, conjugating an upper triangular matrix by a matrix with decreasing entries is a contraction for all upper triangular matrices.

Combined with the decompositions, we can now prove the wavefront lemma.

Proof of the wavefront lemma

Recall the statement: given any neighbourhood of the identity U, we can find a neighbourhood V of the identity such that

$$HVg \subset HgU$$

for any $g \in AK$ First suppose $g \in A$. Then $g \in \exp(\overline{C})$ for some Weyl chamber, so let N be the unipotent subgroup corresponding to it.

Applying the $HMAN$ decomposition (in particular, that it is an open mapping in a neighbourhood of the identity), we get neighbourhoods in $V_m \subset M$, $V_n \subset N$ and $V_a \subset A$ such that $V_mV_aV_n \subset U$. By contraction, $g^{-1}V_ng \subset V_n$.

Let $V_{\mathcal{C}} = HV_mV_aV_n = HV_aV_mV_n$. By our construction,

$$HV_{\mathcal{C}}g \subset HgU$$

Set

$$V = \bigcap_{\mathcal{C}} V_{\mathcal{C}}$$

the intersection of the $V_{\mathcal{C}}$ over the Weyl chambers. This proves the wavefront lemma in the case that $g \in A$. To generalise to $g \in AK$, we make use of the fact that K is compact.

Claim. There exists $U' \subset U$ such that $k^{-1}U'k \subset U$ for all $k \in K$.

Proof. K is compact. Choose a sufficiently small compact neighbourhood of the identity $U' \subset U$ such that the claim holds.

Given U, and U' according to the claim, we choose V as above such that $HVa \subset HaU'$. Then

$$HVak \subset HaU'k \subset HakU$$

as required.

3.5 Counting and Applications

We turn to some consequences of the equidistribution theorem. Suppose now that we have a sequence of measurable sets in G/H, B_n and let $v = eH \in G/H$. Let

$$F_n(g) = |\Gamma v \cap gB_n|$$

Following [6], the main theorem of this section is

Theorem 3.5.1. *Assume the notation above. Suppose that $m(B_n) \to \infty$. Then for $X = \Gamma\backslash G$, and α is any compactly supported continuous function*

$$\frac{1}{m(B_n)} \int_X F_n(g)\alpha(g)dg \to \frac{m(\Gamma \cap H\backslash H)}{m(\Gamma\backslash G)} \int_X \alpha(g)dg$$

For the proof, we estimate

$$\int_X F_n\alpha$$

in terms of an integral over

$$\int_{(\Gamma\cap H)\backslash Hg} \alpha(h)dh$$

whose convergence as $g \to \infty$ we have understood through the equidistribution theorem.

For this purpose, we compare the following fibrations:

$$(\Gamma \cap H)\backslash\Gamma \to (\Gamma \cap H)\backslash G \xrightarrow{\pi} \Gamma\backslash G$$

and

$$(\Gamma \cap H)\backslash H \to (\Gamma \cap H)\backslash G \xrightarrow{\phi} H\backslash G$$

3.5.1 Integration on the fibrations

We now clearly state the measures and properties of the integrals on each part of the fibrations above. Now Γ, H and G are unimodular, by assumption, since Γ is assumed to be a lattice in G and H, so we have Haar measure on the quotients.

Space	Measure
$(\Gamma \cap H)\backslash\Gamma$	$d\gamma$ (counting)
$(\Gamma \cap H)\backslash H$	dh
$(\Gamma \cap H)\backslash G$	dg_1
$\Gamma\backslash G$	dg_2
$H\backslash G$	dg

We normalise these the measures so that

$$dh = dgdg_1 \text{ and } d\gamma = dgdg_2$$

We can push forward and pullback functions by integration:

$$\pi_*\alpha(g) = \int_{\Gamma\cap H\backslash\Gamma g} \alpha(h)d\gamma$$

If α is integrable, then $\pi_*\alpha$ must be finite almost everywhere and

$$\int_{\Gamma\backslash G} \pi_*\alpha dg_2 = \int_{(\Gamma\cap H)\backslash G} \alpha dg$$

3.5.2 Proof of counting

The function on $\Gamma\backslash G$ that we need to estimate is

$$F_n(g) = |\Gamma v \cap gB_n|$$

We compare to a function on $H\backslash G$

$$\chi_n(g) = \begin{cases} 1 & v \in gB_n \\ 0 & v \notin gB_n \end{cases}$$

Notice that χ_n as a function on $H\backslash G$ is the indicator function of B_n^{-1} since

$$v \in gB_n \Leftrightarrow g \in B_n^{-1}$$

Translation by $\gamma \in \Gamma$ induces a bijection on $|\Gamma v \cap B_n|$, so F_n is a function on $\Gamma\backslash G$, and translation by H fixes $v = eH$.

Now

$$F_n(g) = \sum_{\Gamma\cap H\backslash\Gamma} \chi_n(\gamma g) = \pi_*\phi^*(\chi_n)$$

Let α be a compactly supported continuous function on X. The lemma below assembles the considerations from the fibrations

Lemma 3.5.1.

$$\int_X F_n\alpha = \int_{H\backslash G} \chi_n(g) \underbrace{\int_{(\Gamma\cap H)\backslash Hg} \alpha(h)dh}_{:=\beta(g)}$$

Now we can estimate $\beta(g)$ using the equidistribution theorem, since $m(B_n) \to \infty$. Indeed,

$$\beta(g) \to \frac{(\Gamma\cap H)\backslash H}{\Gamma\backslash G} \int_X \alpha dg_2$$

$$\frac{1}{m(B_n)} \int_X F_n(\alpha)$$

is the average of β over B_n^{-1}, which finishes the proof.

We can get point wise convergence of the functions $F_n'(g) = F_n(g)/m(B_n)$ if we add further restrictions to B_n.

Definition 3.5.1. A set system B_n with $m(B_n) \to \infty$ is well rounded if for any $\epsilon > 0$, there exists a symmetric neighbourhood U of the identity with

$$m\left(\bigcap_{u\in U} u(B_n)\right) > (1-\epsilon)m(B_n)$$

We we get the following direct corollary of the theorem above

Corollary 3.5.1. *Let B_n be a a system of well-rounded sets. Then*

$$F'_n(g) \to \frac{m((\gamma \cap H)\backslash H}{m(\Gamma\backslash G)}$$

point wise as $n \to \infty$

Proof. The idea is to compare $F'_n(g)$ to $F_n(a)$ for a constant, and then use continuous bump functions, and the constraints on the measures of well rounded sets to estimate

$$\int \chi_n F_n(a) dg$$

3.6 Applications

We give a basic application of the results to counting integral points on affine symmetric orbits. Further applications will be given in next section on hyperbolic geometry.

Let V be a finite dimensional real (continuous) representation of G (where G is a semisimple, connected Lie group with maximal compact subgroup K and $\|,\|$ a K-invariant Euclidean norm on V). Suppose that the orbit O of v is an affine symmetric space: that is

$$\mathrm{Stab}_G(v) = \mathrm{Fix}^\sigma$$

where σ is some involution. For a lattice Γ, we're interested in the size of the intersection

$$|\Gamma v \cap O_t|$$

where $O_t = O \cap B(0,t)$, the open unit ball of radius t about the origin, and in particular the asymptotic behaviour thereof.

Proposition 3.6.1. The sets $O_t \subset G/H \cong O$ are well rounded: given $\epsilon > 0$, there is a neighbourhood U of the identity such that for all t sufficiently large,

$$\frac{m(U\partial O_t)}{m(O_t)} < \epsilon$$

Proof. Omitted - the technicality is in estimating $m(O_t)$ in terms of t, and the key idea comes from a similar analysis to the previous chapter: $\|.\|$ is K-invariant, and we can making an explicit computation of the integral using the polar decomposition for G to estimate the action along a basis of eigenvectors. The finite dimensionality allows for specific enough estimates to prove the proposition.

Theorem 3.6.1. *Suppose that Γ is a weakly irreducible lattice. Then*

$$|\Gamma v \cap B(0,t)| = |[\Gamma H] \cap O_t| \sim \frac{m((\Gamma \cap H)\backslash H)}{m(\Gamma\backslash G)} m(O_t)$$

This particular approach is useful, not because the quantity $|N(t,O) \cap B(0,t)|$ is difficult to estimate, but because it is in fact readily estimable in many ways.

As an example, if we consider a quadratic form $F(x)$ in n variables, then the group G of automorphisms of $\mathbb{R}^n$ fixing f is isomorphic to $SO(p,q)$ for $p+q=n$. For an orbit, we consider $\Gamma = G(\mathbb{Z})$ (matrices with entries in $\mathbb{Z}$), and $V(\mathbb{Z})$ defined similarly. $G(\mathbb{Z})$ is in fact an irreducible lattice, and there is a theorem stating that $V(\mathbb{Z}) = O_1 \cup \ldots \cup O_n$ a finite union of orbits satisfying the stabiliser condition. For a variety over $\mathbb{Z}$, the quantity

$$W(F) := \frac{1}{\mathrm{vol}(G(\mathbb{Z})\backslash G)} \sum_i \mathrm{vol}(H_i(\mathbb{Z})\backslash H)\,\mathrm{vol}(O_i)$$

is of great value in analytic number theory. The asymptotics of the value of $N(t, O)$ has been classically estimated to satisfy (for $n \geq 5$ at least)

$$N(t, O) \to \mu_\infty(O_t) \prod_p \mu_p$$

where μ_∞ is the induced measure on $\mathbb{R}^n$ by our inner product, and μ_p is computed from the number of p-adic points, and summing over the orbits we get the celebrated result:

$$m(F) = \prod_p \mu_p$$

For further details, see [11].

Section 4

The Surface Subgroup Theorem

This chapter brings together the results of chapter 2 and the ideas of chapter 3 and demonstrates the power of applying an ergodic approach to geometric problems in locally symmetric spaces. The goal will be to understand the proof of the following theorem:

Theorem 4.0.2 (Surface Subgroup Theorem)**.** *Let M be a closed, orientable, irreducible 3-manifold. Then there exists a closed hyperbolic surface S and an immersion $f : S \to M$ such that $f_* : \pi_1(S) \to \pi_1(M)$ is injective.*

The main thing to gain from this chapter, besides an understanding of the proof is to understand how powerful explicit control over mixing gives tight control of geometric constructions using it - and one way of doing that.

The proof is from [9], and for details from preliminary sections, see [4].

4.1 Preliminaries

4.1.1 Hyperbolic Geometry

We first review basic objects of 2 and 3 dimensional hyperbolic geometry, and indicate where the results of section 2 are applicable.

Definition 4.1.1. Hyperbolic 3 space is realised as the open unit ball in $\mathbb{R}^3$

$$\mathbb{H}^3 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < 1\}$$

with the metric

$$g = \frac{dx^2 + dy^2 + dz^2}{x^2 + y^2 + z^2}$$

and we identify $\partial\mathbb{H}^3 = \mathbb{C} \cup \{\infty\}$. It is isometric to the upper half space

$$\mathbb{H}^3 = \{(x, y, z) \in \mathbb{R}^3 : z > 0\}$$

with the metric

$$g = \frac{dx^2 + dy^2 + dz^2}{z^2}$$

An explicit computation shows that:

Proposition 4.1.1. The geodesics in $\mathbb{H}^2$ and $\mathbb{H}^3$ are circle segments whose feet are perpendicular at the boundary. There is a unique geodesic segment between any two points.

Since the metric is conformal, isometries of this metric are conformal maps from the Euclidean ball onto itself. The following proposition and its corollary give a description of the isometries

Proposition 4.1.2. The only conformal maps $\mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ are Mobius transformations. The proof is using basic complex analysis

Proposition 4.1.3. Any conformal map $f : \partial\mathbb{H}^3 = \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ extends uniquely to an isometry of $\mathbb{H}^3$. Any isometry of $\mathbb{H}^3$ extends to a conformal map of the boundary.

Proof. We can check individually for translations, scaling and inversions, which span conjugacy classes of Mobius transformations. To explicitly define the extension, we fix a base point. Since there exists a unique geodesic through each boundary point, we define the image of the extension to be the foot of the corresponding geodesic. That this map is conformal follows from the fact that the metric is conformal.

As described in the basic concepts section, $\mathbb{H}^3$ is a rank 1 symmetric space.

Proposition 4.1.4. $SO(2,1) \cong SL(2,\mathbb{C})$ acts on $\mathbb{H}^3$ through by isometries. The kernel of this action is $\pm\,\mathrm{id}$. This action is transitive on $\mathbb{H}^3$, and the stabiliser of $0 \in \mathbb{R}^3 \cong O(3)$. Thus

$$\mathbb{H}^3 = \mathbb{P}SL(2,\mathbb{C})/SO(3,\mathbb{R})$$

We now define hyperbolic manifolds

Definition 4.1.2. A manifold M of dimension n is hyperbolic if each point has a neighbourhood isometric to $\mathbb{H}^n$.

The fundamental importance of our analysis of the rank 1 group $SL(2,\mathbb{C})$ to studying hyperbolic geometry comes from the following theorem

Theorem 4.1.1. *Let M be a complete, simply connected, hyperbolic manifold. Then M is isometric to $\mathbb{H}^n$*

Sketch. The condition for M being a hyperbolic manifold can be rewritten as M having an atlas of charts with image in $\mathbb{H}^n$, such that the transition maps are in $SO(n,1)$. First, it's shown that if M is any simply connected such manifold, then any coordinate patch $M \supseteq U_0 \to \mathbb{H}^n$ can be extended to a local isometry $M \to \mathbb{H}^n$. We show that if $h : M \to \mathbb{H}^n$ is a local isometry, then h is in fact a covering map and thus a homeomorphism. Then, it is shown that if local isometries agree on some open set, they agree on all of $\mathbb{H}^n$.

As a result of the theorem, the universal cover of any closed hyperbolic manifold M is $\mathbb{H}^n$. Thus for any closed hyperbolic 3-manifold $\mathbf{M}$, we have that

$$\mathbf{M} \cong \pi_1(M)\backslash\mathbb{P}SL(2,\mathbb{C})/SO(3,\mathbb{R})$$

We have a final preliminary definition that enables us to talk about geodesics in more detail on $\mathbf{M}$.

Definition 4.1.3. Let $v \in T^1_p\mathbf{M}$ and $w \in T^1_q M$ be tangent vectors, and γ a geodesic segment joining them. Then we define the complex length

$$\mathbf{d}_\gamma(v,w) = d(p,q) + i\theta(v@q, w)$$

where $v@q$ denotes the parallel transport along γ. and θ denotes the signed angle between tangent vectors, taken in $(-\pi,\pi)$ Similarly, we define the complex lengths between two geodesics using the common orthogonal.

4.1.2 Pairs of pants and Fenchel-Nielsen

For the rest of this section, fix $\mathbf{M}$ as a closed, oriented, irreducible, hyperbolic 3-manifold.

We now give a brief overview of the property of pairs of pants, which we will use to piece together the surface in the statement of the theorem.

Definition 4.1.4. A topological pair of pants is a three holed sphere, with its boundary elements. The boundary circles are *cuffs*, C_i, $i = 1, 2, 3$

Remark. Fix a point $p \in \Pi^0$ our topological pair of pants. The fundamental group $\pi_1(\Pi^0)$ of a topological pair of pants is generated by the homotopy classes of the cuffs. Indeed, Π^0 is homotopic to $S^1 \wedge S^1$, to see this explicitly.

The proposition below enables us to parametrise hyperbolic pants.

Proposition 4.1.5. There's unique hyperbolic pants with cuff lengths (c_1, c_2, c_3)

Proof. In $\mathbb{H}^2$, there's a unique right angled hexagon with side lengths (c_1, c_2, c_3). We glue together edges to get the required metric.

Definition 4.1.5. We define the linking curves η_i as the the orthogonals between C_{i-1} and C_{i+1}

Remark. When indexing cuffs of pants, we will work modulo 3 for notational ease.

A surface, by cutting appropriately, can be decomposed into pairs of pants:

Proposition 4.1.6. Every closed surface S of genus $g \geq 2$ has a pair of pants decomposition.

Definition 4.1.6. The Fenchel-Nielsen coordinates for a surface with a pants decomposition are the the cuff lengths and the twist $s \in \mathbb{R}$, which corresponds to the complex length of a tangent vector to the cuff in a gluing.

The metric on a surface with a pants decomposition is determined uniquely by its Fenchel-Nielsen coordinates.

We now consider mappings of pairs of pants into **M**. We'll be able to understand the topological properties of a surface in M from the properties of each pants in its decomposition. Eventually, we'll do the reverse: piece together an incompressible surface *from* pants in **M**.

4.1.3 Interlude: Geodesics in M

The essential idea by which we construct our incompressible surface is similar to a solution of the following problem, and the solution is essentially of a similar form.

Question. Let M be a closed hyperbolic manifold. Does there exist a closed geodesic of length close to R? More generally, does there exist a geodesic of length close to R between two points x, y

The answer to this question is completely non-obvious, since we know very little about the topological structure of M.

Our approach depends on the following geometric lemma

Lemma 4.1.1. *Let $p, q \in M$. Suppose there is a geodesic γ with $\gamma(0) = p$ and $d(\gamma(R), q) < \epsilon$. There exists a constant $\hat{\epsilon} > 0$ such that $\epsilon < \hat{\epsilon}$ implies that there is a geodesic λ with $|d_\lambda(p, q) - R| < \epsilon$*

Proof. Since M is closed, if $\hat{\epsilon}$ is sufficiently small, then for any $x \in \mathbb{H}^3$, the projection from the universal cover in a $\hat{\epsilon}$-neighbourhood is injective. Pick an arbitrary lift $\tilde{p}$ of p. Write $y = \gamma(R)$ and $\tilde{y}$ for the lift of y along γ. So there is a lift of q, $\tilde{q}$ of distance at most ϵ from $\tilde{q}$. But then, by the triangle inequality:

$$d(\tilde{p}, \tilde{q}) < d(\tilde{p}, \tilde{y}) + d(\tilde{y}, \tilde{q}) = R + \epsilon$$

Thus the length of the unique geodesic $\tilde{\lambda}$ between $\tilde{p}$ and $\tilde{q}$ is at most $R + \epsilon$, so the closed geodesic λ corresponding to the projection has $l(\lambda) < R + \epsilon$.

Since γ lifts to the unique geodesic segment between $\tilde{p}$ and $\tilde{y}$, we have that

$$d(\tilde{p}, \tilde{y}) < d(\tilde{p}, +\tilde{q}) + d(\tilde{q}, \tilde{y})$$

so $l(\lambda) > R - \epsilon$.

Remark. $\hat{\epsilon}$ does not depend on p since M was closed. Fix M and $\epsilon < \hat{\epsilon}$

Setting $p = q$ gives the corresponding lemma for closed geodesics.

For $p \in M$, we define $N_\epsilon(p) \subset T^1M$ to be the unit tangent spheres of points in an ϵ neighbourhood of p.

Fix a base point $x \in M^n$ and let $f_{\epsilon,x} : T^1\mathbb{H}^n \to \mathbb{R}^+$ be a smooth, positive function supported in $N_\epsilon(X)$, such that

$$\int_{T^1\mathbb{H}^n} f_{\epsilon,x} d\Lambda = 1$$

Suppose $gy = x$ for $g \in SO(n,1)$. Then we define

$$f_{\epsilon,y} = g^* f_{\epsilon,x}$$

Define $\mathbf{a} : T^1\mathbb{H}^n \times T^1\mathbb{H}^n \to \mathbb{R}^+$ by

$$\mathbf{a}(u,v) = \int_{T^1\mathbb{H}^n} g_R^* f_{\epsilon,u} f_{\epsilon,v} d\Lambda$$

where g_R denotes the geodesic flow at time R.

For $u, v \in T^1M$, define

$$\mathbf{a}(u,v) = \sum_{\gamma \in \pi^1(M)} \mathbf{a}(\gamma \tilde{u}_0, \tilde{v}_0)$$

where $\tilde{u}_0$ and $\tilde{v}_0$ are arbitrary initial lifts of u and v to $T^1\mathbb{H}^n$.

Remark. For $u \in T_pM$ and $v \in T_qM$, $\mathbf{a}(u,v) > 0$ if there is some geodesic with tangent u at time 0 satisfying that $d(\gamma(R), q) < \epsilon$.

Since $\epsilon < \hat{\epsilon}$, the functions $f_{\epsilon,x}$ are well defined on T^1M. We can now apply exponential mixing of the geodesic flow on T^1M. That is, for two smooth functions ψ and ϕ,

$$\int_{T^1M} g_t * \psi\overline{\phi} d\Lambda = \left(\int_{T^1M} \psi d\Lambda\right) \left(\overline{\int_{T^1M} \phi d\Lambda}\right) + O(e^{-q|t|})$$

So

$$\mathbf{a}(u,v) = \Lambda(T^1(M)) \int_{T^1M} g_R^* f_{\epsilon,u} f_{\epsilon,v} d\Lambda = \Lambda(T^1(M)) \cdot \frac{1}{\Lambda(T^1M)^2} + O(e^{-qR})$$

Hence, we have shown that $\mathbf{a}(u,v) > 0$ for sufficiently large R. This is remarkable: it implies that between any two points, there is a geodesic of length arbitrarily close to R for sufficiently large R.

Remark. Suppose we wanted better control over the tangent vectors at p and q of the geodesic. We could attempt to take a different N_ϵ, taking perhaps a neighbourhood of a unit tangent vector instead of the whole unit tangent sphere. However, it is *Lemma 5.1.1* that enables us to see the relationship between the geodesic constructed with the geodesic flow, and the geodesic of length close to R. This is more difficult to generalise, and its generalisation is an important lemma in the construction of the incompressible surface in the surface subgroup theorem.

Thus, we have proven the following

Theorem 4.1.2. *Let M be a closed hyperbolic manifold and $p, q \in M$. For $\hat{\epsilon} > \epsilon > 0$, and sufficiently large R, there exists a geodesic γ with $\gamma(0) = p$ and $\gamma(T) = q$ and*

$$|T - R/2| < \epsilon$$

Aside: Dolgopyat's estimate

In [16], the following theorem is proved:

Theorem 4.1.3. *Let X be a closed surface of variable negative curvature. Then the geodesic flow on T^1X is exponentially mixing for all $f, g \in L^2(T^1X)$*

We can apply a very similar technique as before to show that between any two $p, q \in X$, for $\epsilon > 0$, and for sufficiently large R, there is a geodesic of length R at starting at p and ending at a distance of at most ϵ away from q.

Clearly Dolgopyat's estimate is useful for different reasons to the results of section 2. However, for the proof of the surface subgroup theorem, control over the frame flow is essential.

4.2 Surfaces in M

Suppose we have a surface S. Then we determine the metric on it from its pants decomposition by the lengths of each cuff and the twist. In what follows, we aim to gain a similar parametrisation for surfaces *immersed* in $\mathbf{M}$, which has an ambient metric. Our goal is to develop a a system of coordinates, in $\mathbf{M}$, that parameterise surfaces within it. Eventually, we will use the results of previous sections to show that these coordinates are equidistributed, in some sense, and use them to piece together a surface.

4.2.1 Parametrising pants in M

First a lemma regarding the ambient geometry of $\mathbf{M}$.

Proposition 4.2.1. Let $C \subset \mathbf{M}$ be a closed curve. Then there exists a unique geodesic γ_C freely homotopic to it.

Remark. For any path connected space X, any closed curve defines a conjugacy class in $\pi_1(X)$, by concatenating with a path to and from the base point.

Hence we have a bijection between the free homotopy classes of closed curves and closed geodesics, and therefore also between elements of $\pi_1(M)$ up to conjugation and closed geodesics.

Definition 4.2.1. A pair of pants in $\mathbf{M}$ is the conjugacy class of an injective homomorphism

$$\rho : \pi_1(\Pi^0) \to \pi_1(\mathbf{M})$$

(where Π^0 is a topological pair of pants)

The universal cover of $\mathbf{M}$, $\tilde{M} \cong \mathbb{H}^3$, and $\pi_1(\mathbf{M})$ acts on $\mathbb{H}^3$ by orientation preserving isometries. Thus ρ determines a map (up to conjugacy)

$$\rho : \pi_1(\Pi^0) \to \mathbb{P}SL(2, \mathbb{C})$$

In the same way that we can decompose a closed surface into pants, and parameterise them by twists, we want to parameterise pants in $\mathbf{M}$ and determine when we can glue them together to form an incompressible surface S.

Consider a ρ a pair of pants in $\mathbf{M}$. Then ρ determines a map up to homotopy

$$f_\rho : \Pi^0 \to M_\rho := \rho(\pi_1(\Pi^0))\backslash\mathbb{H}$$

By our proposition, we can assume that f_ρ maps the cuffs C_i onto closed geodesics γ_i, and the linking curves a_i onto the orthogonals, η_i. We orient the curves so that the pants is on the left of each cuff.

The 1-complex of the C_i and the γ_i divide $f_\rho(\Pi^0)$ into two orthogonal hyperbolic hexagons, which are determined uniquely up to isometry by the complex lengths of the sides and their orthogonals. We define

$$\mathbf{hl}(\gamma_i) = \mathbf{d}_{\gamma_i}(\eta_{i-1}, \eta_{i+1})$$

A useful picture to have is the following one of the universal cover.

Suppose we have a pair of pants in a manifold ρ. Recall that γ_i is the unique geodesic homotopic to $\rho(C_i)$, and η_i the orthogonal geodesic.

We lift γ_i to the universal cover. WLOG we assume, $\widetilde{\gamma}_i$ (by conjugating ρ) is the geodesic between 0 and ∞. Then η_i has lifts as well, and lifts of η_{i-1} and η_{i+1} alternate along $\widetilde{\gamma}_i$, and we define $\mathbf{d}_{\gamma_i}(\eta_{i-1}, \eta_{i+1})$ as the complex distance between corresponding lifts.

Let $A_{\gamma_i} \in \mathbb{P}SL(2, \mathbb{C})$ be the deck transformation corresponding to $[\gamma_i] \in \pi_1(\mathbf{M})$. Then it corresponds to $z \mapsto e^{\mathbf{l}_{\gamma_i}} z$. By the construction, $\sqrt{A_{\gamma_i}} : z \mapsto e^{\mathbf{hl}_{\gamma_i}} z$ maps a lift of η_{i-1} to a lift of η_{i+1}. Indeed, this could be used to define the half length $\mathbf{hl}(\gamma_i)$.

Remark. We make an obvious observation: $\mathbf{hl}(\gamma_i)$ depends on ρ.

The unit normal bundle $N^1(\sqrt{\gamma}_i)$

For a pair of pants in $\mathbf{M}$, we will always consider the ρ in terms of the closed geodesics γ_i corresponding to cuffs and the orthogonal geodesic linking segments η_i. Recall the definition of A_{γ_i} and $\sqrt{A_{\gamma_i}}$: the former representing the loxodromic transformation corresponding to γ_i and the latter the one corresponding to $\mathbf{hl}(\gamma_i)$.

Definition 4.2.2. For a geodesic in $\mathbb{H}^3 \subset \mathbb{R}^3$, we realise $N^1(\gamma)$ as the space of vectors in $\mathbb{R}^3$ orthogonal to γ and of modulus 1 in the hyperbolic metric

We will see that the natural space for studying how pants glue together is in fact the following:

Definition 4.2.3.

$$N^1(\sqrt{\gamma}_i) = N^1(\tilde{\gamma}_i e)/\langle\sqrt{A_{\gamma_i}}\rangle$$

Clearly, the group

$$\mathbb{C}/2\pi i\mathbb{Z} + \mathbf{hl}(\gamma_i)\mathbb{Z}$$

acts simply transitively on $N^1(\sqrt{\gamma_i})$ via the corresponding loxodromic Mobius transformations.

The foot map

Consider the lift of γ_i, and the unit normal n_i at $\gamma_i \cap \eta_{i-1}$ in the direction of η_{i-1} and n'_i at $\gamma_i \cap \eta_{i+1}$ in the direction of η_{i+1}. Then the element $A_{\sqrt{\gamma_i}}$ swaps these two vectors. So they define the same point in $N^1(\sqrt{\gamma_i})$. Thus we define the foot map:

$$\text{foot} : (\rho, C_i) \to N^1(\sqrt{\gamma_i}) \text{ given by } (\rho, C_i) \mapsto [n_i]$$

In the rest of the proof, we shall compare pants in $\mathbf{M}$ by their feet. That is, we consider

$$\text{foot}_{\gamma_i}(\rho) := \text{foot}(\rho, C_i) \text{ where } \rho(C_i) \simeq \gamma_i$$

4.2.2 Viable representations and the shear map

Suppose we have a closed surface of genus ≥ 2 and S^0 and a pants decomposition by a family of curves C. We orient the cuff of each pants Π such that the pants lies to the left of the cuff.

Definition 4.2.4. A marked pair of pants for S^0 is (Π, C), where C is a cuff.

Notice by the panting of the surface, we have that for each (Π, C) there is a (Π', C') corresponding to the other pants glued on at that cuff.

Definition 4.2.5. Suppose we are given $\rho : \pi_1(S^0) \to \mathbb{P}SL(2, \mathbb{C})$ as before, such that ρ is discrete and faithful restricted to each pants. Then we define, as above, $\mathbf{hl}_\Pi(\gamma)$ for $\gamma \simeq C$ using $f_\rho|_\Pi$. Now $\mathbf{hl}_\Pi(C) = \mathbf{hl}_{\Pi'}(C) + i\pi$ or $\mathbf{hl}_\Pi(C) = \mathbf{hl}_{\Pi'}$, depending on the orientation of gluing.

We put these things together in the following definition: we only want to consider closed surfaces S in $\mathbf{M}$, ρ and pants decompositions of them that are sufficiently well behaved to enable us to understand the image of $\rho(\pi_1(S))$ from the corresponding $\rho|_\Pi(\pi_1(\Pi))$.

Definition 4.2.6. A viable representation of $S^0, \mathcal{C}$ is

$$\rho : \pi_1(S^0) \to \pi_1(\mathbf{M})$$

such that

(a) $\rho|_\Pi$ is discrete and faithful for each pants in the decomposition

(b) $\mathbf{hl}_\Pi(C) = \mathbf{hl}_{\Pi'}(C)$ for each cuff C.

If we're given a viable representation of S^0, then $\mathbf{hl}(C)$ is well defined independently of the pants it's in.

Recall that the foot map assigned to each marked pants (Π, C) a point in $N^1(\sqrt{\gamma})$, on which the group

$$G = \mathbb{C}/2\pi i\mathbb{Z} + \mathbf{hl}(\gamma_i)\mathbb{Z}$$

acted simply transitively.

For a cuff C in a viable representation, we can now parameterise the twist in gluing using the feet:

Definition 4.2.7. For a cuff C, homotopic to geodesic γ in a viable representation, define the shear map $s(C)$ by

$$s(C) = \text{foot}_\gamma(\rho|_\Pi) - \text{foot}_\gamma(\rho|_{\Pi'}) - i\pi$$

Remark. The shear map is well defined. Since $\mathbf{hl}_\Pi(C) = \mathbf{hl}_{\Pi'}(C)$, we have that $N^1(\sqrt{\gamma}) = N^1(\sqrt{\gamma'})$. Furthermore, if we interchange Π and Π', we reverse the orientation of γ and get the same element in G.

4.2.3 Good and Perfect Pants

For a viable surface S^0 in $\mathbf{M}$, for each cuff (corresponding to its pants decomposition $\mathcal{C}$), we have $\mathbf{hl}(C)$ and $s(C)$ that parameterise it.

Definition 4.2.8. We say that a viable surface S^0 in $\mathbf{M}$ with pants decomposition $\mathcal{C}$ is

- a perfect panted surface if

$$\mathbf{hl}(C) = R/2 \text{ and } s(C) = 1$$

- a good panted surface if

$$|\,\mathbf{hl}(C) - R/2| < \epsilon \text{ and } |s(C) - R/2| < \frac{\epsilon}{R}$$

The following proposition explains why we are interested in good panted surfaces

Proposition 4.2.2. Let $S^0, \mathcal{C}$ be a surface with pants decomposition. Suppose $f : S^0 \to M$ is a map such that f_* induces a perfect panted surface in $\mathbf{M}$. Then the induced map

$$\tilde{f} : \mathbb{H}^2 \to \mathbb{H}^3$$

is an isometric embedding

A key ingredient of the proof of the surface subgroup theorem is showing that a *good* panted surface corresponds to a *nearly* locally isometric map between the universal covers.

Proposition 4.2.3. Suppose $f : M \to N$ is a map of closed hyperbolic manifolds such that the induced map on the cover is locally δ nearly isometric. Then $f_* : \pi_1(M) \to \pi_1(N)$ is injective, for $\delta < K$ for a constant K.

Proof. A nontrivial kernel corresponds to two different lifts of a point in $\tilde{M}$ being mapped to the same point in $\tilde{N}$ by the induced $\tilde{f}$. But if this is the case, then $\tilde{f}$ cannot possibly be locally isometric near these.

Perfect panted surfaces are rare: however, using exponential mixing of the frame flow, we'll be able to construct good panted surfaces.

The following is the key theorem of Section 2 of [9]

Theorem 4.2.1. *Suppose $S^0, \mathbb{C}$ is a good panted surface in $\mathbf{M}$. Then $\rho : \pi_1(S^0) \to \mathbb{P}SL(2, \mathbb{C})$ is conjugate to $\rho : \pi_1(S') \to \mathbb{P}SL(2, \mathbb{C})$ where S' is a perfect panted surface in $\mathbf{M}$. In particular, ρ is injective.*

Although a major part of the proof, the details are not based on ergodic techniques and are technical. Key aspects to note are:

- The shear is an *essential* feature of the proof. Indeed, we need tighter control over the shear than on the half lengths. A similar approach for constructing incompressible surfaces was suggested, without the shear, in [17], but the additional control provided by the shear turns out to be *crucial*.

- The idea is that control over the shear allows clear control over how the image of $\rho(S^0)$ in $\mathbb{P}SL(2, \mathbb{C})$ acts as we perturb the pairs of pants. A shear close to 1 corresponds to a geometric progression in the way lengths in a tiling of $\mathbb{H}^2$ (the universal cover of the surface) vary. As such, the map corresponds to a quasiconformal map of the boundaries which turns out to be a small enough perturbation of a conformal one to ensure conjugacy.

- The result is a useful characterisation of how the Fenchel-Nielsen coordinates of a representation affect the image of $S^1 = \partial\mathbb{H}^2$ of the induced map $\partial\mathbb{H}^2 \to \partial\mathbb{H}^3 = S^2$.

The rest of this chapter focusses on constructing a good panted surface - which proves the surface subgroup theorem. In and of itself, it is a demonstration of the extreme level of control over geodesic structures given by exponential mixing of the frame flow.

4.3 Piecing together a surface

Here, we outline the approach to be used to construct a good panted surface. We begin by summarising the constructions and their important features.

4.3.1 Any surface will do

Let's recall from above:

- Viable surfaces in $\mathbf{M}$, with their pants decomposition where the complex half-length of each cuff is the same for each pants it's in

- $N^1(\sqrt{\gamma})$ is a complex torus corresponding to the unit normal bundle of a geodesic associated to cuff. We call the disjoint union of these $N^1(\sqrt{\Gamma})$
- foot_γ is a map defined on all the pants in $\mathbf{M}$, ρ such that $\rho(C) \simeq \gamma$ for some cuff C, and takes values in $N^1(\sqrt{\gamma})$. Alternatively,

  $$\text{foot} : \mathbf{\Pi} \to N^1(\sqrt{\Gamma})$$

  where $\mathbf{\Pi}$ is the set of skew pants in $\mathbf{M}$ (defined below).

We now wish to consider representations of pants up to conjugacy

$$\rho : \pi_1(\Pi) \to \mathbb{P}SL(2, \mathbb{C})$$

and we ask: if we have a collection of mappings ρ, what properties does it require in order to be part of a viable representation (and eventually, a good panted surface)?

The definition below is a preliminary one:

Definition 4.3.1. ρ is an admissible representation if it is faithful, and if $\rho(C)$ is loxodromic for each cuff, and

$$\mathbf{hl}(C) = \frac{\mathbf{l}(C)}{2}$$

We say $[\rho] = \Pi$, a skew pants, and write $\mathbf{\Pi}$ for the set of all skew pants, and we let $\gamma^*(\Pi, C)$ be the oriented geodesic corresponding to $\rho(C)$. The pair (Π, γ^*) (where γ^* corresponds to a cuff of Π) is called a marked skew pants, and the set of all these is denoted by $\mathbf{\Pi}^*$.

We need this definition of skew pants so that control over the complex *length* of the geodesic corresponding to it (which we will gain using exponential mixing) will give us control over the complex *half*-length of the cuff.

Constructing a gluing

Suppose we are given a finite collection of marked skew pants, corresponding to labels $\{1, \ldots, n\}$ and a permutation $\sigma \in S_n$. We could hope to construct a surface by gluing

$$i \to \sigma(i)$$

We need the following things to be able to construct a surface:

- If $j = (\Pi, \gamma^*(\Pi, C_i))$, then there must be a k with $(\Pi, \gamma^*(\Pi, C_{i+1}) = k$; i.e. for each cuff of a pant present in our labels, the other cuffs are also present
- We need obviously that $\sigma^2 = \text{id}$. Moreover, if $i = (\Pi, \gamma^*)$, then $\sigma(i)$ must equal $(\Pi', -\gamma^*)$ (so we only glue cuffs with the same length). We call such a σ an admissible involution

Each collection of cuffs for a marked skew pants determines a topological pants Π^0. σ provides instructions for gluing, and we get a closed surface S^0 with pants decomposition given by the cuffs. Furthermore, we determine the representation $\rho : S^0 \to \mathbf{M}$ since we know the geodesics corresponding to the generators.

Remark. Our surface might not be connected. Furthermore we haven't yet controlled the shear map for the representation, which corresponds to the twist in gluing.

We want to find a suitable collection of pants and a map σ such that the corresponding surface is a good panted surface, for which we need control of the shear.

4.3.2 Perfection

We want tighter control over the gluing above to help us construct a good panted surface.

Recall that $N^1(\sqrt{\gamma}) \equiv \mathbb{C}/(\frac{\mathrm{l}(\gamma)}{2}\mathbb{Z} + 2\pi i\mathbb{Z})$ is a complex torus. We now define a map on $N^1(\sqrt{\Gamma})$, their disjoint union, that will help to characterise the shear. We fix the Euclidean metric on each complex torus and denote it by dis.

Definition 4.3.2. Let γ be a closed geodesic in $\mathbf{M}$, and $\tilde{\gamma}$ the lift to $\mathbb{H}^3$. WLOG, γ is the geodesic between 0 and ∞. For $\zeta \in \mathbb{C}$, define $\mathcal{A}_\zeta$ as the induced map on γ by the loxodromic transformation $z \mapsto e^\zeta z$. Also, we denote by $\mathcal{A}_\zeta$ the induced action on $N^1(\sqrt{\gamma})$ and on the disjoint union $N^1(\sqrt{\Gamma})$

Remark. It is a simple but important observation that the $\mathcal{A}_\zeta$, which are just translations on each torus, preserve Lebesgue measure on each torus. Indeed, the only measure that is invariant under all of the $\mathcal{A}_\zeta$ is one which restricted to each torus is a multiple of Lebesgue measure.

The significance of these maps comes from the following proposition, coupled with the previous remark.

Proposition 4.3.1. Let ρ be a viable surface in $\mathbf{M}$. Then for a cuff C in pants Π and Π',

$$s(C) = 1 \Leftrightarrow \mathrm{foot}(\Pi, C) = \mathcal{A}_{1+i\pi}\,\mathrm{foot}(\Pi', C)$$

Furthermore, there exists a $\delta > 0$ such that $\mathrm{dis}(\mathrm{foot}(\Pi, C), \mathrm{foot}(\Pi', C)) < \delta$ implies that

$$s(C) = \mathrm{dis}(\mathrm{foot}(\Pi, C), \mathrm{foot}(\Pi', C))$$

The idea is to quantify equidistribution of pants, by essentially studying the pushforward of measures on marked skew pants onto $N^1(\sqrt{\Gamma})$. We'll construct a measure on skew pants such that the induced measure is 'close' to the Lebesgue measure, and hence show that there are 'enough' pants to construct a good panted surface.

Definition 4.3.3. Let (X, d) be a metric space, and μ and ν Borel measures with compact support. We say that μ and ν are δ-equivalent and write $\mu \overset{\delta}{=}$ if: $\mu(X) = \nu(X)$ and

$$\mu(A) \leq \nu(\mathcal{N}_\delta(A)) \text{ for all measurable } A$$

where $\mathcal{N}_\delta(A)$ is the thickening of A to a δ-neighbourhood

Lemma 4.3.1 (Matching lemma)**.** *Let (X, d) be a metric space, and $f, g : \{1, \ldots, n\} \to X$ a map. Denote by p the counting measure on X. Suppose*

$$f_* p \overset{\delta}{=} g_* p$$

Then there exists a bijection $H : \{1, \ldots, n\} \to \{1, \ldots, n\}$ such that

$$d(f \circ H(j), g(j)) \leq \delta$$

The proof is a simple application of Hall's Marriage theorem.

Remark. We can define an involution $\sigma_H \in S_n$ by setting $\sigma_H(i) = H(i)$ and $\sigma^2 = \mathrm{id}$.

Suppose we have a finite measure on pants μ, which is only nonzero on marked pants with cuff lengths close to $R/2$ and such that the push-forward by the foot map restricted to each complex torus is close to a multiple of Lebesgue measure on it (with the Euclidean metric on each complex torus). We'll use a rationalisation procedure to get a map $f : \{1, \ldots, n\} \to \mathbf{\Pi}^*$ such that $f_* p = \mu$. Then, $\mathrm{foot}_* f_* p$ will be close to a multiple of Lebesgue measure on each complex torus. In particular, $\alpha = \mathrm{foot}_* f_* p$ and $\mathcal{A}_{1+i\pi *}\,\mathrm{foot}_* p$ will be δ-close. Then, we use the *matching lemma* to define an involution σ such that

$$d(\mathrm{foot}(\sigma(i)), \mathcal{A}_{1+i\pi}\,\mathrm{foot}\, f(i)) < \delta$$

and we'll be able to piece together a good panted surface using the technique above.

We now give the details of the above procedure. Constructing the measures described is the main power for the proof of the surface subgroup theorem.

Theorem 4.3.1 (Measures)**.** *There exists $q > 0$ and $D_1, D_2 > 0$ so that for every $1 \geq \epsilon > 0$ and R sufficiently large, there is a finitely supported measure μ on $\mathbf{\Pi}$ such that $\mu(\Pi) > 0$ means that*

$$|\,\mathbf{hl}(C) - \frac{R}{2}| \leq \epsilon$$

and

$$\text{foot}_* \,\mu|_{N^1(\sqrt{\gamma})} \text{ and } K\lambda \text{ are } D_2\left(Re^{-qR} + e^{-\frac{R}{8}}\right)\text{-equivalent}$$

where λ denotes the standard Lebesgue measure on each complex torus $N^1(\sqrt{\gamma})$

The proof of this theorem is the focus of the next section and the heart of the proof of the surface subgroup theorem. Indeed, the challenging part of the proof will be to show the equivalence to Lebesgue measure, and it will require the full power of exponential mixing of this report. As the proof is very technical, we aim to focus on developing an understanding of it instead of verifying the technical detail.

4.4 Constructing the measures

The example to bear in mind is the construction of geodesics from the first part of this chapter. However, we need control over the complex lengths of the geodesic curves formed, and hence need to look at the frame flow, instead of the geodesic flow.

Recall the frame bundle of $\mathbf{M}$ is $\mathcal{F}(\mathbf{M})$, and its Liouville measure Λ, which is preserved by the frame flow and isometries. Equivalently, it's $\pi_1(\mathbf{M})\backslash F(\mathbb{H}^3)$. We'll refer to elements of $\mathcal{F}(\mathbf{M})$ using (p, u, n). To piece together pants, we will use the following

Definition 4.4.1. A tripod in $\mathbf{M}$ is the triple of frames

$$(F, \omega(F), \omega^2(F))$$

where $\omega(p, u, n) = (p, \omega(u), n)$ where ω is rotation clockwise by $2\pi/3$ perpendicular to n. Write $\bar{\omega} = \omega^{-1}$ An antitripod is the triple

$$(F, \bar{\omega}(F), \bar{\omega}^2(F))$$

A bipod is the pair

$$(F, \omega(F))$$

and an antibipod is defined similarly.

We now outline the strategy for constructing the measures. Refer to the example of geodesics from the beginning of this chapter to motivate the construction: it is particularly insightful to do so.

- Define an affinity function on pairs of frames in $\mathbb{H}^3$ and use it to define an affinity function on pairs of frames in $\mathbf{M}$.
- The affinity function will be greater than 0 only if we can use the tripods determined by the frames are to construct a pair of pants with cuff lengths ϵ-close to $R/2$ (such tripods are called *well-connected*, and the bipods attained by forgetting corresponding legs are also well-connected as bipods). We use this to determine a measure on pants μ, realised as the pushforward of the product of Liouville measures on the frames, scaled by the affinity, by the map sending a pair of well connected tripods to the skew pants it determines.

- Construct a foot predicting map, which given a pair of bipods, predicts the foot of the pair of pants and use it to construct a measure β realised as the pushforward by this foot predicting map of the bipods attained by forgetting a leg of well-connected tripods. We'll be able to bound the Radon-Nikodym derivative of this measure with respect to the Lebesgue measure using properties of the affinity function, which we'll control using exponential mixing of the frame flow.
- Control over the Radon-Nikodym derivative will give, through a basic computation, that the measure is β restricted to each torus δ-equivalent to Lebesgue measure.
- The foot predicting map will have the property that, given a pair of well-connected tripods, the predicted foot gained by ignoring one leg is $O(e^{ar})$ close to the actual foot. Thus the pushforward by the foot map of μ will be $O(e^{ar})$ close to β when restricted to each individual complex torus.
- Transitivity of the measure equivalence relation will prove the theorem.

Remark. There are only finitely many pants with cuff lengths ϵ close to $R/2$. In particular, the measure on skew pants will be finite and atomic.

4.4.1 The affinity function on $\mathbb{H}^3$

We denote by r a function of R tending to ∞ as $\mathbb{R} \to \infty$. It's used to notationally simplify the estimates. At this stage, its precise value is not important.

We fix a frame in $\mathcal{F}(\mathbb{H}^3)$, say F_0. We have a natural distance function (not a metric!) on $\mathcal{F}(\mathbb{H}^3)$ given by

$$D((p_1, u_1, n_1), (p_2, u_2, n_2)) = d(p_1, p_2) + \Theta(u_1 @ p_2, u_2) + \Theta(n_2 @ p_2, n_2)$$

Where $\Theta(x, y)$ denotes the unsigned angle between x and y and $u_1 @ p_2$ denotes the parallel transport of u_1 along the geodesic segment joining p_1 and p_2.

Let $N_\epsilon(F_0)$ denote the ϵ neighbourhood with respect to D. Since $\mathbb{P}SL(2, \mathbb{C})$ acts by isometries, it preserves D. Hence if $\alpha(F_0) = F_1$ then $\alpha(N_\epsilon(F_0)) = N_\epsilon(F_1)$. Let f_ϵ denote a smooth function, supported in N_ϵ and normalised such that

$$\int_{\mathcal{F}(\mathbb{H}^3)} f_\epsilon(X) d\Lambda = 1$$

Remark. $\mathbf{M}$ is compact, so we can find an ϵ_0 such that the projection $\mathbb{H}^3 \to \mathbf{M}$ and hence $\mathcal{F}(\mathbb{H}^3) \to \mathcal{F}(\mathbf{M})$ is injective. Assume $\epsilon < \epsilon_0$ from now on.

Suppose $gF = F_0$ for $F_0, F \in \mathcal{F}(\mathbb{H}^3)$ we define

$$f_\epsilon(F) = g^* f_\epsilon$$

Definition 4.4.2. The affinity function $\mathbf{a} : \mathcal{F}(\mathbb{H}^3) \times \mathcal{F}(\mathbb{H}^3) \to \mathbb{R}$ is given by

$$\mathbf{a}(F_1, F_2) = \int_{\mathcal{F}(\mathbb{H}^3)} g^*_{r/2} f_\epsilon(g_{r/4} F_1) \cdot f_\epsilon(\pi g_{r/4} F_2) d\Lambda$$

where π denotes the reflection of the frame (p, u, n) in the plane perpendicular to u.

Remark. There is something present here that wasn't in the example of geodesics. We have an initial flow of $\frac{r}{4}$. Clearly, this doesn't affect mixing, but the initial flow will be used to avoid awkward situations with negative lengths.

4.4.2 The induced affinity function on M and tripods

Given frames $F, G \in \mathcal{F}(\mathbf{M})$, and a geodesic segment joining them, we write $\tilde{F}$ for a lift of $\mathcal{F}$ and $\tilde{G}$ a lift of G along γ. We define

$$\mathbf{a}_\gamma(F, G) = \mathbf{a}(F, G)$$

Remark. $\tilde{F}$ depends only on the homotopy class of γ, so it is enough to consider only geodesics.

Definition 4.4.3. The affinity function on $\mathcal{F}(M)$ is given by

$$\mathbf{a}(F, G) = \sum_\gamma \mathbf{a}_\gamma(F, G)$$

where the sum is taken over all geodesic segments joining the base points of F an G. For tripod pairs, $T_1 = (F_1, \omega F_1, \omega^2 F_1)$ and $T_2 = (F_2, \bar{\omega} F_1, \bar{\omega}^2 F_1)$ we define the affinity for a triple of geodesics $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ connecting them by

$$b_\gamma = \prod_i a_{\gamma_i}(\omega^i F_1, \bar{\omega}^i F_2)$$

and similarly for bipod pairs

Remark. By considering a fundamental region D for $\mathbf{M}$, and $\mathcal{F}(D)$, we see that only finitely many γ provide nonzero affinity

The point of the affinity functions is that if the affinity between tripods is nonzero, then there are geodesic segments γ_i joining each one, such that $\mathbf{l}(\gamma_i)$ is small, and both $\Theta(\omega^i(u_1), \gamma_i')$ and $\Theta(\bar{\omega}^i(u_2), (-\gamma)')$ are exponentially small. Basic hyperbolic geometry will give us a way of creating a pants from well connected tripods, which is briefly described below. But, significantly, we have the following crucial proposition

Proposition 4.4.1. For $F_p, F_q \in \mathcal{F}(\mathbf{M})$ (frames based at p and q respectively)

$$\mathbf{a}(F_p, F_q) = \frac{1}{\Lambda(\mathcal{F}(\mathbf{M}))} + O(e^{-qr})$$

Proof. By our choice of ϵ (so that the projection from $\mathbb{H}^3$ was injective), we have that

$$\int_{\mathcal{F}(\mathbf{M})} f_\epsilon d\Lambda = \frac{1}{\Lambda(\mathbf{M})}$$

By exponential mixing of the frame flow for closed hyperbolic manifolds (*Corollary 2.1.2*), we have

$$[\Lambda(\mathcal{F}(\mathbf{M}))]^2 \int_{\mathcal{F}(\mathbf{M})} (g_r^* f_\epsilon) f_\epsilon d\Lambda = 1 + O(e^{-qr})$$

where $q \geq -1/H(M)$ where $H(M)$ is as defined in section 2 depending on the Cheeger constant of $\mathbf{M}$.

Since each geodesic linking p and q corresponds to deck transformations, we have that

$$\sum_\gamma \mathbf{a}_\gamma(F_p, F_q) = \sum_\gamma \mathbf{a}(\widetilde{F_p}, \gamma \widetilde{F_q}) := \int_\Gamma \int_{\mathcal{F}(\mathbb{H}^3)} \widetilde{\Phi}(\gamma \widetilde{F_q}) d\Lambda d\gamma = \Lambda(\mathcal{F}(\mathbf{M})) \int_{\mathcal{F}(M)} \Phi(F_q) d\Lambda$$

where

$$\Phi(F_q) = g_{r/2}^* f_\epsilon g_{r/4} F_p \cdot f_\epsilon(\pi g_{r/4} F_q)$$

Combining the two equations gives the result.

4.4.3 Well connected tripods

We give a brief outline of the construction of a good pants from a well-connected tripod, and state without proof some of the key estimates.

The key idea comes from the following lemma and its corollaries (which will not be stated here

Lemma 4.4.1 (Chain Lemma)**.** *For $a, b, c \in \mathbb{H}^3$ and $v \in T^1_a(\mathbb{H}^3)$, we have the following inequalities*

$$\Theta(v@b@c@a, v) \leq Area(abc) \leq |bc|$$

where Area(abc) is the area of the triangle abc.

Proof. The Gauss-Bonnet theorem with boundary components implies the first inequality and a computation in $\mathbb{H}^2$ the second.

The idea is that if we have a chain of geodesic segments, such that the angle at each link in the chain is small, then we have good control over how close the resulting chain is to a geodesic.

Remark. This lemma and its corollaries are technical because distances in hyperbolic space between geodesics grow exponentially. To see just how crucial this lemma is, the reader is encouraged to review the remarks following the example of the geodesics and attempt to generalise the proof.

Given a well pair of tripods connected tripod, we construct a skew pants by taking the representation determined up to conjugation by the homotopy classes of the geodesics as shown in the diagram below

Control over the parallel transport of the normal vector, as mentioned above, is control over the complex half lengths of the geodesic segments. Combining these facts, we get:

Lemma 4.4.2. *For two well-connected tripods, T_1, T_2, connected by the triple γ. Then there is a well defined skew pants $\rho(T_1, T_2, \gamma)$ corresponding to them. Moreover,*

$$|\,\mathbf{hl}(C) - r + \log\frac{4}{3}| \leq D\epsilon$$

for some constant $D > 0$.

Remark. The chain lemma is crucial for this.

Remark (Some perspective)**.** Let T_p and T_q be tripods corresponding to frames F_p and F_q. (Perhaps un at this stage of this report) Surprisingly, from the fact that

$$b(T_p, T_q) \geq a(F_p, F_q) \cdot a(\omega F_p, \bar{\omega} F_q) \cdot a(\omega^2 F_p, \bar{\omega}^2 F_q) = \frac{1}{[\Lambda(\mathcal{F}(\mathbf{M})]^3} + O(e^{-q'r}) > 0$$

we have shown (modulo a few computations from hyperbolic geometry) that $\mathbf{M}$ has *some* good pants - in fact, many (since our tripods were arbitrary). This is highly non-obvious! It has required the full power of the representation theory of $\mathbb{P}SL(2, \mathbb{C})$.

Let the map that assigns to a pair of well connected tripods the corresponding skew pants, each with half lengths of cuffs $D\epsilon$-close to $R = 2(r - \log \frac{4}{3})$ be

$$\pi : X = \mathcal{F}(\mathbf{M}) \times \mathcal{F}(\mathbf{M}) \times \Gamma^3 \to \mathbf{\Pi}_{D\epsilon,R}$$

We define the measure $\tilde{\mu}$ on X by

$$d\tilde{\mu}(T_p, T_q, \gamma) = \beta_\gamma(T_p, T_q) d\Lambda(F_p) \times d\Lambda(F_q) \times d\gamma^3(\gamma)$$

where $d\gamma$ is just counting measure on discrete set.

Remark. Expressing the measures in this way provides a particularly convenient way of computing Radon-Nikodym derivatives

Finally, we can now define μ on $\mathbf{\Pi}$ by $\pi_* \tilde{\mu}$.

What is left is to show that $\mathrm{foot}_* \mu$ is close to the Lebesgue measure on each torus $N^1(\sqrt{\delta})$

4.4.4 Interlude

How might we approach the problem of showing that the measure we constructed is close to (a multiple of) the Lebesgue measure on the torus, and why might we expect it to be?

Exponential mixing has given us an abundance of well-connected tripods - and we need to capture the symmetry given that it develops, so we compare it to the 'maximally symmetric' measure on each torus. The key property that we need of the Lebesgue measure is that it's the unique measure (up to scaling) invariant under the torus action. Already above, we showed that there were many good pants in $\mathbf{M}$ and the way we did it was by 'pulling tight' the corresponding geodesics.

If we fix one of the geodesics δ from a well connected tripod, the feet of the resulting pants in $N^1(\sqrt{\delta})$ are really determined by the positioning of frames in relation to δ. But, given such a δ we can only really say that: if it came from a pair of tripods, then they are in a specific torus-shaped region around it. We want to really characterise this 'entropic' behaviour using the exponential mixing.

4.4.5 Predicted feet

Recall that for a well connected pair of tripods, with geodesic segments γ_i linking $\omega^i(F_p)$ and $\bar{\omega}^i(F_q)$, the skew pants $\pi(T_p, T_q, \gamma)$ has cuffs C_i in the conjugacy classes of geodesics δ_i homotopic to $\gamma_{i-1} \cup \gamma_{i+1}$.

We'll now define a 'geometric foot map' as follows. Consider

$$B_p = (F_p, \omega(F_p)) \text{ and } B_q = (F_q, \bar{\omega}(F_q))$$

This is a well connected pair of bipods. Define the geodesic δ as the one in the conjugacy class of $\gamma_0 \cup \gamma_1$ as shown in the diagram below.

Let $f_t^i \in N_1(\delta)$ be the normal to δ as shown above, and let f^i be the value for the limiting geodesic. Where h^i is is the foot of the orthogonal along δ to δ_i, and by construction

$$\mathbf{d}_\delta(h^0, h^1) = \mathbf{hl}(\delta_1)$$

Lemma 4.4.3. $\mathbf{d}_\delta(f^0, f^1) = \mathbf{hl}(\delta)(f^0, f^1)$

Definition 4.4.4. For a well connected (by triple γ), tripod T_p and T_q, the predicted foot along δ is

$$\mathbf{f}_\delta(T_p, T_q, \gamma) = [f^0] \in N^1(\sqrt{\delta})$$

The point is, that since we have a well connected tripod, we know where $\beta^{r/4}$ is: in particular, it's close by a constant to δ_2 . Using the fact that distances between geodesics grow exponentially in the hyperbolic plane, we get the following estimate:

Lemma 4.4.4. *Let* $(\pi(T_p, T_q, \gamma), \delta^*) \in \mathbf{\Pi}^*$. *Then for* r *large and* ϵ *small,*

$$\mathbf{d}(\text{foot}_\delta(\pi(T_p, T_q, \gamma), \mathbf{f}_\delta(T_p, T_q, \gamma)) < De^{-r/4}$$

We denote by S_δ well-connected bipods such that the corresponding closed geodesic is δ; that is

$$S_\delta = \{(F_p, F_q, \gamma_0, \gamma_1) : (B_p, B_q) \text{ is well connected along } \gamma = (\gamma_0, \gamma_1) \text{ and } [\gamma_0 \cup \gamma_1] = \delta\}$$

and thus we have the map $\mathbf{f}_\delta : S_\delta \to N^1(\sqrt{\delta})$. Denote by χ the following map:

$$(T_p, T_q, (\gamma_0, \gamma_1, \gamma_2)) \mapsto (B_p, B_q, (\gamma_0, \gamma_1))$$

where, keeping with our slight abuse of notation, B_p is the bipod taken by forgetting $\omega^2 F_p$ and B_q forgetting $\bar{\omega}^2 F_q$. We define $C_\delta = \chi^{-1} S_\delta$

Remark. The predicted foot map satisfies

$$|\mathbf{f}_\delta \circ \chi(T_p, T_q, \gamma) - \text{foot}_\delta(T_p, T_q, \gamma)| < De^{-r/4}$$

This will imply that $\text{foot}_* \mu$ and β_δ (defined below) are $O(e^{-r/4})$ equivalent.

Define $\mathbb{T}_\delta$ to be the solid torus that covers δ (i.e. $\langle\delta\rangle \backslash \mathbb{P}SL(2, \mathbb{C})$). Each well connected bipod in S_δ has a unique lift to $\mathbb{T}_\delta$, and so the automorphism group of $\mathbb{T}_\delta$, which is the complex torus

$$G = \mathbb{C}/(2\pi i\mathbb{Z} + l(\delta)\mathbb{Z})$$

acts on $\mathbb{T}_\delta$. Now $\mathbb{N}^1(\sqrt{\delta})$ has a natural G action by isometries, from its construction. So we define the action on S_δ so that the predicted foot map $\mathbf{f}_\delta$ becomes equivariant.

We define ν_δ a measure on S_δ by

$$d\nu_\delta(B_p, B_q, \gamma_0, \gamma_1) = \mathbf{a}_{\gamma_0}(F_p, F_q)\mathbf{a}_{\gamma_1}(\omega(F_p), \bar{\omega}(F_q))d\lambda_B(B_p, B_q, \gamma_0, \gamma_1)$$

where λ_B is the obvious measure, and we define the collection of measures ζ_δ as

$$\zeta_\delta = \mathbf{f}_{\delta *}\nu_\delta$$

The construction of the affinity functions means that they are equivariant with respect to the action by isometries, and hence so is λ_B. Thus ζ_δ is equivariant with respect to the torus action on $N^1(\sqrt{\delta})$ and thus by uniqueness of Haar measure, we have that

$$\zeta_\delta = E_\delta \lambda_\delta$$

where λ_δ denotes the Lebesgue measure on the torus $N^1(\sqrt{\delta})$.

Now from the construction we have

$$\left|\frac{d\chi_*(\tilde{\mu}|_{C_\delta})}{d\nu_\delta}\right| = \mathbf{a}(\omega^2(F_p), \bar{\omega}^2(F_q))$$

From the exponential mixing, we have that the total affinity between those frames

$$\mathbf{a}(\omega^2(F_p), \bar{\omega}^2(F_q)) = K + O(e^{-qr})$$

and so

$$\left|\frac{d\chi_*(\tilde{\mu}|_{C_\delta})}{d\nu_\delta} - K\right| < Ce^{-qr}$$

And so

$$K(1 - Ce^{-qr})\nu_\delta \leq \chi_*(\tilde{\mu}|_{C_\delta}) < K(1 + Ce^{-qr})\nu_\delta$$

And in particular:

$$KE_\delta(1 - Ce^{-qr})\lambda_\delta \leq \underbrace{\mathbf{f}_{\delta*}\chi_*\tilde{\mu}|_{C_\delta}}_{\beta_\delta} \leq KE_\delta(1 + Ce^{-qr})\lambda_\delta$$

To finish, we need the following technical lemma, which is really an easy computation

Lemma 4.4.5. *Let μ be a Borel measure on $\mathbb{C}/(a\mathbb{Z} + b\mathbb{Z})$ with $\frac{d\mu}{d\lambda}$ a continuous function such that*

$$K \leq \left|\frac{d\mu}{d\lambda}\right| < K(1 + \delta)$$

Then μ and λ are $4\delta(|a| + |b|)$ equivalent measures.

Thus we have that β, given by $\beta|_{N^1(\sqrt{\delta})} = \beta_\delta$ and λ, which denotes Lebesgue measure restricted to each $N^1(\sqrt{\delta})$ are $O(e^{-qr})$ equivalent measures.

But now, by the remark following the construction of the predicted foot map, we have that $\text{foot}_* \, \mu|_{N^1(\sqrt{\delta})}$ and β_δ are $O(e^{-r/4})$ equivalent, which proves that the measure μ is as required in the measures theorem.

4.4.6 Summary of the proof

We now have a finite measure on skew pants μ, such that

$$\text{foot}_* \, \mu|_{N^1(\sqrt{\delta})} \text{ and } \lambda_\delta$$

are $O(e^{-ar})$ equivalent.

Proposition 4.4.2. μ determines a labelling from $\mathcal{L} : [n] = \{1, \ldots, n\} \to \mathbf{\Pi}$. Furthermore, there is a bijection $h : [n] \to [n]$ such that

$$\text{dis}(\mathcal{L}(k), A_{1+i\pi}L(h(k)) < Ke^{-ar}$$

This bijection satisfies the combinatorial conditions required to make a valid permutation of the pants to construct a surface. In particular, we have constructed a good panted surface.

The construction of a labelling from a measure is a standard rationalisation procedure, which is possible since μ is an atomic measure.

There is a minor technicality that have been omitted in our exposition. Chiefly, there are some additional basic symmetry conditions that the measure μ must satisfy in order that the permutation we defined works combinatorially. However, these symmetries can be restated in terms of actions of the form $\mathcal{A}_\zeta$ on each complex torus of $N^1(\sqrt{\Gamma})$, and being $O(e^{-ar})$ close to Lebesgue measure when restricted to each torus gives these additional symmetries to us for free. For details, see [9, Section 3].

The matching lemma (*Lemma 5.3.1*) and the characterisation of the shear (*Proposition 5.3.1*) shows that the resulting pants have a shear $O(e^{-ar})$ close to 1. Furthermore, by the nature of the construction of μ, it is nonzero on only those pants whose cuff lengths were ϵ close to $R/2$. Thus by *Theorem 4.2.1* (and observing that $Ke^{-ar} < \epsilon/r$ for large r, we have proven the surface subgroup theorem.

Bibliography

[1] S. Helgason. *Differential Geometry and Symmetric Spaces.* Elsevier, Volume XII, 1962.

[2] Y. Caiza. *Analysis on Manifolds via the Laplacian.* Lecture notes, http://www.math.harvard.edu/ canzani/docs/Laplacian.pdf

[3] G. Warner. *Harmonic Analysis on Semisimple Lie Groups.* Springer-Verlag, 1972.

[4] M. Lackenby *Hyperbolic Manifolds* Lecture notes, Oxford University, http://people.maths.ox.ac.uk/lackenby/, 2000.

[5] G. Warner. *Ergodic theory with a view towards number theory.* Graduate Texts in Mathematics, 259. Springer-Verlag, 2011.

[6] A. Eskin, C. McMullen. *Mixing, counting, and equidistribution in Lie groups.* Math. J. 71, no. 1, 181-209, 1993.

[7] C. Moore. *Exponential decay of correlation coefficients for geodesic flows.* Group representations, ergodic theory, operator algebras, and mathematical physics. Berkeley, Calif. 1984), 163-181, Math. Sci. Res. Inst. Publ., 6, Springer, New York, 1987.

[8] J-H Eschenburg. *Lecture notes on symmetric spaces.* Uni-Augsburg, www.math.uni-augsburg.de/ eschenbu/symspace.pdf, 2009.

[9] J. Kahn and V. Markovic. *Immersing almost geodesic surfaces in a closed hyperbolic three manifold.* Ann. of Math. , Volume 175 , pp 1127-1190, 2012.

[10] M. Ledoux *A simple analytic proof of an inequality by P. Buser.* P.N.A.S. 121, 3, 1994.

[11] A. Eskin, Z. Rudnick, P. Sarnak. *A proof of Siegel's Weight Formula.* International Mathematics Research Notices, No. 5, 1991

[12] V. Markovic. *The Good Pants Homology and the Ehrenpreis Conjecture.* Presentation slides, http://www.maths.dur.ac.uk/events/Meetings/LMS/2011/GAL11/talks/markovic.pdf

[13] I.M. Gel'fand, M.I. Graev, I. I. Pyatetskii-Shapiro. *Representation Theory and Automorphic Forms.* Translated from the Russian by K.A. Kirsch. W.B. Saunders Company, 1969.

[14] V. Markovic. *Lectures on the Ehrenpreis conjecture.* Aarhus University, 2012.

[15] M. Pollicott *Exponential mixing for the geodesic flow on hyperbolic 3 manifolds.* Journal of Stat. Phys, Vol 67, Nos. 3/4, 1992.

[16] D. Dolgopyat. *On decay of correlations in Anosov Flows.* Ann. of Math. (2), 147(2):357-390, 1998.

[17] W. Bowen *Weak Forms of the Ehrenpreis conjecture and the Surface Subgroup conjecture.* arXiv:math/0411662, 2004