# EQUIDISTRIBUTION UNDER THE GAUSS MAP

MANUEL W. LUETHI

Abstract. Notes on our seminar

## 1. Definitions and notation

For notational purposes we assume that 0 is *not* a natural number.

## 2. Continued Fraction expansion, Gauss map and the geodesic flow

By $\frac{p}{q} \in \mathbb{Q}$ we always mean an element $r \in \mathbb{Q}$ and a choice $p \in \mathbb{Z}$, $q \in \mathbb{N}$ such that $\gcd(p,q) = 1$ and $r = \frac{p}{q}$.

### 2.1. Continued fraction expansion.

Let $\mathcal{F}(\mathbb{N})$ denote the set of finite words over the alphabet $\mathbb{N}$. We define a *continued fraction expansion* as a map from $\mathbb{Z} \times \mathcal{F}(\mathbb{N})$ to $\mathbb{Q}$. Given a word $[a_0 : a_1 : \cdots : a_n]$ of length $n+1$, $n \in \mathbb{N} \cup \{0\}$, with letters $a_0, \ldots, a_n \in \mathbb{N}$, we define

$$Q([a_0; a_1 : \cdots : a_n]) = a_0$$

if $n = 0$ and inductively set

$$Q([a_0; a_1 : \cdots : a_n]) = a_0 + \frac{1}{Q([a_1; \cdots : a_n])}$$

if $n \in \mathbb{N}$.

**Definition 2.1.** *A continued fraction expansion (cfe) is a sequence of partial fractions* $[a_0; a_1 : \cdots : a_n]$.

We will write $[a_0; a_1 : \cdots]$ to denote the cfe with partial fractions $[a_0; a_1 : \cdots : a_n]$. Note that strictly speaking rational numbers do not have a continued fraction expansion. We will say more about this after having proven certain properties of the *partial fractions* associated with a continued fraction expansion.

In what follows, we will omit mentioning $Q$, i.e. we identify $[a_0; a_1 : \cdots : a_n]$ with its image under $Q$.

**Lemma 2.2** ([EW11, Lem. 3.1])**.** *Let $[a_0; a_1 : a_2 : \cdots]$ be a cfe. Given $n \in \mathbb{N} \cup \{0\}$ let $p_n \in \mathbb{Z}$ and $q_n \in \mathbb{N}$ coprime such that*

$$\frac{p_n}{q_n} = [a_0; a_1 : \cdots : a_n]$$

*for all $n \in \mathbb{N}$. Set $p_{-1} = 1$, $q_{-1} = 0$, $p_{-2} = 0$, $q_{-2} = 1$. Then the sequence $(p_n, q_n)_{n \in \mathbb{N}}$ is given inductively by*

$$\begin{pmatrix} p_{n+1} & p_n \\ q_{n+1} & q_n \end{pmatrix} = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} \begin{pmatrix} a_{n+1} & 1 \\ 1 & 0 \end{pmatrix} \qquad (n \in \mathbb{N} \cup \{0\}).$$

The proof is not hard but a bit delicate in terms of bookkeeping, as the induction tries to deduce the statement for $[a_0; a_1 : \cdots : a_n]$ from $[a_1; a_2 : \cdots : a_n]$. We refer the reader to [EW11].

*Date*: November 24, 2019.

**Corollary 2.3.** *Let* $[a_0; a_1 : \cdots]$ *be a cfe. Then*

$$p_{n+1} = a_{n+1}p_n + p_{n-1},$$
$$q_{n+1} = a_{n+1}q_n + q_{n-1}.$$

**Corollary 2.4.** *Let* $[a_0; a_1 : \cdots]$ *be a cfe. Then*

$$\forall n \in \mathbb{N}\, q_n \geq 2^{\frac{n-1}{2}}.$$

*If* $a_0 \geq 0$*, then*

$$\forall n \in \mathbb{N}\, p_n \geq 2^{\frac{n-1}{2}}.$$

In what follows we will assume that $a_0 \geq 0$.

**Corollary 2.5.** *For all* $n \in \mathbb{N}$ *we have*

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n+1}.$$

**Corollary 2.6.** *For all* $n \in \mathbb{N}$ *we have*

$$[a_0; a_1 : \cdots : a_n] = a_0 + \sum_{\ell=1}^{n} \frac{(-1)^{\ell+1}}{q_{\ell-1}q_\ell}$$

*Proof.* The statement is clearly true for $n = 1$. Assume that the statement is true for $n \in \mathbb{N}$. Then

$$\frac{p_{n+1}}{q_{n+1}} = \frac{p_n}{q_n} + \frac{(-1)^{n+2}}{q_n q_{n+1}} = a_0 + \sum_{\ell=1}^{n+1} \frac{(-1)^{\ell+1}}{q_{\ell-1}q_\ell}.$$

□

**Corollary 2.7.** *Let* $[a_0; a_1 : \cdots]$ *be a cfe. The sequence* $\alpha_n = [a_0; a_1 : \cdots : a_n]$ *converges.*

**Definition 2.8.** *Let* $\alpha \in (0, \infty)$*. A cfe* $[a_0; a_1 : \cdots]$ *is a* continued fraction expansion of $\alpha$ *if*

$$\alpha = \lim_{n\to\infty} [a_0; a_1 : \cdots : a_n].$$

**Proposition 2.9.** *Let* $[a_0; a_1 : \cdots]$ *be a cfe. Then the limit* $\alpha = \lim_{n\to\infty}[a_0; a_1 : \cdots : a_n]$ *is irrational. We have* $\alpha \in (0, 1)$ *if and only if* $a_0 = 0$*. For every* $\alpha \in (0, 1) \setminus \mathbb{Q}$ *there exists a cfe* $[0; a_1 : \cdots]$ *such that*

$$\alpha = \lim_{n\to\infty} [a_0; a_1 : \cdots : a_n].$$

*Moreover, it is uniquely determined by* $\alpha$*.*

**Corollary 2.10.** *Let* $\alpha \in (0, \infty) \setminus \mathbb{Q}$*. Then*

$$\alpha = a_0 + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{q_{n-1}q_n},$$

*where* $(p_n, q_n)$ *is the sequence corresponding to the cfe* $[a_0; a_1 : \cdots]$ *of* $\alpha$*.*

Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Then the cfe $[a_0; a_1 : \cdots]$ of $\alpha$ gives rise to a sequence of rational numbers $\frac{p_n}{q_n} = [a_0; a_1 : \cdots : a_n]$ which approximate $\alpha$. We will see that these approximations are optimal in some sense. As $\alpha$ is irrational, it follows from general properties of alternating series that

$$\frac{p_n}{q_n} < \alpha < \frac{p_{n+1}}{q_{n+1}}$$

for all even $n \in \mathbb{N}$.

Let $n \in \mathbb{N}$, then

$$\left|\alpha - \frac{p_n}{q_n}\right| \leq \frac{1}{q_n q_{n+1}} \leq 2^{\frac{1}{2}-n}.$$

One advantage of cfe is that we can give pretty precise bounds on the error term in the sense that we can also bound it from below. This needs a bit more (elementary) preparation.

Let $\ell \in \mathbb{N} \cup \{0\}$. Then

$$\frac{(\ell+1)p_n + p_{n-1}}{(\ell+1)q_n + q_{n-1}} - \frac{\ell p_n + p_{n-1}}{\ell q_n + q_{n-1}} = \frac{p_n q_{n-1} - p_{n-1}q_n}{((\ell+1)q_n + q_{n-1})(\ell q_n + q_{n-1})},$$

i.e. the sign of the above expression solely depends on $n$. In particular, we get that

$$\frac{p_{n+1}}{q_{n+1}} = \frac{a_{n+1}p_n + p_{n-1}}{a_{n+1}q_n + q_{n-1}} > \cdots > \frac{p_{n-1}}{q_{n-1}}$$

for odd $n$ and similarly

$$\frac{p_{n+1}}{q_{n+1}} = \frac{a_{n+1}p_n + p_{n-1}}{a_{n+1}q_n + q_{n-1}} < \cdots < \frac{p_{n-1}}{q_{n-1}}$$

for even $n$. In particular

$$\left|\alpha - \frac{p_n}{q_n}\right| > \left|\frac{p_n + p_{n+1}}{q_n + q_{n+1}} - \frac{p_n}{q_n}\right| \geq \frac{1}{q_n(q_n + q_{n+1})}.$$

**Definition 2.11** (Best approximation of real numbers)**.** *Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. A rational number $\frac{p}{q} \in \mathbb{Q}$ is called a* best approximation *(of the second kind) of $\alpha$ if for any $\frac{r}{s} \in \mathbb{Q}$ with $1 \leq s \leq q$ we have*

$$|q\alpha - p| \geq |s\alpha - r| \implies \frac{p}{q} = \frac{r}{s}.$$

**Remark 2.12.** Let $\frac{p}{q}$ be a best approximation and let $(r, s) \in \mathbb{Z} \times \mathbb{N} \setminus \{(p, q)\}$ such that $1 \leq s \leq q$. Then

$$\left|\alpha - \frac{r}{s}\right| = \frac{1}{s}|s\alpha - r| \geq \frac{1}{q}|q\alpha - p| > \left|\alpha - \frac{p}{q}\right|.$$

**Theorem 2.13** ([Kh64, Thm. 16])**.** *Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ with partial fractions $\frac{p_n}{q_n}$ associated to its cfe. Let $r \in \mathbb{Q}$ be a best approximation. Then there exists some $n \in \mathbb{N}$ such that*

$$r = \frac{p_n}{q_n}.$$

*If $n \in \mathbb{N}$, then $\frac{p_n}{q_n}$ is a best approximation.*

**Remark 2.14.** A similar discussion can be carried out for rational numbers. In this case one asks how well a rational number with possibly large denominator can be approximated by rational numbers of small denominator. In this case the sequence of approximations eventually stabilizes. We omit this discussion here for the simplicity of exposition.

2.2. **The Gauss map and the Gauss measure.** We use the following notation. Given $\alpha \in \mathbb{R}$, we let

$$[\alpha] = \sup\{k \in \mathbb{Z} : k \leq \alpha\}$$

denote the integral part of $\alpha$. We call $\{\alpha\} = \alpha - [\alpha]$ the fractional part of $\alpha$.

**Definition 2.15** (Gauss map)**.** *Define $\mathcal{G} : [0, 1) \to [0, 1)$ by*

$$\alpha = \begin{cases} 0 & \text{if } \alpha = 0, \\ \{\frac{1}{\alpha}\} & \text{otherwise.} \end{cases}$$

**Lemma 2.16.** *Let $\alpha \in [0, 1]$. The following are equivalent.*

1. *There exists $n \in \mathbb{N} \cup \{0\}$ such that $\mathcal{G}^n(\alpha) = 0$.*
2. $\alpha \in \mathbb{Q}$.

*Proof.* We can without loss of generality assume that $\alpha \neq 0$. In order to check that (2) implies (1), note that $\mathcal{G}(\alpha) \in \mathbb{Q} \cap [0,1)$ for any $\alpha \in \mathbb{Q} \cap (0,1)$ and either $\mathcal{G}(\alpha) = 0$ or the denominator of $\mathcal{G}(\alpha)$ is strictly smaller than the denominator of $\alpha$. Hence the desired implication.

For the opposite implication, note that for any $\alpha \in [0,1) \setminus \mathbb{Q}$ we also have $\mathcal{G}(\alpha) \notin \mathbb{Q}$. Iterating this argument shows that $\mathcal{G}^n(\alpha) \neq 0$ for all $n \in \mathbb{N}$. □

We denote by $\nu$ the Gauss measure on $[0,1)$ given by

$$\nu(a,b) = \frac{1}{\log 2} \int_a^b \frac{1}{1+x} \mathrm{d}x$$

for all $0 \leq a < b < 1$.

**Theorem 2.17** ([EW11, Lem. 3.5 and Prop. 9.25])**.** *$\nu$ is $\mathcal{G}$ invariant and ergodic.*

*Proof of invariance.* It suffices to show that $\nu(\mathcal{G}^{-1}(0,b)) = \nu((0,b))$ for all $b \in (0,1)$. As $[0,1) \setminus \mathbb{Q}$ is a $\mathcal{G}$-invariant set of full $\nu$-measure, it suffices to check that the statement is true for $(0,b) \setminus \mathbb{Q}$. For $\alpha \in (0,1) \setminus \mathbb{Q}$ one checks that

$$0 < \mathcal{G}(\alpha) < b \iff \exists n \in \mathbb{N}\ \tfrac{1}{n+b} < \alpha < \tfrac{1}{n}.$$

Hence $b \in (0,1)$ yields

$$\nu(\mathcal{G}^{-1}(0,b)) = \sum_{n \in \mathbb{N}} \nu\big((\tfrac{1}{n+b}, \tfrac{1}{n})\big).$$

One calculates

$$\nu\big((\tfrac{1}{n+b}, \tfrac{1}{n})\big) = \frac{1}{\log 2} \log\left(\frac{n+1}{n} \frac{n+1}{n+1+b}\right)$$

and therefore

$$\nu\big(\mathcal{G}^{-1}(0,b)\big) = \frac{1}{\log 2} \log\left(\prod_{n \in \mathbb{N}} \frac{n+1}{n} \frac{n+b}{n+1+b}\right) = \frac{\log(1+b)}{\log 2} = \nu\big((0,b)\big).$$

□

**Remark 2.18.** The Gauss map and the continued fraction expansion are closely related. Let $\alpha \in [0,1) \setminus \mathbb{Q}$ and let $[0; a_1 : \cdots]$ be the cfe of $\alpha$. Then $a_1 = [\frac{1}{\alpha}]$, $a_2 = [\frac{1}{\mathcal{G}(\alpha)}]$, $a_3 = [\frac{1}{\mathcal{G}^2(\alpha)}]$ and more generally

$$a_n = \left[\frac{1}{\mathcal{G}^{n-1}(\alpha)}\right]$$

for all $n \in \mathbb{N}$.

2.3. **Continued fractions and the geodesic flow on the space of lattices.** We let $G = \mathrm{SL}_2(\mathbb{R})$ and $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. Denote $X = G / \Gamma$. Recall that $X$ identifies with the space of unimodular lattices in $\mathbb{R}^2$ by the map

$$g\Gamma \mapsto g\mathbb{Z}^2.$$

In what follows we will use the following notation.

$$A = \left\{ a_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} : t \in \mathbb{R} \right\}$$
$$U = \left\{ u_\alpha = \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix} : \alpha \in \mathbb{R} \right\}.$$

Note that

$$u_\alpha \mathbb{Z}^2 = \left\{ \begin{pmatrix} m - n\alpha \\ n \end{pmatrix} : n, m \in \mathbb{Z} \right\}$$

and

$$U\Gamma \cong {}^{\mathbb{R}}/_{\mathbb{Z}}.$$

**Definition 2.19.** *Let* $\Lambda = g\mathbb{Z}^2$ *be a lattice in* $\mathbb{R}^2$*. A vector* $v = \left(\begin{smallmatrix} v_1 \\ v_2 \end{smallmatrix}\right) \in \Lambda$ *is a* best approximation *if there are no* $w = \left(\begin{smallmatrix} w_1 \\ w_2 \end{smallmatrix}\right) \in \Lambda \setminus \{0, \pm v\}$ *such that* $|w_i| \leq |v_i|$ *for* $i = 1, 2$*.*

Fix a matrix $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in G$. The standard action of $G$ on $\mathbb{R}^2$ corresponds to simultaneous evaluations of the linear forms $(x, y) \mapsto ax + by$ and $(x, y) \mapsto cx + dy$. Therefore, being a best approximation corresponds to an integer vector which gives a "good approximation of 0" for both linear forms simultaneously.

**Lemma 2.20.** *Let* $\alpha \in \mathbb{R}$*. Then* $v \in u_\alpha \mathbb{Z}^2 \setminus \mathbb{Z} \times \{0\}$ *is a best approximation if and only if* $v = \pm u_\alpha \left(\begin{smallmatrix} p_n \\ q_n \end{smallmatrix}\right)$ *for a partial fraction* $\frac{p_n}{q_n}$ *of* $\alpha$*.*

*Proof.* A vector $v = u_\alpha \left(\begin{smallmatrix} m \\ n \end{smallmatrix}\right)$ is a best approximation if and only if for all $\left(\begin{smallmatrix} m' \\ n' \end{smallmatrix}\right) \in \mathbb{Z} \setminus \mathbb{Z} \times \{0\}$ with $|n'| \leq |n|$ and $|m' - n'\alpha| \leq |m - n\alpha|$ we have $\left(\begin{smallmatrix} m' \\ n' \end{smallmatrix}\right) = \pm \left(\begin{smallmatrix} m \\ n \end{smallmatrix}\right)$. In particular, $\frac{m}{n}$ is a best approximation for $\alpha$ and hence the claim. □

**Theorem 2.21** (Minkowski's convex body theorem, cf. [Ne99, Ch. I, Thm. 4.4])**.** *Let* $\Lambda \subseteq \mathbb{R}^2$ *be a unimodular lattice and let* $C \subseteq \mathbb{R}^2$ *a centrally symmetric, convex subset. Suppose that* $\mathrm{vol}(C) > 4$*. Then*

$$\Lambda \cap C \neq \{0\}.$$

*Proof.* We show that there exist distinct $v, w \in \Lambda$ such that $\frac{1}{2}C + v \cap \frac{1}{2}C + w \neq \emptyset$. Assuming this, there are $x_1, x_2 \in C$ such that $\frac{1}{2}(x_1 - x_2) = v - w \in \Lambda \setminus \{0\}$. Thus $v - w \in C$ as $x_1, x_2 \in C$ and as $C$ is centrally symmetric, convex.

So assume that the statement was not true and let $F$ be a fundamental parallelogram for $\Lambda$, i.e.

$$F = \{t_1 v_1 + t_2 v_2 : 0 \leq t_1, t_2 < 1\}$$

for a basis $v_1, v_2 \in \mathbb{R}^2$ of $\Lambda$. Then we have in particular that

$$\big(F \cap (\tfrac{1}{2}C + v)\big) \cap \big(F \cap (\tfrac{1}{2}C + w)\big) = \emptyset$$

for all pairs of distinct elements $v, w \in \Lambda$ and hence

$$1 = \mathrm{vol}(F) \geq \sum_{v \in \Lambda} \mathrm{vol}\big(F \cap (\tfrac{1}{2}C + v)\big) = \sum_{v \in \Lambda} \mathrm{vol}\big((F - v) \cap \tfrac{1}{2}C\big).$$

As

$$\mathbb{R}^2 = \bigsqcup_{v \in \Lambda} (F - v),$$

this implies $4 \geq \mathrm{vol}(C)$, which is absurd. □

Minkowski's theorem implies the following.

**Proposition 2.22.** *Let* $\Lambda$ *be a lattice in* $\mathbb{R}^2$*. Then one of the following is true.*

(1) *The set of best approximations in* $\Lambda$ *is infinite.*
(2) $\Lambda \cap (\{0\} \times \mathbb{R}) \neq \{0\}$.

*Proof.* Assume that $\Lambda \cap (\{0\} \times \mathbb{R}) = \{0\}$. Denote by $2\varrho$ the length of a shortest vector in $\Lambda$. Let $S \subseteq \Lambda$ be a finite (possibly empty) set of best approximations. Denote by $S_1 \subseteq \mathbb{R}$ the projection of $S$ to the first coordinate and by $|S_1|$ the set consisting of the absolute values of its elements. Define

$$r = \min\{|S_1| \cup \{\varrho\}\}.$$

Given $s \in (0, \infty)$ let

$$C_{r,s} = \{v = \left(\begin{smallmatrix} v_1 \\ v_2 \end{smallmatrix}\right) \in \mathbb{R}^2 : |v_1| \leq \tfrac{r}{2}, |v_2| \leq \tfrac{s}{2}\}.$$

Then $C_{r,s}$ is centrally symmetric and convex, $C_{r,s} \cap S = \emptyset$ and $\lim_{s \to \infty} \mathrm{vol}(C_{r,s}) = \infty$. By Minkowski's convex body theroem, there exists a minimal $s$ such that $C_{r,s} \cap \Lambda \neq \{0\}$. Note that by definition of $r$ and by the assumption that $\Lambda$ intersects the vertical axis trivially, for any element $\binom{v_1}{v_2}$ in the intersection the vector $\binom{v_1}{-v_2}$ is not contained in the intersection. In particular, any element in this intersection is a best approximation. □

**Lemma 2.23.** *Let $X' \subseteq X$ be the set of lattices intersecting the vertical axis trivially. Then $X'$ has full measure with respect to the $G$-invariant probability measure on $X$.*

*Proof.* Assume that $\Lambda \subseteq \mathbb{R}^2$ is a lattice that contains a non-trivial vector in $\{0\}\times\mathbb{R}$. We claim that $\Lambda$ can be represented by a lower-triangular matrix. So let $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$ such that $g\mathbb{Z}^2 = \Lambda$. If $b = 0$, there is nothing to show. If $a = 0$, let $k = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and note that $gk$ is a lower-triangular matrix such that $\Lambda = gk\mathbb{Z}^2$. Hence we are left with the case that $ab \neq 0$. By assumption there exist $m, n \in \mathbb{Z}$, $mn \neq 0$, such that $ma + nb = 0$. We can assume without loss of generality that $m, n$ are coprime. Let $k, \ell \in \mathbb{Z}$ such that $kn - \ell m = 1$. Then $\gamma = \begin{pmatrix} k & m \\ \ell & n \end{pmatrix} \in \Gamma$, $g\gamma$ is lower-triangular, and $g\gamma\mathbb{Z}^2 = \Lambda$.

Let $B$ denote the group of lower-triangular matrices in $G$. Then clearly $B\Gamma / \Gamma \subseteq X \setminus X'$.

Combining the two cases, we have show that $X \setminus X'$ equals $B\Gamma / \Gamma$. This is a lower dimensional subset and in particular has measure zero. □

We note the following Corollary of the proof.

**Corollary 2.24.** *$X'$ is $A$-invariant.*

For what follows it will be useful to introduce the following notation. Let $\Lambda \subseteq \mathbb{R}^2$ be a lattice. Then $\Lambda_+$ denotes the elements in $\Lambda$ whose second coordinate is at least 1. In what follows we let

$$Y = \{y \in [1,\infty)^{\mathbb{N}} : y_1 = 1, \forall n \in \mathbb{N}\, y_n < y_{n+1}, \liminf_{n\to\infty} y_n = \infty\}$$

equipped with the restriction of the product topology. We define a function $f_1 : X' \to Y$ by attaching to a lattice $\Lambda \in X'$ the sequence (in increasing order) of second coordinates of best approximations contained in $\Lambda_+$, rescaled so that the first entry in the sequence equals 1.

**Lemma 2.25.** *Let $\Lambda \in X'$ and let $y \in Y$. Then*

$$\forall n \in \mathbb{N}\, y_{n+2} = a_{n+2}y_{n+1} + y_n$$

*for some sequence $a \in \mathbb{N}^{\mathbb{N}}$.*

*Proof.* We refer to [Ka13, Ch. 10]. □

**Lemma 2.26.** *Let $\Lambda \subseteq \mathbb{R}^2$ a unimodular lattice and $v \in \Lambda \setminus \{0\}$. Let $t \in \mathbb{R}$. The following are equivalent.*

(1) *$v$ is a best approximation in $\Lambda$.*
(2) *$a_t v$ is a best approximation in $a_t\Lambda$.*

*Proof.* This is immediate. □

**Corollary 2.27.** *Let $T : Y \to Y$ denote the left-shift, i.e. $T(y)_i = y_{i+1}$ for all $i \in \mathbb{N}$. For every $t \geq 0$ and for every $\Lambda \in X'$ there exists some $n \in \mathbb{N} \cup \{0\}$ such that $f_1(a_t\Lambda) = T^n \circ f_1(\Lambda)$.*

*Proof.* The second coordinates of best approximations in $a_t\Lambda$ are of the form $e^{-t}y$, where why is the second coordinate of a best approximation in $\Lambda$. Choose $n \in \mathbb{N} \cup \{0\}$ minimal such that $f_1(\Lambda)_n \geq e^t$. □

**Lemma 2.28.** *$f_1$ is continuous outside of*

$$\{\Lambda \in X' : \Lambda \cap (\mathbb{R} \times \{0,1\}) \neq \},$$

*i.e. on a full measure set.*

2.4. **The theorem and its proof.** Let us state the main

**Theorem 2.29** ([SW19, Thm. 13.1])**.** *Fix $\alpha \in (0,1)$ and suppose that the orbit $\{a_t u_\alpha \Gamma : t \geq 0\}$ is equidistributed in $G / \Gamma$. Then the orbit $\{\mathcal{G}^n(\alpha) : n \in \mathbb{N} \cup \{0\}\}$ is equidistributed with respect to the Gauss measure.*

**Remark 2.30.**

- If $\alpha \in \mathbb{Q}$, then $\{a_t u_\alpha \Gamma : t \geq 0\}$ does not equidistribute in $G / \Gamma$. It might behave wildly but once $t$ exceeds the log of the denominator of $\alpha$, it wanders off into the cusp along a vertical line. So there was no harm in the restriction of the discussion of cfe to irrational $\alpha$. Note also that for irrational $\alpha$ the lattice $u_\alpha \mathbb{Z}^2$ intersects the vertical axis trivially.
- The cfe of irrational numbers defines a map $(0,1) \setminus \mathbb{Q} \to \mathbb{N}^{\mathbb{N}}$. As argued previously, this map is a bijection. Let $\sigma$ denote the left-shift on $\mathbb{N}^{\mathbb{N}}$, i.e. for all $y \in \mathbb{N}^{\mathbb{N}}$ and for all $n \in \mathbb{N}$ we have

  $$\sigma(y)_n = y_{n+1}.$$

  Then the diagram in (2.1) commutes.

  $$\begin{array}{ccc} (0,1) \setminus \mathbb{Q} & \xrightarrow{\mathcal{G}} & (0,1) \setminus \mathbb{Q} \\ \downarrow{\scriptstyle \mathrm{cfe}} & & \downarrow{\scriptstyle \mathrm{cfe}} \\ \mathbb{N}^{\mathbb{N}} & \xrightarrow[\sigma]{} & \mathbb{N}^{\mathbb{N}} \end{array} \tag{2.1}$$

  The push-forward of the Gauss measure $\nu$ under the cfe hence defines a shift-invariant ergodic probability measure $\tilde{\nu}$ on $\mathbb{N}^{\mathbb{N}}$.
- Courtesy of the previous statement, the orbit of $\alpha$ under the Gauss map equidistributes with respect to the Gauss measure $\nu$ if and only if the orbit of the cfe equidistributes under the left-shift with respect to the induced measure $\tilde{\nu}$.

**Lemma 2.31.** *Let $\Lambda \in X'$ and assume that the forward orbit $\{a_t \Lambda\}$ equidistributes in $X$ with respect to the $G$-invariant probability measure. Then the orbit $\{(T^n \circ f_1)(\Lambda) : n \in \mathbb{N}\}$ equidistributes in $Y$ with respect to some probability measure which is independent of $Y$.*

*Proof.* We need to show that for $F$ in a suitable class of test-functions there is some positive linear functional $\lambda$ such that

$$\frac{1}{n} \sum_{\ell=0}^{n-1} (F \circ T^\ell \circ f_1)(\Lambda) \overset{n \to \infty}{\longrightarrow} \lambda(F)$$

and $c(1) = 1$. As $Y$ is a metric space, the right notion is the notion of weak convergence and the right set of test functions is the set of bounded, continuous funtions on $Y$. Note that $Y$ is not locally compact.

Given a bounded, continuous function $F : Y \to \mathbb{R}$ define $F' : Y \to \mathbb{R}$ by

$$F'(y) = \sum_{\substack{n \in \mathbb{N} \\ y_n < e}} (F \circ T^n)(y).$$

The sum in the definition of $F'$ is finite by definition of $Y$. In fact, the recursive relation implies that on $f_1(X')$ the sum has at most three summands. Therefore

$F'$ is bounded and continuous on $f_1(X')$. Let $\Lambda \in X'$ and let $y = f_1(\Lambda)$. Then $1 \leq e^{-t} y_\ell < e$ if and only if $\log y_\ell - 1 < t \leq \log y_\ell$. Therefore

$$(F' \circ f_1)(a_t \Lambda) = \sum_{\substack{n \in \mathbb{N} \\ 1 \leq e^{-t} y_n < e}} (F \circ T^n)(y)$$

and thus for all $n \in \mathbb{N}$

$$\begin{aligned} \int_0^{\log y_n} (F' \circ f_1)(a_t \Lambda) \mathrm{d}t &= \int_0^{\log y_n} \sum_{\substack{\ell \in \mathbb{N} \\ 1 \leq e^{-t} y_\ell < e}} (F \circ T^\ell)(y) \mathrm{d}t \\ &= \sum_{k=1}^{n} \int_{\log y_k - 1}^{\log y_k} \sum_{\substack{\ell \in \mathbb{N} \\ 1 \leq e^{-t} y_\ell < e}} (F \circ T^\ell)(y) \mathrm{d}t \\ &= \sum_{\ell=1}^{n} (F \circ T^\ell)(y). \end{aligned}$$

Rearranging, we get

$$\frac{1}{n} \sum_{\ell=1}^{n} (F \circ T^\ell)(y) = \frac{\log y_n}{n} \frac{1}{\log y_n} \int_0^{\log y_n} (F' \circ f_1)(a_t \Lambda) \mathrm{d}t. \tag{2.2}$$

As $Y$ is a separable, metric space and as the function $F' \circ f_1$ is continuous outside of a set of measure zero, the equidistribution assumption on the orbit $\{a_t \Lambda : t \geq 0\}$ implies that

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T (F' \circ f_1)(a_t \Lambda) \mathrm{d}t = \int_{G / \Gamma} (F' \circ f_1)(g\Gamma) \mathrm{d}g\Gamma.$$

Hence from considering the special case $F \equiv 1$ constant, one obtains that the limit

$$c = \lim_{n \to \infty} \frac{\log y_n}{n}$$

exists and is independent of $\Lambda$. Moreover, the existence of this limit and (2.2) together with the equidistribution imply the existence of the limit

$$\lambda(F) = \lim_{n \to \infty} \frac{1}{n} \sum_{\ell=1}^{n} (F \circ T^\ell)(y).$$

It is clear that this defines a positive linear functional satisfying $\lambda(1) = 1$. $\square$

**Remark 2.32.** Let us state explicitly that we have obtained

$$\lambda(F) = c \int_{G / \Gamma} F' \circ f_1 \mathrm{d}g\Gamma$$

for all continuous, bounded functions on $Y$. In what follows we identify the functional $\Lambda$ with a probability measure, which we also denote by $\lambda$.

In order to prove the Theorem, we need the following

**Lemma 2.33.** *Let*

$$Y_{\mathrm{int}} = \{y \in Y : \exists n \in \mathbb{N} \ \tfrac{y_{n+1}}{y_n} \in \mathbb{Z}\}.$$

*Then $f_1^{-1}(Y_{\mathrm{int}})$ is a nullset for the $G$-invariant probability measure on $G / \Gamma$.*

*Proof of Theorem 2.29.* Define $f_2 : Y \to \mathbb{N}^{\mathbb{N}}$ by $f_2(y) = ([y_{n+1}/y_n])_{n\in\mathbb{N}}$. Note that the diagram in (2.3) commutes.

$$\begin{array}{ccc} Y & \xrightarrow{T} & Y \\ {\scriptstyle f_2}\downarrow & & \downarrow{\scriptstyle f_2} \\ \mathbb{N}^{\mathbb{N}} & \xrightarrow[\sigma]{} & \mathbb{N}^{\mathbb{N}} \end{array} \tag{2.3}$$

Moreover, the measure $\lambda$ is clearly $T$-invariant and therefore $(f_2)_*\lambda$ is $\sigma$-invariant.

Note that $f_2$ is continuous at $y$ unless there is some $n \in \mathbb{N}$ where $y_{n+1}/y_n \in \mathbb{Z}$, i.e. the set of discontinuities is contained in

$$\{y \in Y : \exists n \in \mathbb{N}\ \tfrac{y_{n+1}}{y_n} \in \mathbb{Z}\}.$$

As of Lemma 2.33 this is a null set with respect to the measure $\lambda$ obtained in Lemma 2.31, i.e. $f_2$ is $\lambda$-a.e. continuous. In particular, if $y \in Y$ equidistributes under $T$ in $Y$ with respect to $\lambda$, then $f_2(y)$ equidistributes under $\sigma$ in $\mathbb{N}^{\mathbb{N}}$ with respect to $(f_2)_*\lambda$.

For $\alpha \in (0,1) \setminus \mathbb{Q}$, we note that by previous remarks $f_1(u_\alpha\mathbb{Z}^2)$ is precisely the sequence of denominators in the partial fractions of the cfe of $\alpha$. Letting $[0; a_1 : \cdots]$ denote the cfe of $\alpha$, the recursive relation for the denominators implies

$$\left[\frac{q_{n+1}}{q_n}\right] = \left[a_{n+1} + \frac{q_{n-1}}{q_n}\right] = a_n,$$

i.e. $(f_2 \circ f_1)(u_\alpha\Gamma) = (a_n)_{n\in\mathbb{N}}$ up to finitely many digits. This shows that the sequence given by the cfe of $\alpha$ equidistributes with respect to $(f_2)_*\lambda$. It remains to show that $(f_2)_*\lambda$ agrees with the push-forward of the Gauss measure under the cfe map.

By the Mautner phenomenon, $\{a_t\Lambda\}_{t\geq 0}$ equidistributes in $G\,/\,\Gamma$ for a.e. $\Lambda$. Note that whenever $a$ is a diagonal matrix and $\{a_t\Lambda\}_{t\geq 0}$ equidistributes, then so does $\{a_t a\Lambda\}_{t\geq 0}$. Moreover, for all $s \in \mathbb{R}$

$$a_t \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ e^{-2t}s & 1 \end{pmatrix} a_t.$$

Let $V = \pm^t U$ denote the group of lower triangular matrices with equal diagonal entries. Then the group of lower triangular matrices in $G$ is of the form $AV$, and thus by the preceding arguments we have that for all $\Lambda$ the following are equivalent:

(1) $\{a_t\Lambda\}_{t\geq 0}$ equidistributes.
(2) $\{a_t g\Lambda\}_{t\geq 0}$ for all loqer triangular $g \in G$.

On a neighborhood of the identity in $G$ the Haar measure decomposes as the product of a Haar measure on $AV$ and the Lebesgue measure on $\mathbb{R}$. In particular, for any subsete $S \subseteq \mathbb{R}$ of positive Lebesgue measure, the set

$$\{gu_\alpha\Gamma : \alpha \in S, g \text{ lower triangular}\} \subseteq {}^{G}\!/_{\Gamma}$$

has positive measure. Therefore the Mautner phenomenon and the preceding equivalence implies that for Lebesgue a.e. $\alpha \in \mathbb{R}$ the orbit $\{a_t u_\alpha\Gamma\}_{t\geq 0}$ is equidistributed. On the other hand, ergodicity of the Gauss map with respect to the Gauss measure and the equivalence of the Gauss measure to the Lebesgue measure imply that for Lebesgue almost every $\alpha \in (0,1)$ the cfe of $\alpha$ equidistributes with respect to the Gauss measure. Therefore $(f_2)_*\lambda = \nu$. □

## References

EW11. M. Einsiedler and T. Ward. Ergodic theory with a view towards number theory. Graduate Texts in Mathematics, 259. *Springer-Verlag London, Ltd., London*, 2011.

Ka13. O. Karpenkov, Geometry of Continued Fractions. Algorithms and Computation in Mathematics, 26. *Springer-Verlag Berlin Heidelberg*, 2013.

Kh64. A. Khintchine, Continued Fractions. *The University of Chicago Press*, 1964.

Ne99. J. Neukirch. Algebraic number theory. Grundlehren der Mathematischen Wissenschaften, 322. *Springer-Verlag Berlin*, 1999.

Se85. C. Series. The modular surface and continued fractions. *J. London Math. Soc.* (2) 31 (1985), no. 1, 69–80.

SW19. D. Simmons and B. Weiss. Random walks on homogeneous spaces and Diophantine approximation on fractals. *Invent. Math.* 216 (2019), no. 2, 337–394.