1. Application of mixing: Hyperbolic lattice problems

For more along the lines of this application, see the paper of Eskin and McMullen.

It is easy to see that the number of points in $\mathbb{Z}^2$ within a circle of radius R is $\pi R^2 + O(R)$. (The *Gauss circle problem* is the question of replacing $O(R)$ with $O(R^{1/2+\epsilon})$.

Theorem 1. *The number of* $\gamma := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ *so that* $d(\gamma i, i) \leq T$ *is* $\sim c\exp(T)$. *Equivalently, the number of* (a, b, c, d) *so that* $ad - bc = 1$ *and* $a^2 + b^2 + c^2 + d^2 \leq R^2$ *is* $\sim c' R^2$.

Note that, for $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$, the distance $d(gi, i)$ is given explicitly by $2\log\rho$, where $\rho^2 + \rho^{-2} = a^2 + b^2 + c^2 + d^2$. Therefore, the two claims of the statement are indeed equivalent.

Moreover, the analogous theorem holds for $\mathrm{SL}_n(\mathbb{Z})$: the number of integral $n \times n$ matrices with determinant 1, satisfying $\mathrm{Tr}(M^*M) \leq R^2$, is $\sim c'' R^{n^2-n}$. The proof we will give generalizes to that context.

Why is the *Theorem* harder than the Gauss circle problem? Unlike a Euclidean circle, a hyperbolic circle has comparable area and perimeter! We deduce it from the following fact [1]

Fact. The hyperbolic circle $C_R := \{z \in \mathbf{H} : d(z, i) = R\}$, when projected to$\mathrm{SL}_2(\mathbb{Z})\backslash\mathbf{H}$, becomes uniformly distributed, as $R \to \infty$. (i.e. $f \in C_c(\mathrm{SL}_2(\mathbb{Z})\backslash\mathbf{H})$, the integral $\frac{1}{\mathrm{length} C_R}\int_{C_R} f \to \int_{\mathrm{SL}_2(\mathbb{Z})\backslash\mathbf{H}} f$.

Proof of the theorem, assuming FACT. Let $B_R = \{z \in \mathbf{H} : d(z, i) \leq R\}$ and let $N(R) = \#\mathrm{SL}_2(\mathbb{Z}) \cap B_R$. Note that $\mathrm{length}(C_R) \sim \mathrm{area}(B_R) \sim ce^R$ for a suitable constant R.

Let $f_0 \in C_c(\mathbf{H})$ be non-negative, have integral 1, and be supported in an ϵ-neighbourhood of i. Define $f \in C_c(\mathrm{SL}_2(\mathbb{Z})\backslash\mathbf{H})$ via the rule $f(z) := \sum_{\gamma \in \mathrm{SL}_2(\mathbb{Z})} f_0(\gamma z)$. Then

$$\frac{1}{\mathrm{area}(B_R)} \int_{B_R} f(z) \frac{dx\, dy}{y^2} \to \int_{\mathrm{SL}_2(\mathbb{Z})\backslash\mathbf{H}} f = 1$$

But

$$\int_{B_R} f(z) = \sum_{\gamma in \mathrm{SL}_2(\mathbb{Z})} \int_{z \in B_R} f_0(\gamma z)$$

In particular, this quantity is bounded above by $N(R+\varepsilon)$ and bounded below by $N(R-\varepsilon)$. So $N(R) \sim ce^R$.

Proof of the fact. Let $U = K.\{a(s) : |s| \leq \delta\}.\{n(t) : |t| \leq \delta\} \subset G$. It is a small thickening of K inside G.

Take $f \in C_c(\mathrm{SL}_2(\mathbb{Z})\backslash\mathbf{H})$. Let $\tilde{f}$ be the function on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$ defined by $\tilde{f}(g) = f(gi)$. By mixing, as $s \to \infty$,

$$\frac{1}{\mathrm{vol}(U)} \int_{Ua(s)} \tilde{f} \longrightarrow \int_{\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})} \tilde{f} = \int_{\mathrm{SL}_2(\mathbb{Z})\backslash\mathbf{H}} f(z) \frac{dx\, dy}{y^2}.$$

[1] The idea of this proof is due to the PhD thesis of G. Margulis. It was rediscovered and generalized later by A. Eskin and C. McMullen.

On the other hand, note that:

$$\frac{1}{\mathrm{vol}(U)}\int_{U.a(R)}\tilde{f}\sim\frac{1}{\mathrm{length}(C_R)}\int_{C_R}f(z), \tag{1}$$

Let us see why this is true. By definition, $\int_{Ua(R)}\tilde{f}=\int_{g\in U}f(ga(R)i)$. We claim that that, for every $g\in U$, $ga(R)i$ is within $O(\delta)$ of C_R. In fact, we may write $g=kn(t)a(s)$, where $|s|,|t|\leq\delta$; then $d(ga(R)i,ka(R)i)=d(n(t)a(s+R)i,a(R)i)=O(\delta)$.

The rule $f\mapsto\frac{1}{\mathrm{vol}(U)}\int_{g\in U}f(ga(R)i)$ definesa probability measure on $\mathbf{H}$, which is invariant under the subgroup K (because U is so) and which is supported on a thin neighbourhood of C_R. (1) follows.

2. Ratner's theorem

We now state Ratner's theorem, in increasing levels of abstraction.

2.1. $\mathrm{SL}_2(\mathbb{R})$**-case.** A special case of Ratner's theorem is:

Theorem 2. *Any orbit $x_0n(t)$ on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$ is either dense or periodic. Moreover, in the former case, $\{n(t)x_0\}$ becomes uniformly distributed with respect to the $\mathrm{SL}_2(\mathbb{R})$-invariant probability measure on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$.*

In other terms, the measures defined by $f\mapsto\frac{1}{T}\int_0^T f(x_0n(t))$ converge, weakly, to the $\mathrm{SL}_2(\mathbb{R})$-invariant probability measure on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$.

It should be noted that no such theorem holds if we replace the role of the subgroup $\{n(t)\}$ by $\{a(s)\}$.

2.2. $\mathrm{SL}_n(\mathbb{Z})\backslash\mathrm{SL}_n(\mathbb{R})$**-case.** To work out the theorem, we should first of all work out what is special about the subgroup $n(t)$. The key difference between $\{n(t)\}$ and $\{a(s)\}$ is that $n(t)$ is *unipotent*, i.e. all the eigenvalues of $n(t)$, for any t, are 1.

Thus let $X\in M_n(\mathbb{R})$ and let $u(t)=\exp(tX)\in\mathrm{SL}_n(\mathbb{R})$.

Theorem 3. *Let $x_0\in\mathrm{SL}_n(\mathbb{Z})\backslash\mathrm{SL}_n(\mathbb{R})$. Then the closure of $x_0\exp(tX)$ is of the form x_0S, where $S\supset\{u(t)\}$ is a closed subgroup. Moreover, x_0S supports an S-invariant probability measure, and the trajectory $\{x_0u(t)\}$ becomes uniformly distributed with respect to that measure.*

This implies the previous theorem: if we take $\{n(t)\}\subset\mathrm{SL}_2(\mathbb{R})$, the only possibilities for S are $S=\{n(t)\},S=\mathrm{SL}_2(\mathbb{R})$, and S the group of upper triangular matrices. However, in the latter case x_0S cannot support an S-invariant probability measure: for this to be so, S must be *unimodular* (left Haar measure = right Haar measure).

2.3. **Ratner's theorem – abstract statement.** The most abstract form of Ratner's theorem is:

Theorem 4. *Let G be a Lie group, $\Gamma\leqslant G$ a discrete subgroup. Let H be a subgroup of G generated by Ad-unipotent elements.*

- *(Classification of H-invariant closed sets). The closure of any H-orbit, $\overline{x_0H}$, on $\mathrm{SL}_n(\mathbb{Z})\backslash\mathrm{SL}_n(\mathbb{R})$ is* homogeneous, *i.e. a closed S-orbit x_0S, for a closed subgroup $S\supset H$. Moreover, x_0S admits an S-invariant probability measure.*

- *(Classification of H-invariant measures). Any H-invariant H-ergodic probability measure on $\Gamma\backslash G$ is* homogeneous*, i.e. the S-invariant measure on a closed S-orbit $x_0 S$, for a closed subgroup $S \supset H$.*

3. Proof of Theorem 2 for $\mathrm{SL}_2(\mathbb{R})$: a sketch.

3.1. **Idea of proof in the cocompact case.** Suppose that $\Gamma \leqslant \mathrm{SL}_2(\mathbb{R})$ is a *cococompact* discrete subgroup. We shall show that, for *any* $x \in \Gamma\backslash\mathrm{SL}_2(\mathbb{R})$, the trajectories $xn(t), 0 \leq t \leq T$, become uniformly distributed w.r.t. the $\mathrm{SL}_2(\mathbb{R})$-invariant probability measure on $\Gamma\backslash\mathrm{SL}_2(\mathbb{R})$. In particular, all orbits of $n(t)$ are dense.

Proof. First of all, the action of $SL_2(\mathbb{R})$ on $\Gamma\backslash\mathrm{SL}_2(\mathbb{R})$ is mixing. The same proof as given previously carries over. Mixing shows that, for $f_1, f_2 \in L^2$, we have:

$$\langle a(s)f_1, f_2\rangle \to \int f_1 \int f_2, s \to \infty. \tag{2}$$

Moreover, a simple argument shows that we can take this convergence to be *uniform* if f_1, f_2 are allowed to vary in compact sets of L^2.

The basic strategy of the proof is to use the fact that $a(s)$, for large s, "stretches" in the $n(t)$-direction; in particular:

(3)

A long orbit of $n(t)$ is approximated by the translate of a small ball under $a(s)$.

Combining this with (2), taking for f_1 the characteristic function of a small ball, gives the conclusion.

3.2. **Cocompact quotients of** $\mathrm{SL}_2(\mathbb{R})$**.** How do we construct $\Gamma \leqslant \mathrm{SL}_2(\mathbb{R})$ so that the quotient $\Gamma\backslash\mathrm{SL}_2(\mathbb{R})$ is compact?

They exist in abundance. If M is a Riemann surface of genus ≥ 2, the universal cover of M is $\mathbb{H}$. The fundamental group $\pi_1(M)$ acts on $\mathbb{H}$. The group of holomorphic automorphisms of $\mathbb{H}$ is $\mathrm{PSL}_2(\mathbb{R})$, and, therefore, $\pi_1(M)$ defines a subgroup of $\mathrm{PSL}_2(\mathbb{R})$. It is easily seen to be discrete. Moreover, $\mathbb{H}/\pi_1(M)$ is compact, and so also $\mathrm{PSL}_2(\mathbb{R})/\pi_1(M)$.

3.3. **Proof for** $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$**.** Clearly, the proof we sketched above can't work for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$, because the proof above doesn't allow for periodic $n(t)$-orbits!

Suppose that $x_0 \in \Gamma\backslash\mathrm{SL}_2(\mathbb{R})$. Then, either:

(1) $x_0 n(t)$ is periodic, or;
(2) There exists a compact subset $K \subset \mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$ and an infinite sequence of times $s_1, s_2, \ldots$ so that $x_0 a(s_i) \in K$.

This is "clear by looking at the picture": if we think of orbits as geodesics on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbf{H}$, and in particular reduce to the standard fundamental domain, a geodesic either goes vertically off to ∞, or goes in a big semicircle; in the latter case, it returns to a compact set.