CrossMark

Random Walks in the Hyperbolic Plane and the Minkowski Question Mark Function

Gérard Letac[1] · **Mauro Piccioni**[2]

Abstract Consider $G = SL_2(\mathbb{Z})/\{\pm I\}$ acting on the complex upper half plane H by $h_M(z) = \frac{az+b}{cz+d}$ for $M \in G$. Let $D = \{z \in H : |z| \geq 1, |\Re(z)| \leq 1/2\}$. We consider the set $\mathcal{E} \subset G$ with the nine elements M, different from the identity, such that $\mathrm{tr}\,(MM^T) \leq 3$. We equip the tiling of H defined by $\mathbb{D} = \{h_M(D) : M \in G\}$ with a graph structure where the neighbours are defined by $h_M(D) \cap h_{M'}(D) \neq \emptyset$, equivalently $M^{-1}M' \in \mathcal{E}$. The present paper studies several Markov chains related to the above structure. We show that the simple random walk on the above graph converges a.s. to a point X of the real line with the same distribution of $S_2 W^{S_1}$, where S_1, S_2, W are independent with $\Pr(S_i = \pm 1) = 1/2$ and where W is valued in $(0, 1)$ with distribution $\Pr(W < w) = ?(w)$. Here $?$ is the Minkowski function. If $K_1, K_2, \ldots$ are i.i.d with distribution $\Pr(K_i = n) = 1/2^n$ for $n = 1, 2, \ldots$, then $W = \frac{1}{K_1 + \frac{1}{K_2 + \ldots}}$: this known result (Isola in Appl Math 5:1067–1090, 2014) is derived again here.

Keywords Random continued fractions · Minkowski question mark function · Hyperbolic plane · Modular group

Mathematics Subject Classification (2010) 60J05 · 20H05

1 Introduction

In this paper, we are concerned with random walks in the upper half of the complex plane (hyperbolic plane)

✉ Gérard Letac
  gerard.letac@math.univ-toulouse.fr

1 Laboratoire de Statistique et Probabilités, Université de Toulouse, Toulouse, France

2 Dipartimento di Matematica, Sapienza Università di Roma, 00185 Rome, Italia

🙋 Springer

$$H = \{z \in \mathbb{C} : \Im(z) > 0\}.$$

The random walk is induced by the action on H of the modular group $G = PSL(2, \mathbb{Z})$, which is the multiplicative group of all 2×2 matrices with integer entries and determinant 1, quotiented by identifying two matrices when one is equal to the other multiplied by $-I_2$. The basic properties of this action are presented in Serre [11]. For

$$M = \pm \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{1}$$

with $a, b, c, d \in \mathbb{Z}$ and $\det M = 1$ we let

$$h_M(z) = \frac{az + b}{cz + d}, \quad z \in H. \tag{2}$$

It is readily verified that h_M maps H into H since $\Im(h_M(z))$ has the same sign as $\Im(z)$, for any $M \in G$. Moreover,

$$h_{MM'}(z) = h_M(h_{M'}(z)),$$

which proves that G acts on H. Furthermore, this action is faithful, as a consequence of the quotient made above.

The discrete nature of G implies the existence of fundamental domains for the action of G on H. Roughly speaking, a fundamental domain is a subset of H which contains one element for each G-orbit. The traditional choice of a fundamental domain for the above action is the region D of all z's in H such that $|z| \geq 1$ and $|\Re(z)| \leq 1/2$. Each point in the interior of D (and the point i) belongs to a different orbit, whereas the remaining orbits intersect two different points on the boundary of D. As a consequence, the set of all images $\mathbb{D} = \{h_M(D), M \in G\}$ covers the whole space, with two distinct sets overlapping at most their boundaries. As long as $z \in D$ does not lie on the boundary of a set in $\mathbb{D}$, there is a natural "projection" of H onto $\mathbb{D}$. The image $h_M(D)$ intersects D only for $M = \pm E_i$, for some $i = 0, 1, \ldots, 8$, where

$$E_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad E_2 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \quad E_3 = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \quad E_4 = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}$$

$$h_1(z) = 1 + z \quad h_2(z) = -1 + z \quad h_3(z) = 1 - \frac{1}{z} \quad h_4(z) = -1 - \frac{1}{z}$$

$$E_5 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad E_6 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \quad E_7 = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} \quad E_8 = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$$

$$h_5(z) = \frac{z}{1 + z} \quad h_6(z) = \frac{z}{1 - z} \quad h_7(z) = -\frac{1}{1 + z} \quad h_8(z) = \frac{1}{1 - z}$$

$$E_0 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, h_0(z) = -\frac{1}{z},$$

with $h_i = h_{E_i}$, for $i = 0, 1, \ldots, 8$: see the well-known picture at page 128 of Serre [11]. These matrices (and their sign change) are characterized in G as having the trace of MM^T less or equal to three, aside from the identity.

These matrices generate the whole group G, since already $\pm E_0$ and $\pm E_1$ together have this property. More generally, $h_M(D)$ intersects $h_{M'}(D)$ if and only if $M' = \pm M E_i$ for some $i = 0, 1, \ldots, 8$. As a consequence, one can endow $\mathbb{D}$ with a graph structure, joining intersecting regions with an edge, getting in this way a regular graph with vertices in $\mathbb{D}$ of degree 9 (incidentally, this graph is isomorphic to the Cayley graph of G, taking the $\pm E_i$'s as a set of generators).

In the following, we are going to investigate the asymptotic behaviour of the H-valued processes

$$Z_n^z = h_{M_n} \circ \cdots \circ h_{M_1}(z), \ V_n^z = h_{M_1} \circ \cdots \circ h_{M_n}(z), \ z \in H, \qquad (3)$$

where $(M_k, k = 1, \ldots)$ is a sequence of i.i.d. random matrices, taking each of the 9 possible values E_i, $i = 0, 1, \ldots, 8$, with the same probability $1/9$ (from now on, we take for granted the identification of M with $-M$). For each $z \in H$ and each $n = 1, 2, \ldots$, the complex random variables Z_n^z and V_n^z have the same law, but whereas $(Z_n^z)_{n=0}^{\infty}$ is always a homogeneous Markov chain, the process $(V_n^z)_{n=0}^{\infty}$ is not. More specifically, in most of the cases it remains a Markov chain, but with transition probabilities depending on z. In fact, if z and $h_M(z)$ are known, one deduces M, except when the stabilizer of z is non-trivial, which happens only for a denumerable subset of z in H (Serre, page 129). Whenever $z \in h_M(\mathrm{int}\,D)$, for some $M \in G$, both processes can be "projected" on the vertices of the graph. Of the greatest interest is the fact that, for z in the interior of D, V_n^z "projected" on the graph induces a simple nearest-neighbour random walk on it.

In order to discuss the asymptotic behaviour of these processes, it is necessary to extend the action of G on the boundary $\partial H = \mathbb{R} \cup \{\infty\}$, the extended reals, and to study the processes in ∂H

$$X_n^x = h_{M_n} \circ \cdots \circ h_{M_1}(x), \ Y_n^x = h_{M_1} \circ \cdots \circ h_{M_n}(x), \ x \in \partial H. \qquad (4)$$

The following result can be found in the second chapter of the book by Bougerol, Lacroix [1] (see Benoist and Quint [2] for a more recent presentation). It refers to random compositions of Mobius transformations of the form (2) as they appear in (3) and (4), with general unimodular matrices with real coefficients.

Theorem 1.1 *Let $(M_k, k = 1, \ldots)$ be an i.i.d. sequence of 2×2 unimodular matrices, and let $\mathcal{G}$ be the smallest closed subgroup which contains the support of their law. Suppose that:*

1. *$\mathcal{G}$ is not compact;*
2. *There does not exist a subset L in $\mathbb{R}^2$ which is a finite union of one-dimensional subspaces which is invariant under all matrices in $\mathcal{G}$.*

Then, the following hold:

1. *For any $z \in H$, with probability 1, (V_n^z), as defined in (3), converges to a random variable $Z \in \partial H$, as $n \to \infty$;*

2. *With probability* 1, *for any bounded and continuous function* f *defined on* ∂H, $\int f(Y_n^x) \nu(dx)$ *converges a.s. to* $f(Z)$ *as* $n \to \infty$, *for any probability measure* ν *on* ∂H;

3. *The law* λ *of* Z *is the unique stationary measure for the Markov chain* (X_n^x) *on* ∂H, *and it is atomless.*

The assumptions of the theorem are clearly satisfied when the law of M is supported by the nine values E_i, $i = 0, 1, \ldots, 8$. The different kinds of convergence stated by Theorem 1.1 in the complex and in the real case is due to the fact that the product of matrices $M_1 \circ \cdots \circ M_n$, properly normalized, converges to a matrix of rank one, that has a non-trivial null space which is avoided w.p. 1 by the "initial" vector $(x, 1)^t$ since the distribution of this one-dimensional null space is atomless (Corollary 4.8 in Benoist and Quint [2]).

Thus, from the identity in law of (Z_n^z) with (V_n^z) and of (X_n^x) and (Y_n^x), and the fact that convergence a.s. implies convergence in law, one can deduce the following corollary:

Corollary 1.2 *Under the assumption of the previous theorem:*

1. *For any* $z \in H$, (Z_n^z), *as defined in* (3), *converges weakly to* λ , *as* $n \to \infty$;

2. *For any atomless probability measure* ν *on* ∂H, *the process* (X_n^x) *as defined in* (4), *with* x *taken to be* ν-*distributed, converges weakly to* λ *as* $n \to \infty$.

In the present paper, the main goal is to identify the unique stationary distribution λ for the chain (X_n^x). In order to achieve this goal, we start in Sect. 2 with the observation that the transition kernel of the chain is equivariant under the action of the four elements group Γ on $\mathbb{R} \cup \{\infty\}$, generated by the mappings $g_0(x) = h_0(x) = -1/x$ and $g_1(x) = -x$. This has the consequence that initial laws which are invariant under this group keep this property with the iterations of the Markov chain. Moreover, for any function C which is constant on the orbits of this group one obtains that $(C(X_n^x))$ is by itself a Markov chain. By choosing $C(x) = \min\{|x|, \frac{1}{|x|}\}$, we project the dynamics of (X_n^x) from the extended reals to the unit interval $[0, 1]$, and we characterize the stationary distribution of this projected Markov chain with the two properties of symmetry w.r.t. $1/2$ and invariance under a certain "tent" map of the interval. In Sect. 3, we reformulate these invariances in terms of continued fraction expansions, leading to identify the stationary distribution function for $(C(X_n^x))$ as the Minkowski's question mark function **?** (Minkowski [8]). A definition of **?** can be found in (18). The paper by Chassaing, Letac, Mora [4] can be also consulted for the links between **?** and the sequences of Farey–Brocot. This function is a remarkable example of a continuous singular distribution function on $[0, 1]$, of which we are going to review some of its properties. By "lifting" this law on the extended reals to enforce the desired invariance under Γ, the unique stationary distribution λ for (X_n^x) is finally obtained. It turns out that its survival function is a symmetrized version of the so-called Denjoy–Minkowski function of parameter $1/2$ (see Denjoy [5]). Finally, we have to mention that the interest for the Minkowski function **?** has been recently revived by the proof by Jordan and Sahlsten [7] of the 1943 Salem conjecture $\lim_n \int_0^1 e^{i2\pi nx} d\textbf{?}(x) = 0$.

2 Group Invariance Properties of the Markov Chain (X_n)

The first observation we are going to perform concerns a certain equivariance property of the Markov chain (X_n^x) defined in (4) and its consequences. We will use extensively the notation $X \sim \alpha$ when X has the distribution α and $X \sim Y$ when two random variables X and Y have the same law.

Lemma 2.1 *Let* $g_0(x) = -\frac{1}{x}$ *and* $g_1(x) = -x$ *for* $x \in \mathbb{R} \cup \{\infty\}$. *Let* M_1 *be uniformly distributed on the set of matrices* E_i, *for* $i = 0, 1, \ldots, 8$. *Then*

$$h_{M_1}(g_j(x)) \sim g_j(h_{M_1}(x)), \ j = 0, 1, \tag{5}$$

Proof Notice that

$$h_0(x) = -h_0(-x), \ h_{2i-1}(x) = -h_{2i}(-x), \ i = 1, \ldots, 4,$$

which shows (5) with $j = 1$. Similarly

$$h_0(1/x) = -\frac{1}{h_0(x)}, \ h_i(-1/x) = -\frac{1}{h_{\varphi(i)}(x)}, \ i = 1, 2, 5, 6.$$

with $\varphi(1) = 3$, $\varphi(2) = 4$, $\varphi(5) = 8$, $\varphi(6) = 7$. From this, formula (5) with $j = 0$ is obtained. □

It is clear that the statement of the previous lemma holds for any g belonging to the group Γ generated by g_0 and g_1 (shortly, also for $g(x) = 1/x$). The consequences of the previous lemma are important.

Proposition 2.2 *Let* X_0 *be a random variable on the extended reals with the property* $X_0 \sim g(X_0)$, *for any* $g \in \Gamma$. *Define* $X_1 = h_{M_1}(X_0)$, *where* M_1 *assumes the values* E_i *with probability* $1/9$, *for* $i = 0, 1, \ldots, 8$, *independently of* X_0. *Then* $X_1 \sim g(X_1)$. *Furthermore, if* f *is a bounded measurable function on* $\mathbb{R} \cup \{\infty\}$ *such that* $f(x) = f(g(x))$ *for any* $g \in \Gamma$, *then the function* $s(x) = \mathbb{E}(f(h_{M_1}(x)))$ *has again the property* $s(x) = s(g(x))$, *for any* $x \in \mathbb{R} \cup \{\infty\}$.

Proof For the first statement, it is enough to notice that for any $g \in \Gamma$

$$g(X_1) = g(h_{M_1}(X_0)) \sim h_{M_1}(g(X_0)) \sim h_{M_1}(X_0) = X_1.$$

As far as the second is concerned notice that similarly

$$f(h_{M_1}(g(x))) \sim f(g(h_{M_1}(x))) = f(h_{M_1}(x)).$$

□

The last statement in the above proposition suggests to introduce a function on the extended reals $\mathbb{R} \cup \{\infty\}$, whose values distinguish among the orbits of Γ. A convenient function with this property is the function $C : \mathbb{R} \cup \{\infty\} \to [0, 1]$ defined by

$$C(x) = \min\left\{|x|, \frac{1}{|x|}\right\}.$$

Given a random variable X with extended real values, the distribution of $C(X)$ can be immediately computed. When $X \sim -X$ and $X \sim -1/X$, this transformation can be easily inverted, "lifting" the law of $C(X)$ to $\mathbb{R} \cup \{\infty\}$, as stated in the next lemma.

Lemma 2.3 *Let X be an extended real-valued random variable. Then $X \sim g(X)$ for any $g \in \Gamma$ if and only if the conditional distribution of X given $W = C(X)$ is μ_W, where*

$$\mu_w = \frac{1}{4}\left(\delta_w + \delta_{-w} + \delta_{1/w} + \delta_{-1/w}\right), \quad w \in [0, 1]. \tag{6}$$

In this case

$$X \sim S_2 W^{S_1}, \tag{7}$$

where $S_i, i = 1, 2$ are independent random variables with $\Pr(S_i = \pm 1) = 1/2$, independent of W.

Proof First observe that if f is any bounded measurable function on the extended reals, $x \in \mathbb{R} \cup \{\infty\}$ and $w = C(x)$ one has

$$\frac{1}{4}\left(f(x) + f(-x) + f(-1/x) + f(1/x)\right)$$
$$= \frac{1}{4}\left(f(w) + f(-w) + f(-1/w) + f(1/w)\right).$$

Next observe that $X \sim g(X)$ for any $g \in \Gamma$ if and only if for any f as above and any a bounded measurable function on the unit interval $[0, 1]$, it holds

$$\mathbb{E}(f(X)a(W)) = \frac{1}{4}(\mathbb{E}(f(X)a(W) + f(-X)a(W)$$
$$+ f(-1/X)a(W) + f(1/X)a(W)))$$
$$= \frac{1}{4}(\mathbb{E}(f(W)a(W) + f(-W)a(W))$$
$$+ f(-1/W)a(W)) + f(1/W)a(W)))$$

which yields the first statement of the lemma. For the last distributional representation, it is clear that the r.h.s. of (7) satisfies the assignment of the conditional distributions (6). □

To continue, define the following mappings of the unit interval [0, 1] into itself, namely

$$H_0(x) = x, \ H_1(x) = \frac{1}{1+x}, \ H_2(x) = 1 - x,$$

$$H_3(x) = \min\left\{\frac{x}{1-x}, \frac{1-x}{x}\right\}, \ H_4(x) = \frac{x}{1+x}. \tag{8}$$

It is immediately verified that

$$C(h_i(x)) = C(H_i(x)), \ \ i = 0, 1, \ldots, 4, \tag{9}$$
$$C(H_i(x)) = C(h_{\psi(i)}(x)), \ \psi(1) = 7, \ \psi(2) = 8, \ \psi(3) = 6, \ \psi(4) = 5. \tag{10}$$

Thus, given the sequence (M_n) of independent random matrices with uniform distribution on $E_i, i = 0, 1, \ldots, 8$, define $I_n = 0$ for $M_n = E_0$ and $I_n = i$ for $M_n = E_i$ or $M_n = E_{\psi(i)}, i = 1, 2, 3, 4$. One obtains an i.i.d. sequence (I_n) with values in $\{0, 1, 2, 3, 4\}$ with distribution ρ such that $\rho(\{0\}) = 1/9$ and $\rho(\{i\}) = 2/9$, for $i = 1, 2, 3, 4$.

We are now ready to prove the following

Proposition 2.4 *Let (X_n^x) be defined in* (4). *The process $(W_n^w = C(X_n^x))$, with $w = C(x)$, is a Markov chain with values in the unit interval* [0, 1] *which evolves in the following way*

$$W_{n+1}^w = H_{I_{n+1}}(W_n^w), \ \ n = 0, 1, \ldots, W_0^w = w \in [0, 1]. \tag{11}$$

Moreover, if $x = X_0$ with $X_0 \sim g_i(X_0)$, for $i = 0, 1$, then, for any positive integer n, the conditional distribution of $X_n^{X_0}$ given $W_n^{W_0} = w$, where $W_0 = C(X_0)$, is given by μ_w, as defined in (6).

Proof Collecting together definition (8), relations (9) and (10), Proposition 2.2 and Lemma 2.3, the result is readily obtained. □

As a consequence, we have the following

Corollary 2.5 *Let ν be a stationary distribution for the process $(W_n^w)_{n=1}^\infty$ defined in* (11) *and let $W \sim \nu$. Define the random variable X as in* (7). *Then the distribution of X is stationary for the chain $(X_n^x)_{n=1}^\infty$.*

Thus, for each stationary distribution for the process $(W_n^w, n = 1, 2, \ldots)$, we can construct a corresponding stationary distribution for the original process $(X_n^x, n = 1, 2, \ldots)$ by the operation of "lifting" described above. Now we reduce the construction of a stationary distribution for the Markov chain (W_n^w) to the existence of a law which is invariant under two transformations of the unit interval, the symmetry transformation H_2 around $1/2$ and the tent-like map H_3. The basic point is that the inverse graph of the latter is the union of the graphs of H_1 and H_4.

Proposition 2.6 *Let H_i, $i = 0, 1, \ldots, 4$ be defined as in ([8]), and let W_0 be a random variable with values in $[0, 1]$ with the properties*

$$W_0 \sim 1 - W_0 = H_2(W_0),\ W_0 \sim \min\left(\frac{W_0}{1 - W_0}, \frac{1 - W_0}{W_0}\right) = H_3(W_0). \qquad (12)$$

Then $W_1 = H_I(W_0) \sim W_0$, I having the law ρ, independent of W_0.

Proof It is trivially $W_0 \sim H_0(W_0)$ and by assumption $W_0 \sim H_2(W_0)$ and $W_0 \sim H_3(W_0)$. Therefore, the result holds if one proves that the assumptions on the law of W_0 imply that $W_0 \sim H_J(W_0)$, J being a random variable assuming the values 1 and 4 with the same probability. It is not too complicate to realize that the law of W_0 cannot have an atom at $1/2$. This comes from the fact that the invariance of the law under H_3 implies that 1 has an atom with the same weight. But this is impossible since 0 and 1 are both sent to 0 by H_3, which contradicts the invariance of the distribution of W_0 by H_3.

To continue the proof of Proposition 6, we need the following observation, which for later use is collected as a lemma. □

Lemma 2.7 *If $W_0 \sim 1 - W_0$, W_0 not having an atom at $1/2$, then the law of $H_3(W_0)$ conditional to $W_0 < 1/2$ coincides with the law of $H_3(W_0)$ conditional to $\{W_0 > 1/2\}$. Thus, both coincide with the unconditional law of $H_3(W_0)$.*

Proof Being $H_3(w) = H_3(1 - w)$ for any $w \in [0, 1]$, the law of $H_3(W_0)$ conditional to $W_0 < 1/2$ is equal to the law of $H_3(1 - W_0)$ conditional to $W_0 < 1/2$. Next, replace W_0 with $1 - W_0$, these being equal in law, to obtain that the law of $H_3(W_0)$ conditional to $1 - W_0 < 1/2$ is still the same. The last conditioning being the same as $W_0 > 1/2$, the proof of the first statement of the lemma is finished. The second is obtained from the law of total probabilities. □

Proof of Proposition [2.6], continued Next observe that $H_4(H_3(w)) = w$ for $w < 1/2$ and $H_1(H_3(w)) = w$ for $w > 1/2$. As a consequence for $w > 1/2$, from $H_3(W_0) \sim W_0$ one gets that

$$\Pr(H_J(W_0) > w) = \frac{1}{2}\Pr\left(H_1(W_0) > w\right) = \frac{1}{2}\Pr\left(H_1(H_3(W_0)) > w\right)$$
$$= \Pr\left(W_0 > \frac{1}{2}, H_1(H_3(W_0)) > w\right) = \Pr(W_0 > w),$$

and for $w < 1/2$

$$\Pr(H_J(W_0) < w) = \frac{1}{2}\Pr(H_4(W_0) < w) = \frac{1}{2}\Pr(H_4(H_3(W_0)) < w)$$
$$= \Pr\left(W_0 < \frac{1}{2}, H_4(H_3(W_0)) < w\right) = \Pr(W_0 < w).$$

These two together easily imply that $H_J(W_0) \sim W_0$. □

3 Minkowski's Question Mark Function and Denjoy–Minkowski Distribution on the Real Line

The goal of this section is to deduce from the invariance properties assumed in (12) a unique law for W_0, whose distribution function is the question mark function introduced by Minkowski. This characterization is well known (see Isola [6], Lemma 4.1), but here we give a probabilistic proof of it. For this purpose, the continued fraction representation of irrational numbers in the unit interval [0, 1] is required (see Olds [9]). On this interval, we define the function $A_k(w) = \frac{1}{k+w}$ for $k \in \mathbb{N}^+$. Likewise define for $k_1, \ldots, k_n, \ldots$ in $\mathbb{N}^+$

$$A_{k_1,\ldots,k_n}(w) = A_{k_1} \circ \ldots \circ A_{k_n}(w) = \cfrac{1}{k_1 + \cfrac{1}{k_2 + \cfrac{1}{\ddots + \cfrac{1}{k_n + w}}}}, \quad n \in \mathbb{N}^+. \tag{13}$$

Then

$$x = \lim_n A_{k_1,\ldots,k_n}(w) \overset{\text{def}}{=} [k_1, \ldots, k_n, \ldots] \tag{14}$$

always exists and does not depend on $w \in [0, 1]$. Such an x is necessarily an irrational number. Conversely for any irrational number $x \in (0, 1)$, there exists a unique sequence $(k_n)_{n \in \mathbb{N}^+}$ such that (14) holds. This is called the continued fraction expansion of x: its definition implies the recursion

$$x = [k_1, k_2, \ldots] = \frac{1}{k_1 + [k_2, k_3, \ldots]}. \tag{15}$$

The above construction allows to associate with any probability distribution p on the positive integers an atomless law $\mu(p)$ on the interval [0, 1] in the following way. Let $(K_n, n = 1, 2, \ldots)$ be a sequence of i.i.d. p-distributed random variables: then $W = [K_1, K_2, \ldots]$ has the law $\mu(p)$. The function $p \mapsto \mu(p)$ is clearly injective, since K_1 is the integer part of W^{-1}. The distribution $\mu(p)$ can be characterized as the unique stationary distribution for the Markov chain $(U_n^u, n = 1, 2, \ldots)$, where

$$U_{n+1}^u = A_{K_{n+1}}(U_n^u) = \frac{1}{K_{n+1} + U_n^u}, n = 0, 1, \ldots, U_0^u = u \in [0, 1].$$

This is an instance of a general principle (see Chamayou and Letac [3], Proposition 1). An equivalent way of stating this property is the following: for W and K independent random variables, with values in [0, 1] and $\mathbb{N}^+$, respectively, it holds

$$K \sim p, W \sim \frac{1}{K + W} \implies W \sim \mu(p). \tag{16}$$

Now we are in a position to prove the following result.

Theorem 3.1 *Let W have an atomless law on the interval $[0, 1]$: Then the following are equivalent:*

1. $W \sim 1 - W$ *and* $W \sim \min\left\{\frac{W}{1-W}, \frac{1-W}{W}\right\}$.
2. $W \sim \mu(p)$, *with*

$$p(n) = 2^{-n}, n = 1, 2, \ldots. \tag{17}$$

3. *The distribution function of W at irrational points is given by*

$$P(W < [k_1, k_2, \ldots]) = 2 \sum_{n=1}^{\infty} (-1)^{n+1} 2^{-\sum_{j=1}^{n} k_j} \overset{\text{def}}{=} \mathbf{?}([k_1, k_2, \ldots]) \tag{18}$$

for $k_j = 1, 2, \ldots, j = 1, 2, \ldots$.

The function **?** defined in (18) on the irrational numbers is called the Minkowski's question mark function. Being continuous, it can be uniquely extended to the whole unit interval. In fact, it is known since the work of Salem [10] that

$$|\mathbf{?}(x) - \mathbf{?}(x')| \leq C|x - x'|^{\alpha} \text{ with } \alpha = \frac{\log 2}{2 \log \theta}$$

where $\theta = \frac{1+\sqrt{5}}{2}$ is the golden ratio and C is a constant. The function **?** is strictly increasing but it is singular w.r.t. the Lebesgue measure, since its derivative is zero a.s. (Salem [10], Viader, Bibiloni and Paradis [12]). A direct proof of the characterization (2) of the distribution function **?** stated in the previous theorem can be found in Isola [6], Lemma 4.6.

Proof (1) implies (2). Since W has an atomless law, we can assume that it takes values in the set of irrationals, and write $W = [K_1, K_2, \ldots]$, with the law of the process $(K_n, n = 1, 2, \ldots)$ to be determined. Next observe that for any $w = [k_1, k_2, \ldots] \in [0, 1]\backslash\mathbb{Q}$ one has for $k_1 > 1$ (i.e. $w < 1/2$) and $k_1 = 1$ (i.e. $w > 1/2$), respectively

$$\frac{w}{1-w} = [k_1 - 1, k_2, \ldots], \quad \frac{1-w}{w} = [k_2, k_3, \ldots]. \tag{19}$$

We now show by induction the following facts

- $(A)_n$ $\Pr(K_1 = k) = \frac{1}{2^k}, k = 1, \ldots, n$,
- $(B)_n$ $[K_2, K_3 \ldots] | \{K_1 = n\} \sim W$,
- $(C)_n$ $[K_1 - n, K_2, \ldots] | \{K_1 > n\} \sim W$.

For $n = 1$, $(A)_1$ is a consequence of the symmetry of the law of W around $1/2$, whereas $(B)_1$ and $(C)_1$ are obtained from Lemma 2.7: indeed there we established that both the law of $H_3(W) = \frac{W}{1-W}$ conditional to $W < \frac{1}{2}$ and the law of $H_3(W) = \frac{1-W}{W}$ conditional to $W > \frac{1}{2}$, are equal to the unconditional law of $H_3(W)$, which in turn is equal to the law of W.

Now assume that $(A)_n, (B)_n, (C)_n$ are true and proceed by induction. Since

$$\Pr(K_1 = n + 1) = \Pr(K_1 > n)\Pr(K_1 = n + 1 | K_1 > n)$$

and the first factor by the induction assumption $(A)_n$ is equal to $1/2^n$, we have to prove that $\Pr(K_1 = n + 1 | K_1 > n) = \Pr(K_1 - n = 1 | K_1 - n > 0) = \frac{1}{2}$. This is a consequence of $(C)_n$ and $(A)_1$. Hence $(A)_{n+1}$ is proved.

To prove $(B)_{n+1}$, we condition the l.h.s. of $(C)_n$ by $\{K_1 = n + 1\}$. We get

$$[K_1 - n, K_2, \ldots]|\{K_1 = n + 1\} \sim W|\{K_1 = 1\} = [1, K_2, K_3, \ldots]$$

so that

$$[K_2, K_3, \ldots]|\{K_1 = n + 1\} \sim [K_2, K_3, \ldots]|\{K_1 = 1\} \sim W$$

from $(B)_1$. Hence $(B)_{n+1}$ is proved.

Finally in order to prove $(C)_{n+1}$, we condition the l.h.s. of $(C)_n$ by $\{K_1 > n + 1\}$. We get

$$[K_1 - n, K_2, \ldots]|\{K_1 > n + 1\} \sim W|\{K_1 > 1\} = [K_1, K_2, \ldots]|\{K_1 > 1\}$$

and this in turn implies that

$$[K_1 - n - 1, K_2, \ldots]|\{K_1 > n + 1\} \sim [K_1 - 1, K_2, \ldots]|\{K_1 > 1\} \sim W.$$

from $(C)_1$. Hence $(C)_{n+1}$ is proved. Finally notice that $(A)_n$ and $(B)_n$, for any positive integer n, are equivalent to the l.h.s. of (16), with p given in (17). This establishes (2).

(2) implies (3). Consider the representation $W \sim [K_1, K_2, \ldots]$, where $(K_n, n = 1, 2, \ldots)$ is an i.i.d. sequence of random variables with the same distribution (17). The survival function of K_1 being equal to $P(K_1 \geq k_1) = 2 \times 2^{-k_1}$, the events

$$E_n = \{K_1 = k_1, \ldots, K_{n-1} = k_{n-1}, K_n \geq k_n\}, \quad n = 1, 2 \ldots$$

have probabilities

$$\Pr(E_n) = 2 \times 2^{-\sum_{j=1}^{n} k_j}. \tag{20}$$

Next another sequence (F_n) is constructed by means of the following recursion, starting from $F_1 = E_1$,

$$F_{2n} = F_{2n-1} \backslash E_{2n}, \quad F_{2n+1} = F_{2n} \cup E_{2n+1}, n = 1, 2, \ldots.$$

The fundamental property is that, for any positive integer n

$$F_{2n} \subset \{W < [k_1, k_2, \ldots]\} \subset F_{2n-1},$$

since the functions $A_{k_1, \ldots, k_n}$ are decreasing for n odd and increasing for n even, and the range of $A_{k_1, \ldots, k_{n-1}, k_n+1}$ is an interval adjacent to the right (left) to the range of $A_{k_1, \ldots, k_{n-1}, k_n}$ if n is even (odd). By the properties of continued fraction expansions, both

the sequences (F_{2n-1}) and (F_{2n}) converge (from above and from below, respectively) to the event $\{W < [k_1, k_2, \ldots]\}$. Since for any positive integer n

$$\Pr(F_{2n}) = \Pr(F_{2n-1}) - \Pr(E_{2n}), \Pr(F_{2n+1}) = \Pr(F_{2n}) + \Pr(E_{2n+1})$$

it suffices to substitute the expressions (20) to get the desired (18).

(3) implies (1). It consists in a simple verification. Since for $k_1 > 1$, we have

$$1 - [k_1, k_2, \ldots] = [1, k_1 - 1, k_2, \ldots],$$

in order to prove that **?** corresponds to a probability measure which is symmetric around $1/2$, it is enough to verify

$$?([k_1, k_2, \ldots]) + ?([1, k_1 - 1, k_2, \ldots]) = 1, k_1 > 1,$$

which is straightforward. The second invariance property is deduced from (19) and from the fact that $K - 1|\{K > 1\} \sim K$ when K has the distribution (17). □

Next, by Proposition 2.6 and Theorem 3.1 we have the following

Corollary 3.2 *The function **?** is a stationary distribution function for the Markov chain (W_n^w) defined in (11).*

*By Lemma 2.3, for completing our program we need to compute the law of $S_2 W^{S_1}$, where W, S_1, S_2 are independent, W has the distribution function **?** on $[0, 1]$ and S_1 and S_2 are two random variables which assume the values -1 and $+1$ with the same probability $1/2$. This law is stationary for the process (X_n^x) described in (4) because of Proposition 2.2 and Corollary 3.2, and it is unique by Theorem 1.1.*

As a first step, we prove that the distribution of W^{S_1} is the so-called Denjoy–Minkowski function $\chi_{1/2}$ of order $1/2$ (Chassaing et al. [4] page 41). In order to define it, we write positive irrational numbers y in the form

$$y = [k_0; k_1, k_2, \ldots] = k_0 + \cfrac{1}{k_1 + \cfrac{1}{k_2 + \ldots}},$$

where $k_0 = [y]$ and $y - [y] = [k_1, k_2, \ldots] \in (0, 1)$. Now define

$$\chi_{1/2}(y) = \chi_{1/2}([k_0; k_1, k_2, \ldots]) = \sum_{n=0}^{\infty} (-1)^n 2^{-\sum_{j=0}^{n} k_j}. \tag{21}$$

*As for the function **?**, it is observed that $\chi_{1/2}$ is a continuous function, and thus, it extends uniquely to the whole non-negative real line.*

Proposition 3.3 *Let W have the distribution function **?**. Let $Y = W^{S_1}$. Then the survival function $\Pr(Y > y)$ of Y is the function $\chi_{1/2}(y)$.*

Proof By comparing (18) with (21), it is immediately verified that for y irrational

$$\Pr(Y > y) = 1 - \Pr(S_1 = 1, X < y) = 1 - \frac{?(y)}{2}, 0 < y < 1,$$

and

$$\Pr(Y > y) = \Pr\left(S_1 = -1, \frac{1}{X} > y\right) = \frac{1}{2}?\left(\frac{1}{y}\right), y > 1.$$

Now it remains to verify that the r.h.s. of the above expressions coincide with $\chi_{1/2}(y)$, for all irrationals y. For the former, this is immediately verified. For the latter, we conclude with the observation that, for $k_0 \geq 1$ it is

$$\frac{1}{[k_0; k_1, k_2, \ldots]} = [0; k_0, k_1, \ldots]. \tag{22}$$

$\square$

Here is a noteworthy property of $\chi_{1/2}$.

Proposition 3.4 *Let Y be a positive random variable with the survival function $\chi_{1/2}$. Then, $[Y] + 1$ has the geometric distribution (17), and it is independent of $Y - [Y]$, which has the distribution function* ?. *In other words, if $K_0, K_1, \ldots$ are i.i.d. with distribution (17) then*

$$Y \sim [K_0 - 1; K_1, K_2, \ldots]. \tag{23}$$

Proof If W have distribution function ?, we know that one can construct $W = [K_1, K_2, \ldots]$, where (K_n) is an i.i.d. sequence of random variables with the distribution (17). Moreover, let S_1 independent of W such that $\Pr(S_1 = \pm 1) = 1/2$. From Proposition 3.3 we write $Y = W^{S_1}$. Thus, the law of Y is a mixture, with equal weights, of the law of $[0; K_1, K_2, \ldots]$ and, from (22), of that of $[K_1; K_2, \ldots]$. From this one obtains (23). $\square$

The last step that ends the determination of the unique stationary distribution λ of the Markov chain (X_n^x) defined in (4) is a simple symmetrization of the Denjoy–Minkowski function.

Proposition 3.5 *Let $X \sim \lambda$, the unique stationary distribution of the chain (X_n^x) defined in (4). Then, for any $x \geq 0$*

$$\Pr(X > x) = \Pr(X < -x) = \frac{1}{2}\chi_{1/2}(x)$$

Proof It is immediately obtained from the representation $X = S_2 Y$, where Y has the survival function $\chi_{1/2}$ and S_2 is an independent random variable such that $\Pr(S_2 = \pm 1) = 1/2$. $\square$

References

1. Bougerol, P., Lacroix, J.: Products of Random Matrices with Applications to Schrodinger Operators. Birkhauser, Boston (1985)
2. Benoist, Y., Quint, J.F.: Random Walks on Reductive Groups. Springer, Berlin (2016)
3. Chamayou, J.-F., Letac, G.: Explicit stationary distributions for composition of random functions and products of random matrices. J. Theor. Probab. **4**, 3–36 (1991)
4. Chassaing, P., Letac, G., Mora, M.: 'Brocot sequences and random walks in $SL(2, \mathbb{R})$. Springer Lect. Notes Prob. Groups IX **1084**, 37–50 (1984)
5. Denjoy, A.: Sur une fonction réelle de Minkowski. J. Math. Pures Appl. Sér. 17 **IX**, 105–151 (1938)
6. Isola, S.: Continued fractions and dynamics. Appl. Math. **5**, 1067–1090 (2014)
7. Jordan, T., Sahlsten, T.: Fourier transforms of Gibbs measures for the Gauss map. Math. Ann. **364**, 983–1023 (2016)
8. Minkowski, H.: Zur geometrie der Zahlen. Verhandlungen des III internationalen Mathematiker-Kongress in Heidelberg, Berlin (1904)
9. Olds, C.D.: Continued Fractions. The Mathematical Association of America, Washington, DC (1963)
10. Salem, R.: On some singular monotonic functions which are strictly increasing. Trans. Am. Math. Soc. **53**, 427–439 (1943)
11. Serre, J.-P.: Cours d'arithmétique', deuxième édition revue et corrigée. Presses Univ. de France, Paris (1977)
12. Viader, P., Paradis, J., Bibiloni, L.: A new light on Minkowski's ?(x) function. J. Number Theory **73**, 212–227 (1998)