HOW TO COUNT LATTICE POINTS USING DYNAMICS?

RONGHAN YUAN

Abstract. In this week, we learned the proof of Howe-Moore vanishing theorem for matrix coefficients of unitary representations of $SL(n,\mathbb{R})$. As a application, this note presents how this is applied to the problem of counting lattice points in hyperbolic plane. The main reference of this note is [1].

Contents

1. Statement of the problem 1
2. Counterpart of the problem in dynamic system 1
3. Proof of the main theorem 3
4. bibliography 5
References 5

1. Statement of the problem

The famous Gauss circle problem is the following

Problem 1.1. *Let $N(R)$ denote the number of lattice points $\mathbb{Z}^2$ in the circle $x^2+y^2 \leq R^2$, then*

$$N(R) \sim \pi R^2 \tag{1.2}$$

as $R \to \infty$. The Gauss problem is about the estimation of error term.

We can ask the same question under the setting of hyperbolic plane, that is

Question 1.1. For a torsion free lattice Γ in $\mathbb{H}^2$, so that $\Sigma = \Gamma\backslash\mathbb{H}^2$ is a surface of finite area. Fix a point $p \in \mathbb{H}^2$, let $B(p,R)$ denote the ball centered at p with radius R, let $N(R,q)$ denote the number of points in $\Gamma q \cap B(p,R)$, then how does $N(R,q)$ grow as $R \to \infty$?

We will finally show that

$$N(R,q) \sim \frac{area(B(p,R))}{area(\Sigma)} \tag{1.3}$$

2. Counterpart of the problem in dynamic system

In this section, we prove a theorem about equidistribution of a series of measures, which plays a key role in the proof of main theorem. We fix the notation $G = SL(2,\mathbb{R})$, and identify it as the two-fold covering of $PSL(2,\mathbb{R})$, which is just the unit tangent bundle of $\mathbb{H}^2$, and denote $K = SO(2)$. Recall that we have the Iwasawa decomposition $G = KAN$. Let $e = I_2$ denote the identity.

First we will define such a series of measure. First we define the 'circle' on the surface Σ.

Definition 2.1. For a point $p \in \mathbb{H}^2$, let $[p]$ denote its image in Σ. With out lost of generality, we may assume $p = I_2$. Then denote

$$S([p], t) = \{[kg_t \cdot p] | k \in K\} = \{[kpg_t] | k \in K\}, \tag{2.2}$$

where $g_t = diag\{e^{\frac{t}{2}}, e^{-\frac{t}{2}}\}$

The dot in $g_t \cdot p$ means the action of this geodesic flow on the point p, and is just pg_t, which is the right multiplication by g_t. In the book [2], the notation is slightly different, the g_t here is denoted as a_t^{-1} in [2], so the geodesic flow under this notation becomes right multiplication by a_t^{-1}.

This is not exactly the circle on the surface Γ, but the image of the circle of radius t under quotient. Because under a left invariant metric, kpg_t is the trajectory of a geodesic starting from p, with initial velocity of k, in the unit tangent bundle. (For simplicity, we assume $p = I_2$, at this time, left multiplication of K exactly gives all unit tangent vectors in the tangent space at I_2. For other points, we may need to choose some conjugation of K. Also recall the fact that the transitive action of $PSL(2, \mathbb{R})$ on unit tangent bundle is given by left multiplication, while the geodesic flow acts by right multiplication.)

It is clear the image of this set under projection $G \to G/K$ describes exactly the geodesic circle on hyperbolic plane centered at p with radius t. We are interested in its projection in $\Gamma\backslash G/K$, because this image will become equidistributed in Σ as $t \to \infty$. It is natural to guess such a behavior will occur. Just by imagine the image of circle in $\mathbb{R}^2/\mathbb{Z}^2$, the image also becomes very dense as the radius tends to infinity.

Strictly speaking, we can define a series of measures λ_t supported on the set $S([p], t)$, in the following way

$$\int_\Sigma f d\lambda_t = \int_K f([kpg_t]) dk \tag{2.3}$$

So the measure λ_t is actually a one-dimensional measure. The following theorem says, as $t \to \infty$, $S([p], t)$ fills the surface Σ evenly.

Theorem 2.4. *The above measures λ_t are equidistributed, that is, for all $f \in C_c(\Sigma)$,*

$$\lim_{t\to\infty} \int_{S([p],t)} f \lambda_t = \frac{1}{area(\Sigma)} \int_\Sigma f(x) d\mu(x), \tag{2.5}$$

where μ is the measure on Σ induced from left Haar measure μ on G. We normalize the left Haar measure such that it is induced from $dx = dkd(an)$, where dk is probability measure on K.

Proof. Let $f \in C_c(\Gamma\backslash G/K)$, let $\tilde{f}$ be the function f lifted to $\Gamma\backslash G$. Because Γ is torsion free, and $K \cap \Gamma$ is a finite subgroup of Γ due to compactness, $K \cap \Gamma = \{e\}$ must hold. We may choose a small enough open neighborhood U of e such that $KU \cap \Gamma = \{e\}$ still holds.

By uniform continuity of f, thus also $\tilde{f}$, upon shrinking U, we may assume

$$|\tilde{f}(gu) - \tilde{f}(g)| < \epsilon, \ \forall g \in G, u \in U$$

Using the Iwasawa decomposition, we can choose neighborhoods U_1 and U_2 of e in A and N respectively, such that $U_1U_2 \subset U$. From Iwasawa decomposition, we know $KU_1U_2 \subset KU$ is a open neighborhood of e in G. Denote $V = U_1U_2$.

Because for all elements n in N, $g_t^{-1}ng_t$ tends to e as $t \to \infty$, upon shrinking U_2, we may assume

$$g_t^{-1}U_2g_t \subset U_2, \ \forall t \geq 0$$

Due to U_1 and g_t commute, since A is abelian, we have

$$Vg_t \subset g_tV \ \forall t \geq 0$$

This implies

$$|\tilde{f}(\Gamma kvg_t) - \tilde{f}(\Gamma kg_t)| = |\tilde{f}(\Gamma kg_t(g_t^{-1}vg_t)) - \tilde{f}(\Gamma kg_t)| < \epsilon \ \forall v \in V, k \in K, t \geq 0 \tag{2.6}$$

Integral the above inequality over K and V in the space $\Gamma\backslash G$, we have

$$\begin{aligned} &|\frac{1}{\mu(\Gamma KV)}\int_{\Gamma KV} \tilde{f}(\Gamma kvg_t) - \tilde{f}(\Gamma kg_t)dkdv| \\ =&|\frac{\mu_{AN}(V)}{\mu(\Gamma K)}\int_K \tilde{f}(\Gamma kg_t)dk - \frac{1}{\mu(\Gamma KV)}\int_{\Gamma KV} \tilde{f}(\Gamma kg_t)dkdv| \\ =&|\int_K \tilde{f}(\Gamma kg_t)dk - \frac{1}{\mu(\Gamma KV)}\int_{\Gamma KV} \tilde{f}(\Gamma kg_t)dkdv| < \epsilon \end{aligned} \tag{2.7}$$

The second equation uses Fubini's theorem, and decompose $KV = K \times V$, due to the fact K is unimodular, the measure $d\mu = dkdv$. Because $KV \cap \Gamma = \{e\}$, integration on $\Gamma\backslash G$ can be identified as integration on $KV \subset G$. The third equation is due to the fact $\mu(\Gamma KV) = \mu_K(K)\mu_{AN}(V) = \mu_{AN}(V)$, which can be derived using Fubini's theorem.

Let χ be the characteristic function of the open subset ΓKV of $\Gamma\backslash G$. By Moore's Ergodicity Theorem, the action of geodesic flow is strong mixing, we have

$$\begin{aligned} &\lim_{t\to\infty} \frac{1}{\mu(\Gamma KV)}\int_{\Gamma KV} \tilde{f}(\Gamma kg_t)dkdv \\ =&\frac{1}{\mu(\Gamma KV)}\int_{\Gamma\backslash G} \tilde{f}(xg_t)\chi(x)dx \\ =&\frac{1}{\mu(\Gamma\backslash G)}\int_{\Gamma\backslash G} \tilde{f}(xg_t)dx \end{aligned} \tag{2.8}$$

The last step uses strong mixing. Hence for t large enough,

$$\begin{aligned} &|\int_K f([kpg_t])dk - \frac{1}{area(\Sigma)}\int_\Sigma f(x)d\mu(x)| \\ =&|\int_K \tilde{f}(\Gamma kg_t)dk - \frac{1}{\mu(\Gamma\backslash G)}\int_{\Gamma\backslash G} \tilde{f}(x)d\mu(x)| < 2\epsilon \end{aligned} \tag{2.9}$$

The last step combines (2.7) and (2.8), and completes the proof of this theorem. □

3. Proof of the main theorem

This theorem before shows circle in hyperbolic space is equidistributed in the $\Gamma\backslash G/K$ as radius tends to infinity. Now we want to relate it to counting lattice points.

Theorem 3.1. *For any* $q \in \mathbb{H}^2$*, one has*

$$N(R,q) \sim \frac{area(B(p,R))}{area(\Sigma)} \tag{3.2}$$

as $R \to \infty$*.*

Proof. Fix $q \in \mathbb{H}^2$, choose $\epsilon > 0$ small enough such that $B(q,\epsilon)$ is isometric to the ball $B([q],\epsilon) \subset \Sigma$.

First, we want to relate the lattice counting with integral.

Choose α to be a positive bump function with integral 1 on Σ, and supported in $B([q],\epsilon)$, $\tilde{\alpha}$ be its lift to $\Gamma\backslash G$. Since $x \mapsto N(R,x)$ is a function invariant under Γ, thus is a function on Σ.

We have

$$N(R-\epsilon,q) \le N(R,x) \le N(R+\epsilon,q) \;\forall x \in B(q,\epsilon) \tag{3.3}$$

Because if $d(\gamma q, p) \le R - \epsilon$,

$$d(\gamma x,p) \le d(\gamma x,\gamma q) + d(\gamma q,p) = d(x,q) + d(\gamma q,p) \le R \tag{3.4}$$

So $N(R-\epsilon,q) \le N(R,x)$, $N(R,x) \le N(R+\epsilon,q)$ can be proved similarly.

Thus

$$N(R-\epsilon,q) \le \int_\Sigma \alpha(x)N(R,x)dx \le N(R+\epsilon,q) \tag{3.5}$$

Denote χ the characteristic function of $B(p,R)$, the integral in the middle can be rewritten as

$$\int_\Sigma \alpha(x)N(R,x)dx = \int_\Sigma \alpha(x)\sum_{\gamma\in\Gamma}\chi(\gamma x)dx = \int_{B(p,R)} \tilde{\alpha}(x)dx \tag{3.6}$$

Using the expression of $\mathbb{H}^2$ under geodesic polar coordinate, which is $\frac{dxdy}{y^2} = \sinh t dt d\theta$. (Recall the hyperbolic metric can also be given as $dt^2 + \sinh^2 t d\theta$. Also recall here $p = I_2$ is chosen as the center of this geodesic polar coordinate.)

$$\begin{aligned}\int_{B(p,R)} \tilde{\alpha}(x)dx &= 2\pi\int_0^R\int_K \tilde{\alpha}(kpg_t)\sinh t dk dt \\ &= 2\pi\int_0^R(\int_\Sigma \alpha d\lambda_t)\sinh t dt\end{aligned} \tag{3.7}$$

Using the equidistribution theorem proved in section 2, and the fact α has integral 1, we have

$$\int_\Sigma \alpha d\lambda_t \to \frac{1}{area(\Sigma)} \;\; as\ t \to \infty \tag{3.8}$$

Since $\sinh t$ grows very fast as $t \to \infty$, the main part of the integral is contributed by $\int_\Sigma \alpha d\lambda_t$ with large t. So we have the estimation

$$\int_{B(p,R)} \tilde{\alpha}(x)dx \sim \frac{area(B(p,R))}{area(\Sigma)} \;\; as\ R \to \infty \tag{3.9}$$

We quote the fact that

Lemma 3.10.

$$area(B(p,R)) = 4\pi\sinh^2(\frac{R}{2}) \tag{3.11}$$

Then $area(B(p,R)) \sim \pi e^R$. Let

$$a(R) = \frac{N(R,q)area(\Sigma)}{\pi e^R} \tag{3.12}$$

Due to (3.5)

$$e^{-\epsilon} a(R-\epsilon) \leq \frac{area(\Sigma)}{\pi e^R} \int_{B(p,R)} \tilde{\alpha} dx \leq e^{\epsilon} a(R+\epsilon) \tag{3.13}$$

For any limit point a of $a(R)$ when $R \to \infty$, since the middle term tends to 1, we have

$$e^{-\epsilon} a \leq 1 \leq e^{\epsilon} a \tag{3.14}$$

Since ϵ can be arbitrarily small, this shows $\lim_{R\to\infty} a(R)$ exists and equal to 1, and the main theorem is proved. □

4. BIBLIOGRAPHY

REFERENCES

[1] M.Bachir Bekka and Matthias Mayer. Ergodic Theory and Topological Dynamics of Group Actions on Homogeneous Spaces. Cambridge University Press. 2000.

[2] Manfred Einsiedler and Thomas Ward. Ergodic Theory: with a view towards Number Theory. Springer-Verlag London Limited. 2011.

[3] Anthony W.Knapp. Lie Groups Beyond an Introduction. Second Edition. Birkhauser. 2002.