AN OVERVIEW OF RATNER'S THEOREMS

ROBERT C. RHOADES

Abstract. We give a brief overview of Ratner's Measure Classification Theorems. We motivate the theorems by some number theoretic results that are proved using them. Much of the text is adapted from the first chapter of Dave Morris's notes on Ratner's Theorems [3]. I highly recommend these notes to anyone that is seriously interested in the subject. Bekka and Mayer's [1] book is another good source.

Contents

1. Applications

We start by listing a few applications of Ratner's Theorems. We do not give proofs of any of these things, but we will indicate places where proofs can be found.

1.1. Oppenheim's Conjecture.

A quadratic form is a homogeneous polynomial of degree 2 with real coefficients and any number of variables. An *indefinite* form Q is one that takes both positive and negative values as the variables run over all real values. Further, the form Q is *non-degenerate* if there does not exist a nonzero vector $x \in \mathbb{R}^n$ such that $Q(v + x) = Q(v - x)$.

Remark. Say $Q = Q(x_1, \cdots, x_n)$ now $Q(v+e_n) = Q(v-e_n)$ for all $v \in \mathbb{R}^n$ if and only if there exists a Q' on $\mathbb{R}^{n-1}$ with $Q(x_1, \cdots, x_n) = Q'(x_1, \cdots, x_{n-1})$ for all $x_j \in \mathbb{R}$.

Theorem 1.1 (Margulis). *Let Q be a real, indefinite, non-degenerate quadratic form in $n \geq 3$ variables. If Q is not a scalar multiple of a form with integer coefficients , the $Q(\mathbb{Z}^n)$ is dense in $\mathbb{R}$.*

For the proof of this theorem see [1] or [3].

Example. $Q(x, y, z) = x^2 - \sqrt{2}xy + \sqrt{3}z^2$ for each $r \in \mathbb{R}$ and $\epsilon > 0$ there are $a, b, c \in \mathbb{Z}$ with $|Q(a, b, c) - r| < \epsilon$.

Date: February 24, 2007.

1.2. **Local-To-Global.** Given two quadratic forms Q and Q' with rank n and m, with $m < n$, can we find linear forms $L_1, \cdots, L_n$ in m variables such that $Q(L_1, \cdots, L_n) = Q'$? If so we say that Q' is represented by Q in $\mathbb{Z}$.

Theorem 1.2 (Ellenberg and Venkatesh [2]). *Let Q be a positive definite quadratic form on $\mathbb{Z}^n$. Then there exists $c := c(Q)$ such that Q represents all quadratic forms Q' in $m \leq n - 7$ variables that are everywhere locally representable, have square-free discriminant, and minimum $\geq c(Q)$.*

The discriminant of a form Q is $\det(Q(e_i - e_j) - Q(e_i) - Q(e_j))_{i,j}$ with e_i a basis for $\mathbb{Z}^n$. The minimum of Q is the smallest nonzero element of $Q(\mathbb{Z}^n)$.
Example. $Q_1 = x^2 + 55y^2$ and $Q_2 : 11x^2 + 5y^2$ are isomorphic over $\mathbb{Z}_p$ for every p and $\mathbb{R}$. But Q_1 does not represent 5 while Q_2 does.
Example. $L_1^2 + L_2^2 + 10L_3^2 = x^3 + 2y^2$ has no solutions in linear forms L_1, L_2, L_3. Let $Q_1 : x^2 + y^2 + 10z^2$ and $Q_2 : x^2 + 2y^2$. Does Q_1 represent Q_2 over every $\mathbb{Q}_p$ and $\mathbb{R}$? No. Not in $\mathbb{R}$.

1.3. **Bounds for Coefficients of Modular Forms.** Let $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$ contain $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ be a lattice. Let $f(z)$ be a holomorphic modular form of weight 2 with respect to Γ. The fourier expansion of f is $\sum_{n \geq 1} a_n e^{2\pi i n z}$. Hecke proved that $|a_n| \leq Cn$. Much later Good proved that $|a_n| \leq Cn^{5/6}$. We can use ergodic methods to prove that $|a_n| \leq Cn^{1-\delta}$ for some $\delta > 0$.

Remark. Technically these results require a quantitative version of Ratner's Equidistribution theorem, but I think number theorist should be excited about this example so I list it.

This example is explained in the introduction of [5]. Venkatesh's paper [5] has many more applications of quantitative ergodic theorems.

2. Torii as Examples

Let $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$ and for $x \in \mathbb{R}^n$, let $[x] = x + \mathbb{Z}^n$. Then for any $v \in \mathbb{R}^n$ we can define a C^∞ flow ϕ_t on $\mathbb{T}^n$ by

$$\phi_t([x]) := [x + tv] = x + tv + \mathbb{Z}^n.$$

The proof that this is a C^∞ flow follows by checking the following three simple properties

(1) ϕ_0 is the identity
(2) $\phi_{s+t} = \phi_s \circ \phi_t$
(3) $\phi : \mathbb{T}^n \times \mathbb{R} \to \mathbb{T}^n$ by $\phi([x], t) = \phi_t([x])$ is C^∞.

Questions. What does the orbit of a point x look like? What does $\{\phi_t([x])\}$ look like inside $\mathbb{T}^n$? More precisely what is the closure look like $\overline{[x + \mathbb{R}v]}$?

Remark. We will be asking and answering these types of questions throughout this note.

To understand these questions lets look at a few more specific questions. Let us start with the $n = 1$ case of the scenario given above. Given $\alpha \in \mathbb{R}$ we have

$$\overline{[x + \mathbb{R}\alpha]} = \begin{cases} \{[x]\} & \alpha = 0 \\ [x + \mathbb{R}\alpha] = \mathbb{T}^1 & \alpha \neq 0 \end{cases}.$$

The proof of this is quite simple. Suppose that $\alpha \neq 0$ and let $\beta \in \mathbb{R}$ and we want to find an $r \in \mathbb{R}$ such that $x + r\alpha = \beta$. So we can take $r = \frac{\beta - x}{\alpha} \in \mathbb{R}$.

The case $n = 2$ is slightly more complex. Let $v = (\alpha, \beta) \in \mathbb{R}^2$. Then we have

$$\overline{[x + \mathbb{R}v]} = \begin{cases} \{[x]\} & \alpha = \beta = 0 \\ [x + \mathbb{R}v] & \beta = 0 \text{ or } \alpha/\beta \in \mathbb{Q} \\ \mathbb{T}^2 & \text{else} \end{cases}.$$

The proof of this result is a little bit more involved, but is a good exercise.

In general there exists a vector space S of $\mathbb{R}^n$, such that

(1) $v \in S$,so $tv \in S$ for all t and hence $\phi_t([x]) \subseteq [x + S]$
(2) the image of $[x+S]$ in $\mathbb{T}^n$ is compact, hence $\underline{[x+S]}$ is diffeomorphic to $\mathbb{T}^k$ for some $0 \leq k \leq n$.
(3) the orbit $\phi_t([x])$ is dense in $[x + S]$, hence $\overline{\{\phi_t([x])\}} = \overline{\{[x + tv]\}} = [x + S]$.

Therefore the closure of every orbit is a "nice" geometric subset of $\mathbb{T}^n$. Ratner's theorem generalizes this! The following are the building blocks of what we have just seen.

(1) $\mathbb{R}^n$ is a Lie group. That is it is a group under vector addition and a manifold and the group operation is C^∞.
(2) The subgroup $\mathbb{Z}^n$ is discrete. That is, it has no accumulation points.
(3) The quotient is compact.
(4) The map $t \mapsto tv$ is a one-parameter subgroup of $\mathbb{R}^n$, it is C^∞ group of homomorphisms from $\mathbb{R}$ to $\mathbb{R}^n$.

3. Upper-Half Plane Example

Here we will describe a second example, namely that of quotients of the upper-half plane. For some more about the ergodic theory of this situation see my notes about the geodesic and horocycle flows [4].

Let $G = \mathrm{SL}_2(\mathbb{R})$, $u(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$, and $a(t) = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$. One can easily check that $u(t+s) = u(t)u(s)$ and $a(t+s) = a(t)a(s)$. Assume that Γ is a discrete subgroup so that $\Gamma\backslash G$ is compact.

(1) $\Gamma\backslash G$ is the unit tangent bundle of a compact surface of constant negative curvature. We think of $\Gamma\backslash\mathbb{H}$ where $\mathbb{H}$ is the upper half-plane. So $T^1(\Gamma\backslash\mathbb{H}) = \{(p, v) : p \in \mathbb{H} \text{ and } v \text{ a unit tangent vector}\}$. Then $T^1(\Gamma\backslash\mathbb{H}) \simeq \Gamma\backslash\mathrm{PSL}_2(\mathbb{R})$.
(2) We have two flows on this surface: the horocycle flow defined by

$$\eta_t(\Gamma x) = \Gamma x u(t)$$

and the geodesic flow defined by

$$\gamma_t(\Gamma x) = \Gamma x a(t).$$

We will consider orbits of these flows but first lets build some geometric intuition for these flows

3.1. Geometric Interpretation.

Geodesics are locally length minimizing, that is $z, w \in \mathbb{H}$ then the geodesic through the two points defines a curve through w and z with the least length.

Proposition 3.1. *The geodesics in $\mathbb{H}$ are the half circles with center on the real axis or the half lines parallel to $i\mathbb{R}$.*

So the flow γ_t corresponds to flowing along a geodesic.

The horocycle flows are a little more complicated. Let $(z, \zeta) \in T^1\mathbb{H}$ and σ the geodesic through z. So $\sigma(0) = z$ and $\dot{\sigma}(0) = \zeta$. Let C_t be the hyperbolic circle (circle in hyperbolic geometry) with center $\sigma(t)$ and passing through z. As $t \to \infty$, C_t converges to a circle tangent to the real axis or to a line parallel to $\mathbb{R}$. (Note that hyperbolic circles are Euclidean circles, but with different radii and centers.)

The circle C_∞ is the positive horocycle.

Example. Let (i, ζ_0) be a point of $T^1\mathbb{H}$ with ζ_0 the unit vector in direction of $i\mathbb{R}^+$ (straight up). Then the C_∞ is the line through i that is parallel to the real axis. Also, the horocycle flow carries the point z to z_t which is the point distance t from z on C_∞. So in our example, z_t is a point with imaginary part equal to i.

Remark. Under the identification $T^1\mathbb{H} \simeq \mathrm{PSL}_2(\mathbb{R})$ we have

$$\eta_t(z) = z \cdot u(t)$$
$$\gamma_t(z) = z \cdot a(t).$$

3.2. **Horocycle and Geodesic Flow.** The horocycle flow has dynamics that we can understand. To see this let Γ be a lattice in $\mathrm{PSL}_2(\mathbb{R})$, then the orbit $\{xu(t)\}$ is dense in $\Gamma\backslash\mathrm{PSL}_2(\mathbb{R})$ or is a period. Further, it is periodic if and only if x is identified with a cusp. ($\gamma \in \Gamma$ and p upper triangular with $\gamma x = g_{i_0} p$ where $g_{i_0} \cdot \infty$ is a cusp see [1] page ? for more.) Note: If $\Gamma\backslash G$ has no cusps, so $\Gamma\backslash G$ is compact, then every orbit is dense.

The geodesic flow is not as nicely behaved. It can be shown that the closure of some orbits of γ_t are very far from "nice". The closure of some orbits, say C, can be homeomorphic to $C' \times \mathbb{R}$ where C' is a cantor set.

So what is the difference between these flows? The answer is that the horocycle flow is unipotent, while the geodesic flow is not. A matrix is called unipotent if 1 is the only complex eigenvalue it has. In other words, a matrix T which is $n \times n$ is called unipotent if $(T - I)^n = 0$.

From these examples what might we guess as theorem about general flows on spaces? Well we should need a space which is homogeneous (a quotient of a group) and we should need the flow to be nicely behaved (perhaps unipotent) and then we might hope to be able to prove a theorem about the closure of a general orbit.

4. Ratner's Theorems

The first theorem of Ratner's that we give is her Orbit Closure Theorem.

Theorem 4.1 (Ratner's Orbit Closure Theorem)**.** *Assume that G is a lie group, Γ is a lattice in G, and that ϕ_t is any unipotent flow on $\Gamma\backslash G$, say $\phi_t([x]) = [\Gamma u(t)]$, then for each $x \in G$, there is a connected, closed subgroup S of G with*

1. $\{u(t)\}_{t \subset \mathbb{R}} \subset S$
2. *$[xS]$ in $\Gamma\backslash G$ is closed and has finite S-invariant volume (that is $x^{-1}\Gamma x \cap S$ is a lattice in S)*
3. *the ϕ_t orbit of x is dense in $[xS]$.*

Example. Let us return the example from the previous section. Let $G = \mathrm{SL}_2(\mathbb{R})$, $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, $u(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$. The only connected subgroups of G that contain the $u(t)$ are $\{u(t)\}$, the upper-triangular matrices, and the entire group. It turns out there are no lattices in the set of upper-triangular matrices. So S must be $\{u(t)\}$ or the entirety of G.

Example. $G = \mathrm{SL}_3(\mathbb{R})$, $\Gamma = \mathrm{SL}_3(\mathbb{Z})$, $u(t) = \begin{pmatrix} 1 & 0 & 0 \\ t & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Some orbits of the flow ϕ_t are closed, others are dense, but there can be intermediate cases as well.

As an exercise you might try see that $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$ naturally embeds into $\mathrm{SL}_3(\mathbb{Z})\backslash\mathrm{SL}_3(\mathbb{R})$ (say into the upper left corner). The image is a submanifold. Furthermore, it is the closure of certain orbits of the $u(t)$ flow.

We have seen that the unipotent flows yield orbits which are dense in some nice geometric subsets. In particular, the orbit $\{\phi_t([x])\}$ is dense in $[xS]$ for some $S \subset G$ with $\Gamma\backslash G$ the space. In fact more is true. The orbit is uniformly distributed in $[xS]$.

Returning to our basic example of $\mathbb{T}^n$, we had the flow ϕ_t on $\mathbb{T}^n$ defined by $\phi_t([x]) = [x + tv]$ for $v \in \mathbb{R}^n$. Furthermore, with μ the normalized Lebesque measure so that μ satisfies $\mu(\mathbb{T}^n) = 1$.

Let $n = 2$ and $B \subset \mathbb{T}^2$ any open subset. Then the amount of time that the flow spends in B is proportional to its measure. That is

$$\frac{1}{T}\left|\{t \in [0,T] : \phi_t(x) \in B\}\right| \to \mu(B)$$

as $T \to \infty$, when $\frac{a}{b} \notin \mathbb{Q}$ and $v = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$. Equivalently, if f is a continuous function on $\mathbb{T}^2$, then

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

For $n = 3$, consider the vector $v = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \in \mathbb{R}^2$ with $\frac{a}{b} \notin \mathbb{Q}$. The orbits are not dense in $\mathbb{T}^3$, but are dense in $\mathbb{T}^2 \times \{x_3\}$. But in fact more is true, as the orbits are uniformly distributed in the subtorus $\mathbb{T}^2 \times \{x_3\}$ in $\mathbb{T}^2$ (for $x = (x_1, x_2, x_3)^T$. Let μ_2 be the Haar measure on the horizontal 2-torus $\mathbb{T}^2 \times \{x_3\}$ that contains $x = (x_1, x_2, x_3)$. Then

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

as $T \to \infty$.

In general for any n and $v \in \mathbb{R}^n$ and $x \in \mathbb{T}^n$, there is a subtorus $S \subset \mathbb{T}^n$ with the Haar measure with

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

as $T \to \infty$.

We have the general theorem:

Theorem 4.2 (Ratner's Equidistribution Theorem)**.** *If G is a lie group and Γ is a lattice in G and ϕ_t is any unipotent flow on $\Gamma\backslash G$, then for any $x \in G$ we have as before $S \subset G$ connected, closed with $\{u(t)\} \subset S$, $[xS]$ closed in $\Gamma\backslash G$ with finite S-invariant volume and ϕ_t-orbit of x dense in $[xS]$. Let μ_S be the unique S-invariant probability measure on $[xS]$ and $f \in C_c(\Gamma\backslash G)$, then*

$$\frac{1}{T}\int_0^T f(\phi_t(s))dt \to \int_{[xS]} f d\mu_S$$

as $T \to \infty$.

This theorem leads us to a classification of ϕ_t-invariant measures.

Definition 4.3. X a metric space, ϕ_t a continuous flow on X, μ a measure on X. We say μ is ϕ_t invariant if $\mu(\phi_t(A)) = \mu(A)$, for all borel subsets A and every $t \in \mathbb{R}$.

We say that μ is ergodic if μ is ϕ_t invariant and every ϕ_t invariant Borel function on X is essentially constant (f is ϕ_t invariant if $f(\phi_t(x)) = f(x)$ for all $t \in \mathbb{R}$).

To understand all invariant measures it suffices to understand the ergodic ones. This is because every invariant probability measure is a convex combination of ergodic ones (not trivial).

This leads us to Ratner's third big theorem in this field.

Theorem 4.4 (Ratner's Measure Classification Theorem). *If G is any Lie group, Γ is any lattice in G, and ϕ_t is any unipotent flow on $\Gamma\backslash G$ then every ergodic ϕ_t-invariant probability measure is of the form μ_S, for some x and some subgroup S as in the previous problem.*

Remark. In practice the Measure Classification Theorem is a step toward proving the Equidistribution Theorem. But it is also a corollary of it.

As a second remark, the work of Venkatesh [5] relies heavily on quantative versions of the Equidistribution Theorem.

References

[1] M.B. Bekka and M. Mayer, *Ergodic Theory and Topological Dynamics of Group Actions on Homogeneous Spaces*, London Math Society Lecture Notes 269.
[2] J. Ellenberg and A. Venkatesh, *Local-global principles for representations of quadratic forms*, preprint.
[3] D.W. Morris, *Ratner's Theorems on Unipotent Flows*, arxiv.
[4] R. C. Rhoades, *Ergodicity of Geodesic and Horocycle Flow*, http://www.math.wisc.edu/ rhoades/Notes/notes.html
[5] A. Venkatesh, *Sparse equidistribution problems, period bounds, and subconvexity*, preprint, arxiv.

Department of Mathematics, University of Wisconsin, Madison, WI 53706
E-mail address: rhoades@math.wisc.edu