IMPERIAL

Imperial College London

Department of Mathematics

Second-year Group Research Project

Dynamics on Homogeneous Spaces: Ratner's Theorems and Applications

Authors:
Ariff Jazlan Johan (CID: 02204042)
Chenyang Zhao (CID: 02203280)
Meiyi Li (CID: 02214767)
Zeming Hao (CID: 02040512)

Supervisor:
Dr. Marie-Amelie Lawn

June 17, 2024

Abstract

In this paper, we study homogeneous dynamics with a focus on the behaviour of flows on finite volume homogeneous spaces which arise from lattices. We provide an elementary introduction to manifolds, Lie theory, and homogeneous spaces. We build towards Ratner's Theorems which establish the rigidity of unipotent flows by classifying ergodic measures. They serve as an interface between homogeneous dynamics and Diophantine approximation, which we illustrate with a proof of the Oppenheim-Davenport Conjecture on quadratic forms.

Acknowledgements

We would like to thank our supervisor Dr. Lawn and her PhD student Diego for their consistent support and guidance throughout this project. We would also like to acknowledge the use of ChatGPT to write code, run elementary computations, and deal with formatting issues.

Contents

1 Introduction

Classically, dynamical systems are studied in Euclidean space $\mathbb{R}^n$ where we can readily define differential structures. The most natural extension of this idea is to the setting of manifolds – smooth spaces that locally resemble Euclidean space. We face significant obstructions to studying manifold dynamics without a firm background in differential geometry, but this can be ameliorated if we can algebraically describe motions in the manifold. When this is the case, we are working in the setting of **homogeneous spaces** which are manifolds on which a group acts transitively.

In the latter half of the 20th century, there was a surge of interest in the dynamics of unipotent flows on homogeneous spaces, largely motivated by Margulis' striking proof of the Oppenheim-Davenport Conjecture on Diophantine approximation by quadratic forms in 1986 [1]. The core idea that motivated the field was that **unipotent flows** are so well-behaved as to be 'rigid', and studying the ergodic probability measures invariant under this flow revealed a surprisingly large amount of information about the 'spatial' and 'temporal' niceness of their orbits. Largely independently of the work of other mathematicians, Marina Ratner would go on to prove a sweeping generalisation of this rigidity result on unipotent flows [2], confirming conjectures of Raghunathan from the 1970s [3]. Ratner Theory has gone on to inspire Fields Medal-winning work such as that of Marzakhani and Eskin who established analogous results on *moduli spaces* – which are profoundly more complicated in that they are totally inhomogeneous [4].

In Section 2, we formally introduce Lie groups G by endowing smooth manifolds with a group structure and use these to construct homogeneous spaces G/Γ. These have finite volume when Γ is a lattice, which we articulate by introducing the language of invariant measures. In Section 3, we extend the classical theory of dynamical systems to the setting of homogeneous spaces, with an emphasis on geodesic and unipotent flows, and their ergodic properties. In Section 4, we state Ratner's Theorems using the torus as a motivating example. In Section 5, we present a proof that the Measure Classification Theorem implies the Orbit Closure and Equidistribution Theorem, which relies on understanding the 'polynomial behaviour' of unipotent flows. In Section 6, we (partially) complete this proof by proving Measure Classification in the special case of $G = SL(2, \mathbb{R})$ by understanding the direction in which two nearby orbits diverge from one another. In Section 7, we apply Ratner's Theorems to prove the Oppenheim-Davenport conjecture on quadratic forms.

2 Preliminaries

2.1 Manifolds

To the average human, it is tempting to believe that the Earth is flat as wherever you go the ground looks flat. If you were to walk across a continent, using a map to chart your journey in each country, you could easily be lead to believ that patching the (flat) maps together would imply that the whole Earth is flat. Regardless, viewing the Earth from on high reveals the spherical nature of our planet. This apparent disagreement in the local and global dimensions lies at the heart of manifold theory. We refer to [5] for the following:

Definition 2.1.1. A **(smooth) manifold** of dimension n is a Hausdorff topological space M with a collection of pairs (U_α, ϕ_α) where U_α is an open subset of M and $\phi_\alpha : U_\alpha \to \mathbb{R}^n$ so that:

(a) Each ϕ_α is a homeomorphism of U_α onto an open subset V_α of $\mathbb{R}^n$.

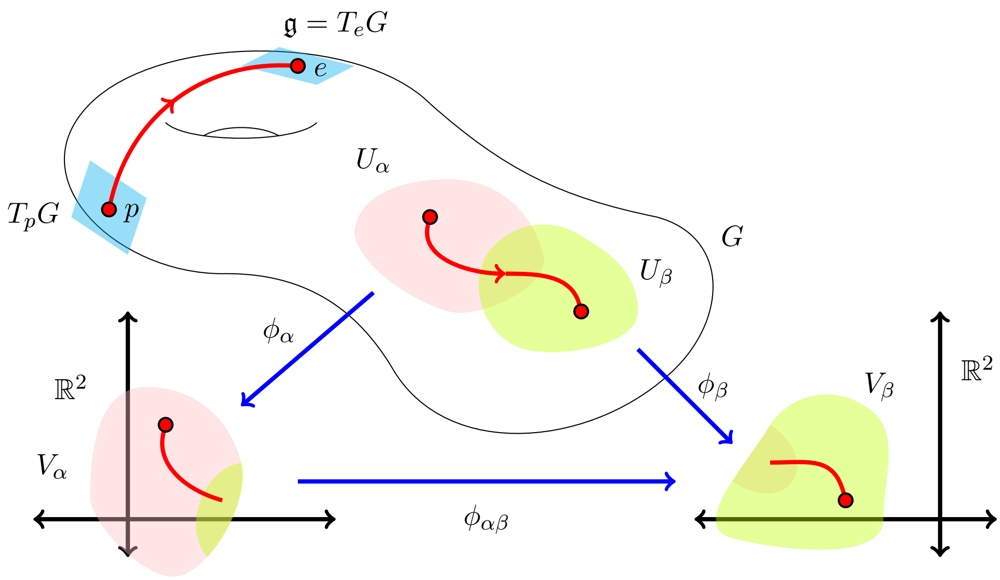


Figure 1 The torus $\mathbb{T}^2$ is both a group and a 2-manifold.

(b) $\cup_\alpha U_\alpha = M$.

(c) For every α, β, the transition functions $\phi_{\alpha\beta} = \phi_\beta \circ \phi_\alpha^{-1}\colon \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$ are smooth, in the sense of smooth (i.e. differentiable) functions between subsets of $\mathbb{R}^n$. In this case the charts (U_α, ϕ_α) and (U_β, ϕ_β) are called compatible.

(d) The family (U_α, ϕ_α) is maximal relative to the conditions (b) and (c).

Example 2.1.2. *The* ***general linear group*** $GL(n, \mathbb{R})$ *is the set of all* $n \times n$ *real matrices* A *with* $\det A \neq 0$. *We can naturally embed this (entry-wise) into the ambient space* $\mathbb{R}^{n^2}$ *which we identify with the set of all* $n \times n$ *real matrices* $M(n, \mathbb{R})$. *Since* $\det A$ *is a polynomial of degree* n *in the coordinates, it is a smooth function on* $M(n, \mathbb{R})$. *Since the set* $\mathbb{R} \setminus \{0\}$ *forms an open set in* $\mathbb{R}$ *and since the inverse image of an open set under a continuous map is open, the set* $GL(n, \mathbb{R})$ *is an open subset of* $M(n, \mathbb{R})$ *which we know is an* n^2*-manifold. Indeed, any open subset of a manifold is itself a manifold, inheriting the smooth structure. Thus,* $GL(n, \mathbb{R})$ *is an* n^2*-manifold.*

A (real-valued) function on a manifold is **smooth** if it is smooth in the appropriate charts in the Euclidean sense. This lets us define tangents and differentials:

Definition 2.1.3. Let p be a point of a manifold M and $F(M)$ be the set of all smooth real-valued functions on a manifold M. A **tangent vector** to M at p is a real-valued function $v : F(M) \to \mathbb{R}$ that satisfies:

(a) **Linearity:** $v(af + bg) = av(f) + bv(g)$

(b) **Product Rule:** $v(fg) = v(f)g(p) + f(p)v(g)$

for all $a, b \in \mathbb{R}$ and $f, g \in F(M)$. The set of all tangent vectors at p is called the **tangent space** T_pM, and the disjoint union $TM = \sqcup_p T_pM$ is called the **tangent bundle** of M.

Definition 2.1.4. Let $f : M \to N$ be a smooth function. Then, for each $p \in M$, the **differential** of f is the function:

$$df_p : T_pM \to T_{f(p)}N, \quad df_p(g) = v(g \circ f)$$

for all $v \in T_pM$ and $g \in F(N)$.

If $M = \mathbb{R}^n$ and $N = \mathbb{R}^m$, then df_p is just the usual Jacobian. As seen in Example 2.1.2, it is typically easier to construct manifolds as subsets of larger manifolds rather than by working through the definition.

Definition 2.1.5. A manifold P is a **submanifold** of the manifold M if

(a) P is a topological subspace of M.

(b) The inclusion map $j : P \hookrightarrow M$ is smooth and at each point $p \in P$ its differential dj_p is one-to-one.

To define a natural notion of distance on a manifold, we first define a metric on the tangent spaces (see Definition 3.1.1 for how this can be made into a metric on the manifold itself):

Definition 2.1.6. A **Riemannian metric** on a smooth manifold M is a correspondence which associates to each point $p \in M$ an inner product $g_p = \langle\ ,\ \rangle$ on the tangent space T_pM which varies differentiable in the following sense: for every pair of smooth vector fields X, Y in a neighborhood of p, the map $p \mapsto \langle X_p, Y_p\rangle_p$ is smooth. A smooth manifold with a Riemannian metric is called a **Riemannian manifold**, and is denoted by $\langle M, g\rangle$.

Example 2.1.7. *Let $M = \mathbb{R}^n$ with $\frac{\partial}{\partial x_i}$ identified with $e_i = (0, ..., 1, ..., 0)$. The metric is given by $g(e_i, e_j) = \delta_{ij}$. In this case $\mathbb{R}^n$ is called the Euclidean space of dimension n.*

Example 2.1.8. *Immersed manifolds. Let $f : M \to N$ be an immersion (that is smooth, with df_p one-to-one for all $p \in M$). If N has a Riemannian metric g', then f induces a Riemannian metric g on M by defining $g_p(u, v) = g'_{f(p)}(df_p(u), df_p(v))$ where $u, v \in T_pM$. This metric on M is called the metric induced by f, and f is called an isometric immersion.*

For example, the metric on the sphere $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n \mid x_1^2 + \cdots + x_n^2 = 1\}$ induced from the Euclidean metric from $\mathbb{R}^n$ is called the standard metric on $\mathbb{S}^{n-1}$.

We can think of the Riemannian metric (2.1.6) as assigning a magnitude to the tangents, giving infinitesimal distances. Combining these distances can give an *intrinsic* notion of volume that does not need to be inherited from an ambient space (e.g. Lebesgue measure).

Definition 2.1.9. Let V be a real vector space of dimension n with a given orientation and an inner product. The corresponding **volume form** is the unique element ω of the space $\bigwedge^n V$ of n-forms (namely alternating multilinear maps $L : V^n \to \mathbb{R}$), such that $\omega(v_1, ..., v_n) = 1$ for each orthonormal (with respect to the given inner product) positively oriented basis of V.

Definition 2.1.10. If (Σ, g) is an oriented Riemannian manifold, then the **volume form** ω on Σ is the unique element of $\bigwedge^n M$ such that, for each $x \in M$, $\omega|_x$ is the volume form on the vector space T_xM relative to the inner product g_x and the orientation chosen. One often writes **dvol**, for the volume form on Σ.

Example 2.1.11. *Consider a local coordinate patch U on Σ. Let $x_1, ..., x_n$ be the corresponding coordinates and denote by $g_{ij}(x)$ the inner product at the point x of the vector fields $\frac{\partial}{\partial x_i}$ and $\frac{\partial}{\partial x_j}$, namely,*

$$g_{ij}(x) = g_x(\tfrac{\partial}{\partial x_i}, \tfrac{\partial}{\partial x_j}).$$

*Let $\det g(x)$ be the determinant of the matric $(g_{ij}(x))_{ij}$. We then have the following formula for the volume form **dvol** on U.*

$$\textbf{\textit{dvol}} = \epsilon\sqrt{\det g(x)}dx_1 \wedge ... \wedge dx_n$$

where $\epsilon = 1$ if $\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_n}$ is positively oriented, and -1 otherwise.

2.2 Lie Theory

Definition 2.2.1. [5] G is called a **Lie group** if:

(a) G is both a group and a manifold, and

(b) The group operations $G \times G \to G$, $(x, y) \mapsto xy$ and $G \to G$, $x \mapsto x^{-1}$ are smooth functions.

In the sequel, group actions may be considered implicitly to be smooth. We also note that a **homomorphism** on a Lie groups must be both a diffeomorphism and group isomorphism to preserve the manifold and group structure respectively.

Definition 2.2.2. [5] A **Lie subgroup** of a Lie group G is a Lie group H that is an abstract subgroup and an immersed submanifold of G. A **closed subgroup** of a Lie group G is an abstract subgroup K that is a (topologically) closed subset of G.

Proposition 2.2.3. *[5] If H is a closed subgroup of a Lie group G, then H is a submanifold of G and hence a Lie subgroup of G. In particular, it has the induced topology.*

Example 2.2.4 (Examples of Lie Groups [5])**.**

(i) The additive groups $\mathbb{R}^n$ and $\mathbb{C}^n$ since they are trivially n-manifolds.

(ii) The multiplicative groups $\mathbb{R} \setminus \{0\}$ and $\mathbb{C} \setminus \{0\}$. The unit circle $\mathbb{S}^1$ may be seen as a closed subgroup of $\mathbb{C}^*$ so it is a Lie group by Proposition 2.2.3.

(iii) The product $G \times H$ of two Lie groups is itself a Lie group with the product manifold structure, and multiplication $(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2)$ where $g_1, g_2 \in G$ and $h_1, g_2 \in H$.

(iv) The General Linear Group $GL(n, \mathbb{R})$,: The group of all invertible $n \times n$ matrices with real entries. This is a manifold by Example 2.1.2. Both multiplication and inverse are definitely smooth for $GL(n, \mathbb{R})$, meaning that it indeed is a Lie group.

(v) The Special Linear Group $SL(n, F)$ which consists of the $n \times n$ matrices with entries in F of determinant 1. We claim that it is a closed subgroup of $GL(n, \mathbb{R})$ as it is defined by the polynomial equation $\det = 1$ (see Appendix B). Thus, by Proposition 2.2.3, it is a Lie group.

(vi) The Orthogonal Group $O(n)$: The group of all $n \times n$ orthogonal matrices.

(vii) The Special Orthogonal Group $SO(n)$: The group of all $n \times n$ orthogonal matrices with determinant 1. The method to verify that it is a Lie group is similar to (v).

(viii) The Unitary Group U(n): The group of all $n \times n$ unitary matrices.

(ix) The Special Unitary Group SU(n): The group of all $n \times n$ unitary matrices with determinant 1. The method to verify that it is a Lie group is similar to (v).

(x) The n-torus $\mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$ for n times is a Lie group of dimension n because of (ii) and (iii).

Since Lie groups are manifolds, we can study them topologically. For example, it is often convenient to restrict to the **identity component** G° [5] which is the set of all $g \in G$ that are (path-)connected to e. We also have a subtly different notion of isomorphic that is more 'local':

Definition 2.2.5. [6] A **local Lie group** is a smooth manifold M together with a base point e, its neighbourhood U in M and a differentiable map (multiplication):

$$U \times U \to M, \quad (x, y) \mapsto xy$$

satisfying the conditions $ex = xe = x$ and $(xy)z = x(yz)$ for $x, y, z, xy, yz \in M$. This implies the existence of a neighborhood of the identity $W \subset V$ and a differentiable map (inversion):

$$W \to W, \quad x \mapsto x^{-1}$$

such that $xx^{-1} = x^{-1}x = e$ for $x \in W$. Every Lie group G can be viewed as a local Lie group by taking $U = W = G$.

Definition 2.2.6. [6] Two Lie groups are said to be **locally isomorphic** if they are isomorphic as local Lie groups.

The only element that is guaranteed to exist in a Lie group G is the identity, whose tangent space is the 'easiest' to study.

Definition 2.2.7. [5] The **Lie algebra** $\mathfrak{g}$ of a Lie group G is the tangent space at the identity element e of G equipped with the **Lie bracket**:

$$\mathfrak{g} = T_e G$$

The Lie bracket is an operation $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ that satisfies the following properties:

- **Bilinearity**: $[aX + bY, Z] = a[X, Z] + b[Y, Z]$ for all $a, b \in \mathbb{R}$ and $X, Y, Z \in \mathfrak{g}$.
- **Antisymmetry**: $[X, Y] = -[Y, X]$ for all $X, Y \in \mathfrak{g}$.
- **Jacobi Identity**: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ for all $X, Y, Z \in \mathfrak{g}$.

Example 2.2.8. *[5] The Lie algebra of the general linear group $GL(n, \mathbb{R})$ is canonically isomorphic to $M(n, \mathbb{R})$, the set of all $n \times n$ matrices. The Lie bracket is given by the commutator $[A, B] = AB - BA$.*

Example 2.2.9. *[5] Consider the Lie group $SL(2, \mathbb{R})$. Let its Lie algebra be $\mathfrak{sl}(2, \mathbb{R})$. We claim the Lie algebra $\mathfrak{sl}(2, \mathbb{R})$ consists of all 2×2 matrices with trace zero. Indeed, let $X \in \mathfrak{sl}(2, \mathbb{R})$. Thus, $e^{tX} \in SL(2, \mathbb{R})$, meaning that $\det e^{tX} = 1$. We have $\det e^{tX} = e^{t \cdot tr(X)}$ where X is diagonalisable. Therefore, $tr(X) = 0$.*

The basis elements are:

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

Indeed, $[H, E] = HE - EH = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = 2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = 2E$. *Similarly,* $[H, F] = -2F$ *and* $[E, F] = H$.

Proposition 2.2.10. *[5] Let $\phi : G \to H$ be a Lie group homomorphism. Then the map $d\phi_e : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism. Furthermore,*

$$\phi(\exp X) = \exp\left(d\phi_e(X)\right)$$

The structural properties of Lie algebras are neatly summarised in **Lie's Theorems** which we state without proof:

Theorem 2.2.11 (Lie's Theorems [5]).

(i) *For any Lie algebra $\mathfrak{g}$, there is a Lie group G whose Lie algebra is $\mathfrak{g}$.*

(ii) *Let G be a Lie group with Lie algebra $\mathfrak{g}$. If H is a Lie subgroup of G with Lie algebra $\mathfrak{h}$, then $\mathfrak{h}$ of $\mathfrak{g}$, there exists a unique connected Lie subgroup H of G which has $\mathfrak{h}$ as its Lie algebra.*

(iii) *Let G_1, G_2 be Lie groups with corresponding Lie algebras $\mathfrak{g}_1, \mathfrak{g}_2$. Then if $\mathfrak{g}_1 \cong \mathfrak{g}_2$,then G_1, G_2 are locally isomorphic. If the Lie groups G_1, G_2 are simply connected (i.e. their fundamental groups are trivial,) then G_1 is isomorphic to G_2.*

Definition 2.2.12. [7] In algebra, a **simple Lie algebra** is a Lie algebra that is non-abelian (that is, there exists $x, y \in \mathfrak{g}$, such that $[x, y] \neq 0$) and contains no nonzero proper ideals. A direct sum of simple Lie algebras is called a **semi-simple Lie algebra**.

Definition 2.2.13. [7] A Lie group is called **semi-simple** if its Lie algebra is semi-simple.

Example 2.2.14. *[8] $SL(2, \mathbb{R})$ is semi-simple. To see that, we need to show its Lie algebra $\mathfrak{sl}(2, \mathbb{R})$ is semi-simple. Indeed, it is non-abelian obviously since the Lie bracket for $\mathfrak{sl}(2, \mathbb{R})$ is $XY - YX$ where $X, Y \in \mathfrak{sl}(2, \mathbb{R})$. For it contains no nonzero proper ideals, we can prove this by contradiction: suppose I is a nonempty proper ideal. Then it contains some nonzero element $A = aH + bE + cF$ where $a, b, c \in \mathbb{R}$ and H, E, F are the same as those in Example 2.2.9. Then, $[H, A] = b[H, E] + c[H, F] = 2bE - 2cF$. And we do the bracket operation once again: $[H, 2bE - 2cF] = 4bE + 4cF$. Thus, either $a = 0$ or I contains H.*

- *If I contains H, then it contains $[H, E] = 2E$ and $[H, F] = 2F$. Hence, $I = \mathfrak{sl}(2, \mathbb{R})$, which is incorrect.*
- *Otherwise, $a = 0$. However, $[E, A] = -cH$. Either $c = 0$ or I contains H. The latter leads to a contradiction as shown above. Thus, $c = 0$. I must contain bE. So it contains $[E, F] = H$. Contradiction appears.*

Definition 2.2.15. [5] A **one-parameter subgroup** of a Lie group G is (the image of a) smooth homomorphism $\phi : (\mathbb{R}, +) \to G$. In particular,

$$\phi : \mathbb{R} \to G, \quad \phi(s + t) = \phi(s)\phi(t), \quad \phi(0) = e, \quad \phi(-t) = \phi(t)^{-1}$$

We may write $\{u^t\}$ to mean $\{u^t\}_{t \in \mathbb{R}}$.

Example 2.2.16. *[5] The map $\phi(t) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$ is a one-parameter subgroup in $U(2)$.*

Corollary 2.2.17. *[5] For each $X \in \mathfrak{g}$ there exists a unique one-parameter subgroup $\phi_X : \mathbb{R} \to G$ such that $\phi_X'(0) = X$.*

Definition 2.2.18. [5] The exponential map $\exp : \mathfrak{g} \to G$ is defined by $\exp X = \phi_X(1)$, where ϕ_X is the unique one-parameter subgroup of X.

Corollary 2.2.19. *[5] The curve $\gamma(t) = \exp tX$ where $X \in \mathfrak{g}$ is the unique homomorphism in G with $\gamma'(0) = X$. Also, since ϕ_X is a homomorphism, it follows that $\exp((s + t)X) = \exp(sX) \cdot \exp(tX)$ and $(\exp tX)^{-1} = \exp(-tX)$.*

Example 2.2.20. *[9] In the case where $G = GL(n, \mathbb{R})$ and $\mathfrak{g} = M(n, \mathbb{R})$ (see Example 2.2.8), the exponential map coincides with the usual exponential map for matrices, that is for $X \in \mathfrak{g}$,*

$$\exp(X) = \sum_{n=0}^{\infty} \frac{1}{n!} X^n$$

Proposition 2.2.21. *[5] There is a neighbourhood of $\mathbf{o} \in \mathfrak{g}$ which is mapped diffeomorphically by $\exp$ to a neighbourhood of the identity $e \in G$.*

This can be seen as a consequence of the inverse function theorem. If we are working with points $h \approx e \in G$, we can work with *exponential coordinates* i.e. choose the unique $\underline{h} \approx \mathbf{o} \in \mathfrak{g}$ such that $\exp(\underline{h}) = h$.

2.3 Representations of Lie Groups

It is generally prohibitively difficult to work with groups when presented abstractly. A concrete (albeit inefficient) solution is offered by **Cayley's Theorem** [10] which says that every finite group of size n is isomorphic to some subgroup of the symmetric group S_n. Hence, it can be represented (faithfully) as a group of $n \times n$ permutation matrices which can be tackled using linear algebra. We formalize this for general groups as follows:

Definition 2.3.1. [5] A (finite-dimensional) **representation** of a group G is a homomorphism $\phi : G \to \mathrm{Aut}(V)$, where V is a (finite-dimensional) vector space. The dimension of the representation is the dimension of the vector space V.

Groups that can be represented in this way are called **linear groups** as we may think of $\mathrm{Aut}(V)$ as being the general linear group $GL(n, F)$ over some field F. There are (infinitely) many possible representations, but the most natural choice for a Lie group G is a representation with V as the corresponding Lie algebra $\mathfrak{g}$. To pass from G to $\mathfrak{g}$ (a tangent space), we can look at the differential of a diffeomorphism on G associated to g. The most natural choice is the **conjugation** map $I_g : x \mapsto gxg^{-1}$ for all $x \in G$ which is a diffeomorphism since multiplications in a Lie group are smooth.

Definition 2.3.2. [5] The **adjoint representation** of a Lie group G is the (group) homomorphism $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$ given by $\mathrm{Ad}(g) = (dI_g)_e$ (i.e. the differential evaluated at e).

Similarly, the structure endowed by the Lie bracket makes studying $\mathfrak{g}$ concretely fairly difficult unless we work with matrices. The main difference with Definition 2.3.1 is that we pass to $\mathrm{End}(V)$ instead, which we may think as passing to the vector space of matrices $M(n, F)$ with the commutator as the Lie bracket. We can obtain a canonical representation by differentiating Ad in Definition 2.3.2:

Definition 2.3.3. [5] The **adjoint representation** of $\mathfrak{g}$ is the (Lie algebra) homomorphism $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ given by $\mathrm{ad}(X) = (d\,\mathrm{Ad})_e(X)$.

Proposition 2.3.4 (Properties of Adjoint Representations [5]). *idk if i have the will to prove*

(i) If G is a matrix group, then $\mathrm{Ad}(g)(X) = gXg^{-1}$ for all $g \in G,\ X \in \mathfrak{g}$ (where the multiplication is matrix multiplication).

(ii) $\mathrm{ad}(X)(Y) = [X, Y]$ for all $X, Y \in \mathfrak{g}$. Since the Lie bracket is bilinear (from Definition 2.2.7), the ad is linear i.e. $\mathrm{ad}(tX_1 + sX_2) = t\,\mathrm{ad}(X_1) + s\,\mathrm{ad}(X_2)$ for all $s, t \in \mathbb{R}$.

(iii) $\exp(\operatorname{ad}(X)) = \operatorname{Ad}(\exp X)$ *for all* $X \in \mathfrak{g}$ *where* $\exp : \operatorname{End}(\mathfrak{g}) \to \operatorname{Aut}(\mathfrak{g})$ *is the exponential map (2.2.18)*

Example 2.3.5. *[8] Consider the Lie group $SL(2, \mathbb{R})$ and its Lie algebra $\mathfrak{sl}(2, \mathbb{R})$. The adjoint representation Ad of an element $g \in SL(2, \mathbb{R})$ on $\mathfrak{sl}(2, \mathbb{R})$ is given by conjugation:*

$$Ad(g)(X) = gXg^{-1} \text{ where } X \in \mathfrak{sl}(2, \mathbb{R})$$

We will use notations and results from Example 2.2.9.

See the proofs of Proposition 6.1.2 and Proposition 6.2.1 for how adjoint representations and exponential maps are used in practice.

2.4 Homogeneous Spaces

In general, it is difficult to move between two points in a manifold without some involved differential geometry. We dispense with these technicalities by examining only those manifolds where movement can be described by smooth group actions. To make this precise, we fix notation:

Definition 2.4.1. [5] Let G be a Lie group acting on a manifold.

(a) An action is called **transitive** if for any $m, n \in M$, there exists a $g \in G$ such that $gm = n$.

(b) Let $m \in M$. The set $G_m = \{g \in G : gm = m\}$ is called the **isotropy subgroup** at m.

(c) The **orbit** of a point $m \in M$ is the set $Gm = \{gm \in M : g \in G\}$.

Since $M = Gm$ for some $m \in M$, the following can be seen in light of the **Orbit-Stabiliser Theorem**:

Proposition 2.4.2. *[5] Let $G \times M \to M$ be a transitive action of a Lie group G on a manifold M, and let $K = G_m$ be the isotropy subgroup of a point $m \in M$. Then:*

(i) The subgroup K is a closed subgroup of G.

(ii) The canonical map $G/K \to M$ given by $gK \mapsto gm$ is a diffeomorphism.

(iii) The dimension of G/K is $\dim G - \dim K$.

Definition 2.4.3. [5] A **homogeneous space** is manifold M with a transitive action of a Lie group G. Equivalently, it is a manifold of the form G/K, where G is a Lie group and K is a closed subgroup of G (denote $[x] = Kx$ to be the left cosets).

Example 2.4.4 (Examples of Homogeneous Spaces [5])**.**

(i) ***Lie Groups:*** *Every Lie group G is a homogeneous space since G is a manifold that acts transitively on itself by left multiplication. Consider the point $g \in G$. The only element that fixes g is the identity element e, so we write $G = G/\{e\}$. Alternatively, we can think of $G \times G$ acting on G by left and right multiplication which is transitive. Since $h_1 g h_2 = g \iff h_1 = g h_2^{-1} g^{-1}$, every choice of left multiplication uniquely determines the right multiplication if we want to fix g. We say that G is diagonally embedded in $G \times G$, so we write $G = G \times G/G$.*

(ii) ***Spheres:*** *Consider the unit sphere* $\mathbb{S}^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$ *which is a 2-manifold. The group* $SO(3)$ *acts transitively on* $\mathbb{S}^3$ *by 3D rotations, hence preserving the quadratic form* $x^2 + y^2 + z^2$ *(cf. Definition 7.1.6). Consider the point* $v = (0, 0, 1)$*. The elements of* $SO(3)$ *that fix* v *are precisely the rotations about the* z*-axis (see Figure 2). Thus, the isotropy subgroup at* v *is*

$$\{\begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix} : A \in SO(2)\} \cong SO(2)$$

We write $\mathbb{S}^2 = SO(3)/SO(2)$*. A similar construction gives* $\mathbb{S}^n = SO(n+1)/SO(n)$*.*

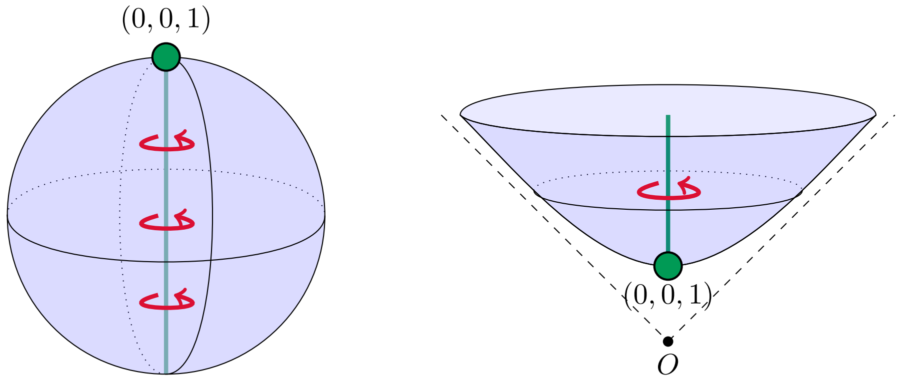


Figure 2 The sphere $\mathbb{S}^2$ and the upper sheet of the hyperboloid $\mathbb{H}_{-1}$ rotating about the z-axis.

(iii) ***Hyperboloids:*** *Consider the hyperboloid* $\mathbb{H}_c = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 - z^2 = c\}$*. We get one continuous sheet when* $c > 0$ *and two separate sheets when* $c < 0$*. The critical value* $c = 0$ *gives a conical surface, known to physicists as the* ***light cone*** *[11]. Analogously to the situation with spheres, we can think of the* $SO(2, 1)$ *as the group of* 3×3 *matrices that preserve the quadratic form* $x^2 + y^2 - z^2$ *(cf. Definition 7.1.6). The identity component* $SO(2, 1)^\circ$ *preserves the sign of* z *i.e. it is* ***orthochronous*** *if we think of* z *as time. This acts transitively on the upper sheet of* $\mathbb{H}_{-1}$ *by matrix multiplication and the isotropy subgroup of* $(0, 0, 1)$ *is* $SO(2)$ *as in Figure 2. Thus, we write* $\mathbb{H}_{-1} = SO(2, 1)^\circ/SO(2)$*. Hyperboloids are one of many models of* ***hyperbolic space*** *(compare with Example 2.5.7).*

(iv) ***Tori:*** *Construct the (flat) torus* $\mathbb{T}^2$ *as in Figure 3. Since we loop back onto ourselves in the horizontal and vertical directions, we may write* $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$ *so it is a Cartesian product of homogeneous spaces (or Lie groups). Alternatively,* $\mathbb{R}^2$ *acts naturally on* $\mathbb{T}^2$ *by translation mod 1 i.e.* $v \in \mathbb{R}^2$ *acts on* $(x_1, x_2) \in \mathbb{R}^2$ *by* $(x_1 + v_1 \mod 1, x_2 + v_2 \mod 1)$*. Translating any point in* $\mathbb{T}^2$ *by an integer vector (mod 1) returns us to the same point, so the isotropy subgroup is* $\mathbb{Z}^2$*. We write* $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$*. A similar construction gives* $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$ *(*n *times). We will return to this example many times throughout this paper as the torus is the simplest toy model for Ratner's Theorems (see Section 4).*

(v) ***Grassmannians:*** *Consider the Grassmannian* $\mathrm{Gr}(k, n)$ *which consists of the* k*-dimensional subspaces* $\mathbb{R}^n$*.* $O(n)$ *acts on* $\mathrm{Gr}(k, n)$ *by left-multiplying each basis vector of a* k*-subspace by an element of* $O(n)$*. This action is transitive as we can directly construct* $A = [e'_1|\cdots|e'_k]$ *where* $\{e'_1, \ldots, e'_k\}$ *is an orthonormal basis of some* k*-subspace* W*, so* $AV = W$ *where* V *is spanned by the first* k *standard basis vectors. The isotropy subgroup at* V *is given by*

$$\{\begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix} : B \in O(k), C \in O(n-k)\} \cong O(k) \times O(n-k)$$

We write $\mathrm{Gr}(k, n) = O(n)/O(k) \times O(n-k)$*.*

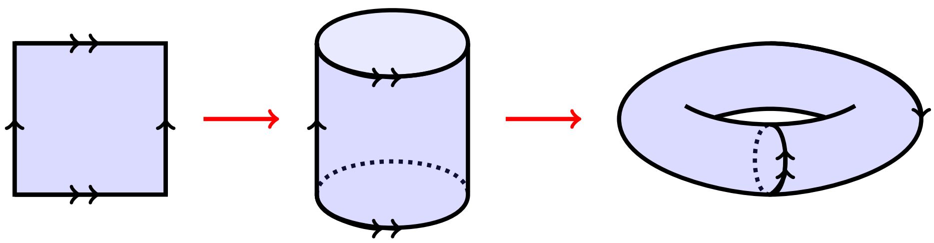


Figure 3 Start with a unit square and identify opposite edges together (without twisting). Visually, we think of rolling a piece of paper up into a tube then joining the circular ends together to get a torus $\mathbb{T}^2$.

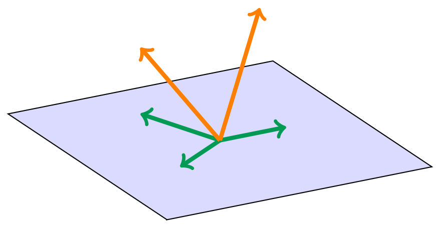


Figure 4 The Grassmannian $\mathrm{Gr}(3,5)$. Think of the three green vectors as a basis of the 'plane', and the two orange vectors as orthogonal to this 'plane'. The green and orange vectors are rotated and reflected by $O(3)$ and $O(2)$ respectively, but stay within their respective spans.

Remark 2.4.5. It can be shown that any connected manifold is homogeneous [12], so it is unsurprising that most manifolds we typically consider are homogeneous spaces. Refer to [13] for an example of a manifold that is **not** homogeneous.

2.5 Lattices and Invariant Measures

We want to generalise the properties of the torus $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. To make sense of this, we must first understand $\mathbb{Z}^n$ as a subgroup of $\mathbb{R}^n$. Since Lie groups possess a manifold structure, we can meaningfully describe $\mathbb{Z}^n$ as a **discrete subgroup** of $\mathbb{R}^n$ – one with no limit points. It has the additional property that it visually appears to be comprised of repeating hypercube 'units' of 'finite volume'. The most natural way to articulate this 'finiteness' is by considering a 'nice' measure which respects the group action:

Definition 2.5.1. A measure μ on a Lie group G is **left-invariant** if $\mu(gA) = \mu(A)$ for all $g \in G$ and all measurable $A \subseteq G$. Likewise, μ is **right-invariant** if $\mu(Ag) = \mu(A)$.

Invariant measures generalise the notion of an *incompressible fluid* from fluid dynamics

Definition 2.5.2. [14] Let (X, τ) be a topological space, and consider measure μ on (X, Σ) where Σ is a σ-algebra on X. Then μ is **regular** if it can be approximated from above by open measurable sets, and from below by compact measurable sets. More precisely, it is

- **Inner regular:** $\mu(A) = \sup\{\mu(F) : F \subseteq A,\ F \in \Sigma,\ F \text{ compact}\}$, and
- **Outer regular:** $\mu(A) = \inf\{\mu(F) : G \supseteq A,\ G \in \Sigma,\ G \text{ open}\}$

Definition 2.5.3. [14] Any Lie group G has a left-invariant regular Borel measure on G which is unique up to a scalar multiple, called the **(left) Haar measure** μ.

The existence and uniqueness of the Haar measure is a highly nontrivial result which we omit from this text for brevity. Refer [15] for a proof. **Unless otherwise stated, we always use the left convention.**

Example 2.5.4 (Examples of Haar measures [16]).

(i) *In a discrete group G, the compact subsets are precisely the finite subsets of G. Up to scalar multiplication, the left and right Haar measures on G are both the* ***counting measure*** *– so G is unimodular.*

(ii) *The* ***Lebesgue measure*** *on $\mathbb{R}^n$ is known to be regular and translation-invariant. Hence, its restriction to the Borel subsets of $\mathbb{R}^n$ is the Haar measure such that $[0,1]$ has measure 1.*

(iii) *For the general linear group $G = GL(n, \mathbb{R})$, we identify the set of $n \times n$ matrices with $\mathbb{R}^{n^2}$ which we endow with the Lebesgue measure λ^{n^2}. We obtain a Haar measure:*

$$\mu(S) = \int_S |\det(X)|^{-n}\, d\lambda^{n^2}(X)$$

where S is a Borel subset of G and $g \in G$. Invariance follows from change of variables.

Definition 2.5.5. A **lattice** Γ in a Lie group G is a discrete subgroup of G such that G/Γ carries a finite invariant regular Borel measure.

If Γ is a lattice in G, then there is a unique G-invariant probability measure μ_G on G/Γ which we call the **(left) Haar measure** on G/Γ. It turns out that μ_G can be represented by a smooth volume form on the manifold G/Γ (see Definition 2.1.10), so this measure-theoretic notion agrees with the differential-geometric notion of **finite volume.**

In fact, if a discrete subgroup Γ is cocompact (i.e. G/Γ is compact), then Γ has cofinite volume [17] i.e. it is a lattice, in which case we say that Γ is a **uniform lattice** (see 2.5.7 for a non-uniform lattice). This includes our motivating example of the torus $\mathbb{T}^n$ as any unit square centered at each integer point can be taken as the **fundamental domain** (i.e. a minimal region $\mathcal{F}$ such that $\Gamma\mathcal{F} = G$). These squares are clearly compact and have finite volume.

In $\mathbb{R}^n$, every lattice is isomorphic to $\mathbb{Z}^n$ [18]. Analogously, one would hope that finding lattices in (algebraic) Lie groups is as straightforward as replacing $\mathbb{R}$ with $\mathbb{Z}$. This 'arithmeticity' is not a given in general and the classification (and even existence) of lattices remains a difficult open problem. There are, however, special cases where this is possible:

Theorem 2.5.6 (Existence of a Lattice [19]).

(i) ***(Borel)*** *Let G be a semi-simple Lie group. Then G contains a lattice. If G is connected, then it contains a uniform lattice.*

(ii) ***(Borel)*** *Let M be a simply connected Riemannian symmetric manifold and let $I(M)$ be its group of isometries. Then $I(M)$ contains a uniform lattice.*

(iii) ***(Borel-Harish-Chandra)*** *Let $G \subseteq GL(n, \mathbb{R})$ be a semisimple linear algebraic group defined over the rational numbers $\mathbb{Q}$ (see Appendix B). Then $G \cap GL(n, \mathbb{Z})$ is a lattice in G.*

Example 2.5.7 (Upper Half Plane [20]). *$SL(2, \mathbb{R})$ is linear, semisimple (2.2.14) and $\mathbb{Q}$-algebraic (i.e. it arises as the roots of the irreducible polynomial $\det - 1$). Thus, the Borel-Harish-Chandra Theorem 2.5.6(iii) immediately tells us that $SL(2, \mathbb{Z})$ is a lattice in $SL(2, \mathbb{R})$. To see this directly, it is best to think of $SL(2, \mathbb{R})$ acting on the* ***upper half plane*** *$\mathfrak{H} = \{x + iy \in \mathbb{C} \mid y > 0\}$ by Möbius transformations:*

$$z \mapsto \frac{az+b}{cz+d} \quad \textit{for} \quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL(2, \mathbb{R}) \quad \textit{and} \quad z \in \mathfrak{H}.$$

It is straightforward to show that $SL(2,\mathbb{Z})$ is generated by

$$S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

which correspond to an inversion $z \mapsto -1/z$ and a horizontal shift $z \mapsto z+1$. A further computation shows that every 'cell' in Figure 5 can be arrived at by successive shifts and inversions. The Haar measure in this case can be computed to be $|y|^{-2}dxdy$ which equips $SL(2,\mathbb{R})/SL(2,\mathbb{Z})$ with finite volume even though it is clearly not compact.

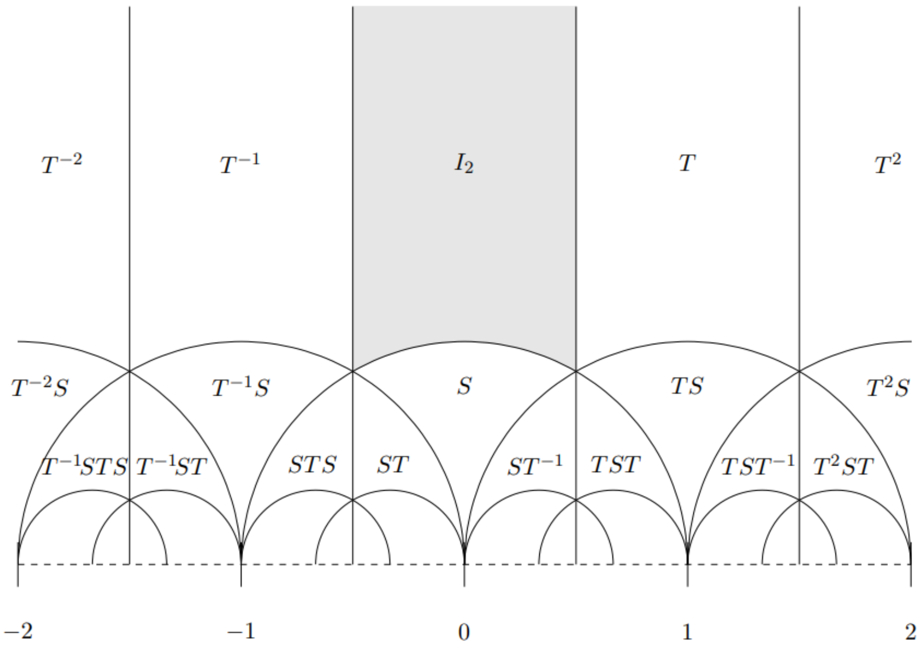


Figure 5 The fundamental domain $\mathcal{F} = \{z \in \mathfrak{H} \mid |\operatorname{Re} z| \leq 1/2,\ |z| \geq 1\}$ of the upper half plane $\mathfrak{H}$ shaded in grey. Figure taken from [20].

Remark 2.5.8. Identifying A with $-A$ makes the action by Möbius transformations faithful, and $SL(2,\mathbb{R})$ turns into the **modular group** $PSL(2,\mathbb{R})$ which can be shown to be isomorphic to $SO(2,1)^{\circ}$. This illustrates the utility of different models of hyperbolic space: On the hyperboloid (Example 2.4.4), the isotropy subgroup $SO(2)$ is clear but on the upper half plane $\mathfrak{H}$, the transitive group action by $SO(2,1)^{\circ}$ is clear.

Despite its apparent simplicity, $SL(2,\mathbb{R})$ is found in remarkably unrelated contexts – appearing here in hyperbolic geometry, but later featuring heavily in the context of Diophantine approximation. We will return to this group many times throughout this paper.

3 Dynamical Systems on Homogeneous Spaces

For completeness, we first state the 'obvious' fact that solutions to (well-behaved) differential equations on a manifold exist. Since a solution can be constructed locally, this essentially reduces to the Euclidean case.

Theorem 3.0.1 (Manifold 'Picard-Lindelöf' [5])**.** *Let X be a smooth vector field (see [5] for a definition) on a smooth manifold M, and let $p \in M$. Then there exists an open neighborhood U of p, an open interval I around 0 , and a smooth mapping $\phi : I \times U \to M$ such that the curve*

$\alpha_q : I \to M$ given by $\alpha_q(t) = \phi(t, q)$ where $q \in U$ is the unique curve that satisfies $\frac{\partial \phi}{\partial t} = X_{\alpha_q(t)}$ and $\alpha_q(0) = q$.

We can formally define flows on a manifold as a certain family of diffeomorphisms that satisfies a differential equation at different initial conditions but this requires some cumbersome analysis. However, when working on a homogeneous space, we can avoid this by simply *defining* a flow by the properties that it must satisfy:

Definition 3.0.2. [21] The action of a one-parameter subgroup $\{u^t\}$ of a Lie group G on the homogeneous space G/H is known as a **flow**. In particular, it satisfies:

$$\phi_t([x])\phi_s([x]) = \phi_{t+s}([x]), \quad \phi_0 = e, \quad \phi_{-t}[x] = \phi_t([x])^{-1}$$

Often we do not make a distinction between ϕ_t and u^t (cf. Definition 2.2.15).

This has the convenient consequence that the ϕ_t-**orbit** of $x \in G$ given by $\{[\phi_t x] \in G/H \mid t \in \mathbb{R}\}$ agrees with the group-theoretic orbit of $\{u^t\}$ on x.

3.1 Geodesic Flows

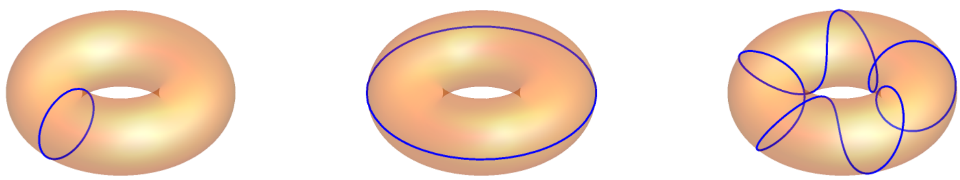
One of the most natural flows to describe on a manifold are those that take the most uniform 'path of least resistance'. To make this precise, we require the notion of distance on our manifolds given by a Riemannian structure (2.1.6):

Definition 3.1.1. [22] Given $x, y \in M$ we can define the **path metric**

$$d(x,y) = \inf_{\gamma} \left\{ \int_a^b \|\dot{\gamma}(t)\| dt \right\}$$

where the norm $\|\cdot\|$ is given by the appropriate Riemmannian metric and the infimum is taken over all smooth curves $\gamma : [0,1] \to M$ which start at x and finish at y.

Definition 3.1.2. A parameterised curve $\gamma : [0,1] \to M$ is a **geodesic** if it (locally) minimises the distance (i.e., for sufficiently large N the restriction $\gamma : [i/N, (i+1)/N]$ minimises the distance $\gamma(i/N)$ and $\gamma((i+1)/N))$ in the sense above.



(a) **(b)** **(c)**

Figure 6 There are essentially three types of closed geodesics on the (flat) 2-torus. The geodesics correspond to (a) a straight vertical line, (b) a straight horizontal line, and (c) a line with rational slope.

Now we want to define a flow on the compact three dimensional manifold

$$S = \{v \in TV : \|v\|_\rho = 1\} \qquad \text{("Sphere bundle")}$$

To define a geodesic flow $\phi_t : S \to S (t \in \mathbb{R})$ we can take $v \in S$ and choose the unique (unit speed) geodesic $\gamma_v : \mathbb{R} \to V$ such that $\dot{\gamma}_v(0) = v$. We can then define $\phi_t(v) := \dot{\gamma}_v(0)$. It is straightforward to check that this is a flow in the usual sense.

3.2 Unipotent Flows

Taking the view of representation theory (see Section 2.3), we can gain insight about a one-parameter subgroup $\{u^t\}$ by understanding its representations. The most natural parameter to control for are the eigenvalues of the adjoint representation.

Definition 3.2.1. A Lie group G is **nilpotent** if the adjoint representation $\mathrm{Ad}(g)$ is nilpotent for each $g \in G$ i.e. $\mathrm{Ad}(g)^k = 0$ for some k. If $\mathrm{Ad}(g) - I$ is nilpotent, then we say that G is **unipotent**.

Remark 3.2.2. Some sources define nilpotency in terms of **central series** (which we will not define here), but these definitions are equivalent by **Engel's Theorem** [7]. We may also think of Definition 3.2.1 in terms of eigenvalues i.e. $\mathrm{ad}(X)$ has eigenvalues 0 only or 1 only if G is nilpotent or unipotent respectively.

Definition 3.2.3 (Unipotent Flows [23])**.** We say that a flow η_t is **unipotent** if the associated one-parameter subgroup $\{u_t\}$ is **unipotent**.

Remark 3.2.4. Let $X \in \mathfrak{g}$. If $\mathrm{Ad}(\exp X)$ is unipotent, then $\exp(\mathrm{ad}(X)) - I$ is nilpotent by 2.3.4(iii). Thinking of this as the matrix exponential (2.2.20), this leads to a term of the form $e^\lambda - 1$. This must be zero for nilpotency, so all eigenvalues λ of $\mathrm{ad}(X)$ must be 0.

The following example is especially relevant to our later discussion of Ratner's Theorems (Section 4) and the Oppenheim-Davenport Conjecture (Section 7):

Example 3.2.5. *Consider the Lie group $G = SL(2,\mathbb{R})$. Its Lie algebra $\mathfrak{sl}(2,\mathbb{R})$ has basis $\{H, E, F\}$ as shown in Example 2.2.9. Taking exponential maps, these give the one-parameter subgroups $\{u^t\}$, $\{a^s\}$, $\{v^r\}$ with corresponding flows η_t, γ_s, and ψ_r. Explicitly,*

$$u^t = \exp(Ft) = \begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix}, \quad a^s = \exp(Hs) = \begin{bmatrix} e^s & 0 \\ 0 & e^{-s} \end{bmatrix}, \quad v^r = \exp(Er) = \begin{bmatrix} 1 & r \\ 0 & 1 \end{bmatrix},$$

We now show that u^t and v^r are unipotent. For u^t, we evaluate the adjoint representation $\mathrm{Ad}(u^t)X = u^t X u^{-t}$ (2.3.4) at the basis elements of $\mathfrak{sl}(2,\mathbb{R})$:

$$\begin{aligned} gFg^{-1} &= F - tH - t^2E \\ gHg^{-1} &= H + 2tE \\ gEg^{-1} &= E \end{aligned} \qquad \Longrightarrow \qquad [Ad(g)]_B = \begin{bmatrix} 1 & 0 & 0 \\ -t & 1 & 0 \\ -t^2 & 2t & 1 \end{bmatrix}$$

which only has eigenvalues 1, therefore the adjoint representation is unipotent. Since v^r is the transpose of u^t, by similarity v^r is also a unipotent flow. Proving that a^s is geodesic is more involved, but this is not the focus of our paper. Refer [24] for more details on the computation of geodesics flows.

We present another example of a unipotent flow before we focus (almost) entirely on $SL(2,\mathbb{R})$.

Example 3.2.6. *The **Heisenberg group** G is comprised of the 3×3 upper-triangular matrices with real entries and 1's on its diagonal. The **discrete Heisenberg group** Γ is the subgroup of G consisting of integer entries. It is straightforward to check that Γ is a lattice (cf. Definition 2.5.5). Take a smooth path $A(t)$ such that $A(0) = I$ so $A(t) = I + tB + \mathcal{O}(t^2)$. In the first order, we have $I + tB \in G$ so the tangent B must not have 1's on its diagonal. Thus, the Lie algebra $\mathfrak{g}$ is comprised of the 3×3 strictly upper-triangular matrices with real entries. A basis B is*

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

A routine calculation shows that the one-parameter subgroups are of the form $\{u^t\}$ where

$$u^t = \begin{bmatrix} 1 & a_0 t & c_0 t + \frac{1}{2} a_0 b_0 t^2 \\ 0 & 1 & b_0 t \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{for constants } a_0, b_0, c_0 \in \mathbb{R}$$

The adjoint representation with basis B is given by

$$g = \begin{bmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix} \quad \Longrightarrow \quad [\mathrm{Ad}_g]_B = \begin{bmatrix} 1 & 0 & 0 \\ -y & 1 & 0 \\ 0 & x & 1 \end{bmatrix}$$

which only has eigenvalue 1. Thus, we obtain unipotent flows:

$$\phi_t(g) = u^t g \Gamma = \begin{bmatrix} 1 & x + a_0 t & z + (a_0 y + c_0)t + \frac{1}{2} a_0 b_0 t^2 \\ 0 & 1 & y + b_0 t \\ 0 & 0 & 1 \end{bmatrix} \Gamma$$

3.3 Ergodicity

The main focus of this paper will be on examining 'ergodic' properties of dynamical systems which was first introduced in the context of statistical mechanics by Boltzmann (see [25] for a survey). It is important to recognise that ergodicity is a general principle which we can define for a variety of mathematical objects:

Definition 3.3.1. [14]

- The *action* of a group G on a space X is **ergodic** if every measurable G-invariant subset of X is null or conull.
- A measure-preserving *flow* ϕ_t on a is **ergodic** if every measurable ϕ_t-invariant subset of X is null or conull.
- A *measure* μ on a metric space X with measure-preserving flow ϕ_t is **ergodic** if ϕ_t is ergodic. Equivalently, μ is an extreme point on the space of ϕ_t-invariant probability measures on X (see Appendix A).

Example 3.3.2. *The theory of ergodic flows is especially rich in the context of manifolds.*

(i) Hopf [26] showed that geodesic flows on certain negatively curved Riemann surfaces (i.e. a connected one-dimensional complex manifold) are ergodic. Notable examples of this include the Hadamard and Artin billiard tables – dynamical systems where a particle moves in straight lines and bounces off the boundaries of the sides of the table (see [27] for more details).

*(ii) Moore [28] showed that if G is a connected, simple Lie group with finite center, Γ is a lattice in G, and g^t is a one-parameter subgroup of G with noncompact closure, then $\{g^t\}$ is ergodic on G/Γ (with respect to the Haar measure). Morris [14] calls this the **Moore Ergodicity Theorem**. It is a vast generalisation of the fact that the geodesic flows on $SL(2,\mathbb{R})/\Gamma$ are ergodic.*

It is a classical result of ergodic theory that almost every orbit of a measure-preserving ergodic flow is uniformly distributed. The mantra here is **"Time Average $=$ Space Average"** which we formalise as follows:

Theorem 3.3.3 (Pointwise Ergodic Theorem [14]). *Let*

- μ *is a probability measure on a locally compact, separable metric space* X*,*
- ϕ_t *is an ergodic, measure-preserving flow on* X*, and*
- $f \in L^1(X, \mu)$ *(i.e.* f *is* μ*-integrable)*

Then for almost every $x \in X$*,*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(\phi_t(x))\, dt = \int_X f\, d\mu$$

This is an *equidistribution* result, and the core of Ratner's Theorems in the following section is that this holds for **all** $x \in X$ if ϕ_t happens to be unipotent.

4 Ratner's Theorems

4.1 Ratner's Orbit Closure Theorem

The first of these theorems concerns the 'spatial niceness' of certain orbits of flows in a homogeneous space. The following example is due to Morris [14]:

Example 4.1.1. *Consider the flow* $\phi_t([x]) = [x + tv]$ *on the (flat) torus* $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ *defined by vectors* $x, v \in \mathbb{R}^2$*. Writing* $v = (a, b)$*, we get periodic behaviour when* a/b *is rational, recovering the closed orbits in Figure 6. If instead* a/b *is irrational, the* ϕ_t*-orbit of* x *is not closed, but it is dense in* $\mathbb{T}^2$ *(see Figure 7).*

Indeed, it is always the case that the orbits of unipotent flows are 'nice' geometric subspaces.

Theorem 4.1.2 (Ratner's Orbit Closure Theorem [14]). *Let*

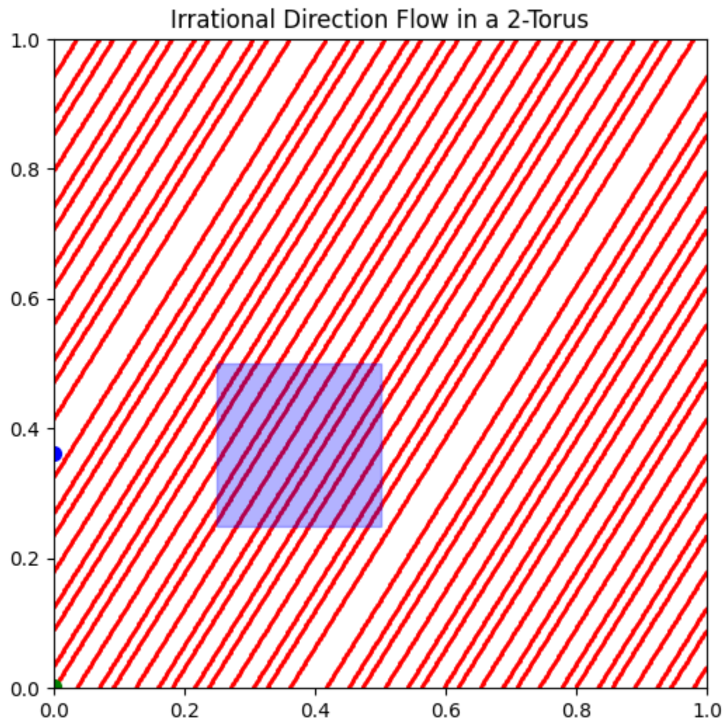
- G *be a Lie group*
- Γ *be a lattice in* G
- ϕ_t *be any unipotent flow on* G/Γ *with corresponding unipotent one-parameter subgroup* u^t

Then the closure of every ϕ_t*-orbit is* ***homogeneous****. More precisely, for each* $x \in G$*, there is a connected, closed subgroup* S *of* G *such that*

- $\{u^t\}_{t\in\mathbb{R}} \subseteq S$*,*
- *the image* $[Sx]$ *of* Sx *in* G/Γ *is closed and has finite* S*-invariant volume, and*
- *the* ϕ_t*-orbit of* $[x]$ *is dense in* $[Sx]$*.*

4.2 Ratner's Equidistribution Theorem

The Equidistribution Theorem strengthens the Orbit Closure Theorem (4.1.2) by further asserting the 'temporal niceness' of unipotent flows. We return to the torus example:



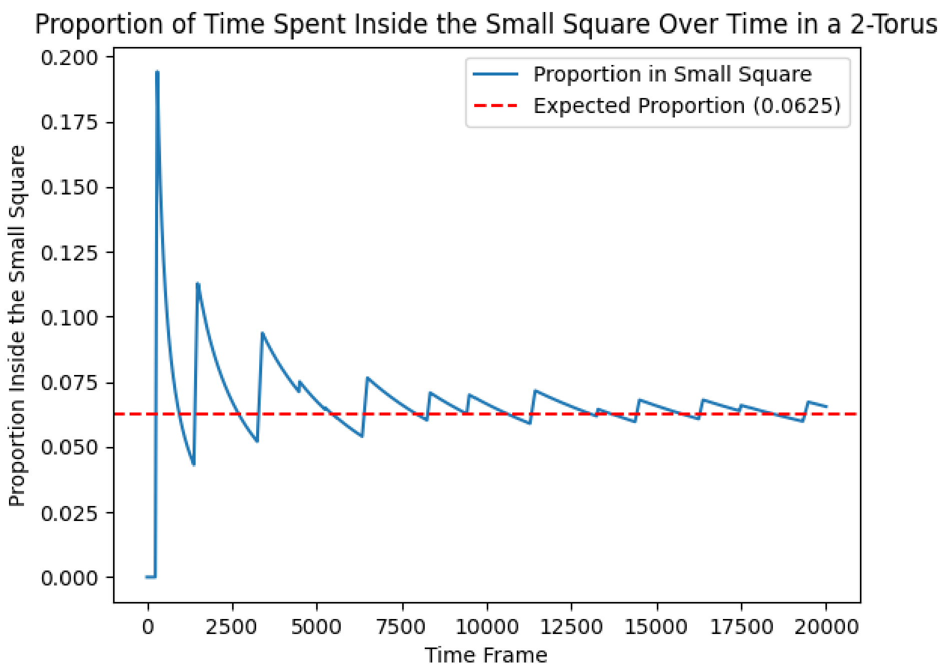


Figure 7 Numerical verification (Python) of Ratner's Equidistribution Theorem (4.2.3) on the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ (cf. Examples 2.4.4, 4.1.1, 4.2.1). We start at the origin and evolve our flow in the direction $v = (1, \phi)$ where ϕ is the golden ratio $(1+\sqrt{5})/2$. As expected, the flow visually tends towards being dense in $\mathbb{T}^2$ (Theorem 4.1.2) and the proportion of time spent in the blue square (whose area/Haar measure is $1/16$) tends towards $1/16$ as time goes to infinity.

Example 4.2.1. *Consider the flow $\phi_t([x]) = [x + tv]$ on the (flat) torus $\mathbb{T}^n$ defined by vectors $x, v \in \mathbb{R}^n$. Let μ be the uniform probability measure.*

1. *Case $n = 2$: Assume $v = (a, b)$ with a/b irrational so every orbit of ϕ_t is dense in $\mathbb{T}^2$. Then every orbit is uniformly distributed in $\mathbb{T}^2$: if B is any 'nice' open subset of $\mathbb{T}^2$, then the amount of time that each orbit spends in B on is proportional to the area of B. In precise terms:*

$$\lim_{T\to\infty} \frac{1}{T}\lambda(\{t \in [0, T] : \phi_t(x) \in B\}) = \mu(B)$$

   We can equivalently state this in terms of integrals over the orbit:

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(\phi_t(x))\, dt = \int_{\mathbb{T}^2} f\, d\mu$$

   for any continuous function f on $\mathbb{T}^2$ so that the RHS is indeed the expectation of f with respect to μ over the torus $\mathbb{T}^2$.

2. *Case $n = 3$: Assume $v = (a, b, 0)$ with a/b irrational. Then ϕ_t is* not *dense in $\mathbb{T}^3$, hence* not *uniformly distributed in $\mathbb{T}^3$ with respect to the usual uniform probability measure. Instead, they are uniformly distributed in some subtorus of $\mathbb{T}^3$: Given $x = (x_1, x_2, x_3) \in \mathbb{T}^3$, let μ_2 be the Haar measure on the horizontal 2-torus $\mathbb{T}^2 \times \{x_3\}$ that contains x. Then*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(\phi_t(x))\, dt = \int_{\mathbb{T}^2\times\{x_3\}} f\, d\mu_2$$

   for any continuous function f on $\mathbb{T}^3$.

3. *General case: There is always a subtorus S of $\mathbb{T}^n$, with Haar measure μ_S, such that*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(\phi_t(x))\, dt = \int_S f\, d\mu_S$$

   for any continuous function f on $\mathbb{T}^n$.

Remark 4.2.2. In Figure 7, we use a very small time step to discretise this continuous flow. Technically, we must be careful with the choice of time step (in view of the **Discrete Kronecker-Weyl Theorem** [29]). Since $\pi \approx 22/7$ choosing $v = (1, \pi)$ causes behaviour which is visually similar to $v = (1, 22/7)$ i.e. at small times we see 22 and 7 lines cross the upper and side boundaries of the square respectively. We instead choose ϕ because when it is presented as a continuous fraction, we get

$$\phi = 1 + \frac{1}{\phi} = 1 + \frac{1}{1 + \frac{1}{1+\cdots}}$$

In this sense, ϕ is the 'most irrational number' and (Diophantine) approximations by rational numbers converge exceedingly slowly. We will discuss in greater detail the interaction between homogeneous dynamics and Diophantine approximations in Section 7.

Generalising examples 4.1.1 and 4.2.1 are the essence of the Kronecker-Weyl Theorem [29] gives:

Theorem 4.2.3 (Ratner's Equidistribution Theorem [14])**.** *Let*

- *G be a Lie group*
- *Γ be a lattice in G*

- *ϕ_t be any unipotent flow on G/Γ*

Then every ϕ_t-orbit is uniformly distributed on its closure. More precisely, for each $x \in G$, there is a connected, closed subgroup S of G such that

- *$\{u^t\}_{t\in\mathbb{R}} \subseteq S$*
- *The image $[Sx]$ of Sx in G/Γ is closed, and has finite S-invariant volume, and*
- *The ϕ_t-orbit of $[x]$ is dense in $[Sx]$.*

Let μ_S be the (unique) S-invariant probability measure on $[Sx]$ (i.e. the Haar measure). Then

$$\lim_{T\to\infty} \frac{1}{T}\lambda(\{t \in [0,T] : \phi_t(x) \in B\}) = \mu_S(B)$$

$$\text{or equivalently} \quad \lim_{T\to\infty} \frac{1}{T}\int_0^T f(\phi_t(x))\, dt = \int_{[Sx]} f\, d\mu_S$$

for any continuous function f on G/Γ with compact support.

4.3 Ratner's Measure Classification Theorem

Arguably the most important of Ratner's theorems is not outright a statement about the behaviour of unipotent flows on G/Γ, but rather a statement about the measures on G/Γ which are invariant under these flows. Consider the following example given by Morris [14]:

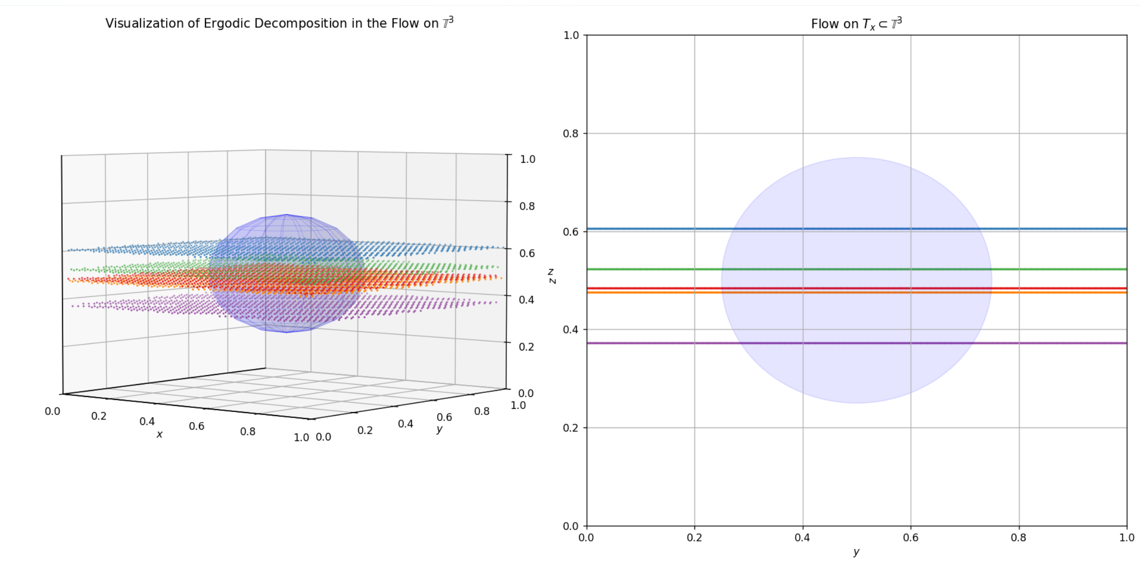


Figure 8 A visualisation (Python) for the ergodic decomposition (4.3.1) in the 3-torus. The $\alpha = \sqrt{2}$ in the visualisation and the five sample points are randomly selected within the reach of the measurable "ball" in the center.

Example 4.3.1. *Let $v = (\alpha, 1, 0) \in \mathbb{R}^3$ for some irrational α, with corresponding flow $\phi_t : [x] \mapsto [x + tv]$ on $\mathbb{T}^3 = \mathbb{R}^3/\mathbb{Z}^3$. Endow $\mathbb{T}^3$ with the Haar measure $\mu^{(3)}$ (which can be thought of as the Lebesgue measure restricted to the unit cube). Then ϕ_t is* **not** *ergodic, because there are nontrivial invariant sets of the form $\mathbb{T}^2 \times A$ where $A \subseteq \mathbb{T}$. However, if we partition $\mathbb{T}^3$ into subtori $T_z = \mathbb{T}^2 \times \{z\}$ for each $z \in \mathbb{T}$, then the restriction of ϕ_t to each T_z is ergodic with respect to the* *appropriate Haar measure μ_z, and clearly μ_x is ϕ_t-invariant. We can then write for all measurable subsets B of $\mathbb{T}^3$:*

$$\mu^{(3)}(B) = \int_{\mathbb{T}^2} \mu_x(B \cap T_x)\, d\mu^{(2)}(x)$$

where $\mu^{(2)}$ is the Haar measure on $\mathbb{T}^2$. We can think of this as saying that $\mu^{(3)}$ is a 'convex combination' of the ergodic measures μ_x. The visualisation of this is shown in (8).

The above example suggests the use of **ergodic decomposition** (see Appendix A) which tells us that a complete classification of invariant measures requires only that we classify those that are ergodic. Ratner provides precisely such a classification:

Theorem 4.3.2 (Ratner's Measure Classification Theorem)**.** *Let*

- *G be a connected Lie group*
- *Γ be a lattice in G*
- *ϕ_t be a unipotent flow on G/Γ*

Then every ergodic ϕ_t-invariant probability measure on G/Γ is homogeneous. More precisely, it is the Haar measure μ_S for some $x \in G$ and some connected, closed subgroup S of G such that

- *$\{u^t\} \subseteq S$, and*
- *The image $[Sx]$ of Sx in G/Γ is closed, and has finite S-invariant volume.*

We will see in the subsequent sections that the other two theorems of Ratner (4.1.2, 4.2.3) are corollaries of Measure Classification. Section 5 will discuss this implication and Section 6 will (partially) prove Measure Classification.

4.4 Applications of Ratner's Theorems

A vast literature has been built up in the wake of Ratner's Theorems, of which we list two applications.

(i) Let $V(k,n)$ be the **Stiefel manifold** i.e. the space of k-tuples of linearly independent vectors in $\mathbb{R}^n$ (not to be confused with the Grassmannian cf. Example 2.4.4). In 2003, Gorodnik [30] proved that dense orbits of a lattice Γ in $SL(n,\mathbb{R})$ are uniformly distributed on $V(k,n)$.

(ii) Brown et al. [31] proved **Zimmer's Conjecture** for $SL(m,\mathbb{Z})$: Let Γ be a finite-index subgroup of $SL(n,\mathbb{Z})$, M be a closed manifold, and $\mathrm{Diff}^2(M)$ be the group of C^2-diffeomorphisms on M i.e. twice-continuously differentiable bijections on M with twice-continuously differentiable inverses. Then

   (a) If $\dim M \leq n-2$, then the image of any group homomorphism $\Gamma \to \mathrm{Diff}^2(M)$ is finite, and

   (b) If $\dim M \leq n-1$, $n > 2$ and vol is a volume form on M, $n > 2$, then the image of any group homomorphism $\Gamma \to \mathrm{Diff}^2(M, \mathrm{vol})$ is finite.

The underlying idea in most of these results is that Ratner's Theorems allow extensive control over the behaviour of certain orbits. Provided that an appropriate lattice can be found to act on some object, it is quite likely that a unipotent flow can be found and density, equidistribution or ergodicity can be established. As we have alluded previously, the most celebrated application of Ratner's Theorems is to Diophantine approximation, where the lattice $SL(n, \mathbb{Z})$ in $SL(n, \mathbb{R})$ can act on a quadratic form in n variables. We will defer a more complete treatment to Section 7. Before proceeding further, we dedicate the next two sections to a (partial) proof of Ratner's Theorems.

5 Equivalence Between Ratner's Theorems

Measure Classification (4.3.2) is logically equivalent to Equidistribution (4.2.3). Historically, Ratner proved Equidistribution using Measure Classification with Orbit Closure (4.1.2) as an immediate corollary. This illustrate a general theme in the homogeneous dynamics – it is usually easier to derive information about the behaviour of a flow by examining the measures which are invariant under that flow (cf. Sections 2.5 and 3.3).

5.1 Polynomial Divergence

The heart of Ratner's Theorems is controlling the behaviour of flows. It is insightful here to distinguish between the 'exponential divergence' of geodesic flows and the 'polynomial divergence' of unipotent flows, which we can exhibit explicitly in the case of lattices Γ in $G = SL(2, \mathbb{R})$. In this subsection, Recall the notation used in Example 3.2.5:

$$u^t = \exp(A_1 t) = \begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix}, \quad a^s = \exp(A_1 s) = \begin{bmatrix} e^s & 0 \\ 0 & e^{-s} \end{bmatrix}, \quad v^r = \exp(A_1 r) = \begin{bmatrix} 1 & r \\ 0 & 1 \end{bmatrix},$$

are one-parameter subgroups with corresponding flows η_t, γ_s, ψ_r respectively. η_t and ψ_r are unipotent whereas γ_s is geodesic in $SL(2, \mathbb{R})$.

Proposition 5.1.1 (Polynomial Divergence). *Nearby points of G/Γ move apart at polynomial speed.*

Proof. Let $x \neq y \in G/\Gamma$. As we are working in a homogeneous space, there exists $g \in G$ such that $y = gx$. If x is close to y, then g is close to the $n \times n$ identity matrix I. We formalise this notion of closeness with a metric d on G/Γ:

$$d(x, y) = \inf\{\|g - I\| : g \in G, \ y = gx\}$$

where $\|\cdot\|$ is a (fixed) matrix norm on $\mathrm{Mat}_{2\times 2}(\mathbb{R})$ (the supremum norm is the easiest to intuit in the sequel). We now evolve x and gx under the unipotent flow η_t:

- Initially, we can get from x to gx by multiplying by g, so $d(x, gx) \leq \|g - I\|$.
- At time t, we now want to compare $u^t x$ and $u^t gx$. We can get from the former to the latter by left-multiplying by $u^t g u^{-t}$, so $d(u^t x, u^t gx) \leq \|u^t g u^{-t} - I\|$.

We can then compute directly:

$$g - I = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \quad \Longrightarrow \quad u^t g u^{-t} - I = \begin{bmatrix} c_{11} - c_{12}t & c_{12} \\ c_{21} + (c_{11} - c_{22})t - c_{12}t^2 & c_{22} + c_{12}t \end{bmatrix}$$

All entries of this matrix are polynomials in t, so the claim follows. □

Remark 5.1.2. A similar analysis for the geodesic flow γ_s yields:

$$g - I = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \quad \Longrightarrow \quad r^t g r^{-t} - I = \begin{bmatrix} c_{11} & c_{12}e^{-2t} \\ c_{21}e^{2t} & c_{22} \end{bmatrix}$$

which we observe in general as well. We draw the following comparison:

1. Two geodesic flows are like a passionate couple. Even if they outwardly appear very close, and even if they spend 90% of their lives in holy matrimony, everything could blow up one day in a massive divorce – their lives diverging (exponentially) away from each other. We may observe fractals as fractured as their once amorous union (see Arnoux-Schmidt's paper for 'Cantor-like' behaviour [32]).

2. Two unipotent flows are like a tepid couple with passive attitudes to life. They may end up disliking each other and drift apart (at polynomial speed), but there's no reason to kick up a fuss so they bide their times. "The first year has been a bit rough, but I'm sure that everything will click in a month's time," and later "We've been together for a decade, I'm sure another year couldn't hurt!" and later "I've spent 90% of my life with this person, there's no point changing things up now!" Such is life.

Remark 5.1.2 in the unipotent case can be formalised in precise terms: Firstly, if two points stay close together for a period of time l, then they should stay reasonably close for an additional length of time that is proportional to l. Indeed, polynomials can be bounded as such:

Lemma 5.1.3. *Given $d \in \mathbb{Z}_{\geq 0}$ and $\delta > 0$, there exists an $\epsilon > 0$ such that if $f(x)$ is any real polynomial of degree $\leq d$, $C \in \mathbb{R}^+$, $[k, k+l]$ is any real interval, and $|f(t)| < C$ for all $t \in [k, k+l]$, then $|f(t)| < (1+\delta)C$ for all $t \in [k, k+l+\epsilon l]$.*

This can be proved by a typical $\epsilon - \delta$ argument. Secondly, we can use Lemma 5.1.3 to show that if two orbits are close most of the time, then they are fairly close all of the time:

Lemma 5.1.4. *For any $\epsilon > 0$, there is a $\delta > 0$ such that if $d(u^t x, u^t y) < \delta$ for 99% of the times t in an interval $[a, b]$, then $d(u^t x, u^t y) < \epsilon$ for **all** of the times t in the interval $[a, b]$.*

Thus, Polynomial Divergence says that we only need good bounds on divergence **most** of the time to bound the divergence **all** of the time. This has profound implications for the asymptotic properties of unipotent flows, as we expect it to 'stay away from infinity' most of the time.

Definition 5.1.5 (Dani-Margulis Theorem [14])**.** Suppose:

- Γ is a lattice in a Lie group G,
- u^t is a unipotent one-parameter subgroup of G,
- $x \in G/\Gamma$,
- $\epsilon > 0$, and
- λ is the Lebesgue measure on $\mathbb{R}$.

Then there is a compact subset K of G/Γ, such that

$$\limsup_{L \to \infty} \frac{\lambda\left(t \in [0, L] \mid u^t x \notin K\right)}{L} < \epsilon.$$

This tells us about the long-term behavior of orbits under unipotent flows: they cannot escape to infinity and are mostly contained within a bounded region.

5.2 Measure Classification Implies Equidistribution

The idea is to describe the proportion of time spent by a flow in some measurable set A as a *measure* μ_L on which we can apply Measure Classification (4.3.2). In the limit, we 'smear' out this proportion and hope to arrive at the volume form μ_∞ of the orbit closure. Although we continue to use the notation from Example (3.2.5), the proof in this section should work in general if we accept that Polynomial Divergence (5.1.1) can be extended beyond the $SL(2, \mathbb{R})$ case (see Remark 6.2.6). We proceed following Morris [14]:

Three Assumptions

To facilitate the proof, we make three assumptions:

1. **Compactness:** Assume that G/Γ is compact. This simplifies the space of probability measures, making it (sequentially) compact in an appropriate weak* topology . This means that every sequence of probability measures has a convergent subsequence (see Appendix A.5).

   Compactness Justification: If G/Γ is not compact, consider its one-point compactification (C.2) $X = (G/\Gamma) \cup \{\infty\}$. The limit measure μ_∞ gives measure 0 to the point at infinity. This result is a consequence of the Polynomial Divergence of Orbits, which ensures that μ_∞ is supported (C.1) on G/Γ. We do not have time for the full details, but the intuitive understanding of this justification is through the Dani-Margulis Theorem (5.1.5). The limit measure μ in the later proof at the point ∞ in the one-point compactification (which represents a state where the orbit might theoretically escape to infinity) is negligible in the limit.

2. **Ergodicity:** Assume that μ is ergodic, which allows us to leverage the Measure Classification Theorem directly. The importance of ergodicity is highlighted in (3.3).

   Ergodicity Justification: If the limit measure μ_∞ is not ergodic, it can be decomposed into ergodic components (see Theorem A.4) which must be homogeneous by Measure Classification (4.3.2). In particular, for each ergodic component μ_e, there exists a point $x_e \in G$ and a closed, connected subgroup S_e of G such that:

   - $\{u^t\} \subset S_e$,
   - The image $[S_e x_e]$ is closed and has finite S_e-invariant volume,
   - $\mu_e = \text{vol}_{[S_e x_e]}$

   The technicalities of this assumption are addressed in [14].

3. **Proper Subgroup:** Assume that there does not exist any connected, closed, proper subgroup S of G such that:

   - $\{u^t\} \subset S$,
   - The image $[Sx]$ of Sx in G/Γ is closed and has finite S-invariant volume.

   This is only a matter of convenience, and we may pass from G to S if need be as we aim to prove that $S = G$.

The Actual Proof

For readability, we modify parts of the proof in Morris's book [14].

Proof of Theorem 4.2.3. We wish to show that the u^t-orbit of $[x]$ is uniformly distributed in all of G/Γ with respect to the G-invariant volume on G/Γ. We define

$$\mu_T(A) = \frac{1}{T}\lambda\{t \in [0,T] \mid \phi_t(x) \in A\}$$

where ϕ_t is the flow induced by the unipotent one-parameter subgroup $\{u^t\}$.

We wish to show that the measures μ_T converge to $\mathrm{vol}_{G/\Gamma}$ as $T \to \infty$, which indicates equidistribution. We have assumed that G/Γ is compact (see Assumption 1). According to the weak* topology on the space of probability measures (A.5), it suffices to show that if μ_{T_n} is any convergent subsequence, then the limit measure $\mu_\infty = \mathrm{vol}_{G/\Gamma}$, which is a functional analysis result that we will not prove. It remains to show the limit measure $\mu_\infty = \mathrm{vol}_{G/\Gamma}$.

We now want to use measure classification, but to use the theorem, we need μ_∞ to be u^t-invariant and ergodic. It is easy to see that μ_∞ is u^t-invariant:

$$\begin{aligned} LHS &= \lim_{T\to\infty} \frac{1}{T}\lambda\left\{t \in [0,T] \mid u^t x \in u^s A\right\} \\ &= \lim_{T\to\infty} \frac{1}{T}\lambda\left\{t \in [0,T] \mid u^{t-s} x \in A\right\} \\ &= \lim_{T\to\infty} \frac{1}{T}\lambda\left\{r \in [-s,T-s] \mid u^r x \in A\right\} \end{aligned}$$

Since λ is the Lebesgue measure here, we can simply translate $[-s, T-s]$ to $[0,T]$ and invariance follows. We have also assumed that μ_∞ is ergodic (see Assumption 2). Hence, we can now use the Measure Classification Theorem (4.3.2): There exists a connected, closed subgroup S of G and some point $x' \in G$ such that:

- $\{u^t\} \subset S$,
- The image $[Sx']$ of Sx' in G/Γ is closed and has finite S-invariant volume,
- $\mu_\infty = \mathrm{vol}_{[Sx']}$.

We are ignoring some technicalities [14], but the main idea is that the orbit of $[x]$ must spend more than 99 percent of its time very close to $[Sx']$. By Polynomial Divergence of Orbits (5.1.4), this implies that the orbit spends all of its life fairly close to $[Sx']$. Since the distance to $[Sx']$ is a bounded polynomial function, the constant must be 0, concluding that $[x] \in [Sx']$. This is similar in spirit to the idea of Theorem 5.1.5.

Since $[x] \in [Sx']$, we have $[Sx] \subset [Sx']$. Now with Assumption 3, we get $S = G$, then $\mu_\infty = \mathrm{vol}_{[Sx']} = \mathrm{vol}_{[Gx']}$. Lastly, since $x' \in G$ and we know G/Γ is a homogeneous space which has transitivity (2.4.1), we can get $\mathrm{vol}_{[Gx']} = \mathrm{vol}_{G/\Gamma}$. Hence, we have $\mu_\infty = \mathrm{vol}_{G/\Gamma}$.

□

6 Towards a Proof of Ratner's Measure Classification Theorem

Historically, the first of Ratner's theorems to be proven is Measure Classification (4.3.2). They were proved by Ratner in their entirety in a series of three papers spanning 178 pages [2, 33, 34]. She would later provide a simpler proof for the $G = SL(2,\mathbb{R})$ case [35] at only 31 pages long, which was shortened still by Einsiedler [36] to 26 pages. Due to time constraints, we will restrict our attention to presenting the key ideas at play in this special case:

Theorem 6.0.1 (Ratner's Measure Classification Theorem for $G = SL(2,\mathbb{R})$)**.** *Let*

- *Γ is any lattice in $G = SL(2,\mathbb{R})$, and*
- *η_t be the unipotent flow on G/Γ with corresponding one-parameter subgroup $\{u^t\}$.*

Then every ergodic η_t-invariant probability measure on G/Γ is homogeneous (cf. Theorem 4.3.2).

We largely follow [14] with modifications for readability. Recall again the notation used in Example 2.2.9 and Example 3.2.5:

$$u^t = \exp(Ft) = \begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix}, \quad a^s = \exp(Hs) = \begin{bmatrix} e^s & 0 \\ 0 & e^{-s} \end{bmatrix}, \quad v^r = \exp(Er) = \begin{bmatrix} 1 & r \\ 0 & 1 \end{bmatrix},$$

with corresponding flows η_t, γ_s, ψ_r respectively. η_t and ψ_r are unipotent and γ_s is geodesic.

6.1 Shearing Property

Polynomial Divergence (5.1.1) offers a quantitative description of how nearby unipotent flows move away from each other. We now take a more qualitative view by describing the direction of this divergence.

Proposition 6.1.1 (Shearing Property)**.** *The fastest relative motion between two nearby points is parallel to the orbits of the flow. More precisely, let*

- *$\{u^t\}$ be a unipotent one-parameter subgroup of G, and*
- *x and y be nearby points in G/Γ.*

Then either

- *there exists $t > 0$ such that $u^t y \approx u^{t\pm 1}x$, or*
- *$y = u^\epsilon x$ for some $\epsilon \approx 0$. $(*)$*

As in Section 5.1, we start by setting $y = gx$ where g is close to the $n \times n$ identity matrix I (see Figure 9). We now evolve the points under the flow η_t to get $u^t x$ and $u^t gx$, and we can get the latter from the former by multiplying by $u^t g u^{-t}$ (so we may take this as a proxy for the distance between the flows at time t). The case $(*)$ in the Shearing Property (6.1.1) above captures the exceptional case when y lies on the η_t-orbit of x, where clearly $u^t g u^{-t} = g$ is fixed for all t since $g = u^\epsilon$. This says that the points $u^t x$ and $u^t gx$ move along parallel orbits with no relative motion at all. More generally, this occurs when g belongs to the **centraliser** of u^t:

$$C_G(u^t) = \{g \in G \mid g = u^t g u^{-t} \text{ for all } t\}$$

A straightforward computation gives $C_{SL(2,\mathbb{R})}(u^t) = \{u^t\} \cup \{-u^t\}$. Since $C_G(u^t)$ can include directions not parallel to $\{u^t\}$ in general, it is important to recognise that Proposition 6.1.1 fails to hold for $SL(n,\mathbb{R})$ when $n \geq 3$. Thus, we must accommodate for motion along the centraliser:

Proposition 6.1.2 (Generalised Shearing Property)**.** *The fastest relative motion between two nearby points is along some direction in the centraliser of u^t. More precisely, if*

- *$\{u^t\}$ is a unipotent one-parameter subgroup of G, and*
- *x and y are nearby points in G/Γ*

then either

- *there exists $t > 0$ and $c \in C_G(u^t)$ such that $\|c\| = 1$ and $u^t x \approx c u^t y$, or*
- *there exists $c \in C_G(u^t)$ with $c \approx I$ such that $y = cx$.*

Proof. We work in exponential coordinates (Proposition 2.2.21). In particular, choose

- $\underline{u} \in \mathfrak{g}$ with $\exp(t\underline{u}) = u^t$, and
- $\underline{g} \in \mathfrak{g}$ with $\exp(\underline{g}) = g$.

Then

$$
\begin{aligned}
\underline{u^t g u^{-t}} &= (\mathrm{Ad}(u^t))(\underline{g}) && \text{(Proposition 2.3.4(i))} \\
&= \exp(\mathrm{ad}(t\underline{u}))(\underline{g}) && \text{(Proposition 2.3.4(iii))} \\
&= \underline{g} + t(\mathrm{ad}(\underline{u}))(\underline{g}) + \frac{1}{2}t^2(\mathrm{ad}(\underline{u}))^2(\underline{g}) + \frac{1}{6}t^3(\mathrm{ad}(\underline{u}))^3(\underline{g}) + \cdots && \text{(Example 2.2.20)}
\end{aligned}
$$

By Remark 3.2.4, $\mathrm{ad}(\underline{u})$ is nilpotent since u^t gives a unipotent flow. For large t, the term that contributes the most to $\underline{u^t g u^{-t}}$ is the one with the highest power of t i.e. the last nonzero term $(\mathrm{ad}\,\underline{u})^k(\underline{g})$. Then

$$
\begin{aligned}
\left[\underline{u}, (\mathrm{ad}\,\underline{u})^k(\underline{g})\right] &= (\mathrm{ad}\,\underline{u})((\mathrm{ad}\,\underline{u})^k \underline{g}) && \text{(Proposition 2.3.4(ii))} \\
&= (\mathrm{ad}\,\underline{u})^{k+1}\underline{g} = 0 && \text{(Nilpotency)}
\end{aligned}
$$

Thus, $(\mathrm{ad}\,\underline{u})^k(\underline{g})$ commutes with $\underline{u}$. Passing back from the exponential coordinates says that fastest relative motion is along some direction in the centraliser of u^t. □

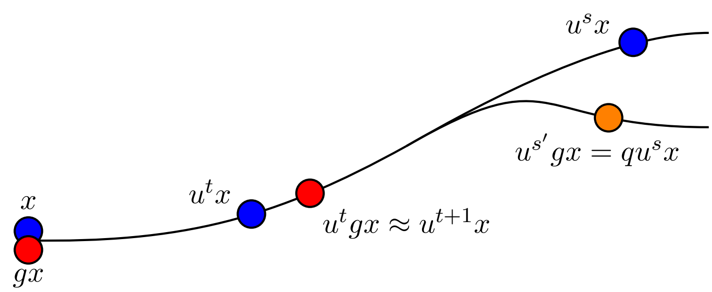


Figure 9 Shearing is 'first-order' whereas transverse divergence is 'second-order'.

6.2 Transverse Divergence

Returning to $G = SL(2,\mathbb{R})$, the current state of affairs is that two nearby points $x \approx y \in G/\Gamma$ will first stay close together by continuity $u^t x \approx u^t y$. The first time we can spot a difference is when we observe 'shearing' (6.1.1) where one point speeds ahead of the other while remaining roughly on the same orbit. The points will continue to diverge at 'polynomial speed' (5.1.1) and eventually we may expect to observe deviations of their trajectories when we 'mod out' the relative motion along the orbit. More precisely, let $u^{s'}y$ be the point on the orbit of y that is closest to $u^s x$. Writing $u^s x = qu^{s'}y$, we say that $q - I$ represents the **transverse divergence** (see Figure 9):

Proposition 6.2.1 (Transverse Divergence)**.** *The fastest transverse motion is along some direction in the normaliser of* $\{u^t\}$.

Proof. Let $\mathfrak{u}$ be the Lie algebra of the unipotent one-parameter subgroup $\{u^t\}$. As in the proof of the Generalised Shearing Property (6.1.2), any term in the expansion of $\underline{u^t g u^{-t}}$ that belongs to $\mathfrak{u}$ represents motion along $\{u^t\}$. Thus, the fastest transverse motion is represented by the last term $(\operatorname{ad}\underline{u})^k \underline{g}$ that is not in $\mathfrak{u}$. Then $(\operatorname{ad}\underline{u})^{k+1}\underline{g} \in \mathfrak{u}$ i.e. $[\underline{u}, (\operatorname{ad}\underline{u})^k(\underline{g})] \in \mathfrak{u}$. Therefore $(\operatorname{ad}(\underline{u}))^k \underline{g}$ normalises $\mathfrak{u}$. ☐

To develop a measure-theoretic version of Transverse Divergence 6.2.1 (e.g. so it works with Measure Classification 4.3.2), we require the following:

Definition 6.2.2. Let G act on a measurable space (X, Σ). We can also think of G acting on the **space of probability measures** $\mathcal{P}(X)$ on X (see Appendix A) by the **pushforward** i.e. $g \in G$ sends $\mu \in \mathcal{P}(X)$ to $g_\sharp\mu$ where $g_\sharp\mu(A) = \mu(g^{-1}A)$ for all $A \in \Sigma$. The **stabiliser** of $\mu \in \mathcal{P}(X)$ is

$$\operatorname{Stab}_G(\mu) = \{g \in G \mid g_\sharp\mu = \mu\}$$

Considering an ergodic u^t-invariant probability measure μ on G/Γ, we get

Remark 6.2.3. We know that $\operatorname{Stab}_G(\mu)$ is an abstract subgroup. Endowing $\mathcal{P}(X)$ with the weak* topology (A.5), it can be shown that the map $f : g \mapsto g_\sharp\mu$ is continuous. Since $\operatorname{Stab}_G(\mu) = f^{-1}\{\mu\}$ and $\{\mu\}$ is closed in the weak* topology (singleton sets are closed in Hausdorff topologies; the weak* topology is Hausdorff [37]), it must be the case that $\operatorname{Stab}_G(\mu)$ is closed subgroup of G. Connectedness (under certain conditions) requires more effort that we could not address due to time constraints.

Corollary 6.2.4 (Transverse Divergence and Measure Invariance 1)**.** *For almost all $x \approx y$, the fastest transverse motion is along some direction in* $\operatorname{Stab}_G(\mu)$.

Corollary 6.2.5 (Transverse Divergence and Measure Invariance 2)**.** *For almost all $x \approx y$, the fastest transverse motion to the U-orbits is along some direction in* $\operatorname{Stab}_G(\mu)$.

Remark 6.2.6. The Shearing Property (6.1.2) was generalised by Morris [38] to what he calls the 'Ratner property', which he used to prove rigidity results closely related to Ratner's Theorems. This is not to be confused with the 'R-property' introduced by Ratner [33, 34] which combines Polynomial Divergence (5.1.1) with Transverse Divergence (6.2.1) for higher-dimensional unipotent subgroups. This plays an essential role in Ratner's full proof, but we omit the statement as it requires substantial machinery.

6.3 Entropy

The last ingredient we'll need to prove Measure Classification (6.0.1) is **entropy** (though there are many more ingredients in Ratner's full proof). Much like ergodicity (see Section 3.3), entropy is a general property that can describe various mathematical objects. At its core, it captures how 'informative', 'predictable' or 'surprising' something is. For the purposes of the proof, we treat entropy as a convenient invariant. We begin with some definitions following [14]:

Definition 6.3.1. Suppose $\mathcal{A} = \{A_1, \ldots, A_m\}$ is a partition of a probability space (Ω, μ) into finitely many measurable sets. The **entropy** of $\mathcal{A}$ is

$$H(\mathcal{A}) = -\sum_{i=1}^{m} \mu(A_i) \log \mu(A_i)$$

where the base of $\log$ is typically 2 or e. By convention, the summand is 0 when $\mu(A_i) = 0$.

We can combine partitions by refining their intersections:

Definition 6.3.2. The **join** of two partitions $\mathcal{A}$ and $\mathcal{B}$ is $\mathcal{A} \vee \mathcal{B} = \{A \cap B \mid A \in \mathcal{A},\ B \in \mathcal{B}\}$. This has entropy

$$H(\mathcal{A} \vee \mathcal{B}) = H(\mathcal{A}) + H(\mathcal{B}|\mathcal{A})$$

where $H(\mathcal{B}|\mathcal{A})$ is the **conditional entropy** of $\mathcal{B}$ given $\mathcal{A}$. Formally,

$$H(\mathcal{B}|\mathcal{A}) = \sum_{A \in \mathcal{A}} H_A(\mathcal{B})\mu(A) \quad \text{where} \quad H_A(\mathcal{B}) = H\left(\frac{\mu(B_1 \cap A)}{\mu(A)}, \ldots, \frac{\mu(B_n \cap A)}{\mu(A)}\right)$$

This can be interpreted as 'performing multiple experiments at once'. In fact, the only function H that satisfies Definition 6.3.2 is of the form given in Definition 6.3.1 for some choice of base of the logarithm. We now adopt a more 'dynamic' view:

Definition 6.3.3. Let $T : \Omega \to \Omega$ be a measurable transformation with a measurable inverse and a T-invariant probability measure μ on Ω. The total amount of information expected to be obtained after k daily experiments is

$$E^k(T, \mathcal{A}) = H(\mathcal{A} \vee T^{-1}(\mathcal{A}) \vee \cdots \vee T^{-(k-1)}(\mathcal{A}))$$

The average information in the limit gives the **entropy** $h(T)$ of T as

$$h(T, \mathcal{A}) = \lim_{k \to \infty} \frac{E^k(T, \mathcal{A})}{k}, \quad h(T) = \sup h(T, \mathcal{A})$$

where the supremum is taken over all partitions $\mathcal{A}$ of Ω into finitely many measurable sets.

Definition 6.3.4. The **entropy** of a flow ϕ_t is defined to be the entropy of its time-one map: that is $h(\{\varphi_t\}) = h(\varphi_1)$.

If it necessary to specify the measure μ, we write $h_\mu(T)$. We now demonstrate various properties of entropy which can be proved straightforwardly:

Proposition 6.3.5 (Properties of Entropy). *Let $\Omega, \mu, T, \mathcal{A}$ be defined as above. Then*

(i) $H(T^l\mathcal{A}) = H(\mathcal{A})$ *for all* $l \in \mathbb{Z}$ *(Notation:* $T^l\mathcal{A} = \{T^l(A_1), \ldots, T^l(A_m)\}$

(ii) ***Inverse Invariance:*** $h(T) = h(T^{-1})$.

(iii) ***Isomorphism Invariance:*** *If* $\psi : (\Omega, \mu) \to (\Omega', \mu')$ *is a measure-preserving map such that* $\psi(T(\omega)) = T'(\psi(\omega))$ *for μ-almost every ω, then* $h_\mu(T) = h_{\mu'}(T')$.

Proposition 6.3.6 (Entropy Estimate). *Let*

- Γ *be a lattice in a connected Lie group* G,
- γ_s *is a geodesic flow with corresponding one-parameter subgroup* $\{a^s\}$,
- $\{u^t\}$ *is a unipotent one-parameter subgroup associated to* γ_s *(see [14] for the technicalities),*
- μ *be an* a^s*-invariant probability measure on* G/Γ.

Then

$$h_\mu(a^s) \leq 2|s| \quad \textit{with equality if and only if } \mu \textit{ is } \{u^t\}\textit{-invariant}$$

6.4 Proof of the $SL(2,\mathbb{R})$ Case

We are now prepared to prove Measure Classifcation (6.0.1) in the $G = SL(2,\mathbb{R})$, $\Gamma = SL(2,\mathbb{Z})$ case. For readability, we state the proof first (following [14]) and defer the technicalities:

Proof of Theorem 6.0.1. Let μ be any ergodic η_t-invariant probability measure on G/Γ and let $S = \mathrm{Stab}_G(\mu)$ (see Definition 6.2.2). By Remark 6.2.3, S is a connected, closed subgroup of G. Further, since μ is η_t-invariant, we know that $\{u^t\} \subseteq S$. Seemingly, S is precisely the subgroup we need for Measure Classification.

We now aim to show that μ is supported on S. If $\{u^t\} = S^\circ$, then S° is unipotent which we handle in Lemma 6.4.1 using Transverse Divergence. Assuming instead that $\{u^t\} \neq S^\circ$, it must be the case that $\{a^s\} \subset S^\circ$ (linear algebra; we will not prove this). This lets us apply Lemma 6.4.2, showing that S must in fact contain $\{v^r\}$ so $G = S$ and μ is the G-invariant Haar measure on G/Γ which is homogeneous as desired. $\square$

Lemma 6.4.1. *Let μ be an ergodic η_t-invariant probability measure on G/Γ. If $U = \mathrm{Stab}_G(\mu)^\circ$ is unipotent, then μ is supported on a single U-orbit.*

Proof. Recall that μ is ergodic if ϕ_t-invariant subsets of G/Γ are null or conull (Definition 3.3.1), so it suffices to find a U-orbit of positive measure. We suppose by contradiction that all U-orbits have measure 0. We will make the stronger assumption that all $N_G(U)$-orbits have measure 0.

For almost every $x \in G/\Gamma$, there exists $y \approx x$ such that $y \notin N_G(U)(x)$ and y is in the support of μ. The former tells us that the U-orbit of y has nontrivial transverse divergence from the U-orbit of x, so $u'y \approx cux$ for some $u, u' \in U$ and $c \notin U$. From Corollary 6.2.5, we know that $c \in \mathrm{Stab}_G(U)$, but this contradicts $U = \mathrm{Stab}_G(\mu)^\circ$. $\square$

The final lemma is most succinctly achieved by using entropy (though the underlying machinery was to cumbersome to include in this paper):

Lemma 6.4.2. *Let Γ be a lattice in $G = SL(2,\mathbb{R})$ and let μ be an ergodic $\{u^t\}$-invariant probability measure on G/Γ. If μ is $\{a^s\}$-invariant, then μ is $SL(2,\mathbb{R})$-invariant.*

Proof.

$$\begin{aligned} 2|s| &= h_\mu(a^s) && \text{(Entropy Estimate 6.3.6)} \\ &= h_\mu(a^{-s}) && \text{(Entropy is inverse invariant (6.3.5(ii))} \end{aligned}$$

There is an automorphism of $SL(2,\mathbb{R})$ that maps a^s to a^{-s} and interchanges $\{u^t\}$ with $\{v^r\}$ (linear algebra; we will not prove this), so the $h_\mu(a^{-s}) = 2|s|$ says that μ is v^r-invariant (in spirit, this is because entropy is isomorphism invariant (6.3.5(iii)). Since $\{u^t\}$, $\{a^s\}$ and $\{v^r\}$ taken together generate the entire $SL(2,\mathbb{R})$, we conclude that μ is $SL(2,\mathbb{R})$-invariant. $\square$

In order to see the proof more concretely, we present a geometric argument (as in [14]) that μ is the Haar measure on G/Γ. Our first observation is that γ_s is not an isometry. Explicitly,

$$\gamma_s(u^t x) = u^{e^{2s}t}\gamma_s(x), \quad \gamma_s(a^t x) = a^t\gamma_s(x), \quad \gamma_s(v^t x) = v^{e^{-2s}t}\gamma_s(x)$$

so infinitesimal distances in u^t-, a^s- and v^r-orbits are multiplied by e^{2s}, 1 and e^{-2s} respectively. Nevertheless, γ_s is volume-preserving because $e^{2s} \cdot 1 \cdot e^{-2s} = 1$ (see Figure 10 below).

Our second observation is that γ_s preserves μ since $\{a^s\} \subset S$. We seek to combine our understanding of the dynamics of η_t and γ_s, so we consider

$$B = \{a^s u^t \mid s, t \in \mathbb{R}\}$$

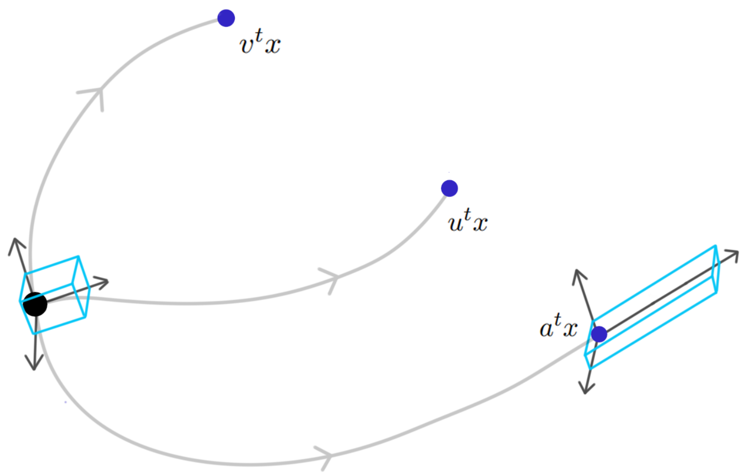


Figure 10

which is a subgroup of G since $\{a^s\}$ normalises $\{u^t\}$. Choose a small (2-dimensional) disk D in some B-orbit. For some (fixed) small $\epsilon > 0$ and each $d \in D$, let $\mathcal{B}_d = \{v^r d : 0 \le r \le \epsilon\}$. Let $\mathcal{B} = \bigcup_{d \in D} \mathcal{B}_d$ so it is the disjoint union of fibers. Then the restriction $\mu|_{\mathcal{B}}$ can be decomposed as an integral of probability measures on the fibers:

$$\mu|_{\mathcal{B}} = \int_D \mu_d \, \nu(d)$$

where μ_d is a probability measure on $\mathcal{B}_d$ (cf. Example 4.3.1). The map γ_s multiplies areas in D by $e^{2s} \cdot 1 = e^{2s}$. Thus, because μ is γ_s-invariant, the contraction along the fibers $\mathcal{B}_d$ must exactly cancel this: for $X \subseteq \mathcal{B}_d$, we have

$$\mu_{\gamma_s(d)}(\gamma_s(X)) = e^{-2s} \mu_d(X)$$

We conclude that the fiber measures μ_d scale exactly like the Lebesgue measure on $[0, \epsilon]$. This implies that μ_d must be precisely the Lebesgue measure, and μ is the Haar measure on G/Γ.

It is important to recognise that we have provided a hugely truncated proof of Measure Classification. We state many key theorems as ‘black boxes’ and omit many additional ideas used by Ratner. Due to time constraints, we now shift our focus to applications of Ratner’s Theorems.

7 Application: Diophantine Approximation

Diophantine approximation [39] is a branch of number theory that deals with the approximation of real numbers by rational numbers. The fact that $\mathbb{Q}$ is dense in $\mathbb{R}$ can be made more quantitative by specifying the range of denominators for an approximation:

Theorem 7.0.1 (Dirichlet’s Theorem on Diophantine Approximation [40]). *For any real number α and any positive integer N, there exist integers p and q (with $1 \le q \le N$) such that:*

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{qN}$$

There is a surprising interaction here with homogeneous dynamics, which has attracted mathematicians of the highest calibre – not least of which is Lindenstrauss, who was awarded the Fields Medal for his work on Quantum Unique Ergodicity and the Littlewood conjecture about approximations to irrational numbers [41]. A landmark result in this area is Margulis' proof of the Oppenheim-Davenport Conjecture (7.3) which practically revived the theory of the geometry of numbers and is tightly related to Ratner's Theorems. The core of this conjecture concerns approximation by quadratic forms, which will be our main focus in the sequel.

7.1 Quadratic Forms

Definition 7.1.1 (Quadratic Form [40])**.** A quadratic form in n variables is a homogeneous polynomial of degree 2, which can be expressed as:

$$Q(x_1, x_2, \ldots, x_n) = \sum_{1 \leq i \leq j \leq n} a_{ij} x_i x_j$$

where a_{ij} are coefficients (e.g. in $\mathbb{Z}$ or $\mathbb{R}$).

Example 7.1.2. *Determining the values that a quadratic form may take when evaluated over the integers is a highly nontrivial result. Classical results from number theory [40] give the following:*

(i) ***Sum of Two Squares Theorem****: Let $Q(x_1, x_2) = x_1^2 + x_2^2$. Then*

$$Q(\mathbb{Z}^2) = \{n \in \mathbb{Z} \mid \textit{Prime factorisation of } n \textit{ does not contain} \\ p^k \textit{ where } p \equiv 3 \ (\textit{mod } 4) \textit{ and } k \textit{ is odd}\}$$

(ii) ***Legendre's Three Square Theorem****: Let $Q(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$. Then*

$$Q(\mathbb{Z}^3) = \{n \in \mathbb{Z} \mid n \neq 4^a(8b + 7) \textit{ for some } a, b \in \mathbb{Z}_{\geq 0}\}$$

(iii) ***Lagrange's Four Square Theorem****: Let $Q(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 + x_3^2 + x_4^2$. Then*

$$Q(\mathbb{Z}^4) = \mathbb{Z}_{\geq 0}$$

If we allow non-integer coefficients, the quadratic form may take irrational values when evaluated over the integers – bringing us into the realm of Diophantine approximation. We now review the theory of quadratic forms.

Definition 7.1.3. [14]

(i) A quadratic form Q is **indefinite** if it takes both positive and negative values.

(ii) A quadratic form Q is **non-degenerate** if there does not exist a nonzero vector $x \in \mathbb{R}^n$ such that $Q(v + x) = Q(v - x)$ for all $v \in \mathbb{R}^n$.

Example 7.1.4. *Consider the quadratic form $Q(x, y) = x^2 - y^2$. It is indefinite since $Q(1, 0) = 1 > 0$ and $Q(0, 1) = -1 < 0$. It is also non-degenerate. Suppose there exists a nonzero vector $x \in \mathbb{R}^n$ such that: $Q(v + x) = Q(v - x)$ for all $v = (v_1, v_2) \in \mathbb{R}^n$. Computation shows $v_1 x_1 = v_2 x_2$. Since this equation to hold for all v, we have $x \neq 0$ which contradicts to our assumption.*

Definition 7.1.5 (Signature of a Non-Degenerate Quadratic Form)**.** Let Q be a non-degenerate quadratic form in n variables. The **signature** [14] of Q is a pair $(k, n - k)$, where k is determined as follows: There exists a unique $k \in \{0, 1, \ldots, n\}$ such that there is an invertible linear transformation T of $\mathbb{R}^n$ with $Q = Q_{k,n-k} \circ T$ [42]. Here, $Q_{k,n-k}$ is the quadratic form in n variables with k positive squares and $n - k$ negative squares.

In order to employ Ratner's Theorems (4), we must construct a group action that complies with the quadratic form. We already know that $SO(n)$ is such a group action (see 2.4.4) as it preserves the sphere. Thus, we define:

Definition 7.1.6 (Orthogonal Group of a Quadratic Form).

(a) If Q is a quadratic form in n variables, then $SO(Q)$ is the orthogonal group of Q. That is,

$$SO(Q) = \{h \in SL(n, \mathbb{R}) \mid Q(vh) = Q(v) \text{ for all } v \in \mathbb{R}^n\}.$$

(Actually, this is the special orthogonal group, because we are including only the matrices of determinant one.)

(b) As a special case, $SO(m, n)$ is a shorthand for the orthogonal group $SO(Q_{m,n})$, where

$$Q_{m,n}(x_1, \dots, x_{m+n}) = x_1^2 + \cdots + x_m^2 - x_{m+1}^2 - \cdots - x_{m+n}^2.$$

Specifically, for the quadratic form of signature $(2, 1)$, $SO(2, 1)$ is defined as:

$$SO(2,1) = \left\{ g \in SL(3, \mathbb{R}) \mid g^T \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} g = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \right\}$$

(c) Furthermore, we use $SO(m)$ to denote $SO(m, 0)$ (which is equal to $SO(0, m)$).

Due to time constraints, now state various Lie-theoretic results which we will require later (see [14]):

The theorems discussed in (7.1.2) deal with positive definite quadratic forms [40]. On the contrary, the Oppenheim-Davenport Conjecture (now called Margulis' Theorem) deals with indefinite quadratic forms.

7.2 Oppenheim-Davenport Conjecture / Margulis' Theorem

Davenport and Heilbronn [43] worked on the problem of representing values of indefinite quadratic forms and provided significant partial results. They demonstrated that for certain forms and under specific conditions, the values could be approximated densely. For an indefinite quadratic form Q in five or more variables, Davenport and Heilbronn showed the values at integer points are dense in $\mathbb{R}$ under certain conditions.

This was strengthened by Grigory Margulis [1] who used results from ergodic theory and the dynamics of unipotent flows to prove a similar result for $n \geq 3$. Margulis' proof involves demonstrating that the action of a unipotent group on the space of lattices is ergodic.

Theorem 7.2.1 (Margulis' Theorem / Oppenheim-Davenport Conjecture [14]). *For an indefinite quadratic form in $n \geq 3$ variables $Q(x_1, x_2, \dots, x_n)$, which is not a multiple of a form with integer coefficients, the set $Q(\mathbb{Z}^n)$ is dense in $\mathbb{R}$. Specifically, for any real number α and any $\epsilon > 0$, there exist integers $x_1, x_2, \dots, x_n$ such that:*

$$|Q(x_1, x_2, \dots, x_n) - \alpha| < \epsilon$$

Example 7.2.2. *[14] If $Q(x, y, z) = x^2 - \sqrt{2}xy + \sqrt{3}z^2$, then Q is not a scalar multiple of a form with integer coefficients ($\sqrt{2}$ and $\sqrt{3}$ are algebraically independent), so Margulis' Theorem tells us that $Q(\mathbb{Z}^3)$ is dense in $\mathbb{R}$.*

The $n \geq 3$ condition cannot be dropped because there is a counterexample for $n = 2$:

Example 7.2.3. *[14] Consider the quadratic form $Q(x, y) = x^2 - (3 + 2\sqrt{2})y^2$. $Q(\mathbb{Z}, \mathbb{Z})$ is not dense in $\mathbb{R}$.*

It turns out that the quadratic forms under consideration can, in a sense, always be reduced to just 3 variables.

Proposition 7.2.4. *Suppose Q satisfies the hypotheses of Theorem 7.2.1. Then there exist integer vectors $v_1, v_2, v_3 \in \mathbb{Z}^n$, such that the quadratic form Q' on $\mathbb{R}^3$, defined by $Q'(x_1, x_2, x_3) = Q(x_1v_1 + x_2v_2 + x_3v_3)$, also satisfies the hypotheses of Theorem (7.2.1).*

Proof. Since Q is indefinite and non-degenerate, select $v_1, v_2 \in \mathbb{Z}^n$ such that $Q(v_1)$ and $Q(v_2)$ have opposite signs and $\frac{Q(v_1)}{Q(v_2)}$ is negative and irrational. Select $v_3 \in \mathbb{Z}^n$ generically to ensure that the new quadratic form $Q'(x_1, x_2, x_3) = Q(x_1v_1 + x_2v_2 + x_3v_3)$ is non-degenerate. The hypotheses of Theorem 7.2.1 for Q' can be verified based on the selections of v_1, v_2, v_3 [14]. $\square$

Morris [14] uses a lot of results in the proof of the Margulis' Theorem (7.3), most of them are quite technical and requires detailed computation that are left as exercises without a detailed justification. These detailed computation are not in the scope of our paper, and instead, we present the general idea of the proof and all the techniques that he used to proof the theorem.

7.3 Proof of Margulis' Theorem

Proof. Let

- $G = SL(3, \mathbb{R})$,
- $\Gamma = SL(3, \mathbb{Z})$,
- $Q_0(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_3^2$, and
- $H = SO(Q_0)^\circ = SO(2, 1)^\circ$.

By Proposition 7.2.4, we can assume Q has exactly three variables without loss of generality. Since Q is indefinite, the signature of Q is either $(2, 1)$ or $(1, 2)$. There exist $g \in SL(3, \mathbb{R})$ and $\lambda \in \mathbb{R}^\times$, such that

$$Q = \lambda Q_0 \circ g.$$

Specifically, we can think consider g acts as a 'change of basis' that reparametrises the quadratic form Q into λQ_0. Note that $SO(Q)^\circ = gHg^{-1}$ ([14]). It can be shown that $H = SO(2, 1)^\circ$ is generated by unipotent elements (see [14] and $SL(3, \mathbb{Z})$ is a lattice in $SL(3, \mathbb{R})$ (Theorem 2.5.6(iii)), we can apply Ratner's Orbit Closure Theorem (4.1.2). The conclusion is that there is a connected subgroup S of G, such that

- $H \subset S$,
- the closure of $[Hg]$ is equal to $[Sg]$, and
- there is an S-invariant probability measure on $[Sg]$.

It can be shown that the only closed, connected subgroups of G that contain H are the two obvious subgroups: G and H [14]. Therefore, S must be either G or H. We consider each of these possibilities separately.

Case 1. Assume $S = G$. This implies that

$$Hg\Gamma \text{ is dense in } G. \tag{*}$$

This is because $[Hg]$ is dense in $[Sg] = [Gg] = G$. We have

$$\begin{aligned} Q(\mathbb{Z}^3) &= Q_0(g\mathbb{Z}^3) && \text{(definition of } g) \\ &= Q_0(g\Gamma\mathbb{Z}^3) && (\Gamma\mathbb{Z}^3 = \mathbb{Z}^3) \\ &= Q_0(Hg\Gamma\mathbb{Z}^3) && (H = SO(Q_0)^\circ \text{ preserves } Q_0) \\ &\simeq Q_0(G\mathbb{Z}^3) && ((*) \text{ and } Q_0 \text{ is continuous}) \\ &= Q_0(\mathbb{R}^3) && (G\mathbb{Z}^3 = \mathbb{R}^3) \\ &= \mathbb{R}, \end{aligned}$$

where "$\simeq$" means "is dense in."

Case 2. Assume $S = H$. This is a degenerate case; we will show that Q is a scalar multiple of a form with integer coefficients. To keep the proof short, we will apply some of the theory of algebraic groups. See Appendix B to fill in the gaps.

Let $\Gamma_g = \Gamma \cap (gHg^{-1})$. Because the orbit $[Hg] = [Sg]$ has finite H-invariant measure, we know that Γ_g is a lattice in $gHg^{-1} = SO(Q)^\circ$. So the Borel Density Theorem (Theorem B.5) implies $SO(Q)^\circ$ is contained in the Zariski closure of Γ_g. Because $\Gamma_g \subset \Gamma = SL(3,\mathbb{Z})$, this implies that the (almost) algebraic group $SO(Q)^\circ$ is defined over $\mathbb{Q}$ (Proposition B.4). Therefore, up to a scalar multiple, Q has integer coefficients (Proposition B.3). This means that **Case 2** cannot happen. This completes the proof.

□

7.4 Quantitative Versions of Margulis' Theorem

There is a quantitative strengthening of Theorem 7.2.1, but the complete proof is beyond the scope of our paper.

Theorem 7.4.1. *Suppose:*

- *Q is a real, nondegenerate quadratic form,*
- *Q is not a scalar multiple of a form with integer coefficients, and*
- *the signature (p, q) of Q satisfies $p \geq 3$ and $q \geq 1$.*

Then, for any interval (a, b) in $\mathbb{R}$, we have

$$\frac{\#\{v \in \mathbb{Z}^{p+q} \mid a < Q(v) < b, \|v\| \leq N\}}{\mathit{vol}\,\{v \in \mathbb{R}^{p+q} \mid a < Q(v) < b, \|v\| \leq N\}} \to 1 \quad \text{as } N \to \infty.$$

We illustrate this result in Figure 11 and numerically verify this in Figure 12.

Remark 7.4.2. The restriction on the signature of Q cannot be eliminated; there are counterexamples of signature $(2, 2)$ and $(2, 1)$. See Eskin's paper [44].

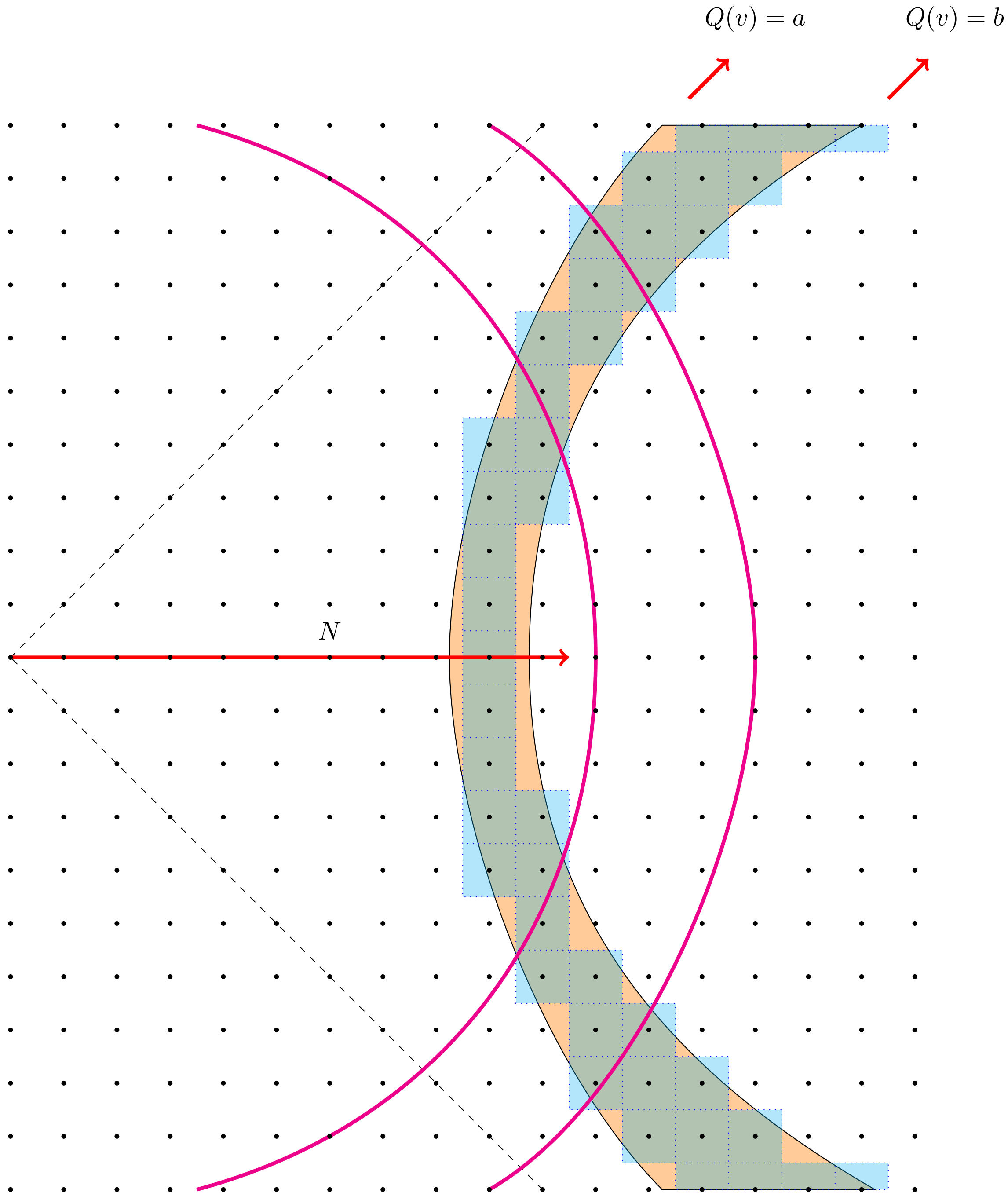


Figure 11 We show a cross-section of the region bounded between $Q(v) = a$ and $Q(v) = b$ which we depict here as a hyperboloid-like shape. If we grow the radius N to expand the bounding sphere, we may include more integer points (depicted here as the points with blue squares). Theorem 7.4.1 says that the amount of points added in this way is approximately the change in volume of the bounded region. The novelty here is that it is indeed possible to describe a set that has nonuniform 'concentrations' of integer points.

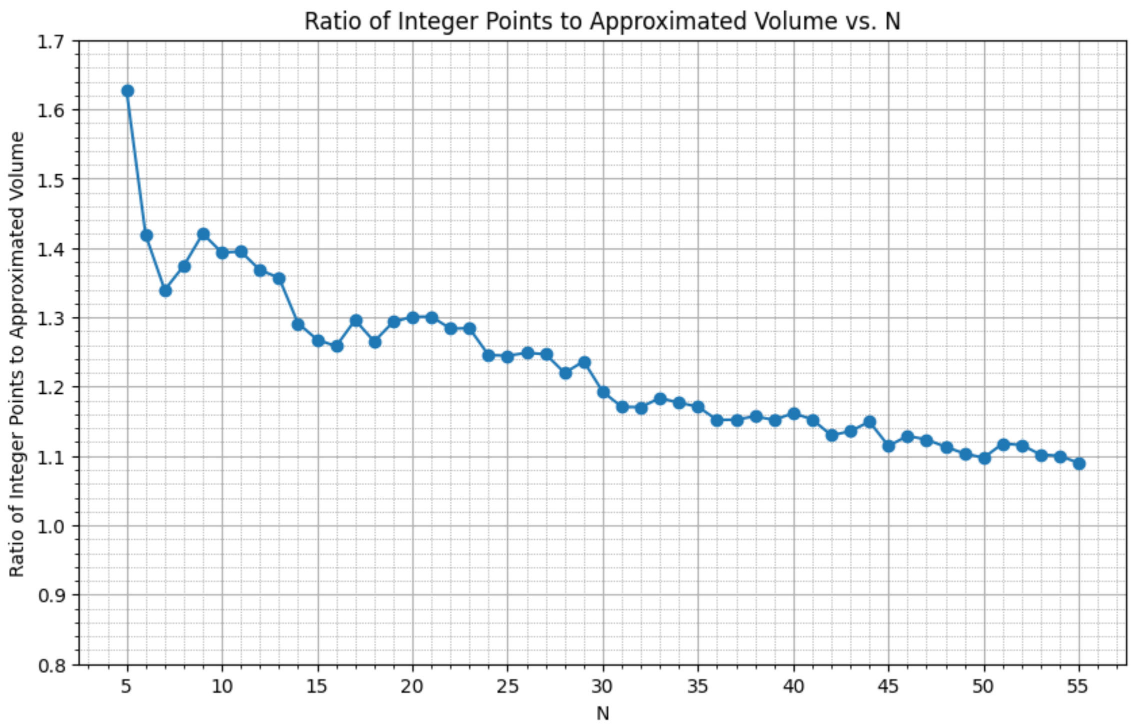


Figure 12 According to Theorem 7.4.1, we can describe the distribution of the number of integer inputs that approximate some real number. Consider the quadratic form $Q(v) = v_0^2 + v_1^2 + \sqrt{2}v_2^2 - v_3^2$ with signature $(p, q) = (3, 1)$ here. We want to approximate the value of 7π, so we let $(a, b) = (7\pi - 1, 7\pi + 1)$ (5% error). To calculate the numerator, we iterate through all possibilities. Instead of calculating the volume in the demonimator analytically, we used a naive Monte Carlo method (see Appendix C) due to time constraints. We indeed observe that the ratio tends towards 1 as $N \to \infty$

.

Appendix A Space of Probability Measures

We will focus on discussing ideas in this appendix rather and make no attempt to define or prove the following results in entirety as these are beyond the scope of this project. The most critical idea here is *ergodicity*, which plays an important role in both the statement of Measure Classification (4.3.2) and in the proof of Equidistribution (5).

Let f be a transformation on X and consider the space of f-invariant Borel probability measures $\mathcal{P}(X)$. Conveniently, $\mathcal{P}(X)$ is convex i.e. if μ_1 and μ_2 are probability measures invariant under the transformation f, then so is $(1-t)\mu_1 + t\mu_2$ for any $0 \leq t \leq 1$. Convexity lets us naturally describe 'extreme' measures $\mu \in \mathcal{P}(X)$ as those measures that cannot be written in the form $(1-t)\mu_1 + t\mu_2$ for any $0 < t < 1$ and $\mu_1, \mu_2 \in \mathcal{P}(X)$. This invites a comparison to:

Theorem A.1 (Carathéodory's Theorem [45])**.** *Any point in a bounded convex polyhedron in $\mathbb{R}^d$ is a convex combination of at most $d+1$ vertices (i.e. the extreme points).*

This captures the idea that a polyhedron can be partitioned into simplices on its vertex set, and indeed this can be readily extended to all convex bodies. This can further be extended to the setting of functional analysis (though this is somewhat beyond the scope of the project):

Proposition A.2 (Krein-Milman Theorem [46])**.** *Every compact convex subset X of a locally convex topological vector space (TVS) is a closed convex hull of its extreme points.*

Recall that a probability measure is extreme in $\mathcal{P}(X)$ if and only if it is ergodic [46], so we are now prepared to formalise the notion that "every invariant probability measure is a convex combination of ergodic measures" in the setting of homogeneous spaces:

Theorem A.3 (Choquet's Theorem). *Let $(X, \mathcal{B}, f)$ be a continuous dynamical system on a compact metric space, and let $\mathcal{E}_f(X)$ denote the extreme points of $\mathcal{P}_f(X)$. Then for each $\mu \in \mathcal{P}_f(X)$, there exists a unique Borel measure γ on $\mathcal{P}_f(X)$ such that $\gamma(\mathcal{E}_f(X)) = 1$ and satisfying for all $\phi \in C(X)$,*

$$\int_X \phi d\mu = \int_{\mathcal{E}_f(X)} \left(\int_X \phi(x) dv(x) \right) d\gamma(v).$$

Each measure $\nu \in \mathcal{E}_f(X)$ is an ergodic invariant probability measure for f.[47]

Theorem A.4 (Ergodic Decomposition Theorem). *Let G be a Lie group acting smoothly on $X = G/\Gamma$ where Γ is a lattice on G. Let μ be a G-invariant measure on X. Then there is a measure space (Y, ν) and a partition of X into G-invariant subsets X_y, $y \in Y$, and measures μ_y on X_y such that:*

1. *for any measurable subset $A \subseteq X$, we have that $A \cap X_y$ is measurable with respect to μ_y for ν-almost every $y \in Y$ and $\mu(A) = \int_Y \mu_y(A \cap X_y)\, d\nu(y)$ (this **barycenter** can be thought of as the convex combination), and*
2. *for ν-almost every $y \in Y$, the action of G on X is ergodic with respect to the measure μ_y*

Finally, we want to ask topological questions about probability measures i.e. limit measures and closed sets so we endow $\mathcal{P}(X)$ with the 'weakest' possible topology that

Definition A.5 (Weak* Topology [48]). In the context of functional analysis, the weak* topology is the weakest topology on the dual space of a normed vector space in which all evaluation maps are continuous.

Appendix B Algebraic Groups and the Zariski Topology

Throughout this paper, we focused on the special linear group $SL(n, \mathbb{R})$. We can construct this geometrically as the subgroup of $GL(n, \mathbb{R})$ that preserves the oriented area when acting on $\mathbb{R}^n$ (rotations, reflections, shears, etc.). This gives us the following algebraic construction: embed $GL(n, \mathbb{R})$ into $\mathbb{R}^{n^2}$ so that $SL(n, \mathbb{R})$ consists of the zeros of the polynomial $p(A) = \det A - 1$ for $A \in GL(n, \mathbb{R})$. It is not difficult to see that p is an *irreducible* polynomial [42], so in the language of *algebraic geometry*, $SL(n, \mathbb{R})$ can be thought of as an **affine variety** [49].

We make this precise by recalling some basics of algebraic geometry. Given a field k, we consider the **affine space** $\mathbb{A}^n$ which consists of points as opposed to k^n which consists of vectors. The set of points in $\mathbb{A}^n$ that arise as the zeros of some collection of polynomials in n variables over k is called an **algebraic set**.

Taking inspiration from Appendix A where we study the space of probability measures by endowing it with the weak* topology, we proceed by endowing the space of varieties with the **Zariski topology**. In particular, the Zariski closed sets are precisely the algebraic subsets of $\mathbb{A}^n$. Their complements are referred to as the **principal open sets**. Note that this topology depends on the field k, so we may refer to k-closed sets to make this explicit.

Example B.1.

(i) If $X = \mathbb{A}^1$, then the proper closed subsets are the finite subsets of X since polynomials only have finitely many zeros. This says that $\mathbb{Z}$ is dense in $\mathbb{R}$.

(ii) If $X = \mathbb{A}^2$, then the proper closed subsets are finite unions of points and curves in X.

(iii) $GL(n, \mathbb{R})$ is a principal open set in $\mathbb{R}^{n^2}$ since it is defined by the non-vanishing of the determinant (which is a polynomial over $\mathbb{Q}$).

Remarkably, the Zariski topology is not Hausdorff: For any two nonempty open sets in a variety V, their intersection is nonempty. Thus, if p and q are distinct points of V, there are never disjoint neighborhoods containing them.

Definition B.2. [19] An **algebraic (k-) group** is a Zariski (k-) closed subgroup of $GL(n, K)$ where $k \subseteq K$ are fields.

Most Lie groups we typically work with are algebraic. The orthonormality conditions on the columns of $A \in O(n)$ give a collection of defining polynomials in the entries of A, so $O(n)$ is algebraic. Likewise, $SO(Q)$ is algebraic for any quadratic form Q. We can directly compute the following result:

Proposition B.3. *The special orthogonal group $SO(Q)$ of a nondegenerate quadratic form Q is defined over $\mathbb{Q}$ if and only if Q is a scalar multiple of a form with integer coefficients.*

As described the start of this section, $SL(n, \mathbb{R})$ is also an algebraic $\mathbb{Q}$-group which is strengthened by the following:

Proposition B.4. *The Zariski closure of every subset of $SL(n, \mathbb{Q})$ is defined over $\mathbb{Q}$.*

Finally, we conclude with a rigidity result on Zariski closures of lattices:

Theorem B.5 (Borel Density Theorem [19])**.** *Let*

- *G be a connected semisimple algebraic $\mathbb{R}$-group.*
- *$H = (G \cap GL(n, \mathbb{R}))^\circ$ and assume H has no compact factors.*
- *Γ be a lattice in H.*

Then Γ is Zariski dense in G.

Appendix C Miscellaneous

Due to the breadth of the material presented, some statements were not defined in full so as to not break the flow of the paper. We have collected these miscellaneous definitions here.

Definition C.1. [14] Let (X, Σ, μ) be a measure space. The **support** of μ on X is defined as

$$\operatorname{supp}(\mu) := \overline{\{A \in \Sigma : \mu(A) > 0\}}$$

Definition C.2 (One-Point Compactification [50])**.** For a locally compact Hausdorff space X, the one-point compactification X^* is compact. Furthermore, in such spaces, **compactness and sequential compactness are equivalent**.

In essence, the one-point compactification X^* makes X compact by "wrapping it up" with the point ∞ at infinity, thereby providing a compact topology for X.

Example C.3. *Consider the real line $\mathbb{R}$. Its one-point compactification is homeomorphic to a circle S^1. By adding the point ∞, we can transform $\mathbb{R}$ into a topological space that behaves like a circle, where ∞ serves as the 'point at infinity' that connects the two ends of the real line.*

Monte Carlo Simulation

To compute the volume in the denominator in Theorem 7.4.1, we generate 100 million random vectors (uniformly) with real entries in the hypercube $[-N, N]^4$ which has volume $(2N)^4$. The set $S = \{v \in \mathbb{R}^4 : 7\pi - 1 < Q(v) < 7\pi + 1,\ \|v\| \leq N\}$ is contained in this hypercube, and the probability of selecting a point in S is $\operatorname{vol} S/(2N)^4$. Thus, we get $\operatorname{vol} S \approx C(2N)^4$ where C is the number of vectors that lies in S. We provide example code for the case $Q(v) = v_0^2 + v_1^2 + \sqrt{2}v_2^2 - v_3^2$, $(p, q) = (3, 1)$, $N = 30$ and $(a, b) = (7\pi - 1, 7\pi + 1)$.

```python
import numpy as np

# Define the quadratic form
def Q(v):
    return v[0]**2 + v[1]**2 + np.sqrt(2) * v[2]**2 - v[3]**2

# Parameters
N = 30  # Bound for norm
a, b = 7*np.pi - 1, 7*np.pi + 1  # Interval
total_vectors = 10**8  # 100 million vectors for approximation

# Generate random uniform vectors
vectors = np.random.uniform(-N, N, (total_vectors, 4))

# Evaluate the quadratic form and count the number of vectors in the
    interval
Q_values = Q(vectors.T)
count_in_interval = np.sum((Q_values > a) & (Q_values < b))

# Calculate the proportion and approximate volume
proportion_in_interval = count_in_interval / total_vectors
volume_of_hypercube = (2 * N)**4
approximated_volume = proportion_in_interval * volume_of_hypercube

approximated_volume
```

Listing 1: Approximating volume by Monte Carlo simulation. Code written with ChatGPT.

8 References

[1] G. A. Margulis. Formes quadratriques indéfinies et flots unipotents sur les espaces homogènes. *C. R. Acad. Sci. Paris Sér. I Math.*, 302(15):489–493, 1986.

[2] M. Ratner. On raghunathan's measure conjecture. *Annals of Mathematics*, 134(3):545–607, 1991.

[3] M. S. Raghunathan. *Discrete subgroups of Lie groups*. Springer-Verlag, 1972.

[4] A. Wright. A tour through mirzakhani's work on moduli spaces of riemann surfaces, 2020.

[5] A. Arvanitoyeorgos. *An Introduction to Lie Groups and the Geometry of Homogeneous Spaces*. American Mathematical Society, 2003.

[6] A. L. Onishchik and E. B. Vinberg, editors. *Lie groups and lie algebras III*. Encyclopaedia of Mathematical Sciences. Springer, Berlin, Germany, 1994 edition, June 1994.

[7] W. Fulton and J. Harris. *Representation theory*. Graduate Texts in Mathematics. Springer, New York, NY, 1 edition, Jan. 1991.

[8] S. Lang. *SL2(R)*. Graduate Texts in Mathematics. Springer, New York, NY, Oct. 2011.

[9] B. C. Hall. *Lie groups, lie algebras, and representations*. Graduate texts in mathematics. Springer International Publishing, Cham, Switzerland, 2 edition, May 2015.

[10] N. Jacobson. *Basic Algebra I*. Dover Books on Mathematics. Dover Publications, Mineola, NY, 2 edition, June 2009.

[11] B. O'Neill. *Semi-Riemannian Geometry with Applications to Relativity*. Academic Press, 1983.

[12] J. M. Lee. *Introduction to Smooth Manifolds*. Springer, 2013.

[13] W. P. Thurston. *Three-Dimensional Geometry and Topology*. Princeton University Press, 1997.

[14] D. W. Morris. *Ratner's Theorems on Unipotent Flows*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, Sept. 2005.

[15] G. B. Folland. *Real Analysis: Modern Techniques and Their Applications*. John Wiley Sons, 1999.

[16] D. Bump. *Lie Groups*. Graduate Texts in Mathematics. Springer, New York, NY, Aug. 2016.

[17] B. Bekka, P. de la Harpe, and A. Valette. *New mathematical monographs: Kazhdan's property (T) series number 11*. Cambridge University Press, Cambridge, England, Apr. 2008.

[18] N. Bourbaki. *General Topology*. Springer, Berlin, Germany, 1 edition, Aug. 1998.

[19] R. J. Zimmer. *Ergodic Theory and Semisimple Groups*. Monographs in Mathematics. Birkhauser Boston, Secaucus, NJ, Jan. 1984.

[20] K. Conrad. Sl 2 (z). 2009.

[21] J. M. Lee. *Introduction to Smooth Manifolds*. Springer New York, New York, NY, 2nd edition, 2012. Web.

[22] M. Pollicott. Lisbon lectures on hyperbolic flows, 2021. Lecture notes.

[23] C. E. de Holanda. Homogeneous dynamics: An introduction to ratner's theorems on unipotent flows.

[24] M. P. do Carmo. *Riemannian Geometry*. Birkhäuser, 1992. Translated by Francis Flaherty.

[25] G. Gallavotti. *Statistical Mechanics: A Short Treatise*. Springer, 1999.

[26] E. Hopf. Ergodic theory and the geodesic flow on surfaces of constant negative curvature. *Bulletin of the American Mathematical Society*, 77(6):863 – 877, 1971.

[27] U. Rozikov. *An Introduction to Mathematical Billiards*. 06 2019.

[28] C. C. Moore. Ergodicity of flows on homogeneous spaces. *American Journal of Mathematics*, 88(1):154–178, 1966.

[29] A. Bailleul. Explicit kronecker–weyl theorems and applications to prime number races. *Research in Number Theory*, 8(3):43, Jul 2022.

[30] A. Gorodnik. Uniform distribution of orbits of lattices on spaces of frames. *Duke Mathematical Journal*, 122, 11 2003.

[31] A. Brown, D. Fisher, and S. Hurtado. Zimmer's conjecture for actions of $SL(m, \mathbb{Z})$. *Inventiones mathematicae*, 221(3):1001–1060, Sep 2020.

[32] P. Arnoux and T. Schmidt. Cross sections for geodesic flows and α-continued fractions. *Nonlinearity*, 26, 01 2013.

[33] M. Ratner. Strict measure rigidity for unipotent subgroups of solvable groups. *Inventiones mathematicae*, 101(1):449–482, Dec 1990.

[34] M. Ratner. On measure rigidity of unipotent subgroups of semisimple groups. *Acta Mathematica*, 165(none):229 – 309, 1990.

[35] M. Ratner. Raghunathan's conjectures for $SL(2, R)$. *Israel Journal of Mathematics*, 80(1): 1–31, Jun 1992.

[36] M. Einsiedler. Ratner's theorem on $SL(2, R)$-invariant measures, 2006.

[37] V. I. Bogachev and O. G. Smolyanov. *Topological vector spaces and their applications*. Springer Monographs in Mathematics. Springer International Publishing, Basel, Switzerland, 1 edition, May 2017.

[38] D. Witte. Rigidity of some translations on homogeneous spaces. *Inventiones mathematicae*, 81(1):1–27, Feb 1985.

[39] J. Cassels. *An Introduction to Diophantine Approximation*. Cambridge University Press, 1957.

[40] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, 5th edition, 1979.

[41] E. Lindenstrauss. Invariant measures and arithmetic quantum unique ergodicity. *Annals of Mathematics*, 163(1):165–219, 2006.

[42] M. Liebeck. Linear algebra, math 50003: Lecture notes, 2022. Lecture notes accessed in June 2023.

[43] H. Davenport and H. Heilbronn. Indefinite quadratic forms in many variables. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 239:215–240, 1946.

[44] A. Eskin, G. Margulis, and S. Mozes. Quadratic forms of signature (2,2) and eigenvalue spacings on rectangular 2-tori. *Annals of Mathematics*, 161(2):679–725, 2005.

[45] G. M. Ziegler. *Lectures on Polytopes*. Graduate Texts in Mathematics. Springer, New York, NY, 1995 edition, Nov. 1994.

[46] R. R. Phelps. *Lectures on Choquet's Theorem*. Lecture Notes in Mathematics. Springer, New York, NY, 2 edition, Oct. 2006.

[47] J. Hawkins. *Ergodic Dynamics: From Basic Theory to Applications*, volume 289. Springer International Publishing, Cham, 2021. Web.

[48] C. Walkden. Ergodic theory, lecture 10. Accessed: 2024-06-16.

[49] W. Fulton. *Algebraic curves*. Benjamin-Cummings Publishing Co., Subs. of Addison Wesley Longman, Reading, PA, Dec. 1974.

[50] S. Willard. *General Topology*. Courier Corporation, 2004.