Equidistribution of Finite Volume Orbits on Homogeneous Spaces

Notes taken by Mike Woodbury

CHAPTER 1

June 1, 2009

1. Michel

These lectures will be less focused on ergodic theory and more on equidistribution and its relation to ergodic theory.

1.1. The Problem of Linnik. We begin with the specific example of the lowest dimensional case of Linnik's problems. This example deals with representations of integers by ternary quadratic forms. We consider the two forms

$$q_1(a,b,c) = a^2 + b^2 + c^2$$
$$q_2(a,b,c) = b^2 - 4ac$$

We say $d \in \mathbb{Z}$ is *represented by a quadratic form* q if there exist $a, b, c \in \mathbb{Z}$ such that $q(a,b,c) = d$. For $d \neq 0$ we define

$$R_q(d) = \{(a,b,c) \in \mathbb{Z}^3 \mid q(a,b,c) = d\}.$$

A natural question is "how does $R_q(d)$ vary?" A more basic question being "when is $R_q(d) \neq \emptyset$?" The answer is

$$\begin{array}{ll} R_q(d) \neq \emptyset \iff d \equiv 0, 1 \pmod 4 & \text{in the case of } q_1 \\ R_q(d) \neq \emptyset \iff d = 4^a(8b+7) & \text{in the case of } q_2 \end{array}$$

In each case the direction $\implies$ is simple to verify by considering congruences modulo 8 (resp. modulo 4) in the q_1 (resp. q_2) case. The other direction is the hard part. They are both special cases of Hasse's principle.

One can show that for $d \to \infty$ such that $R_q(d) \neq 0$, $R_q(d) \to \infty$. Linnik asked "how are these vectors distributed in $\mathbb{R}^3$?" To properly study this question we define for $\epsilon = \pm 1$

$$\begin{aligned} V_{q,\epsilon}(\mathbb{R}) &= \{(x,y,z) \in \mathbb{R}^3 \mid q(x,y,z) = \epsilon\} \\ &= \begin{cases} S^2 & \text{if } q = q_1, \epsilon = 1 \\ \text{a 1-sheeted hyperboloid} & \text{if } q = q_2, \epsilon = +1 \\ \text{a 2-sheeted hyperboloid} & \text{if } q = q_2, \epsilon = -1 \end{cases} \end{aligned}$$

Then

$$\frac{1}{\sqrt{|d|}} R_q(d) \subset V_{q,\epsilon} \qquad \epsilon = \operatorname{sgn} d,$$

Moreover, $V_{q,\epsilon}(\mathbb{R})$ carries a natural measure $\mu_{q,\epsilon}$. Indeed, since $\mathrm{SO}_q(\mathbb{R})$ acts on $V_{q,\epsilon}(\mathbb{R})$, by 'natural' we mean (as described more completely in Section 1.2) an $\mathrm{SO}_q(\mathbb{R})$-invariant measure. For $\Omega \subset V_{q,\epsilon}(\mathbb{R})$, the (unique up to constant) measure is given by defining $\mu_{q,\epsilon}(\Omega)$ to be the Lebesgue measure of the 'cone' in $\mathbb{R}^3$ obtained by taking all line segments between the origin and points of Ω.

With these definitions in hand, we rephrase Linnik's question as "how are $\frac{1}{\sqrt{|d|}}R_q(d)$ distributed on $V_{q,\epsilon}(\mathbb{R})$ with respect to $\mu_{q,\epsilon}$?"

THEOREM 1 (Linnik, Skubenko, Duke). *As $d \to \infty$, $\frac{1}{\sqrt{|d|}}R_q(d)$ become equidistributed on $V_{q,\epsilon}(\mathbb{R})$ with respect to $\mu_{q,\epsilon}$.*

We describe the term *equidistributed*. Fix $f_0 \in C_c(V_{q,\epsilon}(\mathbb{R}))$. Then for any $f \in C_c(V_{q,\epsilon}(\mathbb{R}))$ this means that

$$\frac{\sum_{x\in R_q(d)} f(\frac{x}{\sqrt{|d|}})}{\sum_{x\in R_q(d)} f_0(\frac{x}{\sqrt{|d|}})} \longrightarrow \frac{\int f \; d\mu_{q,\epsilon}}{\int f_0 \; d\mu_{q,\epsilon}} \tag{1}$$

as $d \to \infty$.

Note that because $R_q(d)$ consists of integral points its intersection with any compact set is finite. Hence the sum in (1) is finite, and therefore well defined.

Michel tried to explain why the f_0 is required. I don't see why it is necessary, but if $V_{q,\epsilon}(\mathbb{R})$ is compact obviously one can take $f_0 = 1$. In this case one obtains the 'standard' definition of equidistribution.

1.2. Translation to group theoretic language. Witt's theorem tells us that $\mathrm{SO}_q(\mathbb{R})$ acts transitively on $V_{q,\epsilon}(\mathbb{R})$. So picking $x_\infty \in V_{q,\epsilon}(\mathbb{R})$ gives that

$$V_{q,\epsilon}(\mathbb{R}) \simeq \mathrm{SO}_q(\mathbb{R})/\mathrm{SO}_{q,x}(\mathbb{R})$$

where $\mathrm{SO}_{q,x}(\mathbb{R}) = H$ is the stabilizer of x_∞. The group H is a special orthogonal group in two variables.

So the question of understanding points on $V_{q,\epsilon}(\mathbb{R})$ is equivalent to that of understanding H-orbits in $G = \mathrm{SO}_q(\mathbb{R})$. The H-orbits are the sets

$$\{g_{a,b,c} \in G \mid g_{a,b,c}x_\infty = \frac{(a,b,c)}{\sqrt{|d|}}\}.$$

(At this point we introduce a discrete subgroup into the picture. I'm confused at *why* this is permissible. My assumption is that understanding the distribution of points on G is the same as understanding that on $\Gamma\backslash G$ *if* Γ is a lattice. Michel made no mention of lattices at this point, but Einsiedler does exactly this in his first lecture.)

Let $\Gamma = \mathrm{SO}_q(\mathbb{Z})$ which acts on $R_q(d)$. Hence

$$R_q(d) = \bigsqcup_{[a,b,c]\in\Gamma\backslash R_q(d)} \Gamma(a,b,c).$$

So, we want to understand the distribution of Γ-orbits

$$\bigsqcup(\Gamma g_{a,b,c}H)/H$$

in G/H.

Γ, a priori, does not act transitively on $R_q(d)$, but the number of orbits is finite. This is evident for q_1 because $R_q(d)$ itself is finite. For q_2 this is a consequence of Gauss' reduction theory for binary quadratic forms under the indentification $q_2(a,b,c) = d \leftrightarrow aX^2 + bXY + cY^2$.

There is an equivalence between Γ-orbits on G/H and H-orbits on $\Gamma\backslash G$:

$$\Gamma gH/H \longleftrightarrow \Gamma\backslash\Gamma_g H.$$

This equivalence is valid at the level of measures:

$$\left\{ \begin{array}{c} \Gamma\text{-invariant Radon} \\ \text{measures on } G/H \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} H\text{-invariant Radon} \\ \text{measures on } \Gamma\backslash G \end{array} \right\}$$

In this setting we have the following.

THEOREM 2 (Linnik, Skubenko, Duke). *The finite collection of H-orbits*

$$\{\Gamma\backslash\Gamma g_{a,b,c}H\} \subset \Gamma\backslash G$$

becomes equidistributed on $\Gamma\backslash G$ with respect to the quotient of the Haar measure on G by the counting (Haar) measure on Γ.

1.3. An equivalent formulation for q_2. We take $q = q_2$. The quadratic space $(\mathbb{Q}^3, q)$ is equivalent to the space $M_2^0(\mathbb{Q})$ of traceless 2×2 matrices with form $-\det$ under the linear map

$$(a, b, c) \mapsto \begin{pmatrix} b & -2a \\ 2c & -b \end{pmatrix}.$$

Under this identification SO_q is PGL_2 acting on M_2^0 by conjugation.

In the case $\epsilon = -1$ we take $x_\infty = (\frac{1}{2}, 0, \frac{1}{2})$ which corresponds to $m_\infty = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $H = \mathrm{PSO}_2(\mathbb{R})$. In the case $\epsilon = +1$, we choose $x_\infty = (0, 1, 0)$ for which $m_\infty = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $H = Diag_2(\mathbb{R})/\mathbb{R}^\times = A$. In either case, we have $\Gamma \leftrightarrow \mathrm{PGL}_2(\mathbb{Z}) \simeq \mathrm{GL}_2(\mathbb{Z})/\pm I$.

Thus

$$V_{-1} \leftrightarrow \mathrm{PGL}_2(\mathbb{R})/\mathrm{PSO}_2(\mathbb{R}) \simeq \mathbb{C} \setminus \mathbb{R} = \mathbb{H} \cup \mathbb{H}^{-1},$$

and

$$V_{+1} \leftrightarrow \mathrm{PGL}_2(\mathbb{R})/A$$

under the formulation of the previous section.

Note that

$$\Gamma\backslash G = \mathrm{PGL}_2(\mathbb{Z})\backslash\mathrm{PGL}_2(\mathbb{R}) \simeq \mathrm{PSL}_2(\mathbb{Z})\backslash\mathrm{PSL}_2(\mathbb{R}) \simeq T^1(\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H})$$

which is the unit tangent bundles of the modular surface.

To illustrate that the study of these types of orbits is of interest in number theory, notice that if $d < 0$

$$\Gamma(a, b, c) \leftrightarrow \Gamma\backslash\Gamma g_{a,b,c}H \subset \Gamma\backslash G \twoheadrightarrow \Gamma\backslash G/H \simeq \Gamma\backslash\mathbb{H}$$

because $H = \mathrm{PSO}_2(\mathbb{R})$. The image of $\Gamma(a, b, c)$ in $\Gamma\backslash\mathbb{H}$ is $[z_{a,b,c}] = \left[\frac{-b+i\sqrt{|d|}}{2a}\right]$, the so-called Heegner point.

The case $d > 0$ is also of arithmetic interest. In this case

$$\Gamma(a, b, c)H \longrightarrow \Gamma\backslash\Gamma g_{a,b,c}A \subset \Gamma\backslash G = T^1(\Gamma\backslash\mathbb{H}) \twoheadrightarrow \Gamma\backslash\mathbb{H},$$

and, here the image is the projection onto the modular curve of the geodesic half circle intersecting the real axis perpendicularly at $\frac{-b\pm\sqrt{d}}{a}$. This is the geodesic flow.

2. Einsiedler

In general, we take G to be a 'nice' group, m_G to be the left Haar measure on G and $\Gamma \subset G$ a discrete subgroup. The 'niceness' of G guarantees that there exists a fundamental domain F which is a measurable set such that $G = \sqcup_{\gamma\in\Gamma}\gamma F$. The subgroup Γ is a *lattice* if $m_G(F) < \infty$ for some (all) fundamental domain(s).

To prove the existence of F one uses the following notions. A set $B \subset G$ is Γ-*injective* if $B \cap \gamma B = \emptyset$. A set $B \subset G$ is Γ-*surjective* if $G = \cup_{\gamma\in\Gamma}\gamma B$. (All sets are taken to be measurable.) It is easy to see that to prove the existence of a fundamental domain it suffices to show that there is a Γ-surjective set.

It is also a fact, and not difficult to prove, that any two fundamental domains have the same left Haar measure.

2.1. Lattices in $\mathbb{R}^d$. In the special case that $G = \mathbb{R}^d$, lattices are all of the form

$$\Lambda = \mathbb{Z}v_1 + \cdots + \mathbb{Z}v_d$$

for $\langle v_1, \dots, v_d\rangle$ a basis of $\mathbb{R}^d$. We call Λ *unimodular* if $\det(\begin{bmatrix} v_1 \\ \vdots \\ v_d \end{bmatrix}) = \pm 1$. (The basis vectors are considered as row vectors. Also, the choice of basis can only change the sign of the determinant, so this is well defined.)

We consider

$$\begin{aligned}\Omega_d = X_d =& \{\text{all unimodular lattices in } \mathbb{R}^d\}\\ =& \{\mathbb{Z}^d g \mid g \in \mathrm{SL}_d(\mathbb{R})\}\\ =& \Gamma\backslash \mathrm{SL}_d(\mathbb{R})\end{aligned}$$

where $\Gamma = \mathrm{SL}_d(\mathbb{Z})$.

Two questions: Is Γ a lattice? Is Γ *uniform*, meaning is $\Gamma\backslash G$ compact? The answer to the first question is yes, but in the $d = 2$ case,

$$X_2 = \mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{SL}_2(\mathbb{R}) = T^1(\mathrm{SL}_2(\mathbb{R})\backslash\mathbb{H}).$$

Since $\mathrm{SL}_2(\mathbb{R})\backslash\mathbb{H}$ is not compact neither is X_2. So, the answer to the second question is no.

THEOREM 3 (Minkowski's Theorem on Convex Bodies). *Let $\Lambda \subset \mathbb{R}^d$ be a lattice and $Q \subset \mathbb{R}^d$ a convex symmetric subset. (i.e. $Q = -Q$.) If $\mathrm{vol}(Q) > 2^d \,\mathrm{covol}(\Gamma)$ then $Q \cap \Gamma \neq \{0\}$.*

THEOREM 4 (Minkowski's Theorem on Successive Minima). *Let $\Lambda \subset \mathbb{R}^d$ be a lattice. Define*

$$\lambda_i = \min\{r \mid \overline{B_r}(0) \cap \Lambda \text{ has } i \text{ linearly independant vectors}\}.$$

Then $\lambda_1 \cdots \lambda_d \asymp \mathrm{covol}(\Lambda)$.

The notation $A \asymp B$ here means that $A \ll_d B$ and $B \ll_d A$.

PROOF. In the case $d = 1$, $\lambda_1 = \mathrm{covol}(\Lambda)$. We now proceed by induction. Choose $v_1 \in \Lambda$ such that $\|v_1\| = \lambda_1$. Then $v_1^\perp \simeq \mathbb{R}^{d-1}$. Let $\pi : \mathbb{R}^d \to v_1^\perp$ be the orthonormal projection. So $\pi(\Lambda)$ is a lattice in $\mathbb{R}^{d-1}$. Define $\lambda_i^\perp = \lambda_{i-1}(\pi(V))$. By induction we have $\lambda_2^\perp \cdots \lambda_d^\perp \asymp \mathrm{covol}(\pi(\Lambda))$. Moreover, $\lambda_1 \,\mathrm{covol}(\pi(V)) = \mathrm{covol}(\Lambda)$.

To complete the proof it suffices to show that $\lambda_i^\perp \asymp \lambda_i$. The fact that $\lambda_i^\perp \le \lambda_i$ is immediate because the project decreases (or leaves unchanged) lengths. To prove the other inequality, notice that if $w \in \Lambda$ then $w = v_1^\perp + (a+n)v_1$ where $a \in [-1/2, 1/2)$ and $n \in \mathbb{Z}$. From this we see that we can choose a $\mathbb{Z}$-basis of Λ such that

$$\lambda_1 \le \lambda_i \le \sqrt{(\lambda_i^\perp)^2 + (\frac{1}{2}\lambda_1)^2}. \tag{2}$$

Since $\lambda_1 \le \lambda_i$ this implies that $\lambda_1 \le \sqrt{2}\lambda_i^\perp$. Using this in (2), we find that $\lambda_i \le \sqrt{3}\lambda_i^\perp$. □

In the proof of Minkowski's successive minima theorem we obtained via induction a basis with certain relations between their lengths which lead to the following.

COROLLARY 5. *The set*

$$B = \left\{ \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ n_{ij} & & 1 \end{pmatrix} \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_d \end{pmatrix} k \;\middle|\; \begin{array}{c} n_{ij} \in [-1/2, 1/2), \\ c_1 \ll \cdots \ll c_d, k \in \mathrm{SO}(d) \end{array} \right\}$$

is a surjective domain for the action of $\mathrm{SL}_d(\mathbb{Z})$ *on* $\mathrm{SL}_d(\mathbb{R})$. *In particular,* $\mathrm{SL}_d(\mathbb{R}) = NAK$.

This corollary and the following proposition are useful in proving that $\mathrm{SL}_d(\mathbb{Z})$ is a lattice.

PROPOSTION 6. *Let* $\rho : \mathrm{SL}_d \to \mathrm{SL}_{d'}$ *be a homorphism defined by polynomials with rational coefficients. We put* $H = \{g \in \mathrm{SL}_d \mid \rho(g)v = v\}$ *for some* $v \in \mathbb{Q}^{d'}$. *The orbit* $\mathrm{SL}_d(\mathbb{Z})I \in \mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_d(\mathbb{R})$ *is closed.*

PROOF. To see this, suppose $\Gamma h_n \to \Gamma g$ as $n \to \infty$ for some $h_n \in H(\mathbb{R})$ and $g \in \mathrm{SL}_d(\mathbb{R})$. We need to show that that $g \in \mathrm{SL}_d(\mathbb{Z})H(\mathbb{R})$. The convergence condition is equivalent to saying that there exist $\gamma_n \in \mathrm{SL}_d(\mathbb{Z})$ such that $\gamma_n h_n \to g$. Hence

$$\rho(\gamma_n)v = \rho(\gamma_n h_n)v \to \rho(g)v. \tag{3}$$

Note that $v \in \frac{1}{N_1}\mathbb{Z}^{d'}$. Moreover, the polynomials associated to ρ are defined over $\frac{1}{N_2}\mathbb{Z}$. It follows that $\rho(\gamma_n)v \in \frac{1}{N}\mathbb{Z}^{d'}$ for some integer N independent of n. Since $\frac{1}{N}\mathbb{Z}^{d'}$ is discrete and closed (and ρ is continuous) the points $\rho(\gamma_n)v$ must eventually all be the same. So for n sufficiently large, $\rho(\gamma_n h_n)v = \rho(g)v$. Applying (3), this is equivalent to saying $v = \rho(\gamma_n^{-1}g)v$. Hence $\gamma_n^{-1}g \in H(\mathbb{R})$. □

2.2. Exercises. Mahler's compactness criterion says that a subset $L \subset X$ has compact closure if there exists a $\delta > 0$ such that $\Lambda \in L$ implies $\Lambda \cap B_\delta(0) = 0$. We take this as granted. (It will be proved in tomorrow's lecture.)

We write $\Gamma = \mathrm{SL}_d(\mathbb{Z})$ and $G = \mathrm{SL}_d(\mathbb{R})$ in this section unless otherwise specified.

Let Q be a quadratic form in d variables with rational coefficients. Equivalently $Q(x) = xAx^t$ for a rational symmetric matrix A. Suppose $Q(x) = 0$ for $x \in \mathbb{Q}^d$ implies $x = 0$, i.e. Q does not represent o nontrivially. Let

$$H = \mathrm{SO}(Q) = \{g \in \mathrm{SL}_d \mid Q(xg) = Q(x) \text{ for all } x\} = \{g \in \mathrm{SL}_d \mid gAg^t = A\}.$$

EXERCISE 7. *Prove that the orbit* $\Gamma H(\mathbb{R})$ *is compact in* $\Gamma\backslash G$. *(This is the same as saying that the* H*-orbit of the identity in* $\Gamma\backslash G$ *is compact.)*

SOLUTION. Note that $\rho(g) = gAg^t$ satisfies the criteria of Proposition 6. The fact that H-orbits on $\Gamma\backslash G$ are in continuous bijection with Γ-orbits in G/H then implies closed.

Now we use Minkowski's compactness criterion. Need to show that there exists $\delta > 0$ such that for every $\Lambda \in \mathrm{SL}_d(\mathbb{Z})$, $\Lambda \cap B_\delta(0) = \{0\}$ in $\mathbb{R}^d$. Any $\delta < 1$ works because $\Lambda \subset \mathbb{Z}^d$. Therefore, the closure of the orbit is compact. But, by the above, the orbit is already closed so we are done. □

Let K be a degree d number field. Fix a basis of (an ideal for an order $\mathcal{O}$ in) K. Using this basis we can first identify K with (the row space) $\mathbb{Q}^d$ and also embed K into $Mat_d(\mathbb{Q})$ via the linear map ϕ such that $b\phi(a) = ab$.

Recall that the the norm form $N_{K|\mathbb{Q}}(a) = \det \phi(a)$ which is a polynomial in the coordinates of a (considered as an element of the d-dimensional space.) We define

$$H = \{g \in \mathrm{SL}_d \mid g\phi(a) = \phi(a)g \text{ for all } a \in K\}.$$

EXERCISE 8. *Prove that $\Gamma H(\mathbb{R})$ is compact. Use this to prove Dirichlet's unit theorem.*

SOLUTION. Let $\langle a_1, \cdots, a_d\rangle$ be the basis as above. Then by linearity

$$H = \{g \in \mathrm{SL}_d \mid g^{-1}\phi(a)g = \phi(a) \text{ for all } a \in K\} = \bigcap_{i=1}^{d} H(a_i)$$

where $H(a_i) = \{g \in \mathrm{SL}_d \mid g^{-1}\phi(a_i)g = \phi(a_i)\}$. Since a finite intersection of compact sets is compact, it suffices to show that $H(a)$ is compact for $a \in \{a_1, \ldots, a_d\}$.

To do this define $\rho : \mathrm{SL}_d \to \mathrm{SL}_d$ to be $\rho(g)b = g^{-1}\phi(b)g$. We divide by $\det \phi(a)$ if necessary to make the image SL_d. In any case, H is exactly of the type described in Proposition 6, hence it's closed. By definition, $\phi(a)$. □

EXERCISE 9. *Finish the proof that $\Gamma\backslash G$ has finite volume.*

SOLUTION. It suffices to show is that there is a surjective set C with finite volume. Indeed, we take C as in Corollary 5. To complete the proof one needs to know that the measure on G is essentially the product of the measures on N, A and K respectively: the measure on N (resp. A) being Lebesgue measure on $\mathbb{R}^{(d-1)d/2}$ (resp. R^d), and that on K the standard Haar measure.

Since $C \cap N$ has measure 1 and K is compact, the result now follows by observing that the restriction of on the c_i in C implies that $A \cap C$ has finite volume as well.

(THIS IS WRONG: Check what Haar measure is on A and then what the measure on $NA = B$ is. Since B is not unimodular its Haar measure is not the product of that on A and N, but something twisted by a power of the modular character on B. Moreover, $A \cap C$ does not have finite volume in A, but with this 'corrected' measure it does in B.) □

EXERCISE 10. *Suppose G has a lattice Γ. Show that G is unimodular and that $X = \Gamma\backslash G$ has a right G-invariant measure also called Haar measure.*

SOLUTION. This result follows from my notes on quotient measures. □

EXERCISE 11. *Suppose $H \subset \mathrm{SL}_d(\mathbb{Z})$ is a closed subgroup and x belongs to $\mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_d(\mathbb{R})$. for which xH has finite H-invariant volume. Show that xH is a closed subset of $\mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_d(\mathbb{R})$.*

3. Lindenstrauss

Want to study dynamics in $\Gamma\backslash G$ for G a linear algebraic group. We consider $H \subset G$ a subgroup and its action on G.

- If the acting group H is compact there are no dynamics.
- In the arithmetic case, a compact group can be the shadow of a bigger periodic group.
- If generated by unipotents one gets polynomial dynamics. These are qualitatively (but not quantitatively) understood.
- For diagonalizable groups one has exponential dynamics. This case is only partially understood. It is related to entropy.

Later we'll look at applications to the number theory. In particular, we'll see how entropy relates to a conjecture of Littlewood: $\liminf_{n\to\infty} n\|n\alpha\|\|n\beta\| = 0$.

Also, we'll look at distribution of closed orbits of diagonalizable groups. The case of $Diag_2(\mathbb{R})/\mathbb{R}^\times$ acting on $\mathrm{PGL}_2(\mathbb{Z})\backslash\mathrm{PGL}_2(\mathbb{R})$ is what happens in the situation of $b^2 - 4ac = d$ being discussed in the lectures by Michel.

The general plan is to emphasize how to apply known results of ergodic theory.

3.1. Some elements of ergodic theory. The pointwise ergodic theorem says that if $(X, \mathcal{B}, \mu)$ is a measure space and $T : X \to X$ is a measure preserving transform (i.e. $T * \mu = \mu$) then for any $f \in L^1(\mu)$ there exists $\overline{f}$ (also in $L^1(\mu)$) that is T-invariant, and

$$\frac{1}{N}\sum_{k=1}^{N} f(T^k x) \to \overline{f}(x)$$

almost everywhere.

One can be more explicit about $\overline{f}$. The notion of conditional expectation deals with $\mathcal{A} \subset \mathcal{B}$ a sub σ-algebra. For $f \in L^2(\mu)$, $\mathbb{E}(f \mid \mathcal{A})$ is the projection of f to the space of $\mathcal{A}$-measurable functions. $\mathbb{E}(\cdot \mid \mathcal{A}) : L^2 \to L^2$ is a bounded operator of norm 1. On $L^1 \cap L^2$ it also has norm 1, hence can be extended to L^1.

Suppose that $(X, \mathcal{B}, \mu)$ is a standard Borel space. So, X is locally compact, $\mathcal{B}$ is the Borel σ-algebra and μ is the Borel measure. In this case, any σ-algebra $\mathcal{A}$ is equivalent to a countably generated σ-algebra $\widetilde{\mathcal{A}}$. This means that for any $A \in \mathcal{A}$ there exists $\widetilde{A} \in \widetilde{\mathcal{A}}$ such that $\mu(\mathbb{A} \triangle \widetilde{A}) = 0$ and vice versa. Moreover, $\widetilde{\mathcal{A}}$ is the σ-algebra generated by a countable set $\mathcal{A}_0$. ($A \triangle B$ is symmetric difference of A and B.)

We define the *atom* of $x \in X$ to be

$$[x]_{\widetilde{\mathcal{A}}} = \bigcap_{x \in A \cap \widetilde{\mathcal{A}}} A \quad \left(= \bigcap_{x \in A \cap \mathcal{A}} A \right).$$

(This is up to a set of measure zero which is always implied although sometimes not said.)

If we define $\mu_x^{\mathcal{A}}$ to be the probability measure on $[x]_{\mathcal{A}}$. Then $\mathbb{E}(f \mid \widetilde{\mathcal{A}}) = \int f \, d\mu_x^{\widetilde{\mathcal{A}}}$.

Example 12. *The example to keep in mind with this concept is $X = [0,1]^2$, $\mathcal{B}$ is the standard Borel σ-algebra, and $\mathcal{A}$ is the product $B \times [0,1]$ where B is Borel on $[0,1]$. The atoms of $\mathcal{A}$ are then vertical 'lines.'*

The conditional expectation formulation of the ergodic theorem then says that $\overline{f} = \mathbb{E}(f \mid \Sigma)$ where Σ is the collection of T-invariant sets.

3.2. Information theory. Let X be a random variable with countably many values $\mathcal{S}$,

$$\mathcal{C} : \mathcal{S} \to \text{ words in } \{0,1\}.$$

We say $\mathcal{C}$ is *prefix free* if for all distinct $S, S' \in \mathcal{S}$, $\mathcal{C}(S)$ is not a prefix of $\mathcal{C}(S')$. In other words if $\mathcal{C}(S) = w$ and $S' \neq S$, $\mathcal{C}(S') \neq wv$. Note that, in particular, prefix free implies injectivity. We $\ell_S = \ell(\mathcal{C}(S))$ be the number of bits (i.e. length) of $\mathcal{C}(S)$.

LEMMA 13. *There exists a prefix free code on S if and only if*[1] $\sum 2^{-l_S} \leq 1$.

We define

$$H(X) = \mathbb{E}(-\log Pr(x = Y)) = \sum_{S \in \mathcal{S}} -Pr(x = \mathcal{S}) \log Pr(x = \mathcal{S})$$

This measures, as the following lemma makes precise, how random prefix free codes can be on $\mathcal{S}$.

LEMMA 14. *For any prefix free code $\mathcal{C} : \mathcal{S} \to$ words in $\{0,1\}$*

$$\mathbb{E}(\ell_S) \geq H(X).$$

Moreover, there exists a (prefix free) code $\mathcal{C}$ such that $\mathbb{E}(\ell(\mathcal{C}(S))) \leq H(X) + 1$.

PROOF. Let p_S and q_S be any sequences so that $\sum_{S \in \mathcal{S}} p_S = 1$ and $\sum_{S \in \mathcal{S}} q_S \leq 1$. Then by using Lagrange multipliers as in calculus on has that

$$-\sum_{S \in \mathcal{S}} p_S \log_2 q_S \geq -\sum_{S \in \mathcal{S}} p_S \log_2 p_S. \tag{4}$$

To prove the first statement let $p_S = Pr(x = S)$ and $q_S = 2^{-\ell_S}$. By the previous lemma, (4) applies. To prove the second statement, take $\ell_S = \lceil -\log_2 p_S \rceil$. Then the previous lemma applies again because $\sum 2^{-\ell_S} \leq 1$. It implies the existence of a code $\mathcal{C}$ which satisfies, applying (4) again,

$$\mathbb{E}(\ell(\mathcal{C}(S))) = \sum p_S \ell_S \leq -\sum p_S(\log_2(p_S) + 1) = H(X) + 1.$$

□

3.3. Measure theoretic entropy. We now assume $(X, \mathcal{B}, \mu, T)$ is a measure preserving metric space, and $\mathcal{P}$ is a countable partition of X. Then we define

$$H_\mu(\mathcal{P}) = -\sum_{P \in \mathcal{P}} \mu(P) \log \mu(P).$$

Also define $\mathcal{P}_a^b = \vee_{i=a}^b T^{-i}\mathcal{P}$ where $\mathcal{P} \vee \mathcal{Q}$ is the common refinement of the partitions $\mathcal{P}$ and $\mathcal{Q}$.

Note that

$$H_\mu(\mathcal{P} \vee \mathcal{Q}) \leq H_\mu(\mathcal{P}) + H_\mu(\mathcal{Q})$$

Applying this to $\mathcal{P}_a^b$, we get

$$H_\mu(\mathcal{P}_1^{m+n} \leq H_\mu(\mathcal{P}_1^m) + H_\mu(\mathcal{P}_{m+1}^{m+n})$$

and $H_\mu(\mathcal{P}_1^n) = H_\mu(\mathcal{P}_{m+1}^{m+n})$.

[1] I think it should say "for every $S \in \mathcal{S}$" at this point, but this wasn't what I copied down in my notes.

Let $a_n = H_\mu(\mathcal{P}_1^n)$. The *entropy* is defined to be $h(\mu, T, \mathcal{P}) = \lim a_n/n$ which exists, and $h(\mu, T) = \sup h(\mu, T, \mathcal{P})$.

We say $\mathcal{P}$ is generating if for every $\epsilon > 0$ and $B \in \mathcal{B}$ there exists N and $\widetilde{B}$, a union of atoms of $\mathcal{P}_{-N}^N$ such that $\mu(B \triangle \widetilde{B}) < \epsilon$.

THEOREM 15 (Kolmogorov). *If $\mathcal{P}$ is generating $h(\mu, T) = h(\mu, T, \mathcal{P})$.*

THEOREM 16 (Entropy Theorem, S-M-B). *Suppose $(X, \mu, \mathcal{B}, T)$ as above, $\mathcal{P}$ a countable partition with $H_\mu(\mathcal{P}) < \infty$. Then*

$$h(X, \mathcal{P}) = \lim -\frac{\log \mu(\mathcal{P}_1^n(X))}{n}$$

exists and $\int h(X, \mathcal{P})\ d\mu(x) = h(\mathcal{P})$.

If T is ergodic, can say that the limit of atoms tend toward the entropy.

3.4. Coding reformulation. There exists a function $h(x)$ so that $h(\mathcal{P}) = \int h(x)dx$ that satisfies

(1) For every sequence of prefix free codes $\mathcal{C}^N : \mathcal{P}_1^N \to$ words in $\{0, 1\}$,
$$\liminf \ell(\mathcal{C}_1^N(x))/n \geq h(x) \quad \text{a.e.}$$
(2) There is a sequence of prefix free codes such that
$$\limsup \ell(\mathcal{C}_1^N(x))/n \leq h(x) \quad \text{a.e.}$$

Claim: this is equivalent to the entropy theorem. To see this, define

$$h_n(x) = \frac{1}{n}\mathbb{E}(\log \mu(\mathcal{P}_1^n(\cdot)) \mid T\text{-invariant}).$$

Then $h_n(x) \to h(x)$. The first statement of the entropy theorem gives 1) by taking $\mathbb{C}_1^n$ to be the Shannon code for $\mathcal{P}_1^n$. Then $\ell(\mathcal{C}_1^n(x)) = \lceil -\log \mu(\mathcal{P}_1^n(X)) \rceil$.

To prove 2) let

$$Y_n = \{x \mid \ell(\mathcal{C}_2^n(x)) \leq (h(x) + \delta)n\},$$
$$Z_n = \{x \mid \mu(\mathcal{P}_1^n(x)) \leq 2^{-(h(x)+3\delta)n}\},$$
$$W(\alpha) = \{x \mid \alpha\delta \leq h(x) \leq (\alpha + 1)\delta\}.$$

It is enough to show that for almost every $x \in W(\alpha)$ if n is sufficiently large $x \notin Y_n \cap Z_n$. To do this estimate:

$$\mu(W(\alpha) \cap Y_n \cap Z_n) \leq \sum_{\ell(\mathcal{C}(\mathcal{P})) \leq (\alpha+2\delta)n} \mu(\mathcal{P}) \leq 2^{(\alpha+2\delta)m} \cdot 2^{-(\alpha+3\delta)n} = 2^{-\delta n}.$$

CHAPTER 2

June 2, 2009

1. Philippe

Review: We have $G = \mathrm{SO}_q(\mathbb{R})$, $H = \mathrm{SO}_{q,x_\infty}(\mathbb{R})$ and $\Gamma = \mathrm{SO}_q(\mathbb{Z})$. Understanding $R_q(d)$ is equivalent to understanding the distribution of a set of H-orbits of points $z_{a,b,c} \in \Gamma\backslash G$:

$$\bigsqcup_{[a,b,c]} z_{a,b,c}H \subset \Gamma\backslash G.$$

We have a correspondence

$$z_{a,b,c} \leftrightarrow \Gamma(a,b,c) \leftrightarrow \Gamma\backslash\Gamma g_{a,b,c}H$$

and $g_{a,b,c}x_\infty = \frac{(a,b,c)}{\sqrt{|d|}}$.

We remark that

$$\Gamma\backslash\Gamma g_{a,b,c}H = \Gamma\backslash\Gamma g_{a,b,c}Hg_{a,b,c}^{-1}g_{a,b,c} = \Gamma\backslash\Gamma T_{(a,b,c)}(\mathbb{R})g_{a,b,c}$$

where $T_{(a,b,c)}(\mathbb{R}) = \mathrm{SO}_{q,(a,b,c)}(\mathbb{R})$ is compact by the exercises.

Exercise 17. *Show that $\cup_{[a,b,c]}\Gamma\backslash\Gamma g_{a,b,c}H$ is compact. In particular, that there are only finitely many orbits associated with the representations of d by q. In other words, the set of Γ-orbits in $R_q(d)$ is finite. (This exercise is completely general for q any integral quadratic ternary form.)*

Remark: By Borel and Harish-Chandra a much more general result is true.

Today we will see that the set of H-orbits has an extra structure of a homogeneous space under a group. We'll see that the group is a class group, hence finite.

1.1. $q(a,b,c) = b^2 - 4ac$. We assume that d is squarefree. A solution to $q(a,b,c) = b^2 - 4ac$ corresponds to

$$m_{a,b,c} = \begin{pmatrix} -b & 2c \\ 2a & b \end{pmatrix} \in M_2^0(\mathbb{Z})$$

which has the property that $m_{a,b,c}^2 = d$.

This gives an injection

$$i_{a,b,c} : K = \mathbb{Q}(\sqrt{d}) \to M_2^0(\mathbb{Z}) \quad x + y\sqrt{d} \mapsto x \cdot Id + y \cdot m_{a,b,c}.$$

Then $\Gamma(a,b,c)$ corresponds to a Γ-conjugacy class of $i : K \hookrightarrow M_2(\mathbb{Q})$ along with an integrality property:

$$i(K) \cap M_2(\mathbb{Z}) = i(\mathcal{O}_d) \tag{5}$$

where $\mathcal{O}_d$ is the ring of integers in K.

In (5) we only consider those (a,b,c) that are primitive. If (a,b,c) represents d then (fa,fb,fc) represents f^2d. This is called a *non-primitive* representation. If a representation can not be written in this way for some f it is *primitive*. Then let $R_q^*(d)$ be the set of primitive representations. We don't lose any information in studying primitive solutions because

$$R_q^*(d) = \bigsqcup_{f^2|d} fR_q^*(d/f^2)$$

We have that

$$\left\{ \begin{array}{c} \Gamma\text{-orbits of primitive} \\ \text{representations of } (a,b,c) \end{array} \right\} \leftrightarrow \left\{ \begin{array}{c} \Gamma\text{-conjugacy classes of} \\ \mathcal{O}_d\text{-integral embeddings of} \\ K \hookrightarrow M_2(\mathbb{Q}) \end{array} \right\}$$

Under the embedding $i_{a,b,c}(K^\times) \hookrightarrow \mathrm{GL}_2(\mathbb{Q})$, $T_{a,b,c} = \mathbb{Q}^\times \backslash i(K^\times) \subset \mathrm{PGL}_2(\mathbb{Q})$. So

$$\Gamma\backslash\Gamma T_{a,b,c}(\mathbb{R}) \subset \mathrm{PGL}_2(\mathbb{Z})\backslash\mathrm{PGL}_2(\mathbb{R}) = \left\{ \begin{array}{c} \text{space of lattices} \\ \text{up to homothety} \end{array} \right\} = \left\{ \begin{array}{c} \text{unimodular} \\ \text{matrices} \end{array} \right\}$$

By the exercise, the shortest vector has length bounded below by something independent of a,b,c.

We now study the structure of the space of orbits. Pick $y \in \mathbb{Q}^2 \setminus \{0\}$. Define the $\mathbb{Q}$-vector space homomorphism

$$j_y : K \to \mathbb{Q}^2 \qquad \text{via} \qquad x \mapsto i(x)y.$$

We have $i(x)j_y(z) = j_y(xz)$. Let $I = j_y^{-1}(\mathbb{Z}^3)$. This is a rank 2 $\mathbb{Z}$-module in K. Moreover, it is an $\mathcal{O}_d$-ideal because of the integrality property (5).

If one replaces y by y' then $y' = i(\lambda)y$ from which one deduces that $I_{y'} = \lambda^{\pm 1} I_y$. (It is either λ or λ^{-1}, but Philippe wasn't sure which.) In any case, I_y and I'_y are homothetic in K.

We conclude that to each $i_{a,b,c}$ we obtain a homothety class of $\mathcal{O}_d$-ideals in K. Assuming (a,b,c) primitive implies that the corresponding $\mathcal{O}_d$-ideals are *proper*, meaning $\{\lambda \in K \mid \lambda I \subset I\} = \mathcal{O}_d$. Moreover, if we replace (a,b,c) by $\gamma(a,b,c)$ for some $\gamma \in \Gamma$ this changes $\mathbb{Z}^2$ by γ, hence has no effect. So there is a bijection

$$\left\{ \begin{array}{c} \Gamma(a,b,c)\text{-orbits of} \\ \text{primitive representations} \end{array} \right\} \leftrightarrow \left\{ 0 \neq I \subset K \mid \begin{array}{c} I\mathcal{O}_d = I \\ I \text{ is a proper } \mathcal{O}_d\text{-ideal} \end{array} \right\} / K^\times.$$

This is the classgroup ($Cl(\mathcal{O}_d)$) or the Picard group ($\mathrm{Pic}(\mathcal{O}_d)$.) We conclude that the space of orbits is a principal homogeneous space under $\mathrm{Pic}(\mathcal{O}_d)$.

Remark: I proper implies that there exists I^{-1} such that $II^{-1} = \mathcal{O}_d$.

1.2. $q(a,b,c) = -(a^2+b^2+c^2)$**.** In this case there is a correspondence, not with a matrix algebra, but rather with the Hamilton quaternions. Indeed, the vector space $(\mathbb{Q}^3, q)$ is isomorphic to $(B^0_{\mathbb{Q}}, -N)$ via $(a,b,c) \mapsto ai+bj+ck$. Recall that the quaternions are the 4-dimensional vector space with basis $1,i,j,k$ given a ring structure under the relations

$$i^2 = j^2 = k^2 = -1, \quad ijk = -1$$

with norm $N(z) = N(u+ai+bj+ck) = z\overline{z}$.

Under this vector space isomorphism SO_q corresponds to $PB^\times$. The algebraic group $PB^\times$ has $\mathbb{Q}$-points $PB^\times(\mathbb{Q}) = B^\times(\mathbb{Q})/\mathbb{Q}^\times$ which act by conjugation on $B_{\mathbb{Q}}$.

We choose $x_\infty = (0,0,1)$, so

$$H \leftrightarrow (\mathbb{R} + \mathbb{R}k)/\mathbb{R}^\times \simeq \mathbb{C}^\times/\mathbb{R}^\times = S^1/\{\pm 1\}.$$

The correspondence $\mathrm{SO}_q(\mathbb{Z}) \leftrightarrow PB^\times(\mathbb{Z})$ gives the analog to $M_2(\mathbb{Z})$ of the previous case, and $\Gamma = PB^\times(\mathbb{Z}) = \mathcal{O}_B^\times/\{\pm 1\}$ where $\mathcal{O}_B = \mathbb{Z}[i,j,k,\frac{i+j+k}{2}]$.

As in the previous section we have $z_{a,b,c} = ai + bj + ck$ satisfies $z_{a,b,c}^2 = d$ so we have a conjugacy class of embeddings $i : K = \mathbb{Q}(\sqrt{d}) \hookrightarrow B(\mathbb{Q})$ such that $\mathcal{O}_B \cap i(K) = i(\mathcal{O}_d)$.

Choose $I \subset K$ a proper $\mathcal{O}_d$-ideal. Then $i(I)\mathcal{O}_B$ is a right $\mathcal{O}_B$-module in $B_\mathbb{Q}$. Essentially by Lagrange's theorem on representations of integers by sums of four squares, one deduces that $\mathcal{O}_B$ is a PID. Hence $i(I)\mathcal{O}_B = q\mathcal{O}_B$.

The map

$$\mathrm{Ad}(q) \circ i_{a,b,c} : K \to B(\mathbb{Q})^\times$$

is $\mathcal{O}_d$-integral hence corresponds to $i_{a',b',c'}$ for some $(a',b',c') \in R_q(d)$. Integrality: $q^{-1} i_{a,b,c} q \cap \mathcal{O}_B = i(\mathcal{O}_d)$.

2. Elon

Last time we had a measure preserving transformation $(X, \mu, \mathcal{B}, T)$. We had that $H_\mu(\mathcal{P}) < \infty$,

$$h(x,\mathcal{P}) = \lim_{n\to\infty} \frac{-\log \mu(\mathcal{P}_1^n(x))}{n} \quad \text{and} \quad h(\mathcal{P}) = \int h(x,\mathcal{P})\, d\mu(x).$$

(In the case that T is ergodic, $h(x,\mathcal{P}) = h(x)$ almost everywhere.)

There is also a coding reformulation: there exists a function $h(x,\mathcal{P})$ such that $\int h(x,\mathcal{P}) = h$ for any sequence of codes $\mathcal{C}^N : \mathcal{P}^N \to$ words in $\{0,1\}$, and

(1) $\liminf \frac{\ell(\mathcal{C}^N(x))}{N} \geq h(x)$ almost everywhere, and
(2) there exists some sequence of codes $\mathcal{C}_2^N : \mathcal{P}^N \to$ words in $\{0,1\}$ such that

$$\limsup \frac{\ell(\mathcal{C}^N(x))}{N} \leq h(x)$$

almost everywhere[1].

We'll talk more about 2):

$$h_n(x) = \frac{1}{n}\mathbb{E}(-\log \mu(\mathcal{P}_1^n(y)) \mid T\text{-inv})(x)$$

$$\int h_n(x) = \frac{1}{n}\int -\log \mu(\mathcal{P}_1^n(y)) d\mu(y) = \frac{1}{n} H_\mu(\mathcal{P}_1^n)$$

Building blocks:

- Shannon code for $\mathcal{P}$ (or any code of finite expected length) for $\mathcal{C}'$.
- Shannon code for $\mathcal{P}_1^n$ for $\mathcal{C}^n$. (The Shannon code satisfies $\ell(\mathcal{P}_1^n(x)) = \lceil -\log \mu(\mathcal{P}_1^n(x)) \rceil$.)

[1] I don't know why in 2) he uses '$\leq$' instead of '$=$'

2.1. Entropy in compact metric spaces. Let (X, d) be a compact metric space. Define

$$B_{r,n}(x) = \{y \mid d(T^k x, T^k y) < r \text{ for all } |k| < n/2\}$$

THEOREM 18 (Brin-Katok). *Let*

$$\underline{h}(x,r) = \liminf -\frac{1}{n}\log\mu(B_{r,n}(x)) \quad \text{and} \quad \overline{h}(x,r) = \limsup -\frac{1}{n}\log\mu(B_{r,n}(x))$$

Then

$$\lim r \searrow 0 \underline{h}(x,r) = \lim r \searrow 0 \overline{h}(x,r) = h(x) = h(\mu_x^\epsilon = \sup_{\mathcal{P}} h(x,\mathcal{P}).$$

2.2. exercises.

3. Manfred

Recall

$$X_d = \mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_d(\mathbb{R}) = \{\text{unimodular lattices in } \mathbb{R}^d\}$$

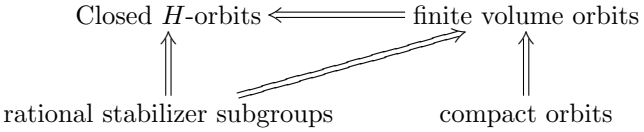
Picture: (He drew a tear drop shaped surface not connecting at the 'tip' because that is the cusp. However, it is a finite volume space.) Define

$$X_d(\epsilon) = \left\{\Lambda \in \mathbb{R}^d \mid \begin{array}{c}\Lambda \text{ is unimodular, and} \\ \Lambda \cap B_\epsilon^{\mathbb{R}^d}(0) = \{0\}\end{array}\right\}$$

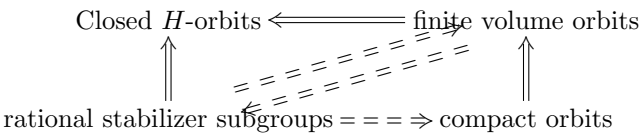
THEOREM 19 (Mahler's compactness theorem). *A subset $L \subset X_d$ is compact if and only if it is closed and there exists $\epsilon > 0$ such that $L \subset X_d(\epsilon)$.*

PROOF. Need to find a compact subset in $\mathrm{SL}_d(\mathbb{R})$ that mas onto L. For any Γg there exists a representative $g = \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix}$ such that $\|v_i\| \asymp \lambda_i$. We know from Minkowski's theorem on successive minima that $\lambda_1\lambda_2 \cdot \lambda_d \asymp 1$. If $\Gamma g \in X_d(\epsilon)$ then $\lambda_1 \geq \epsilon$ hence $\lambda_i \geq \epsilon$ for all i. Therefore, λ_d (as well as $\|g\|$) $\ll \epsilon^{-(d-1)}$. □

The groups $H \subset G$, $\Gamma = \mathrm{SL}_d(\mathbb{Z})$ and $G = \mathrm{SL}_d(\mathbb{R})$ sit in the following picture:



In general one has:



where the dashed lines mean that sometimes you get such an implication. That a finite volume comes from a rational stabilizer subgroups is the Borel density theorem. The opposite implication is a result of reduction theory or non-divergence.

THEOREM 20 (Borel density theorem). *Suppose $H \subset \mathrm{SL}_d(\mathbb{R})$, $\Lambda \subset H$ a lattice. Then the Zariski closure of Λ contains all unipotent elements of H and all elements of H that are diagonalizable with positive eigenvalues.*

This doesn't say anything if H is compact. However, if $H = H(\mathbb{R})$ is the group of $\mathbb{R}$ points of an algebraic group over $\mathbb{R}$ and $\Lambda = \mathrm{SL}_d \cap H$ is a lattice, and H is generated by unipotents and positive diagonalizable elements, then H is defined over $\mathbb{Q}$. This means that there is a $\rho : \mathrm{SL}_d \to \mathrm{SL}_{d'}$ and $v \in \mathbb{R}^d$ such that H is the stabilizer group of v.

Remark: If one can't get a compact orbit ($\to$) for a situation like the exercise, one can do $\nearrow$.

The tools that are used to prove the Borel density theorem come from the most elementary (but nontrivial) results from ergodic theory and the theory of algebraic groups. These are, respectively, the Poincaré and Chevellay theorems.

LEMMA 21. *Suppose $\Lambda \subset H$ is discrete and $X = \Lambda\backslash H$ supports a right H-invariant finite measure μ_X. Then Λ is a lattice.*

PROOF. We need to show that H is unimodular: $\mu_H = m_H^{\text{right}}$. Take $f \in C_c(H)$. Then

$$\sum_{\Lambda h = \Lambda g} f(g) = F(\lambda H) \in C_c(\Lambda\backslash H),$$

and $\int f \, d\mu_H = \int F \, d\mu_X$ is right H-invariant. It is also left Λ-invariant. We want to show that it is left H-invariant.

Let $h \in H$ and $g \approx e$. Then, by Poincaré recurrence $\Lambda g h^{n_k} \to \Lambda g$ for some increasing sequence n_k. Thus $\gamma_k g h^{n_k} \to g$ for some $\gamma_k \in \Lambda$. Apply the modular character to this sequence:

$$mod_H(\gamma_k) mod_H(g) mod_H(h^{n_k}) \to mod_H(g).$$

Since g is arbitrarily close to e we may remove the terms $mod_H(g)$. Hence

$$mod_H(\gamma_k) mod_H(h^{n_k}) \to 1.$$

Since $mod_H(\gamma_k) = 1$ and h is arbitrary, we must have $mod_H(h) = 1$. □

3.1. Other fields. We have always worked with $\mathbb{R}$ so far, but it is possible to use $\mathbb{Q}_p$. It is a fact that

$$\mathbb{Z}[\frac{1}{p}] \hookrightarrow \mathbb{R} \times \mathbb{Q}_p \qquad a \mapsto (a, a)$$

is discrete. So we can play the same game:

$$X_d^{\text{new}} = \mathrm{SL}_d(\mathbb{Z}[\frac{1}{p}])\backslash \mathrm{SL}_d(\mathbb{R} \times \mathbb{Q}_p)$$

is closely related to $\mathrm{SL}_d(\mathbb{Z})\backslash \mathrm{SL}_d(\mathbb{R})$.

By Gauss elimination, one can show that

$$\mathrm{SL}_d(\mathbb{Z}[\frac{1}{p}])\mathrm{SL}_d(\mathbb{R} \times \mathbb{Z}_p) = \mathrm{SL}_d(\mathbb{Z}[\frac{1}{p}])\mathrm{SL}_d(\mathbb{R} \times \mathbb{Q}_p).$$

It follows that

$$\mathrm{SL}_d(\mathbb{Z})\backslash \mathrm{SL}_d(\mathbb{R}) \simeq \mathrm{SL}_d(\mathbb{Z}[\frac{1}{p}])\backslash \mathrm{SL}_d(\mathbb{R} \times \mathbb{Q}_p)/\mathrm{SL}_d(\mathbb{Z}_p).$$

This is saying that X_d^{new} is a compact extension (by $\mathrm{SL}_d(\mathbb{Z}_p)$) of X_d. In particular, it has finite volume.

Besides with $(\mathbb{R} \times \mathbb{Q}_p, \mathbb{Z}_p)$, this can be done with $(\mathbb{A}, \mathbb{Q})$ as well.

THEOREM 22 (Dani). *Let* $X_2 = \mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$, $H = \{\left(\begin{smallmatrix} 1 & t \\ & 1 \end{smallmatrix}\right) \mid t \in \mathbb{R}\}$ *be horocycle subgroup. Any* H*-invariant and ergodic probability measure on* X_2 *is either the natural measure on a periodic orbit or the Haar measure on* X*.*

The periodic orbits in X_2 are (the images of) horizontal line segments in the upper half plane between $\mathrm{Re}(s) = -1/2$ and $\mathrm{Re}(s) = 1/2$. The natural measure on them is thus the Lebesgue measure on $\mathbb{R}$.

THEOREM 23 (Donis-Smillie). *Either* x *is periodic or* $x\left(\begin{smallmatrix} 1 & t \\ & 1 \end{smallmatrix}\right)$ *becomes equidistributed in* X_2 *with respect to* m_{X_2}*.*

CHAPTER 3

June 3, 2009

1. Manfred

We want to know that the measures $\mu_{X,T}$ defined by

$$\int f \, d\mu_{X,T} = \frac{1}{T}\int_0^T f(xu(t))dt$$

(for any 1-parameter subgroup $u(t)$) have only probability measures as weak* limits. But, geodesic flow doesn't have this property. Think of the orbit of the line $\mathrm{Re}(s) = 0$; the measure from this flow won't be probability. For horocycle orbits this behavior doesn't occur as evidenced by the following.

THEOREM 24 (Margulis, Dani). *For any K compact there exists $L \subset X_2$ compact such that if $x \in K$ and $T > 0$ then $\frac{1}{T} m_{\mathbb{R}}(\{t \in [0,T] \mid xu(t) \in L\}) < \epsilon$.*

This is saying (?) that equivalent weak* limits are probability measures.

PROOF. We may assume that $K = X_2(\eta)$. Want to find $L = X_2(\delta)$. Recall that $x \in K \iff x = \Lambda \subset \mathbb{R}^2$ satifies $\Lambda \cap B_\eta(0) = \{0\}$. Let $u(t) = \left(\begin{smallmatrix} 1 & t \\ & 1 \end{smallmatrix}\right)$. If $xu(t) \notin L$ then there exists $v = (v_1, v_2) \in \Lambda$ such that $\|(v_1, v_2)\left(\begin{smallmatrix} 1 & t \\ & 1 \end{smallmatrix}\right)\| < \delta$. The question is for how many values of t does this happen?

If $(v_1, v_2)\left(\begin{smallmatrix} 1 & t \\ & 1 \end{smallmatrix}\right)$ is δ small then $|v_1| < \delta$ and $|v_2 + tv_1| < \delta$. In other words,

$$t \in \Big[-\frac{v_1}{v_2} - \frac{\delta}{|v_1|}, -\frac{v_1}{v_2} + \frac{\delta}{|v_1|}\Big] \tag{6}$$

We denote each of these intervals as $bad(\delta, v)$. We also define intervals $protect(v)$ defined as in (6) except we replace δ with $1/2$. We think of δ as being much smaller than $\eta/2$ so that each bad interval is contained within a protect interval.

So the game is to show that the protect intervals are disjoint. If so, since $\big|bad(v^{(i)})\big| < \epsilon \big|protect(v^{(i)})\big|$, the size of all of the bad sets is less than ϵ times the size of the union of protected sets (which must be less that T.) We claim that if we only consider $\{v\}$ that are not multiples of any smaller vector then the protect sets are disjoint.

Suppose $v, w \in \Lambda$ are $\mathbb{Z}$-linearly independent and $t \in [0,T]$ are such that $vu(t)$ and $wu(t)$ satisfy

$$|v_1| < \delta \le \frac{\eta}{2}, \quad |w_1| < \delta \le \frac{\eta}{2}, \quad |v_2 + tv_1| < \delta \le \frac{\eta}{2}, \quad |w_2 + tw_1| < \delta \le \frac{\eta}{2}.$$

Then $|vu(t)| < \eta$ and $|wu(t)| < \eta$, a contradiction. □

We can now prove the following.

THEOREM 25. *The space X_d has finite volume, and $\mathrm{SL}_d(\mathbb{Z})\backslash \mathrm{SL}_d(\mathbb{Z})\mathrm{SO}(Q)(\mathbb{Z})$ has finite volume if Q is any quadratic form in $d \ge 3$ variables.*

PROOF. We know that X_d supports a right invariant Haar measure, but we don't know that it is finite. Let $K \subset X_d = X$ be compact with $m_X(K) > 0$. Find L for K and $\epsilon = \frac{1}{2}$ of the previous theorem. So

$$f(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \chi_L(xu(t))dt \geq \frac{1}{2}$$

on K. (χ_L is the characteristic function of L.)

Note that the convergence of $f(x)$ is guaranteed by Fatou's lemma. If we define the integrand to be f_T, then we can apply Fatou's lemma to f_T^2. We have

$$\|f_T\|_2 \leq \|\chi_L\| = \sqrt{m_X(L)} < \infty.$$

So,

$$\|f\|^2 = \int (\liminf f_T)^2 = \int \liminf f_T^2 \leq \liminf \int f_t^2$$

which is finite. (Fatou's lemma is applied to get the final inequality.)

Clearly, f is invariant under $u(s)$. The claim is that f is constant. This would then imply that the space has finite measure. The claim, and hence, the proof follows by the Mautner phenomenon. □

PROPOSTION 26 (Mautner phenomenon). *Suppose that* $\mathrm{SL}_d(\mathbb{R})$ *acts by* π *unitarily on a Hilbert space* $\mathcal{H}$*, and* $g \mapsto \langle \pi(g)v, v\rangle$ *is continuous for all* $v \in \mathcal{H}$*. If* $v \in \mathcal{H}$ *is fixed by one unipotent* $u(t)$ *or a positive diagonal matrix* a *(not the identity) then* v *is fixed by* $\mathrm{SL}_d(\mathbb{Z})$*.*

PROOF. Let $v \in \mathcal{H}$, and $\phi(g) = \langle \pi(g)v, v\rangle$. We prove the case $d = 2$. Let $a = \left(\begin{smallmatrix} e & \\ & 1/e \end{smallmatrix}\right)$. Suppose v is fixed by a. Then

$$\phi(a^k g a^l) = \langle \pi(a^k g a^l)v, v\rangle = \langle \pi(g)\pi(a^l)v, \pi(a^k)v\rangle = \langle \pi(g)v, v\rangle = \phi(g).$$

Now let $e = \left(\begin{smallmatrix} 1 & \\ \epsilon & 1 \end{smallmatrix}\right)$ so that $\phi(g) \approx \|v\|^2$. Moreover,

$$= \phi(\left(\begin{smallmatrix} 1 & \\ e^{2k}\epsilon & 1 \end{smallmatrix}\right)) = \phi(a^{-k} g a^k) = \phi(g) \approx \|v\|^2.$$

In other words, $\phi(\left(\begin{smallmatrix} 1 & \\ t & 1 \end{smallmatrix}\right) g) \approx \|v\|^2$ for all t, and the $\approx$ must actually be $=$. Cauchy-Schwartz tells us that

$$\langle \pi(\left(\begin{smallmatrix} 1 & \\ t & 1 \end{smallmatrix}\right) g)v, v\rangle = \|v\|^2 \iff \pi(\left(\begin{smallmatrix} 1 & \\ t & 1 \end{smallmatrix}\right))v = v.$$

One argues similarly to show that $u(t)$ fixes v for all t. Since $\mathrm{SL}_2(\mathbb{R})$ is generated by this subgroups the proposition follows.

In the case that $\pi(u(s))v = v$, we again take g as above and calculate that

$$u^k g u^l = \begin{pmatrix} 1 + k\epsilon & * \\ \epsilon & 1 - l\epsilon \end{pmatrix}$$

where $*$ is a quadratic in k and l which can be made to be as close to zero as one wishes. Now one proceeds as in the diagonal case. □

The following theorem deals with mixing of $\mathrm{SL}_d(\mathbb{R})$ and vanishing of matrix coefficients.

THEOREM 27 (Howe-Moore theorem). *Suppose* $\mathrm{SL}_d(\mathbb{R})$ *acts unitarily on a Hilbert space, and there are no fixed vectors. Then* $\langle \pi(g)v, w\rangle \to 0$ *as* $g \to \infty$*.*

For example, on $X = \Gamma\backslash \mathrm{SL}_2(\mathbb{R})$ and $\mathcal{H} = L_0^2(X)$ the geodesic and horocycle flows are mixing.

THEOREM 28 (Sarnak). *The long periodic orbits for the horocyle flow on X_2 equidistributes on X_2 with respect to M_{X_2}.*

PROOF. As discussed previously, the periodic orbits are line segments from $\mathrm{Re}(z) = -1/2$ to $\mathrm{Re}(z) = 1/2$ at height y. The length of such an orbit is $\frac{1}{y}$. The region

$$Q = \mathrm{SL}_2(\mathbb{Z}) \begin{pmatrix} 1 & [0,1] \\ & 1 \end{pmatrix} B_\delta^A(e) \begin{pmatrix} 1 & 0 \\ [-1,1] & 1 \end{pmatrix} \tag{7}$$

is a neighborhood of the periodic orbit at $\mathrm{SL}_2(\mathbb{Z})$. Multiply (7) on the right by $a = \begin{pmatrix} \sqrt{y} & \\ & 1/\sqrt{y} \end{pmatrix}$:

$$\begin{aligned} Q &= \mathrm{SL}_2(\mathbb{Z}) a a^{-1} \begin{pmatrix} 1 & [0,1] \\ & 1 \end{pmatrix} a a^{-1} B_\delta^A(e) a a^{-1} \begin{pmatrix} 1 & 0 \\ [-1,1] & 1 \end{pmatrix} a \\ &= \mathrm{SL}_2(\mathbb{Z}) a \begin{pmatrix} 1 & [0,1/y] \\ & 1 \end{pmatrix} B_\delta^A(e) \begin{pmatrix} 1 & 0 \\ [-y,y] & 1 \end{pmatrix} \end{aligned}$$

which is a neighborhood around the periodic orbit at $1/y$. Now take $f \in C_c(X_2)$. Continuity of f implies that

$$\int_{\text{long period}} f \, dt \approx \langle f, \frac{1}{m(Q)} \chi_Q \circ a \rangle.$$

Now mixing (the Howe-Moore theorem) implies that this goes to $\int f$. □

Remark: The Haar measure on N^+AN^- *is* the product of the Lebesgue measures on $N^\pm$ and A, so this proof is valid.

2. Elon: Entropy in $\Gamma\backslash G$

THEOREM 29 (Brin-Katok). *Let $B_{r,m}(x) = \{g \mid d(T^k x, T^k y) < r, \ |k| \le n/2\}$. Denote*

$$\underline{h}(x,r) = \liminf -\frac{1}{n} \log \mu(B_{r,m}(x))$$

and $\overline{h}(x,r)$ to be the analogous limsup. Then

$$h(\mu_x^s) = h(x) = \lim_{r \to 0} \underline{h}(x,r) = \lim_{r \to 0} \overline{h}(x,r)$$

3. Philippe

We first clarify the integrality condition discussed at the end of last time. Recall that we are in the case $q(a,b,c) = -(a^2+b^2+c^2)$. A solution to $q(a,b,c) = d$ is equivalent to a map

$$i = i_{a,b,c} : K = \mathbb{Q}(\sqrt{d}) \hookrightarrow B_{\mathbb{Q}} \quad \text{via} \quad \sqrt{d} \mapsto ai + bj + ck.$$

Then $(a,b,c) \in \mathbb{Z}^3 \iff i(K) \cap B(\mathbb{Z}) = \mathcal{O}_B = i(\mathcal{O}_d)$.

Suppose $I \subset K$ is an $\mathcal{O}_d$-ideal. Then $i(I)\mathcal{O}_B = q_I \mathcal{O}_B$ for some $q_I \in B(\mathbb{Q})^\times$. So

$$i' = \mathrm{Ad}\, q_I \circ i : \lambda \in K \mapsto q_I^{-1} i(\lambda) q_I \in B(\mathbb{Q})$$

corresponds to a solution (a',b',c') determined by $i'(\sqrt{d}) = a'i + b'j + c'k$. We claim that $(a',b',c') \in \mathcal{O}_B$ which, by the above comment, happens if and only if

$(a', b', c') \in \mathbb{Z}^3$. We calculate

$$\begin{aligned} i'(\sqrt{d}) = q_I^{-1} i(\sqrt{d}) q_I &\in q_I^{-1} i(\sqrt{d}) q_I \mathcal{O}_B \\ &= q_I^{-1} i(\sqrt{d}) q_I && \in q_I^{-1} i(\sqrt{d}) i(I) \mathcal{O}_B \\ &\subset q_I^{-1} i(\sqrt{d}) q_I && \in q_I^{-1} i(I) \mathcal{O}_B \qquad \text{since } \sqrt{d} I \subset I \\ &= q_I^{-1} q_I \mathcal{O}_B = \mathcal{O}_B. \end{aligned}$$

So $(I, i) \mapsto i'$ defines an action of $\operatorname{Pic} \mathcal{O}_K$ on the points of $R_q(d)$ if and only if $\operatorname{Pic} \mathcal{O}_K$ acts on the Γ-orbit $\Gamma(a, b, c)$. Moreover, this action preserves primitivity: if not, $(a', b', c') = f(a'', b'', c'')$ for some $f \in \mathbb{Z}$. Apply the ideal I^{-1} (which exists by assumption) to i' and this would imply that (a, b, c) is not primitive.

Remark: To show that the composition of maps

$$q = b^2 - 4ac \to m_{a,b,c} = \begin{pmatrix} -b & 2c \\ 2a & b \end{pmatrix} \in \mathrm{PGL}_2(\mathbb{Z}) \backslash \mathrm{PGL}_2(\mathbb{R}) \to \Gamma \backslash \mathbb{H}$$

has image $z = \frac{-b + i\sqrt{|d|}}{2a}$ do not compute $g_{a,b,c}$.

3.1. Adèles. The completions of $\mathbb{Q}$ are $\mathbb{R}, \mathbb{Q}_2, \mathbb{Q}_3 \mathbb{Q}_5, \ldots$. Obviously, $\mathbb{Q}$ embeds into each. Hence

$$\mathbb{Q} \hookrightarrow \mathbb{R} \times \prod_p \mathbb{Q}_p = \prod_v \mathbb{Q}_v.$$

The problem is that under the product topology this group is not locally compact. However, because only finitely many primes divide the denominator of any $\lambda \in \mathbb{Q}$, $\lambda \in \mathbb{Z}_p$ for all but finitely many primes. So

$$\mathbb{Q} \hookrightarrow \prod\nolimits^* \mathbb{Q}_v = \{(q_v) \mid q_p \in \mathbb{Z}_p \text{ for almost all } p\} = \mathbb{A}_{\mathbb{Q}}$$

which is locally compact under the restriction of the product topology. The embedding of $\mathbb{Q} \hookrightarrow \mathbb{A}_{\mathbb{Q}}$ is discrete: a neighborhood of $\{0\}$ intersection $\mathbb{Q}$ trivially is $(-1/2/, 1/2) \times \prod \mathbb{Z}_p$.

There is also an embedding of $\mathbb{Q}_v$ into $\mathbb{A}_{\mathbb{Q}}$ given by

$$\lambda_v \mapsto (0, \cdots, 0, \underbrace{\lambda_v}_{v\text{-th place}}, 0, \cdots)$$

We denote the *finite adèles* $\prod_p^* \mathbb{Q}_p$ by $\mathbb{A}_f$. Inside $\mathbb{A}_f \supset \widehat{\mathbb{Z}} = \prod_p \mathbb{Z}_p$ which is a maximal compact subring of $\mathbb{A}_f$. The fact that $\mathbb{Q} \hookrightarrow \mathbb{A}_f$ is dense is equivalent to the Chinese remainder theorem. Together with the discreteness of $Q \hookrightarrow \mathbb{A}$ this implies that $\mathbb{Q}$ is a lattice. The set $[-1/2, 1/2) \times \widehat{\mathbb{Z}}$ is a fundamental domain.

3.2. Idèles. The *idèles* are $\mathbb{A}^\times$ but the topology on them is not the restriction of that on $\mathbb{A}$. Rather, the following embedding is closed

$$\mathbb{A}^\times \hookrightarrow \mathbb{A} \times \mathbb{A} \qquad \text{via} \qquad x \mapsto (x, x^{-1})$$

and $\mathbb{A}^\times$ is given the topology obtained by restriction the topology of $\mathbb{A} \times \mathbb{A}$.

The maximal compact $\widehat{\mathbb{Z}}^\times$ sits in the *finite idèles* $\mathbb{A}_f^\times$. The inclusion $\mathbb{Q}^\times \hookrightarrow \mathbb{A}_f^\times$ is discrete. It is not dense but close to it: $\mathbb{Q}^\times \widehat{\mathbb{Z}}^\times = \mathbb{A}_f^\times$. More generally, if U is any

compact open subgroup, $\#\mathbb{A}_f^\times/U < \infty$. For instance, for $N \geq 1$, define

$$U_N = \prod_p U_{N,p} \qquad \text{where} \quad U_{N,p} = \begin{cases} \mathbb{Z}_p^\times & \text{if } p \nmid N \\ 1 + p^{v_p(N)}\mathbb{Z}_p & \text{otherwise.} \end{cases}$$

Then $\mathbb{A}_f^\times/\mathbb{Q}U_N \simeq (\mathbb{Z}/N\mathbb{Z})^\times$.

3.3. More generally. If G is an algebraic group over $\mathbb{Q}$, $G(\mathbb{Q}) \hookrightarrow G(\mathbb{Q})$ is discrete, but it need not be dense. However, if G is semisimple or if $G(\mathbb{Q})$ is simply connected then strong approximation says that

$$G(\mathbb{Q})G(\widehat{\mathbb{Z}}) = G(\mathbb{A}).$$

In this case,

$$G(\mathbb{Q})\backslash G(\mathbb{A})/G(\widehat{\mathbb{Z}}) \simeq G(\mathbb{Z})\backslash G(\mathbb{R}). \tag{8}$$

In particular, this holds for $G = \mathrm{SL}_d$ (which is semisimple.) Let $K = G(\widehat{\mathbb{Z}})$, and $\Gamma = G(\mathbb{Q})$. The map in one direction is

$$\Gamma\backslash\Gamma g_\mathbb{R} \mapsto G(\mathbb{Q})\backslash(g_\mathbb{R}, e_f)/K.$$

An element of the double coset can be represented as $(g_\mathbb{R}, g_f)$. Write $g_f = g_\mathbb{Q} g_{\widehat{\mathbb{Z}}}$ for $g_\mathbb{Q} \in G(\mathbb{Q})$ and $g_{\widehat{\mathbb{Z}}} \in G(\widehat{\mathbb{Z}})$. Then the map in the other direction is

$$(g_\mathbb{R}, g_f) = (g_\mathbb{R}, g_\mathbb{Q} g_{\widehat{\mathbb{Z}}}) \equiv (g_\mathbb{Q}^{-1} g_\mathbb{R}, 1) \mapsto g_\mathbb{Q}^{-1} g_\mathbb{R}.$$

CHAPTER 4

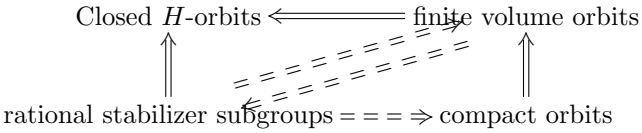
June 4, 2009

1. Elon: Measure classification results that involve entropy

Last time we finally related entropy to dynamics on $\Gamma\backslash G$ and developed sufficiently adequate methods to calculate entropy. Now we're going to start using it.

2. Manfred

Recall that last time we discussed the following diagram:



The implication that algebraic stabilizer groups give finite volume orbits follows if the group is semisimple or if there are no $\mathbb{Q}$-characters.[1] If there are no unipotents in Λ one gets that finite volume orbits are compact.

Recall that we took a neighborhood around the periodic orbit $\mathrm{SL}_2(\mathbb{Z})\left(\begin{smallmatrix}1 & [0,1]\\ 0 & 1\end{smallmatrix}\right)$ and moved it to height y. As $y \to 0$ this equidistributes.

Remark: as a matter of terminology, finite volume orbits are sometimes called *periodic* or *homogeneous* or *algebraic* in the literature.

THEOREM 30 (Furstenberg). *The subgroup* $H = \begin{pmatrix} 1 & * \\ & 1 \end{pmatrix}$ *is uniquely ergodic on compact quotients* $\Gamma\backslash \mathrm{SL}_2(\mathbb{R})$.

A proof of this can be given using the same strategy. This strategy also proves the following.

THEOREM 31 (Dani). *On a non-compact quotient, an ergodic* H*-invariant probability measure is either periodic or Haar.*

Briefly, the strategy is to take a short periodic orbit, thicken it slightly, pull it back to a longer periodic orbit, take the limit, use mixing to get the result. Conceivably, this would not work if the short orbit is far out in the cusp as is the case with a geodesic orbit that goes straight up the cusp. However, this "counterexample" is not a unipotent flow, and the upshot is that unipotent flows never have this type of behavior.

A picture of all ergodic invariant measures on $\Gamma\backslash \mathrm{SL}_2(\mathbb{R})$ in the weak* topology: If $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ you have a line with the Haar measure at one side and 0 at the other (which is not a probability measure) connected by periodic orbits. So it is a ray.

[1] I don't know what this means OR if I corrected copied it down.

Other choices of Γ would give the same picture but with different rays emanating from the same point corresponding to each of the cusps.[2]

One can compare this to the example of the torus T^2. In this case, H is the projection of a line with rational slope. Note that each slope gives a circle of probability measures. (Explicitly, one take H to be given by the line $ay = bx$ with a and b relatively primes integers. Then H fixes the probability measure defined by the projection of $ay = bx + c$ for $c \in [0, 1)$. So there is a circle's worth of invariant measures for each rational number a/b.) For any sequence of rationals that converges to an irrational number, the resulting measures converge to the Lebesgue measure on T^2. So the picture is are countably many circles converging to the point given by Lebesgue measure.

THEOREM 32 (Mozes-Shah). *Suppose m_i are natural Haar measures on a sequence of homogeneous sets in $\Gamma\backslash G$. (So each $m_i = m_{x_i H_i}$ is the Haar measure on a finite volume orbit $x_i H_i$.) Assume that m_i is ergod and invariant under some 1-parameter unipotent subgroup. Then a weak* limit is either 0 (plus additional info) or th elimit is homogeneous and some unipotent acts ergodically (and additional info.)*

In T^2, take $H_n = \mathbb{R}(1, n) \subset \mathbb{R}^2$ and notice that the limiting group (as $n \to \infty$) $H_\infty = \{0\} \times \mathbb{R}$, and the limiting measure is the Lebesgue measure (as discussed above.) Note that the limit group doesn't act ergodically. (This doesn't contradict the Mozes-Shah theorem as H_n are not unipotent.)

Remark: If n is a nilpotent element of the Lie algebra of G, then $n^k = 0$ for some integer k, so $\exp(nt)$ has entries that are polynomials in t. So unipotent dynamics are polynomial dynamics.

THEOREM 33 (Special case of Ratner's classification theorem). *Let G be a Lie group, $H = \mathrm{SL}_2(\mathbb{R}) \subset G$ a subgroup,[3] $\Gamma \subset G$ discrete. The H-invariant and ergodic probability measures μ on $\Gamma\backslash G$ are precisely the homogeneous ones.*

The details of this proof are contained in Einsiedler's survey article available on his website.

SKETCH OF PROOF. Assume μ is H-invariant. We need to show that $\mu = \mu_{xL}$ for some L. The trick will be to guess what L is. Put $S = Stab(\mu)^\circ$ where

$$Stab(\mu) = \{g \in G \mid \mu(Bg) = \mu(B) \text{ for all } B \text{ measurable}\} \supset H.$$

We want to show that μ is concentrated on a single S-orbit. If $\mu(xS) = 1$ for some x we are done by the exercises. Let $\mathfrak{s} = Lie(S) \subset \mathfrak{g} = Lie(G)$. Since ad_H acts on $\mathfrak{s}$, there exists $\mathfrak{s}^\perp$ such that $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{s}^\perp$. We will assume that there does not exist x such that $\mu(xS) = 1$ and from this deduce by Lemma 34 that μ is invariant by more than originally claimed.

Let $x \in \mathrm{supp}(\mu)$ be generic. Since almost all points are generic (see below) there is a point $y \approx x$ such that $y = xg$ is generic. We can take $g = \exp(v)$ for some $v \in \mathfrak{s}^\perp$ with $v \approx 0$. Apply $u(t)$:

$$yu(t) = xgu(t) = xu(t)\underbrace{u(-t)gu(t)}_{\exp(\mathrm{Ad}_{u(t)}\, v)}.$$

[2]In the picture he drew on the board the rays pointed in different directions, but since their limits are the (same) 0 measure, they should all end up at the same point.

[3]In the general case, H is a subgroup of G generated by 1-parameter unitpotent subgroups.

As has been mentioned, $\exp(\mathrm{Ad}_{u(t)}\, v)$ has polynomial divergence.

Since v is small, there exists $T = T_{x,y}$ such that $\sup_{t\in[0,T]}\|\mathrm{Ad}_{u(t)}\, v\| = 1$. We want to know in what direction $\mathrm{Ad}_{u(t)}\, v$ gets big first. We take as an example $v = (c_1, c_2, c_3)^t$:

$$u(t)v = \begin{pmatrix} 1 & t & t^2 \\ & 1 & t \\ & & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} c_1 + tc_2 + t^2c_3 \\ c_2 + tc_3 \\ c_3 \end{pmatrix}$$

Notice that the divergence occurs fastest in the direction of $(1, 0, 0)^t$ which is precisely the eigenspace of $u(t)$.

So we take the limit as y and x get closer and closer, while simultaneously $\mathrm{Ad}_{u(t)}\, v$ gets closer and closer to commuting. In the limit, the lemma below would then apply and we would be done IF the limiting points are generic. To guarantee that the limiting points are generic one has to choose the time moment to stop correctly which can be argued to be possible a certain percentage of the time. (For complete details see the survey.) □

LEMMA 34. *Suppose $x = yg$ are generic for $U = \begin{pmatrix} 1 & * \\ & 1 \end{pmatrix} \subset \mathrm{SL}_2(\mathbb{R}) = H$. If*

$$g \in C_G(U) = \{h \in G \mid hu = uh \text{ for every } u \in U\}$$

then $g \in Stab(\mu)$.

A point x is *generic* for U if

$$\frac{1}{T}\int_0^T f(u(t)x)dt \to \int f\, d\mu \quad \text{as } T \to \infty$$

for every $f \in C_c(X)$. Since the Mautner phenomenon says that $\mathrm{SL}_2(\mathbb{R})$ acts ergodically implies that U acts ergodically, almost every $x \in X$ is generic.

PROOF. Suppose $g \in C_G(U)$ and $y = xg$ is generic. Then

$$yu(t) = xgu(t) = xu(t)u(-t)gu(t) = xu(t)g$$

which gives two parallel orbits. For $f \in C_c(X)$ define $f^g(z) = f(zg)$. Then

$$\int f\, d\mu \approx \frac{1}{T}\int_0^T f(yu(t))dt = \int_0^T f^g(xu(t))dt \approx \int f^g\, d\mu.$$

Hence g preserves the measure. □

3. Philippe

The fact that $\mathbb{Q}^\times \hookrightarrow \mathbb{A}$ is discrete, and that $\mathbb{Q}^\times\widehat{\mathbb{Z}}^\times = \mathbb{A}_f^\times$ gives an example of a much more general fact. If G is any semisimple Q-algebraic group, then $G(\mathbb{Q}) \hookrightarrow G(\mathbb{A})$ is discrete. Moreover, the following holds.

THEOREM 35 (Borel, Harishchandra). *For any subgroup $K \subset G(\widehat{\mathbb{Z}})$ that is open and compact, $G(\mathbb{Q})\backslash G(\mathbb{A})/K$ is finite. In fact, $G(\mathbb{Q})$ is a lattice in $G(\mathbb{A})$.*

This principle is easily demonstrated in the special case that G is simple and simply connected and if $G(\mathbb{R})$ is not compact. In this case, strong approximation says that $G(\mathbb{Q})$ is dense in $G(\mathbb{A}_f)$. So for any $K \subset G(\widehat{\mathbb{Z}})$,

$$G(\mathbb{Q})K = G(\mathbb{R}), \tag{9}$$

and the basic isomorphism theorems of group theory imply

$$G(\mathbb{Q})\backslash G(\mathbb{A})/K \simeq \Gamma\backslash G(\mathbb{R}) \qquad \text{where } \Gamma = G(\mathbb{Q}) \cap K. \tag{10}$$

Some examples of such groups are SL_d and B^1 which is the $\mathbb{Q}$-algebraic group of norm 1 elements of any quaternion algebra such that $B(\mathbb{R}) \simeq M_2(\mathbb{R})$.

Strong approximation does not apply to PGL_2 or $PB_{\mathbb{Q}}^{\times}$. (Note that $B_{\mathbb{Q}}$ is the Hamiltonian quaternions which are not split at ∞.) However, in these cases both (9) and (10) are valid.

3.1. Adèles over number fields. Suppose that K is a degree d extension of $\mathbb{Q}$ with r real embeddings $K \hookrightarrow \mathbb{R}$ and s complex embeddings $K \hookrightarrow \mathbb{C}$. Each of these affords a valuation. On the other hand each prime ideal $\mathfrak{p} \subset \mathcal{O}_K$ determines a (discrete) valuation $v_{\mathfrak{p}}$ and hence a completion $K_{\mathfrak{p}}$ with respect to the valuation. We let $\mathcal{O}_{K,\mathfrak{p}} = \{\lambda \in K_{\mathfrak{p}} \mid v_{\mathfrak{p}}(\lambda) \geq 0\}$.

As before, we let $\mathbb{A}_K = \prod^{*} K_v$ be the restricted product with respect to the rings $\mathcal{O}_{K,\mathfrak{p}}$. Much of what was said about the rational adèles is still valid, and some things are different:

- $K \hookrightarrow \mathbb{A}_K$ is discrete and a lattice.
- $K\backslash\mathbb{A}_K/\mathcal{O}_K \simeq \mathcal{O}_K\backslash K_\infty$ where $K_\infty = \prod_{v|\infty} K_v \simeq \mathbb{R}^d$.
- $K^{\times} \hookrightarrow \mathbb{A}_f^{\times}$ is not dense (as before.)
- In general $K^{\times}\widehat{\mathcal{O}}_K^{\times}$ need not equal $\mathbb{A}_{K,f}^{\times}$. Indeed,

$$\mathbb{A}_{K,f}^{\times}/K^{\times}\widehat{\mathcal{O}}_K^{\times} \simeq \mathrm{Pic}(\mathcal{O}_K) = \left\{ \begin{array}{c} \text{finitely generated} \\ \mathcal{O}_K - \text{modules} \end{array} \right\}/K^{\times}.$$

  The map from left to right is obtained by letting $I_{\mathfrak{p}}$ be the module generated by I in $K_{\mathfrak{p}}$. Since $\mathcal{O}_{K,\mathfrak{p}}$ is a PID, $I_{\mathfrak{p}} = \lambda_\pi \mathcal{O}_{K,\mathfrak{p}}$ for some $\lambda_{\mathfrak{p}} \in K_{\mathfrak{p}}^{\times}$. So the map is $I \mapsto (\lambda_{\mathfrak{p}})$ which is easily seen to be well defined up to $K^{\times}\mathcal{O}_K^{\times}$. In the other direction,

$$(\lambda_{\mathfrak{p}}) \mapsto K \cap \prod_{\mathfrak{p}} \lambda_{\mathfrak{p}} \mathcal{O}_{K,\mathfrak{p}} = I$$

  which is an $\mathcal{O}_K$-module.

Remark: Can replace $\widehat{\mathcal{O}}_K$ with $\widehat{\mathcal{O}}$ where $\mathcal{O} \subset \mathcal{O}_K$ is an order and $\widehat{\mathcal{O}}$ is the completion of $\mathcal{O}$ in $\widehat{\mathcal{O}}_K$ and is a compact subgroup. Then

$$K^{\times}\backslash\mathbb{A}_{K,f}^{\times}/\widehat{\mathcal{O}} \simeq \mathrm{Pic}(\mathcal{O}) = \{\text{f.g. invertible ...}\}.$$

3.2. Idèles of number fields as points of an algebraic group. Let $I \subset K$ be an $\mathcal{O}_K$-ideal (or $\mathcal{O}$-ideal) and choose a $\mathbb{Z}$-basis of I which is a basis of $K/\mathbb{Q}$. This gives a map

$$\begin{aligned} i : K &\longrightarrow M_d(\mathbb{Q}) \\ x &\mapsto \text{ the matrix of multiplication by } x : K \to K \end{aligned}$$

such that $i(K) \cap M_d(\mathbb{Z}) = i(\mathcal{O}_K)$.

An alternative way of describing this is to take $\{x_k\}$ a $\mathbb{Q}$-basis of K which gives an injection $i : K \hookrightarrow M_d(\mathbb{Q})$. Then $i(K) \cap M_d(\mathbb{Z}) = i(\mathcal{O}_I)$ where

$$I = \sum \mathbb{Z} x_i \qquad \mathcal{O}_I = \{\lambda \in K \mid kI \subset I\} \subset \mathcal{O}_K.$$

The image $i(K^\times)$ in $\mathrm{GL}_d(\mathbb{Q})$ is a $\mathbb{Q}$-torus, meaning that it is the $\mathbb{Q}$ rational points of an algebraic group; indeed, the group is $\mathrm{Res}_{K/\mathbb{Q}}(\mathbb{G}_m)$. Call the image $T_K(\mathbb{Q})$. We can think of $i(K)$ as an algebraic group such that

$$i(K)(\mathbb{A}_\mathbb{Q}) = \mathbb{A}_K \quad \text{and} \quad i(K)^\times(\mathbb{A}_\mathbb{Q}) = \mathbb{A}_K^\times.$$

Then $T_K(\mathbb{A}_\mathbb{Q}) = \mathbb{A}_\mathbb{Q}^\times \backslash \mathbb{A}_K^\times$. Hence

$$K^\times \backslash \mathbb{A}_{K,f}^\times / \mathcal{O}_K \simeq \mathrm{Pic}(\mathcal{O}_K) \simeq T_K(\mathbb{Q}) \backslash T_K(\mathbb{A}_f) / T_K(\widehat{\mathbb{Z}})$$

is finite.

3.3. Reinterpretation of Linnik's problem. Fix $b_0^2 - 4a_0c_0 = d$.

Choosing a basis we obtain a groups $T_{a,b,c}$ as in the previous section so that

$$\bigsqcup_{[a,b,c]} \Gamma\backslash\Gamma g_{a,b,c} H = \bigsqcup_{[a,b,c]} \Gamma\backslash\Gamma T_{a,b,c}(\mathbb{R}) g_{a,b,c}$$
$$\subset \mathrm{PGL}_2(\mathbb{Z})\backslash\mathrm{PGL}_2(\mathbb{R}) \simeq \mathrm{PGL}_2(\mathbb{Q})\backslash\mathrm{PGL}_2(\mathbb{A})/\mathrm{PGL}_2(\widehat{\mathbb{Z}}) = X.$$

We can determine the image of the orbit $T_{a_0,b_0,c_0}(\mathbb{A}).(g_{a_0,b_0,c_0}, e_f)$ in X. Let $t \in T_0(\mathbb{A})$, and write $t = t_\mathbb{R} t_f$. By strong approximation $t_f = g_\mathbb{Q} g_{\widehat{\mathbb{Z}}}$. So

$$\begin{aligned} t(g_0, e_f) &= (t_\mathbb{R} g_0, t_f) \\ &= (t_\mathbb{R} g_0, g_\mathbb{Q} g_{\widehat{\mathbb{Z}}}) \\ &\equiv (t_\mathbb{R} g_0, g_\mathbb{Q}) && \text{modulo } \mathrm{PGL}_2(\widehat{\mathbb{Z}}) \text{ on the right} \\ &\equiv (g_\mathbb{Q}^{-1} t_\mathbb{R} g_0, e_f) && \text{modulo } \mathrm{PGL}_2(\mathbb{Q}) \text{ on the left.} \end{aligned}$$

Note that if $t_f \in T_0(\mathbb{A})$ and $m_0 = \left(\begin{smallmatrix} -b_0 & -2c_0 \\ 2a_0 & b_0 \end{smallmatrix}\right)$ then, by definition, $t_f m_0 = m_0 t_f$.

3.3.1. *Linnik Skubenko proof.* Fix $p \geq 5$. Linnik and Skubenko proved that for values of d such that $d \equiv 1 \pmod p$, $R_q(d)$ becomes equidistributed as $d \to -\infty$. In this special case, d is a square modulo p so $p\mathcal{O}_d = \mathfrak{p}\mathfrak{p}'$ splits. This allows Linnik to use the action of $[\mathfrak{p}]^\mathbb{Z} = [\mathfrak{p}']^\mathbb{Z} \subset \mathrm{Pic}(\mathcal{O}_d)$. In this situation m_0 is diagonalizable in $\mathbb{Q}_p$, so $T_0(\mathbb{Q}_p)$ is noncompact. This is essential in the proof.

3.3.2. *Duke proof.* Duke removed the congruence condition by considering something different. In the S^2 case he showed that

$$\frac{1}{R_q(d)} \sum_{x \in R_q(d)} \varphi\left(\frac{x}{\sqrt{|x|}}\right) \to \int_{S^2} \varphi$$

as $d \to \infty$ (d odd[4]) for any $\varphi \in C(S^2)$. By approximation, it suffices to take φ nonconstant harmonic homogeneous polynomials. For such, need the average to approach zero.

Let $d_K = \mathrm{Disc}(K)$. By Waldspurger (this is not exactly Duke's proof)

$$\left|\frac{1}{R_q(d)} \sum_{x \in R_q(d)} \varphi\left(\frac{x}{\sqrt{|x|}}\right)\right|^2 = c\,|d|^{o(1)} \prod_{p \mid d/d_K} \frac{I_p \times L(\varphi, d, \frac{1}{2})}{|d|^{1/2}},$$

$$L(\varphi, d, \frac{1}{2}) = \sum_{n=1}^\infty \frac{\lambda_\varphi(n)\chi_d(n)}{n^s} = \prod_p \left(1 - \frac{\alpha_p \chi_d(p)}{p^s}\right)^{-1} \left(1 - \frac{\beta_p \chi_d(p)}{p^s}\right)^{-1}.$$

[4]If $d = -2^k$, $R_q(d)$ is bounded.

If $L(\varphi, d, \frac{1}{2}) = o(\sqrt{|d|})$ then we would be done. Note that the Riemann hypothesis predicts that $L(\varphi, d, \frac{1}{2}) = o(1)$. In this problem we just need to beat $d^{1/2}$ by a little bit. This is an example of "subconvexity." Iwaniec was the first to do this.

The ratio $L(\varphi, d, \frac{1}{2})/\sqrt{|d|}$ can be realized as a toral object. It has a so-called *Hecke integral representation*:

$$L(\varphi, d, \frac{1}{2}) = \int_{A(\mathbb{Q})\backslash A(\mathbb{A})} \widetilde{\varphi}(an(T_d))\chi(a)d^{\times}a \tag{11}$$

where $A = Diag \subset \mathrm{PGL}_2$, $\widetilde{\varphi}$ is a function on $\mathrm{PGL}_2(\mathbb{Q})\backslash\mathrm{PGL}_2(\mathbb{A})$ and χ is a quadratic character of $\mathbb{A}^{\times}_{\mathbb{Q}}$.

The integral (11) is equal to $\int_{T(\mathbb{Q})\backslash T(\mathbb{A})} \varphi(tg_{a,b,c})dt$, but it is much more convenient, because T was not split (noncompact) at every place, but A is.

CHAPTER 5

June 5, 2009

1. Philippe: The subconvexity problem

The subconvexity problem is about L-functions:

$$L(\pi, s) = \sum_{n \geq 1} \frac{\lambda_\pi(n)}{n^s} = \prod_p L_p(\pi, s), \qquad L_p(\pi, s) = \prod_{i=1}^{d} \left(1 - \frac{\alpha_{\pi,i}(p)}{p^s}\right)^{-1}.$$

An example of such is $\zeta(s)$, the Riemann zeta function: $\lambda_\pi(n) = 1$ for all n and $\zeta_p(s) = (1 - \frac{1}{p^s})^{-1}$.

Often the series converges absolutely uniformly on compact sets of $\mathrm{Re}(s) > 1$. So $L(\pi, s)$ are analytic in that region. In fact, they have analytic (strictly speaking meromorphic) continuation to $\mathbb{C}$. This comes from a functional equation. There are constants $\mu_{\pi,i}$ such that, defining

$$L_\infty(\pi, s) = \prod_{i=1}^{d} \Gamma\left(\frac{s - \mu_{\pi,i}}{2}\right),$$

one has that

$$L_\infty L(\pi, s) = \epsilon(\pi) q_\pi^{1/2-s} \overline{L_\infty L(\pi, 1 - \overline{s})}.$$

The constant $\epsilon(\pi)$ has absolute value 1, while $q_\pi \in \mathbb{Z}$ is the *conductor* of π. Sometimes one writes $L_\infty L(\widetilde{\pi}, 1 - s) = \overline{L_\infty L(\pi, 1 - \overline{s})}$.

The function L is known for $\mathrm{Re}(s) > 1$, so by the functional equation it is also known for $\mathrm{Re}(s) < 0$. The region $\mathrm{Re}(s) \in [0, 1]$ is called the *critical strip*. The *Riemann hypothesis* is that the only zeros of $L_\infty L(\pi, s)$ are on the line $\mathrm{Re}(s) = \frac{1}{2}$. Another question is what is the size of $L(\pi, s)$ on $\mathrm{Re}(s) = \frac{1}{2}$.

The *analytic conductor* is defined to be

$$C(\pi, s) = q_\pi \prod_{i=1}^{d} (1 + |s - \mu_{\pi,i}|).$$

The *convexity bound* says that for $\mathrm{Re}(s) = \frac{1}{2}$,

$$L(s, \pi) \ll_d C(\pi, s)^{\frac{1}{4} + o(1)}.$$

To obtain this one considers the function $f(\sigma)$ where $\sigma = \mathrm{Re}(s)$ and

$$L(s, \pi) \ll_d C(\pi, s)^{f(\sigma) + o(1)}.$$

When $\sigma > 1$, $f(\sigma) = 0$. By looking at the functional equation and Sterling's formula for the gamma function, one finds that $f(\sigma) = \frac{1}{2} - \sigma$ when $\sigma < 0$. Then the Phragman-Lindelof principle imples that $f(\sigma) \leq \frac{1-\sigma}{2}$ in the critical strip. (He then drew the graph of $f(\sigma)$ and pointed out that the Riemann hypothesis would

imply that $f(\sigma) = 0$ for $\mathrm{Re}(s) > \frac{1}{2}$ and $f(\sigma) = \frac{1}{2} - \sigma$ for $\sigma < \frac{1}{2}$ and drew this on the graph with dotted lines.) The Riemann hypothesis would imply that

$$L(s, \pi) \ll_d C(\pi, s)^{o(1)}$$

which is much stronger than the convexity bound. The subconvexity problem is to show that for some $\delta > 0$,

$$L(s, \pi) \ll_d C(\pi, s)^{\frac{1}{2}-\delta}.$$

This is a very hard problem in general.

Duke's theorem is that

$$L(\varphi, d_K, \frac{1}{2}) \ll |d_K|^{\frac{1}{2}-\delta}.$$

Since in this case $q_{0,d} \asymp |d_K|^2$, this is an instance where equidistribution is equivalent to solving the subconvexity problem. To see this, recall that the Waldspurger formula says

$$\left| \frac{1}{R_q(d)} \sum_{x \in R_q(d)} \varphi\left(\frac{x}{\sqrt{|x|}} \right) \right|^2 = \frac{L(\varphi, d, \frac{1}{2})}{\sqrt{|d|}^{1+o(1)}}. \tag{12}$$

Since the left hand side of (12) is a probability measure it is $\ll_\varphi 1$.

The first instance of subconvexity (for $\zeta(s)$) was proven by Weyl in the 20s. Later Burgos (in the 50s) had more results. After a big gap of time, Iwaniec started doing work on this, and since then there has been a flurry of results.

1.1. The L-functions we are talking about. The are L-functions associated to number fields, elliptic curves, motives, etc. They should all come from automorphic forms and automorphic representations. Examples of automorphic forms are Dirichlet characters and modular forms. More generally, automorphic forms are functions on $G(\mathbb{Q})\backslash G(\mathbb{A})$ for G a reductive group.

A typical example is $G = \mathrm{PGL}_2$. On $L^2(G(\mathbb{Q})\backslash G(\mathbb{A}))$, the group $G(\mathbb{A})$ acts unitarily by right multiplication: if $f \in L^2$ and $g \in G$,

$$gf : h \mapsto f(hg).$$

An *automorphic representation* is an irreducible representation of $G(\mathbb{A})$ "occurring" in $L^2(G(\mathbb{Q})\backslash G(\mathbb{A}))$. Formally

$$L^2(G(\mathbb{Q})\backslash G(\mathbb{A})) = \int_{\pi \in \widehat{G(\mathbb{A})}} V_\pi d\mu_P(\pi).$$

The measure μ_P is called the Plancherel measure. It has a point component and a continuous component. An *automorphic form* is a vector inside any one of the spaces V_π.

Remark: Since $G(\mathbb{Q}_v) \hookrightarrow G(\mathbb{A})$ which acts on V_π, this defines a (unitary) representation π_v of $G(\mathbb{Q}_v)$ such that $\pi = \otimes'_v \pi_v$ which is "restricted" tensor product. Being restricted means that, in particular, π_p admits a $G(\mathbb{Z}_p)$ invariant vector for almost all p.

In the case at hand, $G = \mathrm{PGL}_2$, to each π_v can associate a local L-factor

$$L(\pi_v, s) = \begin{cases} (1 - \frac{\alpha_\pi(p)}{p^s})^{-1}(1 - \frac{\beta_\pi(p)}{p^s})^{-1} & \text{if } v = p \text{ is prime,} \\ \prod \Gamma(\frac{s-1}{2}) & \text{if } v \text{ is real.} \end{cases}$$

1.2. The Whittaker model. Let $\varphi \in V_\pi$. (We assume that $V_\pi \subset L^2$ is literally a subrepresentation, then φ is square-integrable. When V_π is not a subrepresentation of L^2, φ is called an *Eisenstein series*.) The function

$$x \in \mathbb{A}_\mathbb{Q} \mapsto \varphi(n(x)g) = g\varphi(n(x)) \qquad \text{where } n(x) = \left(\begin{smallmatrix} 1 & x \\ & 1 \end{smallmatrix}\right)$$

is $\mathbb{Q}$-periodic via the embedding $N(\mathbb{Q}) \subset G(\mathbb{Q})$. Hence there is a Fourier decomposition on $N(\mathbb{Q})\backslash N(\mathbb{A}) \simeq \mathbb{Q}\backslash\mathbb{A}$. Fix a nontrivial character $\psi : \mathbb{Q}\backslash\mathbb{A} \to \mathbb{C}^\times$. Then the Fourier decomposition is

$$\varphi(g) = \sum_{\alpha \in \mathbb{Q}^\times} W_\varphi(a(\alpha)g) \tag{13}$$

where

$$a(\alpha) = \begin{pmatrix} \alpha & \\ & 1 \end{pmatrix} \quad \text{and} \quad W_\varphi(g) = \int_{\mathbb{Q}\backslash\mathbb{A}} \varphi(n(x)g)\overline{\psi}(x)\, dx = W_{g\varphi}(e).$$

(The measure is the Haar probability measure on $\mathbb{Q}\backslash\mathbb{A}$.) The function $g \mapsto W_\varphi$ is a function on $G(\mathbb{A})$ such that

$$W_\varphi(n(x)g) = \psi(x)W_\varphi(g). \tag{14}$$

We describe how the above decomposition is obtained. Having fixed ψ, the group of characters on $\mathbb{Q}\backslash\mathbb{A}$ is precisely

$$\widehat{\mathbb{Q}\backslash\mathbb{A}} = \{\psi_a : \mathbb{Q} \to \mathbb{C}^\times \mid \psi_a(x) = \psi(ax), a \in \mathbb{Q}\} \simeq \mathbb{Q}.$$

So by Fourier inversion

$$\varphi(g) = \sum_{\alpha \in \mathbb{Q}} \int_{\mathbb{Q}\backslash\mathbb{A}} \varphi(n(x)g)\psi(-\alpha x)\, dx.$$

In the case that $\alpha \neq 0$, we make the change of variable $x' = \alpha x$. Then $dx' = |\alpha|_\mathbb{A}\, dx = dx$, and

$$\int_{\mathbb{Q}\backslash\mathbb{A}} \varphi(n(x)g)\psi(-\alpha n(x))\, dx = \int_{\mathbb{Q}\backslash\mathbb{A}} \varphi(n(\frac{x}{\alpha})g)\psi(x)\, dx.$$

Since

$$n(\frac{x}{\alpha}) = \begin{pmatrix} \frac{1}{\alpha} & \\ & 1 \end{pmatrix} \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \begin{pmatrix} \alpha & \\ & 1 \end{pmatrix},$$

this is equal to

$$\int \varphi(n(x)a(\alpha)g)\, dx = W_\varphi(a(\alpha)g).$$

When $\alpha = 0$, by some work, can show that $\int_{N(\mathbb{Q})\backslash N(\mathbb{A})} = 0$. This gives (13).

THEOREM 36. *The mapping $\varphi \mapsto W_\varphi$ is an intertwining map from V_π to an irreducible subspace of functions satisfying* (14)*. The image, $\mathcal{W}(\pi)$, is called the* Whittaker model *of π. Moreover, $V_\pi \to \mathcal{W}(\pi)$ is an isomorphism.*

In the first model $\pi = \otimes'_v \pi_v$ with the π_v local representations. However, one should keep in mind that given $\{\pi_v\}$ there is no guarantee that the resulting tensor product will be $G(\mathbb{Q})$ invariant, so there are many restrictions on the possible choices of π_v. On the other hand, we can define local Whittaker models so that the global model factors as a product of locals, and unlike before, in the Whittaker model there are fewer restrictions on the local representations.

On the Whittaker model can define an inner product

$$\langle W, W' \rangle = \int_{\mathbb{A}^\times} W(a(y))\overline{W'(a(y))}\frac{d^\times y}{|y|}.$$

(This inner product is convenient to work with because one only has to integrate along $A \subset \mathrm{PGL}_2$.) Now the map $V_\pi \to \mathcal{W}(\pi)$ is essentially an isometry[1]:

$$\langle \varphi, \varphi \rangle_{L^2(G(\mathbb{Q})\backslash G(\mathbb{A}))} = c(\pi, \frac{1}{2})^{o(1)} \langle W_\varphi, W_\varphi \rangle_{\mathcal{W}(\pi)}.$$

The constant appearing here is the value of the L-function attached to the adjoint representation at 1.

Within the range of convergence,

$$\frac{L(\pi, s)}{C(\pi, s)^{1/2+o(1)}} \approx \int_{\mathbb{Q}^\times\backslash\mathbb{A}^\times} \varphi_0(a(y))\,|y|^{s-\frac{1}{2}}\,d^\times y$$

for some distinguished vector φ_0 which is sometimes the new vector. (The constant C comes from the c above and some 'laziness' due to estimation.)

Formally,

$$\begin{aligned}
\varphi_0(a(y)) &= \int_{\mathbb{Q}^\times\backslash\mathbb{A}^\times} \sum_{\alpha\in\mathbb{Q}^\times} W_{\varphi_0}(a(\alpha)a(y))\,|y|^{s-\frac{1}{2}}\,d^\times y \\
&= \int_{\mathbb{A}^\times} w_{\varphi_0}(a(y))\,|y|^{s-\frac{1}{2}}\,d^\times y \\
&= \prod_v \int_{\mathbb{Q}_v^\times} W_{\varphi_0,v}(a(y))\,|y|_v^{s-\frac{1}{2}}\,d_v^\times y.
\end{aligned}$$

Note that to get the second line we need that W_{φ_0} be *rapidly decreasing*. The local integrals in the final expression can be computed. From the structure of unitary admissible representations of $\mathrm{PGL}_2(\mathbb{Q}_v)$ which admit a $\mathrm{PGL}_2(\mathbb{Z}_v)$ fixed vector (the so-called *spherical representations*),

$$\int_{\mathbb{Q}_v^\times} W_{\varphi_0,v}(a(y))\,|y|_v^{s-\frac{1}{2}}\,d_v^\times y = L(\pi_v, s).$$

We assume $\mathrm{Re}(s) = \frac{1}{2}$.

$$\begin{aligned}
\left|\int_{\mathbb{A}^\times} W_\varphi(a(y))\,|y|^{s-\frac{1}{2}}\,d^\times y\right|^2 &= \int_{\mathbb{A}^\times}\int_{\mathbb{A}^\times} W(y)W(y')\left|\frac{y}{y'}\right|^{s-\frac{1}{2}} d^\times y d^\times y' \qquad (y'' = \frac{y}{y'}) \\
&= \int_{\mathbb{A}^\times} |y''|^{s-\frac{1}{2}} \int_{\mathbb{A}^\times} W(y)W(y/y'')d^\times y d^\times y'' \\
&= \int_{\mathbb{A}^\times} |y|^{s-\frac{1}{2}} \int_{\mathbb{A}^\times} W(y)a(y')W(y)d^\times y' d^\times y \\
&= \int_{\mathbb{A}^\times} |y|^{s-\frac{1}{2}}\,\langle W_\varphi, a(y')W_\varphi\rangle d^\times y.
\end{aligned}$$

This last expression is good for dynamics.

[1]The restriction of the Whittaker model to the diagonal is called the *Kirillov* model. It is isomorphic to $\mathcal{W}(\pi)$.

1.3. Back to Duke's L-function. Recall that φ was a function on SO_3. Now

$$L(\varphi, d, \frac{1}{2}) = L(\pi \otimes \chi_d, \frac{1}{2})$$

where π is a representation of PGL_2, $\chi_d : \mathbb{Q}^\times \backslash \mathbb{A}^\times \to \{\pm\}$ and

$$\pi \otimes \chi_d = \{g \mapsto \chi_d(\det g)\varphi(g) \mid \varphi \in V_\pi\}.$$

By our previous work,

$$\frac{L(\pi \otimes \chi_d)}{C(\pi \otimes \chi_d)} \approx \int_{\mathbb{Q}^\times \backslash \mathbb{A}^\times} \widetilde{\varphi}_0(a(y))d^\times y$$

for some $\widetilde{\varphi}_0 \in V_{\pi \otimes \chi_d}$. More precisely,

$$\widetilde{\varphi}_0(g) = \chi_d(\det g)n(T)\varphi_0(g)$$

for some $T \in \mathbb{A}$ such that $|T| = C(\pi \otimes \chi_d)^{1+o(1)} = d^{2+o(1)}$.

It is a fact that

$$|n(T)\varphi_0(g)|^2 = n(T)\left|\varphi_0^2\right|(g) = |\varphi_0|^2 (gn(T)).$$

So,

$$\begin{aligned}\left|\int \chi_d(y)n(T)\varphi_0(a(y))d^\times y\right|^2 &\leq \int_{\mathbb{Q}^\times \backslash \mathbb{A}^\times} |n(T)\varphi_0(a(y))|^2 d^\times y \\ &= \int_{\mathbb{Q}^\times \backslash \mathbb{A}^\times} n(T)\,|\varphi_0(a(y))|^2 d^\times y.\end{aligned}$$

By the spectral decomposition in $L^2(G(\mathbb{Q})\backslash G(\mathbb{A})$,

$$\|\varphi_0\|_{L^2}^2 = \sum_{\pi'} \sum_{\psi \in V_{\pi'}} \langle |\varphi_0|^2, \psi\rangle = \sum_{\pi'} \sum_{\psi \in V_{\pi'}} \int n(T)\psi(a(y))d^\times y.$$

This final integral is related to $L(\pi, s)$.

If $\pi' \neq 1$,

$$\left|\int \quad\right|^2 = \int \langle a(y)n(T)\varphi, n(T)\psi\rangle dy = \int \langle n(T)^{-1}a(y)n(T)\varphi, \psi\rangle dy$$

(I'm really not sure where to have φ and where to have ψ in this expression and the next.) As

$$\langle n(T)^{-1}a(y)n(T)\varphi, \psi\rangle \ll \|\mathrm{Ad}(n(T)^{-1}a(y)n(T)\|^{-\beta}\|\varphi_0\|_{L^2}^2$$

for some *beta* $>$, plugging into the above yields that

$$\int \qquad \ll T^{-\delta}$$

where $\delta = \delta(\beta) > 0$. So Duke is finished up to $\pi' = 1$, because this gives $\|\varphi_0\|_{L^2}^2$.

To complete the problem one uses the "amplification method:" replace φ_0 by a linear combination of translates. This still gives L-function with nearly the same decay for $\pi' \neq 1$, but the advantage is that the term for $\pi' = 1$ is small.

(N.B. This method is due to Akshay Venkatesh.)

2. Elon

3. Manfred

The book by Morris Witte "Ratner's Theorems on Unipotent Flows" is a good source for reading up on these things in greater detail.

THEOREM 37 (Ratner's Equidistribution Theorem). *Let $X = \Gamma\backslash G$, Γ a lattice. Let $U \subset G$ be a 1-parameter unipotent subgroup. Let $x \in X$. Then there exists $L \subset G^\circ$ a closed connected subgroup such that xL has finite volume and $xu(t)$ ($u(t) \in U$) equidistributes with respect to m_{xL}: if $f \in C_c(X)$,*

$$\frac{1}{T}\int_0^T f(xu(t))dt \to \int_{xL} f dm_{xL}.$$

Note that this statement is out of the realm of the pointwise ergodic theorem because we are choosing $x \in X$.

THEOREM 38 (Ratner's Orbit Closure Theorem). *Suppose $\Gamma \subset G$ is a lattice, $X = \Gamma\backslash G$, and $H \subset G$ is generated by one-parameter unipotents. Let $x \in X$. Then $\overline{xH} = xL$ for some $L \subset G^\circ$.*

These very general theorems have particular meaning is specific cases. For example, Margulis' Theorem (Oppenheim's conjecture) is a consequence: Q a quadratic form in $d \geq 3$ variables. Then either Q is rational (meaning Q has rational coefficients) or $\overline{Q(\mathbb{Z}^d)} = \mathbb{R}$. One drawback to these theorems is that they don't give an "error rate."

3.1. Effective equidistribution. We discuss, via almost dynamical methods, quantitative results on horocylces. The idea is to replace the analysis of just sets with functions. Returning to our example of X_2, let P_y be horocycle flow in upper half plane of height y. We want to find a constant $\delta > 0$ such that

$$\left|\int_{P_y} f \; dm_{P_y} - \int f \; dm_{X_2}\right| \ll y^\delta S(f). \tag{15}$$

The term $S(f)$ is the Sobolov norm and is really a constant. The value of δ relates to the spectral gap, hence Selberg's theorem.

In order to get (15), you need to know effective decay of matrix coefficients:

$$\left\|\langle gf_1, f_2\rangle - \int f_1 \; dm_X\right\| \ll \|g\|_2^{-\delta'} S(f_1)S(f_2).$$

For bigger groups this is property T, hence easy. (This is shown in the survey article.)

THEOREM 39 (E,Margulis,Venkatesh). *Let $G = \mathrm{SL}_2(\mathbb{R})^4$, $\Gamma = \mathrm{SL}_2(\mathbb{Z})^4 \subset G$, and*

$$H = \{(h, h, h, h) \mid h \in \mathrm{SL}_2(\mathbb{R})\}.$$

Then the finite volume orbits of H on $\Gamma\backslash G$ effectively equidistribute when the volume gets big.

The subgroup $L_1 = \{(h_1, h_1, h_2, h_2) \mid h_i \in \mathrm{SL}_2(\mathbb{R})\}$ also has closed orbits with lots of closed H orbits in it.

This uses a lot more dynamical ideas effective decay but it also uses spectral gap. It uses (uniformity of) property τ, a theorem by Clozel.

3.2. Equidistribution on large spheres in $\mathbb{R}^d$. If $d \geq 4$ this is related to the problem of equidistribution of orbits of an orthogonal group in $d-1$ variables.

$$\mathrm{SO}(d)(\mathbb{R})/\mathrm{SO}(d-1)(\mathbb{R}) \simeq S^{d-1},$$

so $\mathrm{SO}(d-1)(\mathbb{R})$ acts (on the right) on $\mathrm{SO}(d)(\mathbb{Z})\backslash\mathrm{SO}(d)(\mathbb{R})$. (Note that when $d=3$ this gives a compact group acting on another compact space so it's not easy.)

To make things work more nicely consider the compact orbit

$$\mathrm{SL}_d(\mathbb{Z}[\frac{1}{p}])\mathrm{SO}(d)(\mathbb{R}\times\mathbb{Q}_p)$$

in

$$\mathrm{SL}_d(\mathbb{Z}[\frac{1}{p}])\backslash\mathrm{SL}_d(\mathbb{R}\times\mathbb{Q}_p) \simeq \mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R}\times\mathbb{Z}_p). \tag{16}$$

This is interesting because $\mathrm{SO}(d)(\mathbb{Q}_p)$ need not be compact.

Because

$$\Gamma = \mathrm{SO}_d(\mathbb{R}\times\mathbb{Z}_p) \subset \mathrm{SO}_d(\mathbb{R}\times\mathbb{Q}_p)$$

is open, we get that

$$\Gamma\backslash\mathrm{SO}_d(\mathbb{R}\times\mathbb{Q}_p) = \bigsqcup_{i=1}^{h}\Gamma(g_\infty,g_p)\mathrm{SO}_d(\mathbb{R}\times\mathbb{Q}_p).$$

Using (16), can assure that $g_p \in \mathrm{SL}_d(\mathbb{Z}_p)$. We may assume that $(g_\infty,g_p)=(e,e)$ for $i=1$.

Take $m\mathbb{Z}^d$ primitive such that $\|m\|^2 = D$. The action of $\mathrm{SO}(d)(\mathbb{Z})$ gives some more points, but as $D\to\infty$ the number of points obtained in this way is bounded. We would like to find additional solutions 'using' $\mathbb{Q}_p$.

Let $H_m = Stab_{\mathrm{SO}_d}(m)$. Note that $H_m \simeq \mathrm{SO}(d-1)$ but not over the rationals. Mozes-Shah should tell us that $\Gamma H_m(\mathbb{R}\times\mathbb{Q}_p)^+$ become equidstributed in a g bigger space[2]. (Similar to the exercise that $\mathrm{SO}(2,1)(\mathbb{R}) \subset \mathrm{SO}(3)(\mathbb{R})$ is maximal compact.)

This is helpful for finding addtional points. Often ($\frac{1}{h}$ of the time)

$$\Gamma(e,h_p) = \Gamma(e,e)(g_\infty,g_p) \tag{17}$$

for some $h_p \in H(\mathbb{Q}_p)$ large. The claim is that $g_\infty m$

- has norm $\sqrt{D}$, and
- is integral.

The first statement is obviously true because $g_\infty \in \mathrm{SO}$. Equation (17) implies that $(\gamma,\gamma)(e,h_p) = (g_\infty,g_p)$, so

(1) $g_\infty = \gamma \implies g_\infty m \in \mathbb{Q}^d$ (or better $\mathbb{Z}[\frac{1}{p}]^d$) and
(2) $\gamma h_p = g_p \iff \gamma = g_p h_p^{-1} \implies g_\infty m = g_p h_p^{-1} m = g_p m$. Since $g_p \in \mathrm{SO}_d(\mathbb{Z}_p)$ this implies that $g_p m \in (\mathbb{Z}_p)^d$.

Hence, the claim is true.

[2]The + refers to the fact that one actually works in a Spin group instead of SO.