An introduction to flows on homogeneous spaces*

S.G. Dani

December 9, 2013

This article is intended to give a first introduction to the topic of flows on homogeneous spaces of Lie groups, which has developed considerably in the last four decades, and has had interesting applications in various areas including especially Diophantine approximation. We shall aim at conveying the flavour of the results with minimal framework, and not strive to present the most general results. Proofs are included in the simpler cases wherever possible. References are included for the interested reader to pursue the topic further.

1 Homogeneous spaces

We begin with a quick introduction to the spaces involved. While we shall actually be concerned only with Lie groups, it would be convenient to formulate the notions and questions in the more general set up of locally compact groups - these will however be routinely assumed to be second countable.

Let G be a locally compact second countable group. A reader not familiar with the general theory may bear in mind the following examples: $\mathbb{R}, \mathbb{R}^n, n \geq 2$, the general linear group $GL(n, \mathbb{R})$ consisting of all nonsingular $n \times n$ matrices with real entries, closed subgroups of known locally compact groups, quotients of locally compact groups by closed normal subgroups, direct products of locally compact groups, covering groups, etc.

By a homogeneous space X we mean a quotient space $X = G/H$, where H is a closed subgroup of G, consisting of the cosets $gH, g \in G$; the space is considered equipped with the quotient topology, defined by the condition that the map

*This article represents an edited version of the notes of a short course of lectures given at the Workshop in the ICTS program on Groups, Geometry and Dynamics held at Almora, in December 2012.

$g \mapsto gH$ is an open map; then X is a locally compact space. Equivalently, a homogeneous space of G is a topological space with a continuous action of G on it that is transitive, namely such that for any $x, y \in X$ there exists $g \in G$ such that $y = gx$; given a transitive action on X and $x_0 \in X$ the map $g \mapsto gx_0$ is an open map (G being assumed to be second countable), which then implies that X may be realised as the quotient space G/H, where $H = \{g \in G \mid gx_0 = x_0\}$ is the stabiliser of x_0.

Examples: 1. Consider the natural action of $GL(n, \mathbb{R})$ on $\mathbb{R}^n$, $n \geq 2$. Then $\mathbb{R}^n \backslash (0)$ is a single orbit and we get that

$$\mathbb{R}^n \backslash (0) \approx GL(n, \mathbb{R}) / \{g \in GL(n, \mathbb{R}) \mid ge_1 = e_1\};$$

(here and in the sequel we denote by $\{e_i\}$ the standard basis of $\mathbb{R}^n$). The subgroup $SL(n, \mathbb{R})$ of $GL(n, \mathbb{R})$, consisting of matrices with determinant 1, also acts transitively on $\mathbb{R}^n \backslash (0)$ and hence the latter may be viewed as $SL(n, \mathbb{R})/H$, where $H = \{g \in SL(n, \mathbb{R}) \mid ge_1 = e_1\}$.

2. Similarly $\mathbb{P}^{n-1}$ and the Grassmannian manifolds may be realised as homogeneous spaces of $GL(n, \mathbb{R})$.

3. By a lattice in $\mathbb{R}^n$ we mean a subgroup generated by n linearly independent vectors; such a subgroup Λ is discrete and the quotient $\mathbb{R}^n/\Lambda$ is compact. Let Ω_n be the space of all lattices in $\mathbb{R}^n$. Then Ω_n can be realised as $GL(n, \mathbb{R})/GL(n, \mathbb{Z})$, where $GL(n, \mathbb{Z})$ is the subgroup of $GL(n, \mathbb{R})$ consisting of matrices with integral entries; it is the stabilizer of the lattice $\mathbb{Z}^n$ generated by the standard basis $\{e_i\}$. (The topology on Ω_n may be defined to be the one making the bijection a homeomorphism; it can also be defined in terms of admitting bases which are close to each other, which turns out to be equivalent to the former).

2 Measures on homogeneous spaces

We consider a homogeneous space G/H equipped with the G-action on the left: $(g, xH) \mapsto (gx)H$ for all $g, x \in G$. Consider the question whether G/H admits a measure invariant under the G-action. We note that $\mathbb{R}^n/\mathbb{Z}^n$ admits an $\mathbb{R}^n$-invariant (finite) measure, and $\mathbb{R}^n \backslash (0)$, which may be viewed as $SL(n, \mathbb{R})/H$ as seen above, admits an $SL(n, \mathbb{R})$-invariant (infinite, locally finite) measure; viz. the restriction of the Lebesgue measure to $\mathbb{R}^n \backslash (0)$. On the other hand $\mathbb{P}^{n-1}$ does not admit a measure invariant under the usual action of $SL(n, \mathbb{R})$; see [6] for measures on $\mathbb{P}^{n-1}$ invariant under subgroups of $SL(n, \mathbb{R})$.

The following theorem describes a necessary and sufficient condition for G/H to admit a measure invariant under the G-action; by a measure we shall always mean a Radon measure, namely a measure defined on all Borel subsets which assigns finite measure to every compact set. For any closed subgroup H of G, including G itself, we denote by Δ_H the modular homomorphism of H. Then we have the following (see for instance [10], Theorem 2.49):

Theorem 2.1. *G/H admits a G-invariant measure if and only if $\Delta_G(h) = \Delta_H(h)$ for all $h \in H$; when it exists the invariant measure is unique upto scaling.*

Conformity of the examples as above with the theorem may be checked easily - the student readers are encouraged to carry this out. We note also that as a consequence of the theorem, if G is a unimodular locally compact group and Γ is a discrete subgroup of G then G/Γ admits a G-invariant measure; if the quotient G/Γ is compact then automatically the measure is finite (recall our convention that all measures are finite on compact sets), but in general the invariant measure may not be finite. These observations apply in particular to $G = SL(n, \mathbb{R})$. Homogeneous spaces with finite invariant measure are of special interest.

2.1 Lattices

Definition 2.2. Let G be a locally compact group. A closed subgroup Γ of G is called a *lattice* in G if Γ is discrete and G/Γ admits a finite G-invariant measure.

Examples: 1. A subgroup Λ of $\mathbb{R}^n$ is a lattice in $\mathbb{R}^n$ if and only if it is generated (as a subgroup) by n linearly independent vectors; namely a lattice in $\mathbb{R}^n$ in the sense recalled above.

2. $G = SL(n, \mathbb{R})$ and Γ a discrete subgroup such that G/Γ is compact (see above).

3. In $G = SL(n, \mathbb{R})$, $\Gamma = SL(n, \mathbb{Z})$ is a lattice in G; see [16].

A lattice for which the corresponding quotient is compact is said to be *uniform*; otherwise it is said to be *nonuniform*.

Proposition 2.3. *$SL(n, \mathbb{Z})$ is a nonuniform lattice in $SL(n, \mathbb{R})$.*

Proof. Let $G = SL(n, \mathbb{R})$ and $\Gamma = SL(n, \mathbb{Z})$. We have noted that Γ is a lattice in G. Suppose G/Γ is compact. Then there exists a compact subset K of G such that $G = K\Gamma$; (the latter stands for $\{x\gamma \mid x \in K, \gamma \in \Gamma\}$). This implies that $G(e_1) = K\Gamma(e_1) \subset K(\mathbb{Z}^n)$. However, this is not possible, since the norms of all nonzero vectors contained in $K(\mathbb{Z}^n)$ are bounded below by a positive constant that depends only on K, while on the other hand it is easy to see that given any $v \in \mathbb{R}^n \backslash (0)$, which we may choose to be of arbitrarily small norm, we can find $g \in G$ such that $v = ge_1 \in g\mathbb{Z}^n$. □

2.2 Flows

Let G be a locally compact group and Γ a lattice in G. For a closed subgroup H of G the H-action on G/Γ is called the flow induced by H on G/Γ. Typically we shall be interested in actions of cyclic subgroups (equivalently of elements of G), or one-parameter flows, namely actions induced by (continuous) one-parameter subgroups $\{g_t\}_{t\in\mathbb{R}}$ where $g_t \in G$ for all $t \in \mathbb{R}$.

Examples: 1. $G = \mathbb{R}^n$, $\Gamma = \mathbb{Z}^n$, $H = \{tv \mid t \in \mathbb{R}\}$, where $v \in \mathbb{R}^n$.
2. $G = SL(2, \mathbb{R})$, Γ a lattice in G, and $H = \left\{ \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \mid t \in \mathbb{R} \right\}$. This corresponds to what is called the geodesic flow associated with the surface $\mathbb{H}^2/\Gamma$, where $\mathbb{H}^2$ is the Poincaré upper half-plane (see § 9). In particular when $\Gamma = SL(2, \mathbb{Z})$ it corresponds to the geodesic flow associated with the modular surface.

3 Ergodic properties

Let $(X, \mathfrak{M}, \mu)$ be a finite measure space, namely $\mathfrak{M}$ is a σ-algebra of subsets of X and μ is a finite measure defined over $\mathfrak{M}$ - we shall further assume $\mu(X) = 1$.

A transformation $T : X \to X$ is said to be measurable if $T^{-1}(E) \in \mathfrak{M}$ for all $E \in \mathfrak{M}$, and it is said to be measure preserving if $\mu(T^{-1}(E)) = \mu(E)$ for all $E \in \mathfrak{M}$.

Definition 3.1. T is said to be *ergodic* if for $E \in \mathfrak{M}$, $\mu(T^{-1}(E)\Delta\, E) = 0$ holds only when $\mu(E) = 0$ or $\mu(X\backslash E) = 0$; here Δ stands for the symmetric difference of the two sets, namely $A\Delta B = (A\backslash B) \cup (B\backslash A)$.

Exercise: T is ergodic if and only if, for $E \in \mathfrak{M}$, $T^{-1}(E) = E$ implies $\mu(E) = 0$ or $\mu(X\backslash E) = 0$.

The action of a locally compact group G on $(X, \mathfrak{M}, \mu)$ is said to be *measurable* if for all $E \in \mathfrak{M}$ the set $\{(g, x) \mid gx \in E\}$ is a measurable subset of $G \times X$ (with respect to the measurable structure as the Cartesian product, G being equipped with the Borel σ-algebra), and it is said to be measure-preserving if $\mu(gE) = \mu(E)$

for all $g \in G$ and $E \in \mathfrak{M}$. In the sequel all actions shall be understood to be measurable.

The statement analogous to the above mentioned Exercise holds for actions in general (but is a little more technical to prove).

Definition 3.2. A measure preserving transformation $T : X \to X$ is said to be *mixing* if for any two measurable subsets A and B, $\mu(T^{-k}(A) \cap B) \to \mu(A)\mu(B)$, as $k \to \infty$. The action of a noncompact locally compact group, on (X, μ) as above, is said to be *mixing* if for any divergent sequence $\{g_k\}$ in G and any two measurable subsets A and B, $\mu(g_k(A) \cap B) \to \mu(A)\mu(B)$, as $k \to \infty$. (A sequence $\{g_k\}$ in G is said to be *divergent* if for any compact subsets C of G there exists k_0 such that $g_k \notin K$ for $k \geq k_0$.)

Remarks: 1. For the group $\mathbb{Z}$ (with discrete topology) the two definitions of mixing, as above, coincide.

2. If $T : X \to X$ is mixing then it is ergodic, since if $\mu(T^{-1}(E) \Delta E) = 0$ then $\mu(T^{-k}(E) \cap E) = \mu(E)$ for all k, and by the mixing condition we get $\mu(E)^2 = \mu(E)$, so $\mu(E) = 0$ or 1. Similarly the mixing condition implies ergodicity in the case of general groups actions as well.

3. In general ergodicity does not imply mixing; see Remark 4.3.

4. If a G-action on $(X, \mathfrak{M}, \mu)$ is mixing and H is a closed noncompact subgroup of G then the H-action on $(X, \mathfrak{M}, \mu)$, defined by restriction, is also mixing. The corresponding statement however does not hold for ergodicity.

There are a variety of weaker and stronger forms of mixing, for cyclic as well as general group actions, that we shall not go into here.

3.1 Topological implications

The ergodic properties have implications to the topological behaviour of orbits. The first assertion below goes back to the work of Hedlund from 1930's.

Proposition 3.3. *Consider a measure-preserving action of a locally compact group G on a measure space $(X, \mathfrak{M}, \mu)$, with $\mu(X) = 1$. Suppose that $\mu(\Omega) > 0$ for all nonempty open subsets Ω. Then we have the following:*

i) if the action is ergodic then almost all G-orbits are dense in X; that is,

$$\mu(\{x \in X \mid Gx \text{ not dense in } X\}) = 0.$$

ii) if the action is mixing then for every divergent sequence $\{g_k\}$

$$\mu(\{x \in X \mid \{g_k x\} \text{ not dense in } X\}) = 0.$$

Proof. Let $\{\Omega_j\}$ be a countable basis for the topology on X. Then each $G\Omega_j$ is an open G-invariant subset and hence by the ergodicity $\mu(G\Omega_j) = 1$, and in turn $\mu(\cap_j G\Omega_j) = 1$. The assertion now follows, since for all $x \in \cap_j G\Omega_j$ the G-orbit of x intersects each Ω_j and hence is dense in X. This proves (i). The proof of (ii) is analogous. □

3.2 Translation flows on tori

We realise $\mathbb{R}^n$ as the space of n-rowed column vectors, written as $v = (\xi_1, \dots, \xi_n)^t$, with $\xi_1, \dots, \xi_n \in \mathbb{R}$ and the t over the row stands for transpose.

For flows on tori, namely $\mathbb{R}^n/\mathbb{Z}^n$, equipped with the Borel σ-algebra and the Haar measure μ we have the following:

Proposition 3.4. *Let $v = (\alpha_1, \dots, \alpha_n)^t$. The translation of $\mathbb{R}^n/\mathbb{Z}^n$ by v is ergodic if and only if $1, \alpha_1, \dots, \alpha_n$ are linearly independent over $\mathbb{Q}$ (that is, no nontrivial linear combination $\sum_1^n q_i\alpha_i$, with $q_i \in \mathbb{Q}$, is rational).*

Proof. : Let T denote the translation of $\mathbb{R}^n/\mathbb{Z}^n$ by v and let E be a measurable subset such that $\mu(T^{-1}(E)\Delta\, E) = 0$. Let f denote the characteristic function of E. Consider the Fourier expansion of f in $L^2(\mathbb{R}^n/\mathbb{Z}^n)$, say $f = \sum_\chi a_\chi \chi$, where the summation is over the all characters χ on $\mathbb{R}^n/\mathbb{Z}^n$. The invariance of E under T implies that $f \circ T$ and f are equal a.e.. We see that $f \circ T$ has a Fourier expansion as $\sum_\chi a_\chi \chi(v)\chi$, and hence by the uniqueness of the Fourier expansion it follows that $a_\chi \chi(v) = a_\chi$ for all characters χ. The condition as in the hypothesis then implies that $a_\chi = 0$ for all nontrivial characters χ. Hence f is constant a.e., or equivalently $\mu(E) = 0$ or 1. This proves the proposition. □

3.3 More about the translation flows

We note that every orbit of the translation action is a coset of the subgroup of $\mathbb{R}^n/\mathbb{Z}^n$ generated by $v + \mathbb{Z}^n$. When the action is ergodic then there exists a dense coset, and hence the subgroup is dense as well. (On the other hand the latter statement may be proved directly under the condition as in the hypothesis of

Proposition 3.4, and used to deduce ergodicity.) Conversely, when the subgroup generated by $v + \mathbb{Z}^n$ is dense in $\mathbb{R}^n/\mathbb{Z}^n$, *all* orbits are dense in $\mathbb{R}^n/\mathbb{Z}^n$ - note that ergodicity assures only *almost all* orbits to be dense, so what we see here is a rather special situation.

When all orbits are dense the action is said to be *minimal.* Thus for translations of tori ergodicity is equivalent to minimality, and either of them holds if and only if the subgroup generated by v and $\mathbb{Z}^n$ is dense in $\mathbb{R}^n$, or equivalently also if and only if there exists a dense orbit.

By similar arguments it can be seen that the flow induced by $\{tv \mid t \in \mathbb{R}\}$ is ergodic if and only if $\alpha_1, \ldots, \alpha_n$ are linearly independent over $\mathbb{Q}$; equivalently, the flow is ergodic if and only if there exists $t \in \mathbb{R}$ such that the translation action of tv is ergodic.

For one-parameter translation flows also ergodicity is equivalent to minimality, and also to existence of a dense orbit.

4 Unitary representations

Let G be a locally compact (second countable) group. Let $\mathfrak{H}$ be a (separable) Hilbert space and let $\mathfrak{U}(\mathfrak{H})$ be the group of all unitary operators on $\mathfrak{H}$.

A unitary representation π of G over $\mathfrak{H}$ is a homomorphism of G into $\mathfrak{U}(\mathfrak{H})$ which is continuous with respect to the strong operator topology; that is, $g \mapsto \pi(g)\xi$ is continuous for all $\xi \in \mathfrak{H}$.

Let (X, μ) be a measure space with $\mu(X) = 1$ equipped with a G-action and let $\mathfrak{H} = L^2(X, \mu)$. The action induces a unitary representation of G over $\mathfrak{H}$, by

$$\pi(g)f(x) = f(g^{-1}x) \text{ for all } g \in G, f \in \mathfrak{H} \text{ and } x \in X;$$

(following standard abuse of notation we view elements of $\mathfrak{H}$ as pointwise defined functions - while this does involves some technical issues, in the final analysis there is no ambiguity).

The ergodicity and mixing conditions can be translated to the following in terms of the associated unitary representation.

Proposition 4.1. *i) The G-action on X is ergodic if and only if there is no non-constant function in $\mathfrak{H}$ fixed under the action of $\pi(g)$ for all $g \in G$.*

ii) The G-action on X is mixing if and only if for any divergent sequence $\{g_k\}$ in G and all $\phi, \psi \in \mathfrak{H}$,

$$\langle \pi(g_k)\phi, \psi\rangle \to \langle \phi, 1\rangle\langle 1, \psi\rangle \text{ as } k \to \infty.$$

It is convenient to consider the restriction of π to the ortho-complement of constants, $\mathfrak{H}_0 = \{\phi \in \mathfrak{H} \mid \phi \perp 1\}$ (which is an invariant subspace).

Proposition 4.2. *: The G-action is mixing if and only if $\langle \pi(g)\phi, \psi\rangle \to 0$ as $g \to \infty$ for all $\phi, \psi \in \mathfrak{H}_0$.*

Remark 4.3. It is easy to see from the Proposition that the translations of tori, as in the last section, are not mixing. The ergodic translations of tori thus provide examples of ergodic transformations that are not mixing.

4.1 Mautner phenomenon

Let e denote the identity element in G and for $g \in G$ let

$$H_g^+ = \{x \in G \mid g^k x g^{-k} \to e \text{ as } k \to \infty\}$$

and

$$H_g^- = \{x \in G \mid g^k x g^{-k} \to e \text{ as } k \to -\infty\}.$$

These are called the contracting and expanding *horospherical subgroups* corresponding to g.

The following simple observation, known as Mautner phenomenon is very useful in proving ergodicity and mixing properties.

Theorem 4.4. *Let π be a unitary representation of G over a Hilbert space $\mathfrak{H}$. Let $g \in G$ and $\phi \in \mathfrak{H}$ be such that $\pi(g)\phi = \phi$. Then $\pi(x)\phi = \phi$ for all x in the subgroup generated by $H_g^+ \cup H_g^-$.*

Proof. The vector ϕ may be assumed to be nonzero, and by scaling it we may assume $\|\phi\| = 1$. Now let $x \in H_g^+$ be arbitrary. We have

$$\langle \pi(x)\phi, \phi\rangle = \langle \pi(g)\pi(x)\phi, \pi(g)\phi\rangle = \langle \pi(g)\pi(x)\pi(g^{-1})\phi, \phi\rangle = \langle \pi(gxg^{-1})\phi, \phi\rangle,$$

under the given condition $\pi(g)\phi = \phi$. Hence $\langle \pi(x)\phi, \phi\rangle = \langle \pi(gxg^{-1})\phi, \phi\rangle$, and by successive application of the same we get that $\langle \pi(g^k x g^{-k})\phi, \phi\rangle = \langle \pi(x)\phi, \phi\rangle$ for all k. As the sequence on the left hand side converges, as $k \to \infty$, to $\langle \phi, \phi\rangle = 1$ it follows that $\langle \pi(x)\phi, \phi\rangle = 1$. But then $\|\pi(x)\phi - \phi\|^2 = 2 - 2\,\mathrm{Re}\langle \pi(x)\phi, \phi\rangle = 0$, which shows that $\pi(x)\phi = \phi$ for all $x \in H_g^+$.

Similarly, using that $\pi(g^{-1})\phi = \phi$ we conclude that $\pi(x)\phi = \phi$, for all $x \in H_g^-$, and hence for the subgroup generated by $H_g^+ \cup H_g^-$.

5 Flows on $SL(2,\mathbb{R})/\Gamma$

In this section we discuss the ergodic and dynamical properties of the flows on homogeneous spaces of $SL(2,\mathbb{R})/\Gamma$, where Γ is a lattice in $SL(2,\mathbb{R})$. Throughout the section let $G = SL(2,\mathbb{R})$ and Γ be a lattice in G.

Elementary linear algebra shows that every element of G is either conjugate to a diagonal or an upper triangular matrix with ± 1 on the diagonal, or is contained in a compact subgroup of G; in the latter case it acts as a "rotation" of the plane, with respect to a suitable choice of the basis. It is easy to see that the action of a compact subgroup on G/Γ can not be ergodic; by Proposition 3.3 such an action would have to be transitive, which is not possible. Thus it suffices to discuss ergodicity of actions of the diagonal and upper triangular unipotent elements (as far as actions of cyclic subgroups are concerned - for an upper triangular matrix with -1's on the diagonal the square is unipotent, and the study may be reduced to the latter). The action of the one-parameter subgroup consisting of diagonal matrices with positive entries corresponds to the geodesic flow associated with the hyperbolic surface (orbifold, strictly speaking) with fundamental group Γ (see § 9 for some details). This provides a geometric context for discussing the ergodicity of the flow induced by this one-parameter subgroup.

5.1 Ergodicity of the geodesic flow

Let $g = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$, with $0 < |\lambda| < 1$. Then it can be seen that

$$H_g^+ = \left\{ \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \mid t \in \mathbb{R} \right\}$$

and

$$H_g^- = \left\{ \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \mid t \in \mathbb{R} \right\}.$$

We note also that the subgroups H_g^+ and H_g^- as above generate the whole of G. Hence by Theorem 4.4 we have

Corollary 5.1. *Let π be a unitary representation of G over a Hilbert space $\mathfrak{H}$. If $\phi \in \mathfrak{H}$ is fixed by g as above then it is fixed by the G-action.*

Together with Proposition 4.1, and the fact that the action of g is ergodic if and only if the action of g^{-1} is ergodic, this implies the following.

Corollary 5.2. *Let Γ be a lattice in $G = SL(2, \mathbb{R})$. Let g be a diagonal matrix in $SL(2, \mathbb{R})$ other than $\pm I$ (where I is the identity matrix). Then the action of g on G/Γ is ergodic.*

In particular the flow induced by the one-parameter subgroup consisting of positive diagonal matrices, namely the "geodesic flow" is ergodic.

5.2 Ergodicity of the horocycle flow

The action of the subgroup $\left\{ h_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \right\}$ (which played a role in the proof of ergodicity of the geodesic flow) on G/Γ is called the *horocycle flow*.

(The other - lower triangular - subgroup is conjugate to $\{h_t\}$, and need not be considered separately, with regard to dynamical properties; only in dealing with the geodesic flow we need both, in which case we talk of the contracting and expanding horocycle flows.)

We next show that the flow induced by $\{h_t\}$ is ergodic. The proof below can be tweaked to show that in fact the action of every nontrivial element from the subgroup, and hence the action of any nontrivial unipotent element of G, is ergodic; however for simplicity of exposition we shall restrict to considering the action of the full one-parameter subgroup.

Theorem 5.3. *The horocycle flow, viz. the flow induced by $\{h_t\}$ on G/Γ is ergodic.*

Proof. : Let $\mathfrak{H} = L^2(G/\Gamma)$, π the associated unitary representation over $\mathfrak{H}$, and $\phi \in \mathfrak{H}$ be fixed by $\pi(h_t)$ for all $t \in \mathbb{R}$. We have to show that ϕ is constant a.e.. We may assume $\|\phi\| = 1$.

Let F be the function on G defined by $F(g) = \langle \pi(g)\phi, \phi \rangle$, for all $g \in G$. It is a continuous function and $F(h_s g h_t) = F(g)$ for all $s, t \in \mathbb{R}$ and $g \in G$. We next define a function f on $\mathbb{R}^2 \backslash (0)$ by setting, for $v \in \mathbb{R}^n \backslash (0)$, $f(v) = F(g)$ where $g \in G$ is such that $v = ge_1$; such a g exists, and since $F(gh_t) = F(g)$ for all $t \in \mathbb{R}$ it follows that f is a well-defined function. Also, it is continuous and $f(h_s v) = f(v)$ for all $v \in \mathbb{R}^2 \backslash (0)$. For $v \in \mathbb{R}^2$ which are not on the x-axis the orbits $\{h_s v \mid s \in \mathbb{R}\}$ consist of horizontal lines; the points on the x-axis are fixed points of the flow. Thus the invariance property of f as above implies that it is constant along horizontal lines in $\mathbb{R}^2$, except possibly the x-axis. But then by continuity it must be constant along the x-axis as well; that is, $f(te_1) = f(e_1)$ for all $t \neq 0$.

We thus get that for $g = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$, $F(g) = f(ge_1) = f(\lambda e_1) = f(e_1) = F(e) = 1$. Arguing as before we see that this implies that ϕ is fixed by $\pi(g)$. Since the action of g is ergodic it follows that ϕ is a constant function a.e.. Hence the action of $\{h_t\}$ on G/Γ is ergodic. $\square$

6 Flows on $SL(n, \mathbb{R})/\Gamma$

Let me now mention the stronger results with regard to ergodicity and mixing, for flows on the homogeneous spaces $SL(n, \mathbb{R}))/\Gamma$, $n \geq 2$, where Γ is a lattice in $SL(n, \mathbb{R})$.

Theorem 6.1. *Let H be a closed noncompact subgroup of $SL(n, \mathbb{R})$. Then its action on $SL(n, \mathbb{R})/\Gamma$ is mixing. In particular it is ergodic.*

This can be deduced from the theorem of Howe and Moore, which describes necessary and sufficient conditions under which given a unitary representation π of a Lie group G over a Hilbert space $\mathfrak{H}$ the "matrix coefficients" $\langle \pi(g)\phi, \psi \rangle$ converge to 0, as $g \to \infty$, for all $\phi, \psi \in \mathfrak{H}$. For the case of $G = SL(n, \mathbb{R})$ such a convergence holds for all representations that do not admit any nonzero fixed point. (See [1] and [9] for more details).

The action on $SL(n, \mathbb{R})/\Gamma$ by a compact subgroup of $SL(n, \mathbb{R})$ can not be ergodic (in view of Proposition 3.3, as seen in the last section for $n = 2$). Thus the action of a closed subgroup H on $SL(n, \mathbb{R})/\Gamma$ is ergodic if and only if the subgroup is noncompact.

In particular, for $G = SL(2, \mathbb{R})$ given any $g \in G$ which is not contained in a compact subgroup of G and a sequence $\{n_k\}$ of integers such that $|n_k| \to \infty$, for almost all $x \in G/\Gamma$ the sequences $\{g^{n_k}x \mid k \in \mathbb{N}\}$ are dense in G/Γ; this holds in particular for $g = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$, $\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ or $\begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$, for any $t \neq 0$. The cases of orbits $\{g^j x \mid j \in \mathbb{Z}\}$ and forward trajectories $\{g^j x \mid j = 0, 1, \ldots\}$ to which this applies, are of particular interest.

6.1 Orbits of individual points

It is in general a nontrivial matter to determine for which particular x the orbit, or trajectory, is dense under a flow.

It turns out that the horocycle flow is well behaved in this respect, in the sense that the closures of the individual orbits are amenable to a neat description. The following is a classical theorem, proved by Hedlund in the 1930's; see [1] and [9] for details.

Theorem 6.2. *Let $G = SL(2,\mathbb{R})$ and Γ be a lattice in G. Let $\{h_t\}$ be a horocycle flow. Then we have the following:*

i) for any $x \in G/\Gamma$, either $h_s x = x$ for some $s \neq 0$ (viz. the orbit is periodic) or $\{h_t x \mid t \in \mathbb{R}\}$ is dense in G/Γ.

ii) if G/Γ is compact then for all $x \in G/\Gamma$, $\{h_t x \mid t \in \mathbb{R}\}$ is dense in G/Γ; thus, in this case the horocycle flow is minimal.

When $h_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$, for $t \in \mathbb{R}$, and x is a periodic point of $\{h_t\}$ then gx is also a periodic point of $\{h_t\}$ for any diagonal matrix g, as g normalizes $\{h_t\}$. In the sense of hyperbolic geometry, which we shall discuss briefly in § 9, when Γ is the fundamental group of a hyperbolic surface of finite area, to each of the (finitely many) cusps of the surface there corresponds such a cylinder of periodic orbits of the horocycle flow. In particular the set of periodic points in nonempty, and the flow is not minimal when the associated surface is noncompact.

The statement analogous to Theorem 6.2 is also true for actions of cyclic groups generated by unipotent elements, in place of the one-parameter subgroups; every orbit of h_s, $s \neq 0$, which is not contained in a periodic orbit of $\{h_t \mid t \in \mathbb{R}\}$ is dense in G/Γ (when G/Γ is compact only the latter possibility occurs). This is subsumed by the stronger results on uniform distribution that we consider next.

6.2 Uniformly distributed orbits

Apart from being dense the non-periodic orbits of the horocycle flow are also "uniformly distributed" in G/Γ: Let x be a point with a non-periodic orbit under $\{h_t\}$. Then we have the following: for any bounded continuous function f on G/Γ,

$$\frac{1}{T}\int_{t=0}^{T} f(h_t x)dt \to \int_{G/\Gamma} f d\mu \text{ as } T \to \infty,$$

where μ is the G-invariant probability measure on G/Γ, and correspondingly

$$\frac{1}{N}\sum_{k=0}^{N-1} f(h^k x) \to \int_{G/\Gamma} f d\mu \text{ as } N \to \infty,$$

for any $h = h_t$, $t \neq 0$. Thus the "time averages" (for continuous or discrete time) converge to the "space average", in the case of all the dense orbits as above - the latter is evidently a necessary condition for the former to hold. In place of the continuous functions, uniform distribution can also be expressed in terms of sets, but for obvious reasons this can be expected to hold only for "good" sets. It turns out that the condition of "Jordan measurability" suffices in this respect; we recall that a subset Ω is said to be Jordan measurable if $\mu(\partial\Omega) = 0$, where $\partial\Omega$ denotes the topological boundary of the set Ω. Thus if Ω is a Jordan measurable subset of G/Γ then, for x as above, as $k \to \infty$,

$$\frac{1}{k}\#\{0 \leq j \leq k-1 \mid h^j x \in \Omega\} \longrightarrow \mu(\Omega), \text{ as } k \to \infty,$$

where $h = h_t$ with $t \neq 0$ as before, and $\#$ stands for the cardinality of the set in question. This means that the trajectory visits each "good" set with frequency equal to its proportion in the space (with respect to μ).

In the case of $\Gamma = SL(2,\mathbb{Z})$, $\left\{ \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} g\Gamma \mid n \in \mathbb{Z}_+ \right\}$ is dense, and uniformly distributed for $g \in G$, if and only if $g \notin P\Gamma$, where P is the subgroup consisting of all upper triangular matrices.

6.3 Higher dimensional situations

A general result in this direction is the following theorem of Marina Ratner, which has been a path-breaking result in the area; for the sake of simplicity we limit the exposition to the special case of $G = SL(n,\mathbb{R})$, though the result is known in considerable generality.

Theorem 6.3. *Let $G = SL(n,\mathbb{R})$ and Γ be a lattice in G. Let $\{u_t\}$ be a unipotent one-parameter subgroup of G. Then for any $g \in G$ there exists a closed connected subgroup F of G such that the following holds:*

i) $Fg\Gamma$ is closed, $Fg\Gamma/\Gamma$ admits a F-invariant probability measure μ, and

ii) $\{u_t g\Gamma \mid t \geq 0\}$ is dense in $Fg\Gamma/\Gamma$ and uniformly distributed with respect to μ.

For a $g\Gamma$ for which the orbit is dense in G/Γ the above holds with G as the choice for F. The theorem therefore means that for $g\Gamma$ for which the orbit is not dense, it is contained in a closed orbit of a proper closed connected subgroup F - an orbit admitting a finite measure invariant under the F-action - and if one considers the smallest subgroup F for which this would hold, then the orbit-closure of $g\Gamma$ under $\{u_t\}$ is precisely the F-orbit of $g\Gamma$, and moreover, within the latter it is uniformly distributed. A subset $Fg\Gamma/\Gamma$ as above is called a "homogeneous subset", and the F-invariant measure is called a "homogeneous measure". For the horocycle flow, where $G = SL(2,\mathbb{R})$, and $u_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$, the only proper closed connected subgroups F for which there exists $g \in G$ such that $Fg\Gamma$ is closed turn out to be of the form $\{gu_tg^{-1} \mid t \in \mathbb{R}\}$ (same g is involved), and correspondingly the $\{u_t\}$-orbits are either periodic or uniformly distributed in G/Γ as noted earlier in this section.

A major part of the proof of Theorem 6.3 involves classifying the invariant measures of the $\{u_t\}$-action on G/Γ. It suffices to classify the ergodic invariant measures (namely those with respect to which the action is ergodic) and these are shown to be the homogeneous measures as described above; see [15] for an exposition of the ideas involved; the reader may also refer the survey article [4] and the references cited there, for further details.

7 Duality

So far we have been considering actions on homogeneous spaces G/Γ, where G is some locally compact group and Γ is a lattice in G, by subgroups H of G. The results can be readily applied to study the action of a lattice Γ in G, on the homogeneous space G/H, via what is sometimes referred to as "duality".

Proposition 7.1. *Let G be a locally compact group, Γ be a lattice in G and H be a closed subgroup of G. Then we have the following:*

(i) for $g \in G$, the Γ orbit of gH is dense in G/H if and only if the H-orbit of $g^{-1}\Gamma$ is dense in G/Γ.

(ii) the Γ-action on G/H is ergodic if and only if the H-action on G/Γ is ergodic.

[With regard to (ii) it may be clarified that though ergodicity was defined in § 3 only for actions with a finite invariant measure the same condition can be readily adopted in general, and it is indeed considered in the literature when either there is an infinite invariant measure, or there is a "quasi-invariant" measure, namely such that under the action of any element of the group, sets of measure 0 are transformed to sets of measure 0.]

The proof of (i) is immediate from the fact that the respective statements correspond to the subsets $\Gamma g H$ and $H g^{-1}\Gamma$ being dense in G, and these are sets of inverses of each other in the group G.

We shall not go into the proof of (ii), which though not difficult involves some technical details that are outside the scope of the present discussion.

In analogy with the Proposition the invariant measures of the two actions involved can be described in terms of those of the other (see [4] for details and references).

8 Diophantine approximation

Results on flows on homogeneous spaces have had many fruitful applications to questions on Diophantine approximation. To begin with we note that Hedlund's theorem, Theorem 6.2, has the following consequence, via the duality principle.

Corollary 8.1. *Let $\Gamma = SL(2,\mathbb{Z})$ and consider the natural Γ-action on $\mathbb{R}^2$. Then for $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, $\Gamma(v)$ is dense in $\mathbb{R}^2$ if and only if $v_1 \neq 0$ and v_2/v_1 is irrational.*

Thus given an irrational α, and $w_1, w_2 \in \mathbb{R}$, and $\epsilon > 0$, taking $v = \begin{pmatrix} \alpha \\ 1 \end{pmatrix}$ we see that there exist $p, q, r, s \in \mathbb{Z}$, with $ps - qr = 1$ such that

$$|p\alpha + q - w_1| < \epsilon \text{ and } |r\alpha + s - w_2| < \epsilon;$$

note also that we can get the pairs (p, q) and (r, s) can be chosen to be primitive (having gcd 1).

Proof of the corollary: We realise $\mathbb{R}^2\backslash(0)$ as G/N, where $G = SL(2,\mathbb{R})$ and $N = \left\{\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \mid t \in \mathbb{R}\right\}$, the one-parameter subgroup corresponding to the horocycle flow, which is also the stabiliser of e_1. Let $g \in G$ be such that $v = ge_1$. Then by Proposition 7.1 Γv is dense in $\mathbb{R}^2$ if and only if $Ng^{-1}\Gamma/\Gamma$ is dense in G/Γ. This now is an orbit of the horocycle flow and we recall that it is dense if and only if $g^{-1} \notin P\Gamma$, in the notation as before. The latter condition is equivalent to $g \notin \Gamma P$, by inversion, and it holds if and only if $v = ge_1 \notin \Gamma P e_1$. Now, Pe_1 is the x-axis and the condition $v \notin \Gamma P e_1$ can be readily seen to be equivalent to $v_1 \neq 0$ and v_2/v_1 irrational. This proves the 'if' part. The other way is easy to see. This completes the proof. $\square$

An analogous result holds for $\Gamma = SL(n, \mathbb{Z})$, $n \geq 2$, and $v = (v_1, \ldots, v_n)^t \in \mathbb{R}^n$, namely Γv is dense in $\mathbb{R}^n$ if and only if there exist j, k such that $v_j \neq 0$ and v_k/v_j is irrational. More generally the following is true, by a result of this author in collaboration with S. Raghavan (see [4] for some details and references).

Theorem 8.2. *Let $V = \mathbb{R}^n$, $n \geq 2$ and $1 \leq p \leq n-1$. Let $W = V^p = V \times \cdots \times V$ (p copies), and consider the componentwise action of $\Gamma = SL(n, \mathbb{Z})$ on W. Then for $w = (v_1, \ldots, v_p) \in W$, Γw is dense in W if and only if for no nonzero p-tuple $(\lambda_1, \ldots, \lambda_p)$ with $\lambda_i \in \mathbb{R}$, $1 \leq i \leq p$, $\lambda_1 v_1 + \cdots + \lambda_p v_p$ is an integral vector in V.*

In analogy with the above this theorem can be applied to study simultaneous integral solutions of systems of linear inequalities.

8.1 Oppenheim conjecture

A similar question came up concerning quadratic forms, in the form of Oppenheim conjecture, which was settled by Margulis in mid 1980's; the part about primitive solutions as in the following theorem was proved in a later joint work of Margulis and the present author.

Theorem 8.3. *Let $n \geq 3$ and $Q(x_1, \ldots, x_n) = \sum a_{ij} x_i x_j$ be a quadratic form, with $a_{ij} = a_{ji} \in \mathbb{R}$ for all i, j, and $\det(a_{ij}) \neq 0$. Suppose that*

i) there exists a nonzero n-tuple $(v_1, \ldots, v_n)$ such that $Q(v_1, \ldots, v_n) = 0$ and

ii) a_{ij}/a_{kl} is irrational for some i, j, k, l.

Then given $a \in \mathbb{R}$ and $\epsilon > 0$ there exist $x_1, \ldots, x_n \in \mathbb{Z}$ such that

$$|Q(x_1, \ldots x_n) - a| < \epsilon.$$

Moreover the n-tuple $(x_1, \ldots x_n)$ can be chosen to be primitive.

We note that both conditions (i) and (ii) are necessary for the validity of the desired conclusion. Condition (i) means that the quadratic form is "indefinite", namely neither positive definite nor negative definite, which is indeed necessary, and condition (ii) precludes forms which are scalar multiples of those with integer coefficients for which also the conclusion can not hold.

The corresponding statement is not true for $n = 2$, as we shall see in § 9 (see Theorem 9.1), which justifies the condition $n \geq 3$ in the hypothesis of Theorem 8.3; see [3] for a discussion of various issues around the theorem.

The proof of the theorem can be reduced in a routine way to the case of $n = 3$. For $n = 3$ the argument involved may be outlined as follows.

Let H be the subgroup of $SL(3, \mathbb{R})$ consisting of the elements leaving the quadratic form Q invariant, that is, $\{g \in G \mid Q(gv) = Q(v)\ \forall v \in \mathbb{R}^3\}$. Then $Q(\mathbb{Z}^3) = Q(H\Gamma\mathbb{Z}^3)$ on account of the invariances involved, and hence if $H\Gamma$ is dense in $SL(3, \mathbb{R})$ then it follows that $Q(\mathbb{Z}^3)$ is dense in $Q(\mathbb{R}^3 \backslash (0)) = \mathbb{R}$. One shows that under condition (ii) as in the theorem $H\Gamma$ is not closed. Thus the task is to show that every H-orbit on G/Γ which is not closed is dense in G/Γ. This statement may be compared to what we had seen earlier for the horocycle flows, for which the orbits are either periodic or dense in the ambient space, but the proof is more intricate in this case. A proof accessible via elementary arguments may be found in [7] (see also [5]).

8.2 Quantitative version

While existence of solutions to diophantine inequalities can be dealt with via consideration of density of orbits, uniform distribution of orbits enables to get asymptotic results for the number of solutions in large balls. Uniform distribution in the general case is studied via classification of invariant measures.

Using Ratner's classification of invariant measures of unipotent flows, and proving a "uniformized" version of her theorem on uniform distribution of orbits of unipotent flows (Theorem 6.3, supra), quantitative results were obtained regarding solutions of the quadratic forms as in Oppenheim's conjecture, in a paper of this author with Margulis). Let Q be a quadratic form satisfying the conditions as in the Oppenheim conjecture (viz. a quadratic form as in the hypothesis of Theorem 8.3). Then for any $a, b \in \mathbb{R}$, $a < b$ there exists a $c > 0$ such that

$$\#\{x \in \mathbb{Z}^n \mid a < Q(x) < b, \|x\| \leq r\} \geq cr^{n-2}.$$

The choice for a constant c can be explicitly described, but we shall not go into it here; see [13] for more details. The proof involves comparing the number on the left hand side with the volume of the region $\{v \in \mathbb{R}^n \mid a < Q(v) < b, \|v\| \leq r\}$ defined by the corresponding inequalities; these volumes asymptotic to the term on the right hand side in the above inequality.

By a result of Eskin, Margulis and Mozes for $n \geq 5$,

$$\frac{\#\{x \in \mathbb{Z}^n \mid a < Q(x) < b, \|x\| \leq r\}}{\mathrm{vol}\ \{v \in \mathbb{R}^n \mid a < Q(v) < b, \|v\| \leq r\}} \to 1, \text{ as } r \to \infty.$$

For $n = 3$ and 4 there are some situations when the number of solutions can be more than the "expected value" suggested by the denominator as above, and, specifically, the left hand side expression as above can have growth of the order of $\log r$. This phenomenon has turned out to be of some special interest. The reader is referred to [13] for more details on the theme.

9 Geodesic flows and applications

The orbit structure of the geodesic flow is much more complicated, compared to the horocycle flow. To describe the situation in this respect we first recall the geometric form of the flow; see [12] for an introduction to the general area.

The Poincaré upper half plane is

$$\mathbb{H} = \{z = x + iy \in \mathbb{C} \mid y > 0\},$$

equipped with the Riemannian metric, called the Poincaré metric,

$$ds^2 = \frac{dx^2 + dy^2}{y^2}.$$

Thus the distance between any two points z_1, z_2 is given by

$$d(z_1, z_2) = \inf \int y(t)^{-1} \sqrt{(\frac{dx}{dt}(t))^2 + (\frac{dy}{dt}(t))^2} dt,$$

with infimum taken over piecewise C^1 curves $x(t) + iy(t)$ joining z_1 and z_2.

The geodesics in this metric are vertical lines over points of $\mathbb{R}$, viz. the x-axis, or semicircles orthogonal to the x-axis; it may be emphasized that the x-axis is not in $\mathbb{H}$ but may be thought of as its "boundary". Thus each geodesic corresponds uniquely to, and is determined by, a pair of distinct points from $\mathbb{R} \cup \{\infty\}$ (as an ordered pair, when the geodesic is considered oriented in terms of the time direction); we call these points the endpoints of the geodesic (backward and forward endpoints, when we need to be specific.)

We denote by $S(\mathbb{H})$ the "unit tangent bundle", viz. the set of pairs (z, ξ), where $z \in \mathbb{H}$ and ξ is a (unit) tangent direction at the point z.

The "geodesic flow" corresponding to $\mathbb{H}$ is the flow $\{\varphi_t\}_{t\in\mathbb{R}}$ defined on $S(\mathbb{H})$, as follows: let $(z,\xi) \in S(\mathbb{H})$ be given and let $\gamma(t)$ be the geodesic (with unit speed, parametrized by the length parameter) starting at z and pointing in the direction ξ; then we choose $\varphi_t(z,\xi) = (z_t, \xi_t)$, where $z_t = \gamma(t)$ and $\xi = \gamma'(t)$, the unit tangent direction to γ at t.

The group $PSL(2,\mathbb{R}) = SL(2,\mathbb{R})/\{\pm I\}$ has an action on $\mathbb{H}$, with the action of $g \approx \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in PSL(2,\mathbb{R})$ given by

$$g(z) = \frac{az+b}{cz+d} \text{ for all } z \in \mathbb{H}.$$

The action of each $g \in PSL(2,\mathbb{R})$ is an isometry with respect to the Poincaré metric. These isometries together form a subgroup of index 2 in the group of all isometries, consisting of all orientation-preserving isometries.

The action on $\mathbb{H}$ induces also, canonically, an action of $PSL(2,\mathbb{R})$ on $S(\mathbb{H})$. For the latter action it can be readily seen that

$$g \in PSL(2,\mathbb{R}) \leftrightarrow g(i, v_0),$$

where v_0 is the unit direction at i pointing vertically upward, is homeomorphism, and using the correspondence we can identify $S(\mathbb{H})$ with $PSL(2,\mathbb{R})$. Under the identification the geodesic flow corresponds to the flow $\psi_t : PSL(2,\mathbb{R}) \to PSL(2,\mathbb{R})$ given, for all $t \in \mathbb{R}$, by

$$\psi_t(g) = g \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$$

(the product of matrices, viewed modulo $\{\pm I\}$); see [1] and [9] for more details.

Now let M be a surface of constant negative curvature and finite (Riemannian) area. Then M has $\mathbb{H}$ as its universal cover and it can be realised canonically as the quotient $\Gamma\backslash\mathbb{H}$, where Γ is a lattice $PSL(2,\mathbb{R})$, under the action of Γ as above); we note that Γ is also the fundamental group of M. Via identifications as above the unit tangent bundle $S(M)$ of M can be realised as $\Gamma\backslash PSL(2,\mathbb{R})$ and the geodesic flow associated with M, defined on $S(M)$, is given by $\psi_M(\Gamma g) \mapsto \Gamma g \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$, for all $t \in \mathbb{R}$; indeed, ψ_M is the projection of ψ on $S(M) = \Gamma\backslash PSL(2,\mathbb{R})$. The measure on $S(M)$ induced by the Riemannian area on M and angle measure along the directions in the fibers corresponds to the $PSL(2,\mathbb{R})$-invariant measure on $\Gamma\backslash PSL(2,\mathbb{R})$. It follows from Corollary 5.2 that the geodesic flow on $S(M)$ as above is ergodic. (See [1] for more details). In particular almost all of its orbits of the flow are dense in $S(M)$. Projecting to M we see that almost all geodesics in M are dense in M. Though the same holds for the torus $\mathbb{R}^2/\mathbb{Z}^2$ equipped with the flat metric, it may be noted that the situation here is different, and more chaotic in a certain sense, since the lifts of the geodesics in the unit tangent bundle are also dense in $S(M)$, while in the case of the torus each remains confined to a surface.

9.1 Geodesics on the modular surface

We now consider images of the geodesics in $\mathbb{H}$ in the quotient $PSL(2,\mathbb{Z})\backslash\mathbb{H}$.

Recall that a geodesic in $\mathbb{H}$ is determined by two (distinct) points in $\mathbb{R}\cup\{\infty\}$, the "endpoints", as noted earlier. The endpoints have considerable bearing on the behaviour of the geodesic in $PSL(2,\mathbb{Z})\backslash\mathbb{H}$ obtained by projecting it, and the corresponding orbit of the geodesic flow associated to $PSL(2,\mathbb{Z})\backslash\mathbb{H}$.

Theorem 9.1. *For $\alpha,\beta\in\mathbb{R}\cup\{\infty\}$ let $g(\alpha,\beta)$ be the geodesic in $\mathbb{H}$ with endpoints (α,β), and let $\overline{g}(\alpha,\beta)$ be its image in $PSL(2,\mathbb{Z})\backslash\mathbb{H}$. Then we have the following:*

i) $\overline{g}(\alpha,\beta)$ is a closed noncompact subset if and only if $\alpha,\beta\in\mathbb{Q}\cup\{\infty\}$;

ii) $\overline{g}(\alpha,\beta)$ is periodic if and only if α and β are conjugate quadratic numbers, viz. the two (irrational) roots of an irreducible quadratic polynomial with rational coefficients;

iii) $\overline{g}(\alpha,\beta)$ is contained in a compact subset of $PSL(2,\mathbb{Z})\backslash\mathbb{H}$ if and only if α and β are badly approximable.

We recall that a real number θ is said to be "badly approximable" if there exists a $\delta>0$ such that for all $p,q\in\mathbb{Z}$, $q\neq 0$, $|\theta-\frac{p}{q}|>\delta/q^2$.

9.2 Continued fractions

Every real number θ has a "continued fraction" expansion as

$$\theta = a_0 + \frac{1}{a_1+\frac{1}{a_2+\cdots}},$$

where $a_0\in\mathbb{Z}$ and $a_k\in\mathbb{N}$ for all $k\in\mathbb{N}$; the expansion terminates at a finite stage if θ is rational and yields an infinite sequence if θ is irrational (see, for instance, [11] for basic details, and [9] for an introduction from the perspective of Ergodic theory). The sequence $\{a_k\}$ (finite or infinite) is called the continued fraction expansion of θ; for convenience we represent the expansion as above by $(a_0, a_1, \dots)$. Conversely, every such finite expression defines a rational number, and given an infinite sequence $\{a_k\}$ with $a_0 \in \mathbb{Z}$ and $a_k \in \mathbb{N}$ for all $k \in \mathbb{N}$, the (rational) numbers corresponding to the finite sequences $\{a_n\}_{k=0}^{n}$ converge as $n \to \infty$ to an irrational number, with $(a_0, a_1, \dots)$ as its continued fraction expansion.

The a_k's as above are called the partial quotients of the continued fraction expansion. Properties (ii) and (iii) as in Theorem 9.1 have the following analogues in terms of continued fractions:

ii′) θ is a quadratic number if and only if the partial quotients are eventually periodic; i.e. there exist l and m such that $a_{k+l} = a_k$ for all $k \geq m$.

iii′) θ is badly approximable if and only if the partial quotients are bounded; i.e. there exists M such that $a_k \leq M$ for all k.

The last statement in particular tells us that badly approximable numbers exist, since we can construct them starting with bounded sequences in $\mathbb{N}$. The badly approximable numbers form a set of Lebesgue measure 0, but nevertheless constitute a large set in other ways: the Hausdorff dimension of the set is 1, the maximum possible for a subset of $\mathbb{R}$, and they are winning sets of certain games (called Schmidt games); see [2] for more details and references.

9.3 Generic numbers

A real number θ is said to be *generic* if in the continued fraction expansion $(a_0, a_1, \dots, a_k, \dots)$ of θ every finite block of positive integers occurs; that is, given $(b_1, \dots, b_l)$, $b_k \in \mathbb{N}$, there exists m such that $a_m = b_1, \dots, a_{m+l-1} = b_l$. With this we can add the following to the list in Theorem 9.1.

iv) $\overline{g}(\alpha, \beta)$ is dense in $PSL(2, \mathbb{Z})\backslash\mathbb{H}$ if and only if at least one of α and β is generic.

The set of generic numbers is a set of full Lebesgue measure. The ergodicity result mentioned earlier tells us that the set of pairs (α, β) for which the conclusion as in (iv) holds is a set of full measure in $\mathbb{R}^2$. The point of assertion (iv) however is that it gives a specific set of numbers, in terms of their continued fraction expansions, for which it holds. Also, historically this result was proved, by E. Artin, before the ergodicity result came up. Recently in a paper of this author with Nogueira a strengthening of the result was obtained by a different method. We shall not go into the details of the dynamical result here but describe in the next subsection one of its consequences, to Diophantine approximation, involving binary quadratic forms.

9.4 Binary quadratic forms

We consider a binary quadratic form Q defined by $Q(x,y) = (x - \alpha y)(x - \beta y)$, where $\alpha, \beta \in \mathbb{R}$ and $\alpha \neq \beta$; we note that up to scalar multiples and transposition of the variables x and y every binary quadratic form satisfying conditions (i) and (ii) in Theorem 8.3 can be expressed in this way. We have the following.

Theorem 9.2. *i) If α, β are badly approximable then there exists $\delta > 0$ such that $Q(\mathbb{Z}^2) \cap (-\delta, \delta) = \{0\}$, so in particular $Q(\mathbb{Z}^2)$ is not dense in $\mathbb{R}$; (this may be contrasted with the Oppenheim conjecture for $n \geq 3$, in § 8.1).*

ii) If one of α and β is generic then $Q(\mathbb{Z}^2)$ is dense in $\mathbb{R}$.

iii) If one of α and β is a positive generic number then $Q(\mathbb{N}^2)$ is dense in $\mathbb{R}$.

Assertion (ii) is deduced from the result of Artin mentioned in § 9.3. Assertion (iii) is deduced from our (A. Nogueira and this author) result alluded to above.

Recently we (the author in collaboration with A. Nogueira) studied continued fractions for complex numbers, in terms of the Gaussian integers, and proved analogues of the density result for complex binary forms; see [8]. In this case also the result contrasts the analogue of Oppenheim conjecture, which is known, by a result of A. Borel and Gopal Prasad.

9.5 Actions of diagonal subgroups

Even though the orbit structure of the geodesic flows is rather intricate as discussed, for $n \geq 3$ the orbit structure of the corresponding subgroup D_n, consisting of all diagonal matrices with positive entries, on $SL(n,\mathbb{R})/\Gamma$ with Γ a lattice in $SL(n,\mathbb{R})$, is expected to be "nice", in a way similar to the unipotent case. In particular, Margulis has conjectured that all orbits of D_n having compact closure are homogeneous subsets; there is also a more general version of this, but a reader interested in pursuing the topic should also bear in mind the counterexamples given by Maucourant [14]. This conjecture of Margulis implies a well-known conjecture in Diophantine approximation called Littlewood conjecture:

Let α, β be irrational real numbers. Then

$$\liminf_{n \to \infty} \; n\|n\alpha\|\|n\beta\| > 0$$

where for any $\theta \in \mathbb{R}$, $\|\theta\|$ denotes the distance of θ from the integer nearest to it.

Landmark work has been done by Einsiedler, Katok and Lindenstrauss on these questions. It has been shown in particular that the set of pairs (α, β) for which the Littlewood conjecture does not hold is small in the sense that it has Hausdorff dimension 0; see [17] for an exposition of the work. In its full form the conjecture remains open.

Acknowledgement: The author is thankful to the anonymous referee for suggestions found useful in improving the text of the paper.

References

[1] M. Bachir Bekka and Matthias Mayer, *Ergodic theory and topological dynamics of group actions on homogeneous spaces*, London Mathematical Society Lecture Note Series, 269, Cambridge University Press, Cambridge, 2000.

[2] S.G. Dani, On badly approximable numbers, Schmidt games and bounded orbits of flows, *Number Theory and Dynamical Systems* (York, 1987), 69 - 86, London Math. Soc. Lecture Note Ser., 134, Cambridge Univ. Press, Cambridge, 1989.

[3] S.G. Dani, On values of linear and quadratic forms at integral points, *Number theory*, 107 - 119, Trends Math., Birkhuser, Basel, 2000.

[4] S.G. Dani, in *Dynamical Systems, Ergodic Theory and Applications*, Edited and with a preface by Sinai, Encyclopaedia of Mathematical Sciences, 100, Mathematical Physics, I. Springer - Verlag, Berlin, 2000.

[5] S.G. Dani, On the Oppenheim conjecture on values of quadratic forms, Essays on geometry and related topics, Vol. 1, 2, 257 - 270, Monogr. Enseign. Math., 38, Enseignement Math., Geneva, 2001.

[6] S.G. Dani, Dynamical properties of linear and projective transformations and their applications, Indian J. Pure Appl. Math. 35 (2004), no. 12, 1365 - 1394.

[7] S.G. Dani and G.A. Margulis, Values of quadratic forms at integral points: an elementary approach, Enseign. Math. (2) 36 (1990), no. 1-2, 143 - 174.

[8] S.G. Dani and A. Nogueira, Continued fractions for complex numbers and values of binary quadratic forms, Trans. Amer. Math. Soc. (to appear).

[9] Manfred Einsiedler and Thomas Ward, *Ergodic Theory with a View towards Number Theory*, Graduate Texts in Mathematics, 259. Springer-Verlag London, Ltd., London, 2011.

[10] Gerald B. Folland, *A course in abstract harmonic analysis*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1995.

[11] G.H. Hardy and E.M. Wright, *An introduction to the Theory of Numbers*, Sixth Edition. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles. Oxford University Press, Oxford, 2008.

[12] Svetlana Katok, *Fuchsian groups*, Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1992.

[13] G.A. Margulis, Oppenheim conjecture, *Fields Medallists' Lectures*, 272 - 327, World Sci. Ser. 20th Century Math. 5, World Sci. Publ., River Edge, NJ, 1997.

[14] Francois Maucourant, A nonhomogeneous orbit closure of a diagonal subgroup, Ann. of Math. (2) 171 (2010), 557 - 570.

[15] Dave Witte Morris, *Ratner's Theorems on Unipotent Flows*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 2005.

[16] M.S. Raghunathan, *Discrete subgroups of Lie groups*, Springer, 1972.

[17] Akshay Venkatesh, The work of Einsiedler, Katok and Lindenstrauss on the Littlewood conjecture, Bull. Amer. Math. Soc. (N.S.) 45 (2008), 117 - 134.

S.G. Dani
Department of Mathematics
Indian Institute of Technology Bombay
Powai, Mumbai 400076
Indai

E-mail: `sdani@math.iitb.ac.in`