Chapter 6
Unipotent Dynamics and Ratner's Theorems

In this chapter we discuss unipotent dynamics and prove several special cases of Ratner's measure classification. We will not discuss the history in detail, and refer to the survey papers of Kleinbock, Shah and Starkov [86], Ratner [138], Margulis [109], and Dani [24] for that. In particular, the order in which the material is developed is not historical but instead emphasizes a logical development with the benefit of hindsight.

6.1 Unipotent Invariance

Unipotent dynamics and Ratner's theorem (as discussed in greater detail in the next section) are of great importance for various applications. To see this we explain how unipotent invariance may arise naturally in applications of various types.

6.1.1 Factor Rigidity

For simplicity we let $G = \mathrm{SL}_2(\mathbb{R})$ or $G = \mathrm{SL}_3(\mathbb{R})$. Let $\Gamma_1, \Gamma_2 < G$ be lattices and write $X_j = G/\Gamma_j$ for $j = 1, 2$. Moreover, let $H = \{u_t \mid t \in \mathbb{R}\}$ be a one-parameter subgroup.

Definition 6.1. A *factor map* from the action of H on X_1 to the action of H on X_2 is a measurable almost everywhere defined map $\phi\colon X_1 \to X_2$ such that

$$\phi(u_t{\bullet}x) = u_t{\bullet}\phi(x) \tag{6.1}$$

for every $t \in \mathbb{R}$ and almost every $x \in X_1$ and $\phi_* m_{x_1} = m_{X_2}$. Moreover, ϕ is called an *isomorphism* if there exists a measurable set $X_1' \subseteq X_1$ of full measure such that $\phi|_{X_1'}$ is injective.

In general factor maps or isomorphisms are potentially indeed only measurable, of fractal nature, and may not respect (for example) topological dimension. However, if $H = U$ is a unipotent subgroup then factor maps are essentially algebraic, and in particular are smooth maps. That is, the equivariance condition in (6.1) along one direction of G forces good behaviour along all directions of G. This is quite surprising and is an instance of a 'rigidity' phenomenon. The case of $G = \mathrm{SL}_2(\mathbb{R})$ is due to earlier work of Ratner [131] and more general cases have been obtained by Morris [119] as a corollary of the results of Ratner to be discussed in the next section.

Theorem 6.2 (Factor rigidity). *Let X_1, X_2 be as above, let U be a one-parameter unipotent subgroup, and let $\phi\colon X_1 \to X_2$ be a factor map for U. Then there exists an automorphism $\varphi\colon G \to G$ whose restriction to U is the identity map and some element $p \in G$ so that*

$$\phi(g\Gamma_1) = \varphi(g)p\Gamma_2$$

for all $g \in G$. In particular, the lattices are related by $\varphi(\Gamma_1) \subseteq p\Gamma_2 p^{-1}$ and the factor map ϕ is an isomorphism if and only if Γ_1 and Γ_2 have the same covolume.

The proof of Theorem 6.2 starts by encoding the map ϕ in terms of the probability measure

$$\mu = \mu_\phi = (\mathrm{id}, \phi)_* m_{X_1}$$

on $X_1 \times X_2$. We note that μ is concentrated[†] on $\mathrm{Graph}(\phi)$, meaning that

$$\mu\bigl(\mathrm{Graph}(\phi)\bigr) = 1.$$

Moreover, the equivariance of ϕ in (6.1) shows that μ is invariant under the action of (u_t, u_t) on $X_1 \times X_2$ for $u_t \in U$. Ergodicity of this action follows as it is measurably isomorphic to the action of U on X_1 with respect to m_{X_1}. The projection of μ to X_2 is $\phi_*(m_{X_1}) = m_{X_2}$.

Theorem 6.3 (Ratner's joining classification). *Let μ be a probability measure on $X_1 \times X_2$ that projects to m_{X_1} and m_{X_2} under the coordinate projections and that is invariant and ergodic under the action of (u_t, u_t) for $u_t \in U$. Then μ is the L-invariant probability measure on a closed L-orbit for a closed connected subgroup $L < G \times G$. Moreover, either $L = G \times G$ or L is the graph of an automorphism $\varphi\colon G \to G$.*

We will prove Theorems 6.2 and 6.3 in this chapter and use the latter to further understand all possible 'abstract' factors of U acting on X_1.

Exercise 6.4. Try to prove Theorem 6.3 using Theorem 6.2 as a black box.

[†] We avoid saying 'supported' as we do not know whether $\mathrm{Graph}(\phi)$ is closed.

6.1.2 Oppenheim's Conjecture

We note that for a linear form L in $d \geqslant 2$ variables it is easy to determine whether $L(\mathbb{Z}^d) \subseteq \mathbb{R}$ is dense or not. In fact $L(\mathbb{Z}^d)$ is dense in $\mathbb{R}$ if and only if L is not a multiple of a form with rational coefficients. As a generalization Oppenheim [125] conjectured in 1929 that a non-degenerate indefinite quadratic form Q in $d \geqslant 5$ variables that is not a multiple of a form with integer coefficients has $Q(\mathbb{Z}^d)$ dense in $\mathbb{R}$.

Raghunathan [129] noticed(31) in the mid 1970s the connection to homogeneous dynamics and, motivated by this, formulated far-reaching conjectures concerning orbit closures for subgroups generated by unipotent subgroups. Margulis developed these ideas to prove the Oppenheim conjecture in the following stronger form in 1986 [110].

Theorem 6.5 (Margulis' proof of Oppenheim conjecture). *Let Q be a non-degenerate indefinite quadratic form in $d \geqslant 3$ variables that is not a multiple of a form with integer coefficients. Then $Q(\mathbb{Z}^d_{\mathrm{prim}})$ is dense in $\mathbb{R}$, where*

$$\mathbb{Z}^d_{\mathrm{prim}} = \{v = (v_1, \dots, v_d)^{\mathrm{t}} \in \mathbb{Z}^d \mid \gcd(v_1, \dots, v_d) = 1\}.$$

We will prove Oppenheim's conjecture later, but for now let us point out that it also follows quickly from the following dynamical result.

Theorem 6.6 (Orbit closure by Dani–Margulis). *Let*

$$Q_0(x_1, x_2, x_3) = x_2^2 - 2x_1x_3$$

and $H = \mathrm{SO}_{Q_0}(\mathbb{R})^o$. For any $x_0 \in \mathsf{X}_3$ either $H \bullet x_0$ is closed or $H \bullet x_0$ is dense in X_3.

We note that an orthogonal group in 3-dimensions is 3-dimensional. Moreover, $\mathrm{SO}_{2,1}(\mathbb{R})$ is locally isomorphic to $\mathrm{SL}_2(\mathbb{R})$. To see this† consider the adjoint representation of $g \in \mathrm{SL}_2(\mathbb{R})$ on the Lie algebra $\mathfrak{sl}_2(\mathbb{R})$ by $v \mapsto gvg^{-1}$ and the indefinite quadratic form $\det v = -a^2 - bc$ for $v = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$. This shows that $H \cong \mathrm{PSL}_2(\mathbb{R})$ is simple and generated by unipotent one-parameter subgroups.

We have chosen the above quadratic form Q_0 so that its orthogonal group SO_{Q_0} is easy to work with. For instance, $\mathrm{SO}_{Q_0}(\mathbb{R})$ contains the diagonal matrices of the form $\mathrm{diag}(a, 1, a^{-1})$ for $a \in \mathbb{R}^\times$. Moreover, it contains the unipotent one-parameter subgroup

$$\left\{ \begin{pmatrix} 1 & t & \frac{t^2}{2} \\ & 1 & t \\ & & 1 \end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\} \leqslant \mathrm{SO}_{Q_0}(\mathbb{R}).$$

† This is an instance of a low-dimensional accident: In essence there is only one real noncompact simple Lie algebra of dimension 3 up to isomorphism.

To see this we calculate

$$\begin{aligned} Q_0\left(\begin{pmatrix}1 & t & \frac{t^2}{2}\\ & 1 & t\\ & & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}\right) &= Q_0\left(\begin{pmatrix}x_1+tx_2+\frac{t^2}{2}x_3\\ x_2+tx_3\\ x_3\end{pmatrix}\right)\\ &= (x_2+tx_3)^2 - 2\bigl(x_1+tx_2+\tfrac12 t^2x_3\bigr)x_3\\ &= x_2^2+2tx_2x_3+t^2x_3^2-2x_1x_3-2tx_2x_3-t^2x_3^2\\ &= x_2^2-2x_1x_3. \end{aligned} \tag{6.2}$$

Similarly, we also have

$$\begin{pmatrix}1 & & \\ t & 1 & \\ \frac{t^2}{2} & t & 1\end{pmatrix} \in \mathrm{SO}_{Q_0}(\mathbb{R})$$

for $t\in\mathbb{R}$. Finally we note that the Lie algebra elements corresponding to the diagonal group and these two one-parameter unipotent subgroups can be chosen to match the $\mathfrak{sl}_2$-triple (2.5) on page 65.

Proof that Theorem 6.6 implies Theorem 6.5 for $d=3$. Suppose first that Q is a non-degenerate indefinite quadratic form in $d=3$ variables and is not a multiple of a form with integer coefficients. By our discussion of signatures of quadratic forms in Theorem 3.5 it follows that there exist $\lambda\in\mathbb{R}^\times$ and $g_0\in\mathrm{SL}_3(\mathbb{R})$ so that $Q=\lambda Q_0\circ g_0$. In particular, $Q(\mathbb{Z}^3_{\mathrm{prim}})=\lambda Q_0(g_0\mathbb{Z}^3_{\mathrm{prim}})$. We define $x_0=g_0\mathbb{Z}^3\in\mathsf{X}_3$, $H=\mathrm{SO}_{Q_0}(\mathbb{R})^o$, and consider the two cases in Theorem 6.6.

Density implies Density. Suppose that the orbit $H\!\centerdot x_0$ is dense in X_d. Fix some $a\in\mathbb{R}$ and set $g=\begin{pmatrix}1&0&0\\0&1&0\\-a&0&1\end{pmatrix}$ so that $a\in Q_0(g\mathbb{Z}^3_{\mathrm{prim}})$. If now $h\in H$ is such that $h\!\centerdot x_0$ is very close to x then the values in $\lambda Q_0(g_0\mathbb{Z}^3_{\mathrm{prim}})=\lambda Q_0(hg_0\mathbb{Z}^3_{\mathrm{prim}})$ can be used to approximate $\lambda a=\lambda Q_0(ge_1)$ arbitrarily well. In other words $Q(\mathbb{Z}^3_{\mathrm{prim}})$ is dense in $\mathbb{R}$.

Closed implies Rationality. Suppose now that the orbit $H\!\centerdot x_0$ is closed. As $\mathrm{SO}_{3,1}(\mathbb{R})^o\cong\mathrm{PSL}_2(\mathbb{R})$ is non-compact and simple, Theorem 4.18 applies and shows that $H\!\centerdot x$ actually has finite volume. By Borel density (Theorem 3.50) it follows that $g_0^{-1}Hg_0\cap\mathrm{SL}_d(\mathbb{Z})$ is Zariski dense in $g_0^{-1}\mathrm{SO}_{Q_0}g_0=\mathrm{SO}_Q$. We now show that this implies that a multiple of Q has integer coefficients, giving a contradiction to our assumptions.

We start by showing that the linear hull $\langle Q_0\rangle$ of $Q_0((x_1,x_2,x_3)^{\mathrm{t}})=x_2^2-2x_1x_3$ is the subspace of all quadratic forms q satisfying $q\circ h=q$ for all $h\in H$. Indeed for $a=\mathrm{diag}(\mathrm{e},1,\mathrm{e}^{-1})\in H$ every quadratic monomial is an eigenvector but only x_1x_3 and x_2^2 have eigenvalue 1. Hence $q\circ a=q$ implies that $q((x_1,x_2,x_3)^{\mathrm{t}})=\alpha x_2^2+\beta x_1x_3$ for two scalars α and β. Going through the calculation in (6.2) that H contains the unipotent elements u_t again, we see that the required cancellation of term in the expression $q \circ u_t$ only happens if $\beta = -2\alpha$.

By conjugation with g_0 we also see that the linear hull $\langle Q \rangle$ of Q is the subspace of all quadratic forms q satisfying $q \circ h = q$ for all $h \in \mathrm{SO}_Q(\mathbb{R})^o$, or equivalently for all $h \in \mathrm{SO}_Q(\mathbb{R})^o \cap \mathrm{SL}_3(\mathbb{Z})$. As the latter are rational equations that have a nontrivial solution, we obtain that a multiple of Q has integer coefficients. □

One may wonder why it might be advantageous to study three-dimensional orbits inside the eight-dimensional space X_3 in Theorem 6.6 to prove a mere density statement in $\mathbb{R}$ as in Theorem 6.5. As a partial answer to this we note that the set $Q(\mathbb{Z}^3) \subseteq \mathbb{R}$ has very little structure and, in particular, has no invariance properties. However, in the eight-dimensional X_3 applying the orthogonal group to a lattice does not change the values of the quadratic form. Moreover, as the quadratic form is indefinite we obtain in this way the powerful tool of unipotent invariance.

Exercise 6.7. Show that Oppenheim's conjecture cannot hold for binary quadratic forms (that is, forms in two dimensions). For this let $\alpha \in \mathbb{R}$ be badly approximable with $\alpha^2 \notin \mathbb{Q}$ and consider the form $x^2 - \alpha^2 y^2 = (x - \alpha y)(x + \alpha y)$.

6.1.3 Distorted Orbits

In Sections 5.4–5.7 we have seen the importance of studying 'distorted orbits' of the form $gH \boldsymbol{\cdot} \Gamma$ while varying $gH \in G/H$. We wish to explain—under suitable assumptions—why weak* limits of the Haar measures on such distorted orbits often have unipotent invariance.

Lemma 6.8 (Unipotent invariance for limits of distorted orbits). *Let G be a closed subgroup of $\mathrm{SL}_d(\mathbb{R})$, $\Gamma < G$ a lattice, and $H < G$ a closed subgroup with Lie algebra $\mathfrak{h} = \mathrm{Lie}\, H$ so that $\Gamma \cap H < H$ is also a lattice. We assume moreover that (a_n) is a sequence in G with*

$$\lim_{n\to\infty} \big\| \mathrm{Ad}_{a_n}|_{\mathfrak{h}} \big\| = \infty. \tag{6.3}$$

Then any weak limit of the Haar measures on $a_n H \boldsymbol{\cdot} \Gamma$ inside $X = G/\Gamma$ is invariant under a one-parameter unipotent subgroup.*

Proof. We assume without loss of generality that the Haar measures $m_{a_n H \boldsymbol{\cdot} \Gamma}$ on $a_n H \boldsymbol{\cdot} \Gamma$ converge in the weak* topology to a measure μ. By the assumption (6.3) there exists a sequence (v_n) in $\mathfrak{h}$ so that $v_n \to 0$ but $\|\mathrm{Ad}_{a_n} v_n\| = 1$ for all $n \geqslant 1$. By choosing a subsequence once more we may assume that $\mathrm{Ad}_{a_n} v_n$ converges to an element w in the Lie algebra $\mathfrak{g}$ of G. Let $f \in C_c(X)$ and $t \in \mathbb{R}$. Then

$$\int f\big(\exp(tw) \boldsymbol{\cdot} x\big)\, \mathrm{d}\mu(x) = \lim_{n\to\infty} \int f\big(\exp(tw) \boldsymbol{\cdot} x\big)\, \mathrm{d}m_{a_n H \boldsymbol{\cdot} \Gamma}(x)$$

by definition. Moreover, $\mathsf{d}\bigl(\exp(tw), a_n \exp(tv_n)a_n^{-1}\bigr) \to 0$ as $n \to \infty$. Hence uniform continuity of f allows us to replace $\exp(tw)$ by $a_n \exp(tv_n)a_n^{-1}$, giving

$$\int f\bigl(\exp(tw)\boldsymbol{\cdot}x\bigr)\,\mathrm{d}\mu(x) = \lim_{n\to\infty} \int f\bigl(\underbrace{a_n \exp(tv_n)a_n^{-1}}_{\in a_n H a_n^{-1}}\boldsymbol{\cdot}x\bigr)\,\mathrm{d}m_{a_n H\boldsymbol{\cdot}\Gamma}(x)$$

$$= \lim_{n\to\infty} \int f\,\mathrm{d}m_{a_n H\boldsymbol{\cdot}\Gamma} = \int f\,\mathrm{d}\mu,$$

where we have used the fact that $m_{a_n H\boldsymbol{\cdot}\Gamma}$ is invariant under $a_n H a_n^{-1}$. As the function $f \in C_c(X)$ and $t \in \mathbb{R}$ were arbitrary we deduce that μ is invariant under the one-parameter subgroup determined by w. Finally note that the eigenvalues of $w = \lim_{n\to\infty} \mathrm{Ad}_{a_n} v_n$ vanish as $\lim_{n\to\infty} v_n = 0$ and the conjugation Ad_{a_n} does not change eigenvalues. □

We note that Lemma 6.8 and the powerful theorems due to Ratner from the next section can be used for proving equidistribution (and counting) results in situations where the banana mixing argument from Chapter 5 does not apply.

6.1.4 Orbits Arising From Expanding Curves

We wish to explain another way in which unipotent invariance may arise. For this we suppose that $\mathcal{I} \subseteq \mathbb{R}$ is a compact interval and that $\gamma\colon \mathcal{I} \to \mathbb{R}^d$ has continuous second derivative. We also assume that $\gamma'(s) \neq 1$ for $s \in \mathcal{I}$ and that $\mathcal{I} \ni s \mapsto g_s \in \mathrm{GL}_d(\mathbb{R})$ is continuous so that $g_s\gamma'(s) = \mathbf{e}_1$ for all $s \in \mathcal{I}$. We identify $g \in \mathrm{GL}_d(\mathbb{R})$ with

$$\begin{pmatrix} (\det g)^{-1} & \\ & g \end{pmatrix} \in \mathrm{SL}_{d+1}(\mathbb{R})$$

to simplify the notation. For $v \in \mathbb{R}^d$ and $t \in \mathbb{R}$ we also define $u_v = \begin{pmatrix} 1 & \\ v & I_d \end{pmatrix}$ and $a_t = \begin{pmatrix} t^{-d} & \\ & tI_d \end{pmatrix}$.

Lemma 6.9 (Twisting trick, first step). *Let $x_0 \in \mathsf{X}_{d+1}$. Using the above assumptions and notation, any weak* limit of*

$$\frac{1}{|\mathcal{I}|} \int_{\mathcal{I}} \delta_{a_t g_s u_{\gamma(s)} x_0}\,\mathrm{d}s \tag{6.4}$$

for $t \to \infty$ is invariant under the one-parameter subgroup $U = \{u_{r\mathbf{e}_1} \mid r \in \mathbb{R}\}$.

Proof. Let $f \in C_c(\mathsf{X}_{d+1})$, $r \in \mathbb{R}$, and $t \geqslant 0$. We define $\kappa = 1 + \frac{1}{d}$ and note that $a_t u_v a_t^{-1} = u_{\mathrm{e}^{-\kappa t} v}$ for $v \in \mathbb{R}^d$. Then we have

$$\int_{\mathcal{I}} f(u_{r\mathbf{e}_1} a_t g_s u_{\gamma(s)} x_0)\,\mathrm{d}s = \int_{\mathcal{I}} f(a_t u_{\mathrm{e}^{-\kappa t} r\mathbf{e}_1} g_s u_{\gamma(s)} x_0)\,\mathrm{d}s$$
$$= \int_{\mathcal{I}} f(a_t g_s u_{\mathrm{e}^{-\kappa t} r\gamma'(s)+\gamma(s)} x_0)\,\mathrm{d}s$$

by the way a_t and g_s interact with $u_{\mathbf{e}_1}$. Fixing r and thinking of a large $t > 0$, we have

$$\gamma(s) + \mathrm{e}^{-\kappa t} r\gamma'(s) = \gamma\left(s + \mathrm{e}^{-\kappa t} r\right) + \varepsilon(s)$$

for an error term $\varepsilon(s) = \varepsilon_{r,t}(s) = \mathrm{O}\big(\mathrm{e}^{-2\kappa t}\big)$ as γ is assumed to be twice continuously differentiable. Let

$$\mathcal{I}' = \mathcal{I} \cap \left(\mathcal{I} + \mathrm{e}^{-2\kappa t} r\right)$$

which (for fixed $r \in \mathbb{R}$ and large $t > 0$) is basically equal to $\mathcal{I}$. We let

$$s' = s + \mathrm{e}^{-\kappa t} r$$

for $s \in \mathcal{I}$, and obtain

$$\int_{\mathcal{I}} f(u_{r\mathbf{e}_1} a_t g_s u_{\gamma(s)} x_0)\,\mathrm{d}s = \int_{\mathcal{I}'} f(a_t g_s u_{\gamma(s')+\varepsilon(s)} x_0)\,\mathrm{d}s' + \mathrm{O}\big(r\mathrm{e}^{-\kappa t}\big).$$

We now take the error term $\varepsilon(s)$ and again move it across g_s (which will rotate and mildly stretch or contract it) and a_t (which will expand it) to the left. Defining $\varepsilon_{\mathrm{new}}(s) = g_s\varepsilon(s)$, we obtain

$$\int_{\mathcal{I}} f(u_{r\mathbf{e}_1} a_t g_s u_{\gamma(s)} x_0)\,\mathrm{d}s = \int_{\mathcal{I}'} f(u_{\mathrm{e}^{\kappa t}\varepsilon_{\mathrm{new}}(s)} a_t g_s u_{\gamma(s')})\,\mathrm{d}s' + \mathrm{O}\big(r\mathrm{e}^{-\kappa t}\big).$$

As $\varepsilon(s) = \mathrm{O}\big(\mathrm{e}^{-2\kappa t}\big)$, we know that

$$\mathrm{e}^{\kappa t}\varepsilon_{\mathrm{new}}(s) = \mathrm{O}\big(\mathrm{e}^{-\kappa t}\big)$$

is tiny. Using continuity of f, we see that the term $u_{\mathrm{e}^{\kappa t}\varepsilon_{\mathrm{new}}(s)}$ does not change the value of f much. For a weak* limit μ of (6.4) as $t \to \infty$ this shows that

$$\int f(u_{r\mathbf{e}_1} x)\,\mathrm{d}\mu(x) = \int f\,\mathrm{d}\mu.$$

As $f \in C_c(\mathsf{X}_{d+1})$ was arbitrary, it follows that μ is invariant under U, as desired. □

Suppose now we can use unipotent magic (meaning Ratner's theorem and related results) to show that the measures in (6.4) equidistribute. We next explain what this has to do with the expanded curves $a_t u_{\gamma(s)} x_0$ for $s \in \mathcal{I}$.

Lemma 6.10 (Twisting trick, second step). *Suppose now in addition that for any interval $\mathcal{I}$ and any $x_0 \in \mathsf{X}_{d+1}$ we have*

$$\frac{1}{|\mathcal{I}|}\int_{\mathcal{I}} \delta_{a_t g_s u_{\gamma(s)} x_0}\,\mathrm{d}s \longrightarrow m_{\mathsf{X}_{d+1}}.$$

Then for any interval $\mathcal{I}$ and $x_0 \in \mathsf{X}_{d+1}$ we also have that

$$\frac{1}{|\mathcal{I}|}\int_{\mathcal{I}} \delta_{a_t u_{\gamma(s)} x_0}\,\mathrm{d}s \longrightarrow m_{\mathsf{X}_{d+1}}.$$

In other words, the additional matrix g_s that was so helpful above to obtain unipotent invariance can simply be forgotten.

Proof of Lemma 6.10. Let $\varepsilon > 0$ and $f \in C_c(\mathsf{X}_{d+1})$. By the assumed continuity of the map $\mathcal{I} \ni s \mapsto g_s$, the function $\mathcal{I} \times \mathsf{X}_{d+1} \ni (s,x) \mapsto f(g_s^{-1}x)$ is uniformly continuous. Hence there exists some $\delta > 0$ so that

$$\left| f(g_{s_1}^{-1}x) - f(g_{s_2}^{-1}x)\right| < \varepsilon \tag{6.5}$$

whenever $x \in \mathsf{X}_{d+1}$ and $s_1, s_2 \in \mathcal{I}$ satisfy $|s_1 - s_2| < \delta$.

We split $\mathcal{I}$ into finitely many sub-intervals $\mathcal{I}_\ell$ for $\ell = 1, \ldots, L$ of equal length and length less than δ. Using the assumed equidistribution for each of these intervals, the above continuity property of f, and the fact that g_s commutes with a_t, we can now obtain the desired conclusion up to 2ε. Indeed, fix some $s_\ell \in \mathcal{I}_\ell$ for $\ell = 1, \ldots, L$ and apply our assumption to the interval $\mathcal{I}_\ell$ and the function $\mathsf{X}_{d+1} \ni x \mapsto f(g_{s_\ell}^{-1}x)$. As $m_{\mathsf{X}_{d+1}}$ is invariant under g_{s_ℓ}, we therefore have

$$\left|\frac{1}{|\mathcal{I}_\ell|}\int_{\mathcal{I}_\ell} f\left(g_{s_\ell}^{-1} g_s a_t u_{\gamma(s)}\boldsymbol{\cdot} x_0\right)\,\mathrm{d}s - \int_{\mathsf{X}_{d+1}} f\,\mathrm{d}m_{\mathsf{X}_{d+1}}\right| < \varepsilon$$

for $\ell = 1, \ldots, L$ and all sufficiently large t. Now we may use the estimate (6.5) for $x = g_s a_t u_{\gamma(x)}\boldsymbol{\cdot} x_0$, $s, s_\ell \in \mathcal{I}_\ell$, and $\ell = 1, \ldots, L$ to obtain

$$\begin{aligned}
&\Big|\frac{1}{|\mathcal{I}|}\int_I f(a_t u_{\gamma(s)} x_0)\,\mathrm{d}s - \int_{\mathsf{X}_{d+1}} f\,\mathrm{d}m_{\mathsf{X}_{d+1}}\Big| \\
&\qquad = \left|\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{|\mathcal{I}_\ell|}\int_{\mathcal{I}_\ell} f(g_s^{-1} g_s a_t u_{\gamma(s)}\boldsymbol{\cdot} x_0)\,\mathrm{d}s - \int_{\mathsf{X}_{d+1}} f\,\mathrm{d}m_{\mathsf{X}_{d+1}}\right| \\
&\qquad \leqslant \left|\frac{1}{L}\sum_{\ell=1}^{L}\left(\frac{1}{|\mathcal{I}_\ell|}\int_{\mathcal{I}_\ell} f(g_{s_\ell}^{-1} g_s a_t u_{\gamma(s)}\boldsymbol{\cdot} x_0)\,\mathrm{d}s - \int_{\mathsf{X}_{d+1}} f\,\mathrm{d}m_{\mathsf{X}_{d+1}}\right)\right| + \varepsilon \\
&\qquad \leqslant 2\varepsilon
\end{aligned}$$

for all sufficiently large $t > 0$. As $f \in C_c(\mathsf{X}_{d+1})$ and $\varepsilon > 0$ were arbitrary, we obtain the lemma. □

6.1.5 Totally Geodesic Submanifolds in Hyperbolic Quotients

(to come)

6.1.6 Integer Points on Spheres and their Orthogonal Complement

(to come)

6.2 The Main Theorems

We let $X = G/\Gamma$, where G is a connected Lie group and $\Gamma < G$ a lattice. Let

$$U = \{u_s \mid s \in \mathbb{R}\} < G$$

be a one-parameter unipotent subgroup of G. Then the U-invariant probability measures on X can be completely classified. This was conjectured by Dani (in [19, Conjecture I], as an analogue of Raghunathan's conjecture, which will be described below) and proved by Ratner [133], [134], [135]. As it turned out, this was a powerful starting point for the other results that follow.

The classification results will generally take the form of asserting that an initially unknown measure has 'algebraic structure'.

A probability measure μ on G/Γ is called *algebraic*, *homogeneous*, or *periodic* if there exists a closed connected unimodular subgroup L with $U \leqslant L \leqslant G$ such that μ is the L-invariant normalized probability measure on a closed finite volume orbit $L{\scriptstyle\bullet}x_0$ (for any $x_0 \in \operatorname{supp}\mu$).

Theorem 6.11 (Dani's conjecture and Ratner's measure classification). *If $X = G/\Gamma$ and $U = \{u_s \mid s \in \mathbb{R}\} < G$ is a one-parameter unipotent subgroup, then every U-invariant ergodic probability measure μ on X is algebraic.*

In this result (unlike the following ones), it is sufficient to assume that Γ is discrete or even just closed. Theorem 6.11, the theorem of Dani and Smillie [26], its generalization from Section 5.3.2, and the general non-divergence property of unipotent orbits in Chaper 4, suggest other results. Ratner [136] generalized all of these results in the following theorem.

Theorem 6.12 (Ratner's equidistribution theorem). *Let $X = G/\Gamma$ where Γ is a lattice, and let $U = \{u_s \mid s \in \mathbb{R}\} < G$ be a one-parameter unipotent subgroup. Then for any $x_0 \in X$ there exists some closed connected unimodular subgroup $L \leqslant G$ such that $U \leqslant L$,*

- *$L{\scriptstyle\bullet}x_0$ is closed with finite L-invariant volume, and*

- $\displaystyle\frac{1}{T}\int_0^T f(u_s\bullet x_0)\,\mathrm{d}s \longrightarrow \frac{1}{\mathrm{vol}(L\bullet x_0)}\int_{L\bullet x_0} f\,\mathrm{d}m_{L\bullet x_0}$ *as* $T\to\infty$.

It is interesting to note that Theorem 6.12 in particular implies that any point $x \in X$ returns close to itself under a unipotent flow. That is, for any one-parameter unipotent subgroup $\{u_s \mid s \in \mathbb{R}\}$ and any $x \in X$ there is a sequence $(t_k)_{k\geqslant 1}$ for which $t_k \to \infty$ and $\mathsf{d}(x, u_{t_k}\bullet x) \to 0$ as $k \to \infty$. This close return statement is of course incomparably weaker than Ratner's equidistribution theorem, but even this weak statement does not seem to have an independent proof to our knowledge.

Theorem 6.12 also suggests that the closures of orbits under the action of a unipotent one-parameter subgroup should have some algebraic structure. A more general version of that statement is the famous conjecture of Raghunathan(32) that motivated all of the theorems above, and was proved by Ratner [135] using the above results as stepping stones. The orbit closure $\overline{H\bullet x_0}$ of a point x_0 in G/Γ under the action of a closed subgroup H is similarly called *algebraic*, *homogeneous*, or *periodic* if there exists some closed connected unimodular subgroup L with $H \leqslant L \leqslant G$ such that $\overline{H\bullet x_0} = L\bullet x_0$, and $L\bullet x_0$ supports a finite L-invariant measure.

Theorem 6.13 (Raghunathan's conjecture; Ratner's orbit closure theorem). *Suppose that $X = G/\Gamma$, with G a connected Lie group and Γ a lattice. Let $H < G$ be a closed subgroup generated by one-parameter unipotent subgroups. Then the orbit closure of any $x_0 \in X$ is algebraic.*

It is also interesting to ask what the structure of the set of all probability measures that are invariant and ergodic under some unipotent flow really is. This generalizes the theorem of Sarnak (Theorem 5.8) concerning periodic horocycle orbits. At first sight, one might only ask this out of curiosity or to satisfy the urge to complete our understanding of this aspect of these dynamical systems. However, this line of enquiry turns out to be useful for applications to number-theoretic problems. A satisfying answer to this question is given by Mozes and Shah [122].

Theorem 6.14 (Mozes–Shah equidistribution theorem). † *Let X be the homogeneous space G/Γ with G a connected Lie group and Γ a lattice, and let (H_n) be a sequence of subgroups of G generated by unipotent one-parameter subgroups. Let μ_n be an invariant ergodic probability measure for the action of H_n for all $n \geqslant 1$. Assume that*‡ *$\mu_n \to \mu$ in the weak*-topology as $n \to \infty$. Then either $\mu = 0$ or μ is an algebraic measure, where in each case more can be said.*

If $\mu = 0$ then $\operatorname{supp}\mu_n \to \infty$ as $n \to \infty$ in the sense that for every compact set $K \subseteq X$ there is an N with $\operatorname{supp}\mu_n \cap K = \varnothing$ for $n \geqslant N$.

† This version differs from the theorem in the paper, but should follow from it. Awaiting a decision: Will it be proven here from scratch or using their theorem?

‡ By Tychonoff-Alaoglu there always exists a subsequence that converges.

Otherwise $\mu = m_{L\cdot y}$ *is the* L*-invariant probability measure on a closed finite volume orbit* $L\cdot y$ *for the closed connected group* $L = \mathrm{Stab}_G(\mu)^o \leqslant G$*. Moreover,* μ *is invariant and ergodic for the action of a one-parameter unipotent subgroup. Furthermore, suppose that* $x_n = \varepsilon_n \cdot x \in \operatorname{supp}\mu_n$ *for* $n \geqslant 1$ *and some* $x \in X$ *with* $\varepsilon_n \to I$ *as* $n \to \infty$*, and suppose the connected subgroups* (L_n) *satisfy* $\mu_n = m_{L_n\cdot x_n}$ *for* $n \geqslant 1$*. Then* $Lx = Ly = \operatorname{supp}\mu$ *and there exists some* N *with* $\varepsilon_n^{-1} L_n \varepsilon_n \subseteq L$ *for* $n \geqslant N$*.*

The additional information in each case is useful in applying this theorem. According to (1), once we know that for every measure μ_n there exists some point $x_n \in \operatorname{supp}\mu_n$ within a fixed compact set, the limit measure is a probability measure.

In (2), if we know that $H_n = H$ for all $n \geqslant 1$, then L has to contain H and the conjugates $\varepsilon_n^{-1} H \varepsilon_n$ as in (2). Together this often puts severe limitations on the possibilities that $L \leqslant G$ can take, and sometimes forces L to be G. This situtation arises, for example, if we study long periodic horocycle orbits, or orbits of a maximal subgroup $H < G$. In any case, the final claim of (2) says that the convergence to the limit measure $m_{L\cdot x}$ is almost from within the orbit $L\cdot x$. In fact, after modifying the measures in the sequence only slightly by the elements ε_n we get

$$\operatorname{supp}\left((\varepsilon_n)_*^{-1}\mu_n\right) = \varepsilon_n^{-1} L_n \cdot x_n = \varepsilon_n^{-1} L_n \varepsilon_n \cdot x \subseteq L\cdot x = L\cdot y = \operatorname{supp}\mu$$

for $n \geqslant N$.

We will prove special cases of the theorems above.

6.2.1 Rationality Questions

A natural question is to ask which subgroups $L < G$ appear for a certain choice of one-parameter unipotent subgroup $U < G$ and $x \in X = G/\Gamma$. In this section we explain how this kind of question is intimately related to questions of rationality.

This relationship is elementary in the abelian setting of $G = \mathbb{R}^d$, $\Gamma = \mathbb{Z}^d$, and $U = \mathbb{R}v$ for some $v \in \mathbb{R}^d$. In this case L is independent of

$$x \in X = \mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$$

(and one should only expect this independence for abelian Lie groups). Moreover, L is the smallest subspace of $\mathbb{R}^d$ that can be defined by rational linear equations and contains $U = \mathbb{R}v$. This claim follows quickly from the special case where no such $L \neq \mathbb{R}^d$ exists. Under this assumption, $\{tv \mid t \in \mathbb{R}\}$ is equidistributed, as may be shown for example by integrating the characters of $\mathbb{T}^d$.

To start to see the possibilities in the general case, consider the special case

$$U = \left\{ \begin{pmatrix} 1 & s \\ & 1 \end{pmatrix} \,\middle|\, s \in \mathbb{R} \right\} < \mathrm{SL}_2(\mathbb{R})$$

and $\mathsf{X}_2 = \mathrm{SL}_2(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$, which we already understand in some detail (see Section 1.2, Chapter 5, and [45, Sec. 11.7]). If $x = g\Gamma$ for some

$$g = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix},$$

then $L = U$, and otherwise $L = \mathrm{SL}_2(\mathbb{R})$. In order to be able to phrase this in terms of a rationality question, notice that $x \in X$ determines a geodesic in the upper half-plane (where we choose for example the base point in our fundamental domain, as illustrated in Figure 6.1). Then $L = U$ if the forward end point of the geodesic $\alpha \in \mathbb{R} \cup \{\infty\}$ is rational, meaning $\alpha \in \mathbb{Q} \cup \{\infty\}$, and $L = \mathrm{SL}_2(\mathbb{R})$ otherwise. This dichotomy is independent of the chosen representative within the orbit $\mathrm{SL}_2(\mathbb{Z}) \cdot (z, v)$.

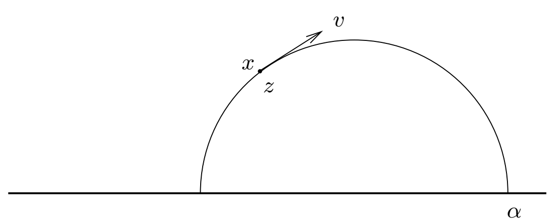

Fig. 6.1: The geodesic determined by x.

In general the answer is given by the following result found by Borel and Prasad [10]. A more general version of this result was obtained more recently by Tomanov [160].

Theorem 6.15. *Let $\mathsf{X}_d = \mathrm{SL}_d(\mathbb{R})/\mathrm{SL}_d(\mathbb{Z})$, $x = g\Gamma \in X$, and $U < G$ a one-parameter unipotent subgroup (or $H < G$ a closed subgroup generated by one-parameter unipotent subgroups). Then the group L appearing in Theorems* 6.11 *and* 6.12 *(respectively Theorem* 6.13*) is the connected component of $g\mathbb{F}(\mathbb{R})g^{-1}$, where $\mathbb{F}(\mathbb{R})$ is the group of $\mathbb{R}$-points of the smallest algebraic group $\mathbb{F} \leqslant \mathrm{SL}_d$ defined over $\mathbb{Q}$ for which $g\mathbb{F}(\mathbb{R})g^{-1}$ contains U (respectively H).*

Similarly, the group L in Theorem 6.14 *is the connected component of $g\mathbb{F}(\mathbb{R})g^{-1}$ where $x = g\Gamma$ and $\mathbb{F}$ is the smallest algebraic group $\mathbb{F} \leqslant \mathrm{SL}_d$ defined over $\mathbb{Q}$ for which $g\mathbb{F}(\mathbb{R})g^{-1}$ contains $\varepsilon_n L_n \varepsilon_n^{-1}$ for $n \geqslant N$, where N is as in Theorem* 6.14.

For this result, one needs some understanding of the mechanisms that make orbits $\mathbb{F}(\mathbb{R})\,\mathrm{SL}_d(\mathbb{Z})$ of $\mathbb{Q}$-groups closed or not closed, and the Borel density theorem. In the setting of $\Gamma = \mathrm{SL}_d(\mathbb{Z}) < G = \mathrm{SL}_d(\mathbb{R})$, which contains *all* other arithmetic quotients even over number fields if we allow d to vary, the connection to algebraic group theory described above puts additional constraints on the possible structure of the subgroup L.

For instance, the algebraic group $\mathbb{F}$ over $\mathbb{Q}$ must have the property that the radical of $\mathbb{F}$ is equal to the unipotent radical of $\mathbb{F}$. In the language of Lie groups this implies that the radical of L, which by definition is only solvable, is nilpotent. Another restriction is, for example, that L cannot be isomorphic to $\mathrm{PSL}_2(\mathbb{R}) \times \mathrm{SO}_5(\mathbb{R})$. This is because the unipotent group has to be contained in $\mathrm{PSL}_2(\mathbb{R})$ and the induced lattice $L \cap g^{-1} \mathrm{SL}_d(\mathbb{Z}) g$ cannot give an irreducible lattice in $\mathrm{PSL}_2(\mathbb{R}) \times \mathrm{SO}_5(\mathbb{R})$ as the direct factors are simple groups of different types in the classification of complex Lie algebras and they cannot be exchanged by a Galois action. On the other hand

$$L = \mathrm{PSL}_2(\mathbb{R}) \times \mathrm{SO}_3(\mathbb{R})$$

is a possibility since $\mathrm{PSL}_2(\mathbb{R}) \cong \mathrm{SO}_{2,1}(\mathbb{R})^o$, and a simple switch in the sign of the quadratic forms (via a Galois automorphism) can interchange these groups. We will discuss this and the required language further in Section 9.7.

6.3 First Ideas in Unipotent Dynamics

The structure of proof of Theorem 6.11 is to study

$$\mathrm{Stab}_G(\mu) = \{g \in G \mid g_*\mu = \mu\}$$

and to show that the measure μ on $X = \Gamma\backslash G$ is supported on a single orbit of this subgroup. This is achieved indirectly; if μ is not supported on a single orbit of a particular subgroup $H < G$ that leaves the measure invariant then one shows that the subgroup can be enlarged to some $H' > H$ so that the new subgroup H' also preserves μ.

We also note that, in the setting of Theorem 6.11, once we have shown that μ gives a single orbit of $\mathrm{Stab}_G(\mu)$ positive measure, we actually obtain that μ is supported on a single closed orbit of $\mathrm{Stab}_G(\mu)^o$.

Lemma 6.16. *Let $X = G/\Gamma$ be a quotient of a Lie group by a discrete subgroup Γ. Let H be a connected subgroup of G and let μ be an H-invariant and ergodic probability measure. If μ gives positive measure to a single orbit of its stabilizer subgroup $\mathrm{Stab}_G(\mu)$, then μ is the Haar measure on a closed orbit of the subgroup $\mathrm{Stab}_G(\mu)^o$.*

Proof. If $\mu\big(\mathrm{Stab}_G(\mu)\cdot x_0\big) > 0$ for some $x_0 \in X$, then $\mu\big(\mathrm{Stab}_G(\mu)\cdot x_0\big) = 1$ by ergodicity. As the index of $\mathrm{Stab}_G(\mu)^o$ in $\mathrm{Stab}_G(\mu)$ is at most countable, there exists a point x_1 so that $\mu\big(\mathrm{Stab}_G(\mu)^o\cdot x_1\big) > 0$. This implies once more that $\mu\big(\mathrm{Stab}_G(\mu)^o\cdot x_1\big) = 1$ as H is assumed to be connected. It follows that μ is the Haar measure on $\mathrm{Stab}_G(\mu)^o\cdot x_1$, which is also closed by Corollary 1.36. □

6.3.1 Generic Points

We present in this section the basic idea for using generic points to show an 'additional invariance', which in a more specialized context goes back to work of Furstenberg on the unique ergodicity of skew product extensions, leading to the equidistribution of the fractional parts of the sequence $(n^2\alpha)_{n\geqslant 1}$ for α irrational.(33)

Recall that $x \in X$ is said to be *generic* with respect to μ and a one-parameter flow $\{u_s \mid s \in \mathbb{R}\}$ if

$$\frac{1}{T}\int_0^T f(u_s{\bullet}x)\,\mathrm{d}s \longrightarrow \int_X f\,\mathrm{d}\mu$$

as $T \to \infty$ for all $f \in C_c(X)$. Using the pointwise ergodic theorem [45, Cor. 8.15] and separability of $C_0(X)$ one can easily show that μ-almost every point is generic if only μ is invariant and ergodic under the one-parameter flow

$$U = \{u_s \mid s \in \mathbb{R}\}$$

(see Lemma 6.20).

Lemma 6.17 (Centralizer Lemma). *If $x, y = h{\bullet}x \in X$ are generic for μ and $h \in C_G(U) = \{g \in G \mid gu = ug \text{ for all } u \in U\}$, then h preserves μ.*

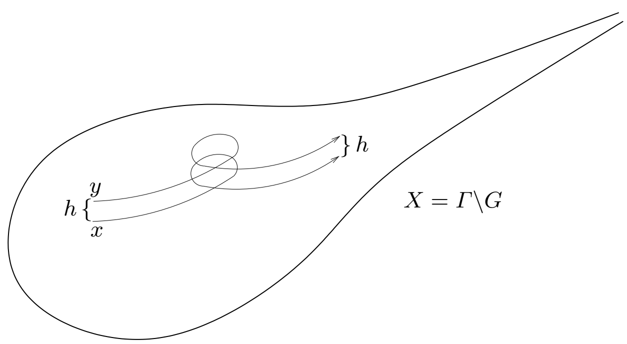

Fig. 6.2: If $y = xh^{-1}$ with $h \in C_G(V)$, then the two orbits are parallel. If in addition both x and y are generic, then the orbits equidistribute (that is, approximate μ), which gives Lemma 6.17.

Proof of Lemma 6.17. We refer to Figure 6.2 for a depiction of the proof. We know that

$$\frac{1}{T}\int_0^T f(u_s{\bullet}y)\,\mathrm{d}s \longrightarrow \int_X f\,\mathrm{d}\mu$$

for any $f \in C_c(X)$. On the other hand

$$
\begin{aligned}
\frac{1}{T}\int_0^T f(u_s \centerdot y)\,\mathrm{d}s &= \frac{1}{T}\int_0^T f(u_s \centerdot (h \centerdot x))\,\mathrm{d}s \\
&= \frac{1}{T}\int_0^T \underbrace{f(h \centerdot (u_s \centerdot x))}_{f^h(u_s \centerdot x)}\,\mathrm{d}s \qquad (\text{since } h \in C_G(\{u_s\})) \\
&\longrightarrow \int_X f^h\,\mathrm{d}\mu = \int_X f(h \centerdot z)\,\mathrm{d}\mu
\end{aligned}
$$

so μ is h-invariant. □

Lemma 6.17 seems (and is) useful, but it can only be applied in very special circumstances as the centralizer is usually very small, and we would need to be extremely fortunate to find two generic points bearing such a special relation to each other.

6.3.2 Polynomial divergence leading to invariance

A much more useful observation, due to Ratner, that leads to additional invariance in more circumstances, is the following observation[†] which is based on the polynomial divergence property of unipotent flows. In fact, as we have seen before, the action of an element $u \in G$ on $\Gamma\backslash G$ is locally described by conjugation and hence can also be described by the adjoint representation of u on the Lie algebra $\mathfrak{g}$ of G. More precisely, if $y = \varepsilon \centerdot x$ is close to x, and $\varepsilon \in G$ is the local displacement between x and y, then $u \centerdot y = u \centerdot \varepsilon \centerdot x = u\varepsilon u^{-1} \centerdot (u \centerdot x)$ and so a displacement between $u \centerdot x$ and $u \centerdot y$ is given by the conjugated element $u\varepsilon u^{-1}$. If the displacement ε was not small enough, then $u\varepsilon u^{-1}$ may not be the smallest displacement between $u \centerdot x$ and $u \centerdot y$. However, if ε is very small, then the calculation leading to the conjugated element as the displacement may be iterated several times. Thus, in order to compare the orbit of points close to x to the orbit of x we will need to study conjugation by u (or equivalently its adjoint representation on the Lie algebra).

If $\{u(t) \mid t \in \mathbb{R}\}$ is a unipotent one-parameter subgroup of G, then $\mathrm{Ad}_{u(t)}$ is unipotent for all $t \in \mathbb{R}$ also, and is a (matrix-valued) polynomial in t. This polynomial structure (as opposed to exponential) of unipotent subgroups has the following consequence. Given a nearby pair of points x and $y = \varepsilon \centerdot x$, let $v = \log \varepsilon$ and consider the $\mathfrak{g}$-valued polynomial $\mathrm{Ad}_{u(t)}(v)$. For very small values of ε, this polynomial is close to zero in the space of all polynomials. However, if we choose a large 'speeding up' parameter T then we may consider the polynomial

$$p(r) = \mathrm{Ad}_{u(rT)}(v)$$

[†] This is often called the H-principle. Our presentation of the idea will be closer to the work of Margulis and Tomanov [111].

in the rescaled variable $r \in \mathbb{R}$. Assuming the original polynomial is non-constant (equivalently, ε does not lie in $C_G(\{u(s)\})$), we can choose T precisely so that the polynomial p above in the variable r belongs to a compact set of polynomials not containing the zero polynomial. In fact, if $T > 0$ is the smallest number with[†] $\|\operatorname{Ad}_{u(T)}(v)\| = 1$, then

$$\sup_{r\in[0,1]} \|p(r)\| = 1.$$

Moreover, p is a polynomial of bounded degree. Notice that this feature—that this acceleration or renormalization of a polynomial is again a polynomial from the same finite-dimensional space—is specific[‡] to polynomials and hence to unipotent flows.

In order to state the principle that gives additional invariance, we will need the following refinement of the notion of genericity.

Definition 6.18. A set $K \subseteq X$ is called a set of *uniformly generic points* if for any $f \in C_c(X)$ and $\varepsilon > 0$ there is some $T_0 = T_0(f, \varepsilon)$ with

$$\left| \frac{1}{T} \int_0^T f(u_s \bullet x)\, \mathrm{d}s - \int_X f \,\mathrm{d}\mu \right| < \varepsilon$$

for all $T \geqslant T_0$ and all $x \in K$.

Proposition 6.19 (Polynomial divergence leads to invariance). *Suppose that $(x_n), (y_n)$ are sequences of uniformly generic points with $y_n = \varepsilon_n \bullet x_n$ for all $n \geqslant 1$ where $\varepsilon_n \to I$ as $n \to \infty$ and $\varepsilon_n \notin C_G(U)$ for $n \geqslant 1$. Define $v_n = \log \varepsilon_n$ and polynomials*

$$p_n(r) = \operatorname{Ad}_{u(T_n r)}(v_n),$$

where the speeding up parameter $T_n \to \infty$ is chosen so that

$$\sup_{r\in[0,1]} \|p_n(r)\| = 1$$

for each $n \geqslant 1$. Suppose that $p_n(r) \to p(r)$ as $n \to \infty$ for all $r \in [0, 1]$, where

$$p\colon \mathbb{R} \longrightarrow \mathfrak{g}$$

is a polynomial with entries in the Lie algebra $\mathfrak{g}$. Then μ is invariant under $\exp(p(r))$ for all $r \in \mathbb{R}_{\geqslant 0}$.

Notice that the assumption that the sequence of polynomials converges is a mild one. The polynomials all lie in a compact subset of a finite-dimensional

[†] It does not matter which norm on $\mathfrak{g}$ is used; for concreteness we use the norm derived from the Riemannian metric.

[‡] In contrast, diagonalizable flows leading in the same way to exponential maps do not have this property, as the acceleration would change the base of the exponential functions involved.

space, so there is a subsequence that converges with respect to any norm on that space. Also the assumption $\varepsilon_n \notin C_G(U)$ is somewhat unproblematic as in the case $\varepsilon_n \in C_G(U)$ one may be able to apply Lemma 6.17. Part of the argument for Proposition 6.19 is illustrated in Figure 6.3.

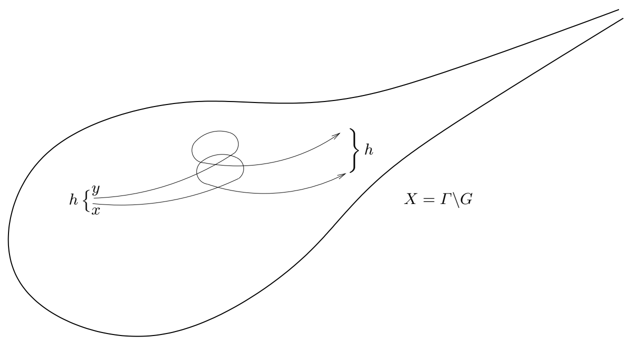

Fig. 6.3: If $y = x\varepsilon^{-1}$ with $\varepsilon \notin C_G(V)$ close to the identity, then the orbits of x and y move away from each other at polynomial speed. If x and y are generic then the last 1% of these pieces of orbits are almost parallel and equidistribute.

PROOF OF PROPOSITION 6.19. Fix $r_0 \in \mathbb{R}_{>0}$, $f \in C_c(X)$, and $\varepsilon > 0$. By uniform continuity of f there exists some $\delta = \delta(f, \varepsilon) > 0$ with

$$\mathsf{d}(h_1, h_2) < \delta \implies |f(h_1 \bullet x) - f(h_2 \bullet x)| < \varepsilon$$

for all $x \in X$. Furthermore, choose $\kappa > 0$ so that

$$\mathsf{d}(\exp p(r), \exp p(r_0)) < \delta/2$$

for $r \in [r_0 - \kappa, r_0]$. Then there is an N such that we also have[†]

$$\mathsf{d}(\exp p_n(r), \exp p(r_0)) < \delta \tag{6.6}$$

for $n \geqslant N$ and $r \in [r_0 - \kappa, r_0]$. We know by the uniform genericity of x_n that

$$\frac{1}{r_0 T_n} \int_0^{r_0 T_n} f(u_s \bullet x_n)\,\mathrm{d}s \longrightarrow \int_X f\,\mathrm{d}\mu$$

as $n \to \infty$, and

$$\frac{1}{(r_0 - \kappa) T_n} \int_0^{(r_0 - \kappa) T_n} f(u_s \bullet x_n)\,\mathrm{d}s \longrightarrow \int_X f\,\mathrm{d}\mu$$

[†] This is the formal version of the statement in Figure 6.3 that the last 1% are parallel.

as $n\to\infty$. Taking the correct linear combination ($\kappa>0$ is fixed) and replacing f by $f^{\exp(p(r_0))}$, we get[†]

$$\frac{1}{\kappa T_n}\int_{(r_0-\kappa)T_n}^{r_0T_n} f^{\exp p(r_0)}(u_s\boldsymbol{\cdot}x_n)\,\mathrm{d}s\longrightarrow\int_X f^{\exp p(r_0)}\,\mathrm{d}\mu$$

as $n\to\infty$ and, by the same argument, we also have

$$\frac{1}{\kappa T_n}\int_{(r_0-\kappa)T_n}^{r_0T_n} f(u_s\boldsymbol{\cdot}y_n)\,\mathrm{d}s\longrightarrow\int_X f\,\mathrm{d}\mu$$

as $n\to\infty$. However, using the definition of v_n and p_n we have

$$u_s\boldsymbol{\cdot}y_n=u_s\exp(v_n)\boldsymbol{\cdot}x_n=\exp\left(\mathrm{Ad}_{u_s}(v_n)\right)u_s\boldsymbol{\cdot}x_n=\exp\left(p_n(s/T_n)\right)u_s\boldsymbol{\cdot}x_n$$

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

We now restrict ourself to the range of $s\in\mathbb{R}$ with $\frac{s}{T_n}\in[r_0-\kappa,r_0]$. Together with (6.6), we deduce that

$$\mathsf{d}(u_s\boldsymbol{\cdot}y_n,\exp p(r_0)u_s\boldsymbol{\cdot}x_n)<\delta,$$

and so

$$|f(u_s\boldsymbol{\cdot}y_n)-f(\exp p(r_0)u_s\boldsymbol{\cdot}x_n)|<\varepsilon$$

for every $s\in[(r_0-\kappa)T_n,r_0T_n]$. Using this estimate in the integrals above gives

$$\left|\frac{1}{\kappa T_n}\int_{(r_0-\kappa)T_n}^{r_0T_n} f^{\exp p(r_0)}(u_s\boldsymbol{\cdot}x_n)\,\mathrm{d}s-\frac{1}{\kappa T_n}\int_{(r_0-\kappa)T_n}^{r_0T_n} f(u_s\boldsymbol{\cdot}y_n)\,\mathrm{d}s\right|<\varepsilon,$$

and so

$$\left|\int_X f^{\exp p(r_0)}\,\mathrm{d}\mu-\int_X f\,\mathrm{d}\mu\right|\leqslant\varepsilon.$$

Since this holds for any $\varepsilon>0$ and $f\in C_c(X)$ we deduce that μ is invariant under $\exp p(r_0)$. As $r_0>0$ was arbitrary, the proposition follows. □

Because of the results above, we are interested in finding large sets of uniformly generic points. It is too much to expect that almost every point with respect to an invariant measure will have this property (due to the requested uniformity), but we can get close to this statement as follows.

Lemma 6.20 (Almost full measure sets consisting of uniformly generic points). *Let μ be an invariant and ergodic probability measure on X for the action of a one-parameter flow $\{u_s\mid s\in\mathbb{R}\}$. For any $\rho>0$ there is a compact set $K\subseteq X$ with $\mu(K)>1-\rho$ consisting of uniformly generic points.*

PROOF. Let $D=\{f_1,f_2,\dots\}\subseteq C_c(X)$ be countable and dense. Then by the pointwise ergodic theorem [45, Cor. 8.15] for every $f_\ell\in D$ we have

[†] In Figure 6.3 we referred to this as the equidistribution of the last 1% of the orbit.

$$\frac{1}{T}\int_0^T f_\ell(u_s\bullet x)\,\mathrm{d}s \longrightarrow \int_X f_\ell\,\mathrm{d}\mu$$

almost everywhere with respect to μ as $T \to \infty$, equivalently for every $\varepsilon > 0$

$$\mu\left(\left\{x \in X \;\middle|\; \sup_{T>T_0}\left|\frac{1}{T}\int_0^T f_\ell(u_s\bullet x)\,\mathrm{d}s - \int_X f_\ell\,\mathrm{d}\mu\right| > \varepsilon\right\}\right) \longrightarrow 0$$

as $T_0 \to \infty$. Now choose, for every $f_\ell \in D$ and for every $\varepsilon = \frac{1}{n}$, a time $T_{\ell,n}$ so that

$$\mu\left(\left\{x \in X \;\middle|\; \sup_{T>T_{\ell,n}}\left|\frac{1}{T}\int_0^T f_\ell(u_s\bullet x)\,\mathrm{d}s - \int_X f_\ell\,\mathrm{d}\mu\right| > \frac{1}{n}\right\}\right) < \frac{\rho}{2^{\ell+n}}.$$

Let $K' \subseteq X$ be the complement of the union of these sets, so that

$$\mu(K') > 1 - \rho$$

by construction. It is clear that the points in K' are uniformly generic for all function $f \in D$. Moreover, since $D \subseteq C_c(X)$ is dense in the uniform norm, this extends to all functions by a simple approximation argument. Finally we may choose a compact $K \subseteq K'$ with $\mu(K) > 1 - \rho$ by regularity of μ. □

The principle outlined above is sufficient to prove the measure classification theorem for 2-step nilpotent groups (see Exercise 6.22; as we will see in the next section with more effort the same holds for more general nilpotent groups). However, in general this use is limited—for example, in the above form it does not even allow us to give a new proof of measure classification for the horocycle flow. This will be discussed again in Section 6.6, where we discuss the second, more powerful, refinement of the use of generic points to show additional invariance. This will lead to a strengthening of Dani's theorem (Theorem 5.7), due to Ratner, and is the key to the general case.

Exercise 6.21. Show that the limit polynomial in Proposition 6.19 takes only values in the centralizer $C_{\mathfrak{g}}(U) = \{v \in \mathfrak{g} \mid \mathrm{Ad}_u(v) = v \text{ for all } u \in U\}$ of U in the Lie algebra $\mathfrak{g}$ of G.

Exercise 6.22. Use the results from Section 6.3.2 to prove the measure classification theorem (Theorem 6.11) under the assumption that G is a 2-step nilpotent group.

6.4 Unipotent Dynamics on Nilmanifolds

In this section we will assume that G is a nilpotent Lie group and $\Gamma < G$ a discrete subgroup. In this case $X = G/\Gamma$ is called a *nilmanifold*.

6.4.1 Measure Classification for Nilmanifolds

Theorem 6.23. *Let $\Gamma < G$ be a discrete subgroup of a connected nilpotent Lie group G and let $X = G/\Gamma$. Let $U \leqslant G$ be a one-parameter subgroup. Then any U-invariant and ergodic probability measure μ on G is algebraic.*

Proof. As we will see, the result follows from a (double) induction argument and Proposition 6.19. First, notice that the theorem is trivial if $\dim G = 1$.

A second special case is obtained by assuming in addition that U belongs to the centre C_G of G. In this case, if $X' = \{x \in X \mid x \text{ is generic for } \mu\}$, $x_0 \in X'$, and $y = g{\boldsymbol\cdot}x_0 \in X'$, then $g \in C_G(U) = G$, so $g \in \operatorname{Stab}_G(\mu)$ by Lemma 6.17 and $y \in \operatorname{Stab}_G(\mu){\boldsymbol\cdot}x_0$ also. It follows that $X' \subseteq \operatorname{Stab}_G(\mu){\boldsymbol\cdot}x_0$ has full measure, and we deduce that μ must be the Haar measure on $\operatorname{Stab}_G(\mu){\boldsymbol\cdot}x_0$ as required.

We assume now that G is a nilpotent connected Lie group of nilpotency degree k, meaning that

$$G_0 = G \geqslant G_1 = [G, G_0] \geqslant \cdots \geqslant G_{k-1} = [G, G_{k-2}] \geqslant G_k = [G, G_{k-1}] = \{I\}.$$

We also assume that $U \leqslant G_j$ for some $j \in \{0, \dots, k-1\}$. We may also assume that $U \nsubseteq C_G$. The inductive hypothesis is then the following statement: The theorem holds for any $X' = \Gamma' \backslash G'$, $U' \leqslant G'$ and any U'-invariant and ergodic probability measure μ' if either

- $\dim G' < \dim G$, or
- $G' = G$, $\Gamma' = \Gamma$, and $U' \leqslant G_{j+1}$.

Now let $K \subseteq X$ be a set of uniformly generic points of measure $\mu(K) > 0.9$ as in Lemma 6.20. Choose some

$$x_0 \in K \cap \operatorname{supp}(\mu|_K). \tag{6.7}$$

We distinguish between two possible scenarios.

It could be that there is some $\delta > 0$ such that $y = h{\boldsymbol\cdot}x_0 \in K$ with $\mathsf{d}(h, I) < \delta$ implies that $h \in C_G(U)$ and so also $h \in \operatorname{Stab}_G(\mu)$ by Lemma 6.17. In this case (6.7) implies that $\operatorname{Stab}_G(\mu){\boldsymbol\cdot}x_0$ has positive measure, and so we may apply Lemma 6.16 to conclude.

In the second case we find a sequence $(y_n = \varepsilon_n{\boldsymbol\cdot}x_0)$ in K with $\varepsilon_n \to I$ as $n \to \infty$ but $\varepsilon_n \notin C_G(U)$ for all $n \geqslant 1$. Choosing a subsequence, we may assume that the sequence of polynomials $(p_n(r))$ from Proposition 6.19 converges to a non-constant polynomial $p\colon \mathbb{R} \to \mathfrak{g}$. By Proposition 6.19 we deduce that μ is invariant under $\exp(p(r))$ for all $r \geqslant 0$.

We claim that $\exp(p(r))$ takes values in G_{j+1}. Indeed, since $U \subseteq G_j$ we have (in the notation of Proposition 6.19)

$$p_n(r) = \operatorname{Ad}_{u(T_n r)}(\log \varepsilon_n) \in \log \varepsilon_n + \mathfrak{g}_{j+1}$$

for all r, where

$$\mathfrak{g}_{j+1} = \operatorname{Lie}(G_{j+1}) = [\mathfrak{g}, \mathfrak{g}_j].$$

Since $\varepsilon_n \to I$ as $n \to \infty$ this gives $p(r) \in \mathfrak{g}_{j+1}$ for all $r \geqslant 0$ as claimed.

The argument above shows that

$$\left(\mathrm{Stab}_G(\mu) \cap G_{j+1}\right)^o$$

is a nontrivial subgroup. Clearly U normalizes this subgroup and its Lie algebra, and since $\mathrm{Ad}_{u(t)}$ is unipotent for all $t \in \mathbb{R}$, it follows that there exists a one-parameter unipotent subgroup

$$U' = \{u'_t \mid t \in \mathbb{R}\} \leqslant \mathrm{Stab}_G(\mu) \cap G_{j+1} \cap C_G(U).$$

We are going to apply the inductive hypothesis to $G' = G$, $\Gamma' = \Gamma$, and U'. However, as μ may not be[†] ergodic with respect to U' we first have to decompose μ into U'-ergodic components. Recall from [45, Th. 6.2, 8.20] that the ergodic decomposition allows us to write

$$\mu = \int_X \mu_x^{\mathcal{E}'} \,\mathrm{d}\mu, \tag{6.8}$$

where $\mu_x^{\mathcal{E}'}$ is the conditional measure for the σ-algebra

$$\mathcal{E}' = \{B \in \mathcal{B}_X \mid \mu(u'_t{\scriptstyle\bullet} B \triangle B) = 0 \text{ for all } t\}$$

and that for μ-almost every x the conditional measure $\mu_x^{\mathcal{E}'}$ is a U'-invariant and ergodic probability measure on X with $x \in \operatorname{supp} \mu_x^{\mathcal{E}'}$.

By applying the inductive hypothesis to μ-almost every $\mu_x^{\mathcal{E}'}$ we obtain a function $x \mapsto L_x$ that assigns to x the connected subgroup L_x for which $\mu_x^{\mathcal{E}'}$ is the L_x-invariant probability measure on the closed orbit $L_x{\scriptstyle\bullet} x$. We claim that there is a connected subgroup L such that $L_x = L$ for μ-almost every x. Indeed, since $U = \{u(t) \mid t \in \mathbb{R}\}$ preserves μ and leaves the σ-algebra $\mathcal{E}'$ invariant (since U' and U commute) we get

$$(u_t)_* \mu_x^{\mathcal{E}'} = \mu_{u_t{\scriptstyle\bullet} x}^{\mathcal{E}'} \tag{6.9}$$

for every $t \in \mathbb{R}$ and μ-almost every x by [45, Cor. 5.24]. Since $\mu_x^{\mathcal{E}'}$ is L_x-invariant, it follows from (6.9) that $(u_1)_* \mu_x^{\mathcal{E}'}$ is $u_1 L_x u_1^{-1}$-invariant, which implies that

$$u_1 L_x u_1^{-1} \subseteq L_{u(1){\scriptstyle\bullet} x}$$

and, by a similar argument for the reverse inclusion,

$$u_1 L_x u_1^{-1} = L_{u_1{\scriptstyle\bullet} x}.$$

Iterating this relationship shows that

[†] In fact U' never acts ergodically with respect to μ.

$$u_1^n L_x u_1^{-n} = L_{u(n)\bullet x} \tag{6.10}$$

for μ-almost every x. Now either L is normalized by u_1, or the sequence of subgroups in (6.10) converges to a subgroup that is normalized by u_1 (to see this, apply the argument from the proof of Lemma 3.51 to any element of $\bigwedge^{\dim L_x}(\mathrm{Lie}(L_x))$. Hence Poincaré recurrence shows that we must have

$$u_1 L_x u_1^{-1} = L_x$$

for μ-almost every x. Notice that for any such x we also get

$$u(t) L_x u(t)^{-1} = L_x$$

for all $t \in \mathbb{R}$. By ergodicity it follows that $L_x = L$ is constant μ-almost everywhere. The cautious reader will have noticed that the argument above has assumed implicitly that the function $x \mapsto L_x$ is measurable, which we will show in Lemma 6.24 below. Equation (6.8) now shows that μ is a convex combination of L-invariant measures and hence is itself L-invariant.

To summarize, we have shown that there exists a nontrivial connected subgroup $L \leqslant \mathrm{Stab}_G(\mu)$ containing U' such that the orbit $L\bullet x$ is for μ-almost every x closed, with finite L-invariant measure and with the property that $U' \leqslant L$ acts ergodically on $L\bullet x$. Since $L \leqslant G$ is nilpotent, simply connected and connected, $M = C_L(L)$ is a nontrivial connected subgroup. We claim that the orbit $M\bullet x$ is compact for μ-almost every x and postpone the proof to Lemma 6.26.

Next we claim that $N_G^1(M)\bullet x$ is a closed orbit for μ-almost every x, see Lemma 6.27. This implies that μ is supported on a single orbit $N_G^1(M)\bullet x_0$ of the unimodular normalizer. In fact we note first that

$$U \leqslant N_G(L) \leqslant N_G(M),$$

and since U is unipotent we also have $U \leqslant N_G^1(M)$. If now x_0 is generic for μ and U, then

$$\operatorname{supp}\mu = \overline{U\bullet x_0} \subseteq N_G^1(M)\bullet x_0.$$

Therefore, without loss of generality we may assume $x_0 = \Gamma$, $G = N_G^1(M)^o$ and hence $M \lhd G$ and that the orbit $M\bullet\Gamma$ is compact.

Let $\pi_M\colon G \to G/M$ denote the canonical projection $\pi_M(g) = gM$. We claim that $\pi_M(\Gamma) \leqslant G/M$ is again discrete. Suppose that

$$\pi_M(\gamma_n) \longrightarrow I$$

in G/M as $n \to \infty$ with $\gamma_n \in \Gamma$, or equivalently $\gamma_n m_n \to I$ as $n \to \infty$ in G for $\gamma_n \in \Gamma$ and $m_n \in M$ for all $n \geqslant 1$. Since $M \cap \Gamma$ is co-compact in M, we may simultaneously modify γ_n and m_n by elements of $M \cap \Gamma$ and assume that m_n lies in a pre-compact fundamental domain for Γ for all $n \geqslant 1$. Choosing a subsequence, we may also now assume that $m_n \to m \in M$ as $n \to \infty$. This implies that $\gamma_n \to \gamma \in \Gamma$ as $n \to \infty$ for some γ, and so $\gamma_n = \gamma$ for all large $n \geqslant 1$.

This shows that $\pi_M(\gamma_n) = \pi_M(\gamma) = I$ for large enough n, and hence that $\pi_M(\Gamma)$ is discrete.

There is also an associated factor map

$$\pi_X \colon G/\Gamma \longrightarrow \pi_M(G)/\pi_M(\Gamma)$$

defined by

$$\pi_X \colon g{\bullet}\Gamma \longmapsto \pi_M(g){\bullet}\pi_M(\Gamma).$$

The fibers of this map are precisely the M-orbits in the sense that

$$\pi_X^{-1}\left(\pi_X(g{\bullet}\Gamma)\right) = \{h{\bullet}\Gamma \mid hM{\bullet}\pi_M(\Gamma) = gM{\bullet}\pi_M(\Gamma)\} = gM{\bullet}\Gamma$$

for all $g \in G$.

We set $G' = \pi_M(G)$, $\Gamma' = \pi_M(\Gamma)$, $U' = \pi_M(U)$, $\mu' = (\pi_X)_*\mu$ and deduce from the inductive hypothesis that μ' is an algebraic measure. Let $H' \leqslant G'$ be a connected subgroup, so that μ' is the H'-invariant probability measure on a finite volume orbit $\pi_M(g)H'{\bullet}\pi_M(\Gamma)$ for some $\pi_M(g) \in G'$. Finally, we claim that μ is the H-invariant probability measure on the closed orbit $gH{\bullet}\Gamma$ where $H = \pi_M^{-1}(H')$.

Since $\pi_M(\Gamma)\pi_M(g)H'$ is closed we also obtain that

$$\Gamma g H = \pi_X^{-1}\left(H'\pi_M(g)\pi_M(\Gamma)\right)$$

is closed. Now let $f \in C(X)$. Then

$$\int_X f(x)\,\mathrm{d}\mu(x) = \int_X f(m{\bullet}x)\,\mathrm{d}\mu(x) \tag{6.11}$$

for all $m \in M$. Now take a Følner sequence (F_n) in M and notice that

$$\frac{1}{m_M(F_n)}\int_{F_n} f(m{\bullet}x)\,\mathrm{d}m_M(m) \longrightarrow \int_{M{\bullet}x} f(z)\,\mathrm{d}m_{M{\bullet}x}(z) = \overline{f}(\pi_X(x))$$

for all $x \in X$, where the expression on the right defines a function $\overline{f}$ in $C(\pi_X(X))$. Applying this convergence to the average of (6.11) over the Følner sequence gives

$$\int_X f(x)\,\mathrm{d}\mu(x) = \int_{\pi_X(X)} \underbrace{\int_{M{\bullet}x} f(z)\,\mathrm{d}m_{M{\bullet}x}(z)}_{\overline{f}(\pi_X(x))}\,\mathrm{d}\mu'$$

Now fix $h \in H$ and define f^h by $f^h(x) = f(h{\bullet}x)$ so that

$$\overline{f^h}(\pi_X(x)) = \int_{M{\bullet}x} f(h{\bullet}z)\,\mathrm{d}m_{M{\bullet}x}(z) = \int_{M{\bullet}(h{\bullet}x)} f(z)\,\mathrm{d}m_H(z) = \overline{f}(h{\bullet}\pi_X(x)),$$

and

$$\int_X f^h \,\mathrm{d}\mu = \int_{\pi_X(X)} \overline{f^h}\,\mathrm{d}\mu' = \int_{\pi_X(X)} \left(\overline{f}\right)^h \,\mathrm{d}\mu' = \int_{\pi_X(X)} \overline{f}\,\mathrm{d}\mu' = \int_X f\,\mathrm{d}\mu.$$

Therefore μ is supported on $H{\scriptstyle\bullet}x$ and is H-invariant. This concludes the induction, and the theorem follows. □

In the course of the proof we made use of several lemmas which we now prove.

Lemma 6.24 (Measurability of stabilizer). *Let G be a Lie group, $\Gamma \leqslant G$ a discrete subgroup, and let $X = \Gamma\backslash G$. Then the map*

$$\mathcal{M}(X) \ni \mu \longmapsto \mathrm{Stab}_G(\mu)^o$$

from the space $\mathcal{M}(X)$ of Borel probability measures on X is measurable.

Implicit in the statement of the lemma is a measurable structure on the space of connected subgroups, and this is achieved as follows. We identify a connected subgroup $L \leqslant G$ with its Lie algebra $\mathrm{Lie}(L)$, and if $L \neq \{I\}$ with the corresponding point of the Grassmannian of G. In other words, we consider the map in the lemma as a map from $\mathcal{M}(X)$ to

$$\{e\} \sqcup \bigsqcup_{\ell=1}^{\dim G} \mathrm{Grass}_\ell(\mathrm{Lie}(G)),$$

which is a compact metric space and hence has a measurable structure via the Borel σ-algebra.

Proof of Lemma 6.24. Let $d = \dim G$, so that

$$\mathcal{M}_d = \{\mu \in \mathcal{M}(X) \mid \dim \mathrm{Stab}_G(\mu) = d\} = \{m_X\}$$

and $\mathcal{M}_d \ni \mu \mapsto \mathrm{Stab}_G(\mu)^o$ is trivially measurable.

Fix k with $0 \leqslant k \leqslant d$ and suppose that we have already shown that the sets

$$\mathcal{M}_\ell = \{\mu \mid \dim \mathrm{Stab}_G(\mu) = \ell\}$$

for $\ell \geqslant k$ and the map $\mathcal{M}_{k+1} \ni \mu \mapsto \mathrm{Stab}_G(\mu)^o$ are measurable.

Let $\mu_n \in \mathcal{M}_{\geqslant k} = \mathcal{M}_k \cup \cdots \cup \mathcal{M}_d$ for $n \geqslant 1$ and suppose that $\mu_n \to \nu$ in the weak*-topology as $n \to \infty$. Let $\mathfrak{h}_n$ be the Lie algebra of $\mathrm{Stab}_G(\mu)^o$. As

$$\{I\} \cup \bigcup_{1 \leqslant \ell \leqslant d} \mathrm{Grass}_\ell(\mathrm{Lie}(G))$$

is compact, we may choose a subsequence and assume also that $\mathfrak{h}_n \to \mathfrak{h} \leqslant \mathfrak{g}$ as $n \to \infty$ with $\dim \mathfrak{h} \geqslant k$. We will prove below that μ is invariant under $\exp(\mathfrak{h})$ and so $\mu \in \mathcal{M}_{\geqslant k}$. It follows that $\mathcal{M}_{\geqslant k}$ is closed and hence measurable, which implies that $\mathcal{M}_k = \mathcal{M}_{\geqslant k} \smallsetminus \mathcal{M}_{\geqslant k+1}$ is also measurable.

The argument above also shows that the assumption $\mu_n \in \mathcal{M}_k$ for all $n \geqslant 1$ and $\mu_n \to \mu \in \mathcal{M}_k$ as $n \to \infty$ implies that $\mathfrak{h}_n \to \mathfrak{h}$ as $n \to \infty$, with $\dim \mathfrak{h} = k$. Therefore

$$\mathcal{M}_k \ni \mu \longmapsto \mathrm{Stab}_G(\mu)^o$$

is actually continuous on the measurable set $\mathcal{M}_k$.

Iterating the argument until we reach $k = 0$ proves the lemma.

It remains to prove the invariance of $\mu = \lim_{n\to\infty} \mu_n$ under $\mathfrak{h} = \lim_{n\to\infty} \mathfrak{h}_n$. For $v \in \mathfrak{h}$ there exists a sequence (v_n) with $v_n \in \mathfrak{h}_n$ for $n \geqslant 1$ with $v_n \to v$ as $n \to \infty$. Then, by uniform continuity,

$$\left\| f^{\exp(v_n)} - f^{\exp(v)} \right\|_\infty \longrightarrow 0$$

as $n \to \infty$ for $f \in C_c(X)$. As μ_n is a probability measure for $n \geqslant 1$ this also shows that

$$\Bigl| \underbrace{\int f^{\exp(v_n)} \,\mathrm{d}\mu_n}_{=\int f \,\mathrm{d}\mu_n} - \int f^{\exp(v)} \,\mathrm{d}\mu_n \Bigr| \leqslant \left\| f^{\exp(v_n)} - f^{\exp(v)} \right\|_\infty \longrightarrow 0$$

as $n \to \infty$. Taking limits gives

$$\int f \,\mathrm{d}\mu = \int f^{\exp(v)} \,\mathrm{d}\mu,$$

so $\exp(v)$ preserves μ. As $v \in \mathfrak{h}$ was arbitrary, the lemma follows. □

Lemma 6.25. *Let G be a σ-compact, locally compact group equipped with a left-invariant metric. Let $\Gamma < G$ be a discrete subgroup and $\eta_1, \dots, \eta_k \in \Gamma$ arbitrary elements. Then $C_G(\eta_1, \dots, \eta_k)\Gamma$ is closed in $X = G/\Gamma$.*

Proof. The proof is similar to the proof of Proposition 3.1 or Proposition 3.11. So suppose that $g_n \cdot \Gamma \to g \cdot \Gamma$ as $n \to \infty$ with $g_n \in C_G(\eta_1, \dots, \eta_k)$ for $n \geqslant 1$ and some $g \in G$. Choose $\gamma_n \in \Gamma$ for $n \geqslant 1$ with $g_n \gamma_n \to g$ as $n \to \infty$. Fix some $i \in \{1, \dots, k\}$ and notice that

$$\Gamma \ni \gamma_n \eta_i \gamma_n^{-1} = \gamma_n g_n \eta_i (\gamma_n g_n)^{-1} \longrightarrow g \eta_i g^{-1}$$

as $n \to \infty$ has to become eventually stable. So assume that

$$\gamma_N \eta_i \gamma_N^{-1} = \gamma_n \eta_i \gamma_n^{-1} = g \eta_i g^{-1}$$

for all $n \geqslant N$ and all i. However, this shows that $\gamma_N^{-1} g \in C_G(\eta_1, \dots, \eta_k)$ and

$$\Gamma g = \Gamma \gamma_N^{-1} g \in \Gamma C_G(\eta_1, \dots, \eta_k)$$

as required. □

Lemma 6.26. *Let $G \leqslant \mathrm{SL}_d(\mathbb{R})$ be a closed linear group and let $\Gamma < G$ be a discrete subgroup. Suppose that $L < G$ is a unipotent subgroup such that xL has finite volume. Then $xC_L(L)$ is compact.*

Proof. Clearly $xL \cong \Lambda\backslash L$ for a lattice $\Lambda < L$, so it suffices to consider the case $G = L$ and $x = \Lambda \in \Lambda\backslash L$. By Borel density (Theorem 3.50; also see the argument on p. 143) there exist elements $\lambda_1, \dots, \lambda_k \in \Lambda$ with

$$C_L(L) = C_L(\lambda_1, \dots, \lambda_k).$$

Thus Lemma 6.25 shows that $\Lambda C_L(L)$ is closed.

Finally, notice that if $\Lambda g_n \to \infty$ for some $g_n \in C_L(L)$ as $n \to \infty$, then the injectivity radius at Λg_n has to approach zero. In fact, by Proposition 1.35 there exist $\lambda_n \in \Lambda\smallsetminus\{I\}$ for which $g_n^{-1}\lambda_n g_n \to I$ as $n \to \infty$. However, for $g_n \in C_L(L)$ we have $g_n^{-1}\lambda_n g_n = \lambda_n \in \Lambda\smallsetminus\{I\}$ which contradicts the stated convergence. Therefore $\Lambda C_L(L)$ is a bounded closed set in ΛL, and so is compact. □

Lemma 6.27. *Suppose that $G \leqslant \mathrm{SL}_d(\mathbb{R})$ is a closed linear group, $\Gamma < G$ is a discrete subgroup, and $M < G$ is a unipotent abelian subgroup. If xM is compact for some $x \in X = \Gamma\backslash G$, then $xN_G^1(M)$ is closed, where*

$$N_G^1(M) = \{g \in G \mid gMg^{-1} = M \text{ and } gm_M g^{-1} = m_M\}$$

is the unimodular normalizer of M in G.

Proof. Let $x = \Gamma g$. By conjugating M with g we may assume without loss of generality that $x = I$. As in the proof of Lemma 6.25, we assume that $\gamma_n g_n \to g$ as $n \to \infty$ for $g_n \in N_G^1(M)$, $\gamma_n \in \Gamma$ and $g \in G$. We wish to show that $\gamma g \in N_G^1(M)$ for some $\gamma \in \Gamma$.

Notice that

$$\Gamma g_n M \cong \Bigl((g_n^{-1}\Gamma g_n) \cap M\Bigr)\backslash M,$$

which is isomorphic to $(\Gamma \cap M)\backslash M$ via conjugation by $g_n \in N_G^1(M)$. This implies that $\Gamma g_n M$ has the same volume as ΓM since conjugation by g_n in $N_G^1(M)$ preserves the Haar measure on M by definition. Moreover, since

$$\Gamma g_n \longrightarrow \Gamma g$$

as $n \to \infty$, we see that the injectivity radius of Γg_n stays bounded away from zero. By Minkowski's theorem on successive minima (Theorem 1.45, equivalently via the argument in the proof of Mahler's compactness criteria in Theorem 1.51) there exist elements

$$\eta_{n,1}, \dots, \eta_{n,\dim M} \in \Gamma$$

such that

$$(\gamma_n g_n)^{-1}\,\eta_{n,i}\,(\gamma_n g_n) \in M \tag{6.12}$$

is of bounded size (independent of n) and gives a basis of $(g_n^{-1}\Gamma g_n) \cap M$ for $i = 1, \dots, \dim(M)$. Therefore, we may choose a subsequence such that for every $i = 1, \dots, \dim(M)$ we have (after renaming the indexing variable in the sequence) that

$$(\gamma_n g_n)^{-1}\,\eta_{n,i}\,(\gamma_n g_n) \longrightarrow m_i \in M. \tag{6.13}$$

Since we also have $\gamma_n g_n \to g$ we may conjugate by $\gamma_n g_n$ in (6.13) to obtain

$$\eta_{n,i} \longrightarrow g m_i g^{-1}$$

as $n \to \infty$. However, since $\eta_{n,i} \in \Gamma$ this shows that we must have

$$\eta_{N,i} = \eta_{n,i} = g m_i g^{-1}$$

for $i = 1, \dots, \dim(M)$ and all $n \geqslant N$ for some large enough N. Conjugating by γ_n we obtain from (6.12) that

$$\underbrace{\gamma_n^{-1} \eta_{n,i} \gamma_n}_{\in M} = \gamma_n^{-1} g \underbrace{m_i}_{\in M} g^{-1} \gamma_n,$$

by the definition of $\eta_{n,i}$ for $i = 1, \dots, \dim(M)$ and all $n \geqslant N$. Since

$$(\gamma_n g_n)^{-1} \eta_{n,i} \gamma_n g_n$$

gives a basis of the lattice

$$\left(g_n^{-1} \Gamma g_n\right) \cap M$$

by definition of $\eta_{n,i}$, and a lattice in M is Zariski dense, it follows that

$$\langle m_1, \dots, m_{\dim M} \rangle$$

is also Zariski dense in M and

$$\gamma_n^{-1} g \in N_G(M)$$

for all $n \geqslant N$.

In particular,

$$\gamma_N^{-1} g \left(\gamma_n^{-1} g\right)^{-1} = \gamma_N^{-1} \gamma_n \in N_G(M)$$

for all $n \geqslant N$. We claim that $\gamma_N^{-1} \gamma_n \in N_G^1(M)$. For if $\eta = \gamma_N^{-1} \gamma_n$ (or its inverse) were to contract the Haar measure on M then $\eta^\ell (\Gamma \cap M) \eta^{-\ell}$ would have to contain shorter and shorter vectors as $\ell \to \infty$ by Minkowski's first theorem (Theorem 1.44). As $\eta^\ell (\Gamma \cap M) \eta^{-\ell} \subseteq \Gamma$ this is impossible, proving the claim.

It follows that

$$\gamma_N^{-1} g = \lim_{n \to \infty} \gamma_N^{-1} \gamma_n g_n \in N_G^1(M)$$

as required. □

6.4.2 Equidistribution and Orbit Closures on Nilmanifolds

Using Theorem 6.23 we can establish the equidistribution theorem (Theorem 6.12) and the orbit closure theorem (Theorem 6.13) on nilmanifolds. In the case of unipotent flows on nilmanifolds this step of the proof is significantly easier due to the following special feature of unipotent flows on nilmanifolds (which we know is false for the horocycle flow on a non-compact quotient, for example).

Corollary 6.28. *Let G be a connected nilpotent Lie group, let $\Gamma < G$ be a lattice in G, and let $X = \Gamma\backslash G$. Let $U \leqslant G$ be a one-parameter subgroup and $x_0 \in X$. Then the orbit closure $\overline{U\bullet x} = L\bullet x$ is algebraic, and the U-action on $L\bullet x_0$ is uniquely ergodic.*

Proof. (to come) □

6.5 Invariant Measures for Semisimple Groups

Using Section 6.3.2 we are also ready to prove the special case of Ratner's measure classification theorem where the acting group is semisimple†. We are going to use the Mautner phenomenon to find an ergodic one-parameter unipotent flow. This is possible due to the results of Chapter 2, but requires that the group H has no compact factors. While almost all of the ideas of the proof certainly go back to the work of Ratner, and in particular to the paper [135], the observation that this particular case has a short and relatively easy proof was made in [39].

Theorem 6.29 (Ratner measure classification; the semisimple case). *Let G be a connected Lie group, $\Gamma < G$ a discrete subgroup, and assume that $H < G$ is a semisimple subgroup without compact factors. Suppose that μ is an H-invariant and ergodic probability measure on X. Then μ is algebraic.*

Proof. Define the closed subgroup $\mathrm{Stab}(\mu) = \{g \in G \mid g_*\mu = \mu\}$, the connected component $L = \mathrm{Stab}(\mu)^o$, and its Lie algebra $\mathfrak{l}$. We need to prove that μ is supported on a single L-orbit. So let us assume (for the purposes of a contradiction) that this is not the case. Then by ergodicity of μ, each L-orbit must have zero μ-measure since $H \leqslant L$.

† This case is interesting as the proof is relatively straightforward, even though there may be a large gap in the dimensions of the acting group and the group that gives rise to the ambient space. Furthermore, due to this gap there may be a large collection of possible intermediate subgroups $H \leqslant L \leqslant G$. However, the use of this special case is limited as the acting group is not amenable and hence it is *a priori* not even clear why we should have any H-invariant probability measure on a given orbit closure $\overline{H\bullet x} \subseteq X$.

There exists a subgroup of H that is locally isomorphic to $\mathrm{SL}_2(\mathbb{R})$, which acts ergodically on X with respect to μ. This follows from the Mautner phenomenon. Indeed, H is by assumption an almost direct product of non-compact simple Lie groups, and each of these contains a subgroup that is locally isomorphic to $\mathrm{SL}_2(\mathbb{R})$. Now consider a diagonally embedded subgroup M that is locally isomorphic to $\mathrm{SL}_2(\mathbb{R})$ and that projects nontrivially to each simple almost direct factor. Furthermore, we let $U \leqslant M$ be the subgroup corresponding to the upper unipotent subgroup in $\mathrm{SL}_2(\mathbb{R})$. By Proposition 2.25 the subgroup U, and hence also M, satisfies the Mautner phenomenon for H. Since H acts ergodically, so does the subgroup. So we may assume that H is locally isomorphic to $\mathrm{SL}_2(\mathbb{R})$.

By the structure theory of finite-dimensional representations of $\mathrm{SL}_2(\mathbb{R})$ (see [48, Th. 4.11], Fulton and Harris [55], or Knapp [89, Th. 1.64], for example), we see that the H-invariant subspace $\mathfrak{l} \leqslant \mathfrak{g}$ (with respect to the adjoint action) has an H-invariant complement $V < \mathfrak{g}$. We note that we have no reason to expect that V is a Lie algebra, and that this step uses crucially the fact that H is semisimple.

Now let $K \subseteq X$ be a set of μ-measure exceeding 0.99 comprising uniformly generic points for $U < H$. We would like to find points $x_n, y_n \in K$ with

$$y_n = g_n \boldsymbol{\cdot} x_n,$$

for some $g_n \neq I$ with $g_n \in \exp(V)$ belonging to the 'transverse' direction for all $n \geqslant 1$, and with $g_n \to I$ as $n \to \infty$. We then may consider the polynomials

$$p_n(r) = \mathrm{Ad}_{u(T_n r)}(\log g_n), \tag{6.14}$$

assume that these converge as $n \to \infty$, and apply Proposition 6.19. By the H-invariance of V all the polynomials p_n would have values in V and so we would then be able to find a polynomial $p\colon \mathbb{R} \to \mathfrak{g}$ taking values in V and with μ invariant under $\exp p(r)$ for all $r > 0$. The existence of such a polynomial contradicts the definition of $L = \mathrm{Stab}(\mu)^o$.

To find x_n, y_n as above, we can apply a relatively simple Fubini argument as follows (crucially, using the fact that μ is invariant under L).

So let $B_\delta^L = B_\delta^L(I)$ be a small open metric ball in L around the identity, and define

$$Y = \left\{ x \in X \;\middle|\; \int_{B_\delta^L} \mathbb{1}_K(\ell \boldsymbol{\cdot} x)\,\mathrm{d}m_L(\ell) > 0.9 m_L(B_\delta^L) \right\}.$$

We claim first that $\mu(Y) > 0.9$, which may be seen by looking at the complement as follows:

$$\begin{aligned}\mu(X\smallsetminus Y) &= \mu\left(\left\{x\in X \;\Big|\; \int_{B_\delta^L} \mathbb{1}_{X\smallsetminus K}(\ell\cdot x)\,\mathrm{d}m_L(\ell) \geqslant 0.1 m_L(B_\delta^L)\right\}\right)\\ &\leqslant \frac{1}{0.1 m_L(B_\delta^L)}\int_X\int_{B_\delta^L} \mathbb{1}_{X\smallsetminus K}(\ell\cdot x)\,\mathrm{d}m_L(\ell)\,\mathrm{d}\mu\\ &= \frac{1}{0.1 m_L(B_\delta^L)}\int_{B_\delta^L}\underbrace{\int_X \mathbb{1}_{X\smallsetminus K}(\ell\cdot x)\,\mathrm{d}\mu}_{=\mu(X\smallsetminus K)}\,\mathrm{d}m_L \qquad \text{(by Fubini)}\\ &= \frac{\mu(X\smallsetminus K)}{0.1} < \frac{0.01}{0.1} = 0.1,\end{aligned}$$

since L preserves μ and $\mu(K) > 0.99$.

We now claim that for any nearby points $x, y \in Y$ we can find $\ell_x, \ell_y \in B_\delta^L$ such that

$$x' = \ell_x\cdot x \in K, \tag{6.15}$$

$$y' = \ell_y\cdot y \in K, \tag{6.16}$$

and

$$y' = \exp(v)\cdot x' \tag{6.17}$$

with $v \in V$. To see this, notice that if δ is sufficiently small, then (by the inverse mapping theorem) the map

$$\begin{aligned}\psi\colon B_{2\delta}^L \times B_{2\delta}^V(0) &\longrightarrow G\\ (\ell, v) &\longmapsto \ell\exp(v)\end{aligned}$$

is a diffeomorphism from $B_{2\delta}^L \times B_{2\delta}^V(0)$ onto an open neighbourhood O of the identity in G. Let now $g \in B_\kappa^G(I)$ be chosen so that $y = g\cdot x$. Then we would like to find $\ell_x, \ell_y \in B_\delta^L$ with $g\ell_x^{-1} = \ell_y^{-1}\exp(v)$, which will give (6.17). This can be done using the local diffeomorphism above: If κ is sufficiently small, then $g\ell_x^{-1} \in O$ and may define ℓ_y and v by

$$\psi^{-1}(g\ell_x^{-1}) = (\ell_y^{-1}, v). \tag{6.18}$$

However, we still have to worry about the conditions (6.15) and (6.16).

For this, we are going to see that most points $\ell_x \in B_\delta^L$ (and the corresponding ℓ_y) will satisfy this. Indeed, by definition of Y, at least 90% of all $\ell_x \in B_\delta^L$ satisfy $x' = \ell_x\cdot x \in K$, and at least 90% of all $\ell_y \in B_\delta^L$ satisfy $y' = \ell_y\cdot y \in K$. However, we need to do this while ensuring that (6.18) (or equivalently, (6.17)) holds. So define the map

$$\begin{aligned}\phi\colon B_\delta^L &\longrightarrow B_{2\delta}^L\\ \ell_x &\longmapsto \ell_y\end{aligned}$$

with ℓ_y as in (6.18). This smooth map depends on the parameter $g \in B_\kappa^G$ and is close to the identity in the C^1-topology if κ is sufficiently small (all maps we deal with are analytic and for $g = e$ we have $\phi = I_{B_\delta^L}$). Therefore ϕ does not distort the chosen Haar measure of L much, and sends B_δ^L into a ball around the identity that is not much bigger than B_δ^L (both with respect to the metric structure and with respect to the measure). In other words, if κ is sufficiently small, then

$$\begin{aligned} m_L\left(\phi\left(\left\{\ell_x \in B_\delta^L \;\middle|\; \ell_x{\cdot}x \in K\right\}\right) \cap B_\delta^L\right) &> 0.9 m_L\left(\phi\left(\left\{\ell_x \in B_\delta^L \;\middle|\; \ell_x{\cdot}x \in K\right\}\right)\right) \\ &> 0.8 m_L\left(\left\{\ell_x \in B_\delta^L \;\middle|\; \ell_x{\cdot}x \in K\right\}\right) \\ &> (0.8)(0.9) m_L\big(B_\delta^L\big) > 0.7 m_L\big(B_\delta^L\big). \end{aligned}$$

Together with

$$m_L\left(\left\{\ell_y \in B_\delta^L \mid \ell_y{\cdot}y \in K\right\}\right) > 0.9 m_L\big(B_\delta^L\big),$$

we see that there are many points $\ell_x \in B_\delta^L$ with $\ell_x{\cdot}x \in K$ for which ℓ_y defined by (6.18) also satisfies $\ell_y{\cdot}y \in K$.

The theorem now follows relatively quickly as outlined earlier. Recall that we may assume that every $L = \operatorname{Stab}(\mu)^o$-orbit has μ-measure zero. Let

$$z \in \operatorname{supp}\left(\mu|_Y\right).$$

Then for every $\kappa = \frac{1}{n}$ there exist $x_n = z, y_n = g_n{\cdot}x_n \in Y$ with

$$g_n \in B_{1/n}^G(I) \smallsetminus L.$$

Applying the procedure above to x_n, y_n (which we certainly may if n is large) then we get

$$x_n', y_n' = \exp(v_n){\cdot}x_n' \in K, v_n \in V, v_n \neq 0, v_n \longrightarrow 0$$

as $n \to \infty$. There are now two cases to consider.

If v_n is in the eigenspace of Ad_{u_s} for infinitely many n (and so let us assume for all n by passing to that subsequence), then we may apply Lemma 6.17 to each v_n and deduce that $\exp(v_n)$ preserves μ. However, since $v_n \to 0$ as $n \to \infty$ and the unit sphere in V is compact, we may assume that $\frac{v_n}{\|v_n\|} \to w$ as $n \to \infty$ by passing to a subsequence again. We conclude that since $\operatorname{Stab}(\mu)$ is closed, $\exp(tw) \in \operatorname{Stab}(\mu)$ for all t. Since V is a linear complement to the Lie algebra of $L = \operatorname{Stab}(\mu)^o$, this is a contradiction.

So assume that v_n is not in the eigenspace for any $n \geqslant 1$ (by deleting finitely many terms). In this case we may define T_n such that the polynomials in (6.14) have norm one. Use compactness of the set of polynomials with bounded degree and norm one to choose a subsequence (again denoted (p_n)) that converges to a polynomial p, and then apply Proposition 6.19 to see that μ is invariant under by $\exp p(t)$ for all $t > 0$. Since p is the limit of $\mathrm{Ad}_{T_n r}(v_n) \in V$, p also takes values in V which again contradicts the definition of V. □

Notes to Chapter 6

(31)(Page 231) Arguably this was a rediscovery of a connection between Diophantine problems of this sort and homogeneous dynamics used earlier by Cassels and Swinnerton-Dyer [13]. We refer to a survey by Margulis [110] for a motivated history.

(32)(Page 238) This appeared in print in the work of Dani [19, Conjecture II].

(33)(Page 242) This is an instance of a more general result due to Weyl [167] giving equidistribution modulo one for the values on the natural numbers of any polynomial with an irrational coefficient. Furstenberg [56] showed that this followed from a general result extending unique ergodicity from irrational circle rotations to certain maps on tori. We refer to [45, Sec. 4.4.3] for a detailed discussion.