Pure and Applied Mathematics Quarterly
Volume 17, Number 3, 933–969, 2021

Products of random matrices: a dynamical point of view

Tien-Cuong Dinh, Lucas Kaufmann, and Hao Wu

This paper is dedicated to Professor Duong Hong Phong on the occasion of his 65^{th} birthday

Abstract: We study products of random matrices in $\mathrm{SL}_2(\mathbb{C})$ from the point of view of holomorphic dynamics. For non-elementary measures with finite first moment we obtain the exponential convergence towards the stationary measure in Sobolev norm. As a consequence we obtain the exponentially fast equidistribution of forward images of points towards the stationary measure. We also give a new proof of the Central Limit Theorem for the norm cocycle under a second moment condition, originally due to Benoist-Quint, and obtain some general regularity results for stationary measures.

1. Introduction and main results

Let G be the group $\mathrm{SL}_2(\mathbb{C})$ of complex 2×2 matrices with determinant one and let μ be a probability measure on G. It is a classical problem to study products of the form $g_n \cdots g_1$ where the g_i are independent and identically distributed (i.i.d.) random matrices with law μ. This is a very rich theory with many beautiful results. A standard reference is the book [6]. For a more recent account that deals with more general Lie group actions, the reader may also consult [4]. The goal of this paper is to revisit this problem using the point of view of holomorphic dynamics. This is inspired by our recent work [9]. We hope that our methods can be applied in higher dimensions and can give a simplified treatment of known results.

The group G acts naturally on the complex projective line $\mathbb{P}^1$. In the standard affine coordinate of $\mathbb{P}^1 = \mathbb{C} \cup \{\infty\}$ a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ acts via the Möbius transformation $z \mapsto \frac{az+b}{cz+d}$. This allows us to identify the group $\mathrm{Aut}(\mathbb{P}^1)$ of holomorphic automorphisms of $\mathbb{P}^1$ with the group $\mathrm{PSL}_2(\mathbb{C}) = \mathrm{SL}_2(\mathbb{C})/\{\pm \mathrm{Id}\}$. In what follows we will also denote this group by G and we keep denoting by μ the probability measure induced on $\mathrm{PSL}_2(\mathbb{C})$ by the measure μ on $\mathrm{SL}_2(\mathbb{C})$. This shouldn't cause any confusion.

Received May 23, 2019.

The probability measure μ defines a positive closed $(1,1)$-current on $\mathbb{P}^1 \times \mathbb{P}^1$ given by

$$[\Gamma_\mu] := \int_G [\Gamma_g]\, \mathrm{d}\mu(g),$$

where $[\Gamma_g]$ is the current of integration along the graph Γ_g of $g \in \mathrm{Aut}(\mathbb{P}^1)$. The reader may consult [8] and [12] for background material on currents on complex manifolds.

The current $[\Gamma_\mu]$ can be seen as the graph of a generalized correspondence, which we will denote by f_μ. When the support of μ is finite f_μ is a correspondence in the usual sense, that is, $[\Gamma_\mu]$ is an effective one-dimensional cycle on $\mathbb{P}^1 \times \mathbb{P}^1$. In this case f_μ can be seen as a multivalued holomorphic map.

This generalized correspondence acts on a current T on $\mathbb{P}^1$ (e.g. a continuous function, a positive measure or a differential form) by the formula

$$f_\mu^*(T) := \int_G g^* T\, \mathrm{d}\mu(g),$$

or equivalently $f_\mu^*(T) = (\pi_1)_*(\pi_2^*(T) \wedge [\Gamma_\mu])$, where $\pi_1, \pi_2 : \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^1$ are the canonical projections. We can also define $(f_\mu)_*$ by interchanging the roles of π_1 and π_2 or, equivalently, by replacing g^* by g_* in the above formula.

For a continuous function φ on $\mathbb{P}^1$ we get

$$f_\mu^*(\varphi)(x) = \int_G \varphi(g \cdot x)\, \mathrm{d}\mu(g), \quad x \in \mathbb{P}^1, \tag{1}$$

which is the standard Markov operator associated with μ. Dually, if m is a probability measure on $\mathbb{P}^1$ we have $(f_\mu)_* m = \mu * m$, the convolution of μ and m (see [6] for more details). A probability measure m on $\mathbb{P}^1$ is called **stationary** with respect to μ if $\mu * m = m$, or equivalently if m is $(f_\mu)_*$-invariant.

For $n \geq 1$ we define f_μ^n to be the correspondence associated with the convolution measure $\mu^{*n} = \mu * \cdots * \mu$ (n times) which is the pushforward of the product measure $\mu^{\otimes n}$ on G^n by the map $(g_1, \ldots, g_n) \mapsto g_n \cdots g_1$. When μ is finitely supported we recover the usual notion of iteration of a correspondence.

We say that μ is **non-elementary** if its support does not preserve a finite subset of $\mathbb{P}^1$ and if the semi-group generated by $\mathrm{supp}(\mu)$ is not relatively compact. It is a result of Furstenberg that a non-elementary measure admits a unique stationary measure (see [6, II.4.1] and Remark 2.12).

Our first main result is the following. See Theorem 2.11 and also Remark 2.12 for the precise statement. See also Definition 5.1 for more on moment conditions on μ.

Theorem 1.1. *Let μ be a non-elementary probability measure on $G = \mathrm{PSL}_2(\mathbb{C})$ and let ν be its unique stationary measure. Assume that μ has a finite first moment, i.e. $\int_G \log \|g\| \mathrm{d}\mu(g) < +\infty$. Then the iterates of the Markov operator associated with μ converge exponentially fast to ν with respect to the Sobolev norm $\|\cdot\|_{W^{1,2}}$ on test functions.*

When the measure μ has a finite exponential moment, that is, when $\int_G \|g\|^\alpha \mathrm{d}\mu(g) < +\infty$ for some $\alpha > 0$, the exponential convergence of the Markov operator towards the stationary measure is a fundamental result of Le Page [17]. The convergence in this case is for test functions in some Hölder space and it has many important consequences such as the Central Limit Theorem mentioned below, the Large Deviation Theorem and other analogues of classical limit theorem for i.i.d. random variables.

The results that follow will be consequences of Theorem 1.1. The first one says that the forward images of any given point $a \in \mathbb{P}^1$ by the generalized correspondence f_μ converge to ν exponentially fast and uniformly in a.

Theorem 1.2. *Let μ be a non-elementary probability measure on $G = \mathrm{PSL}_2(\mathbb{C})$. Assume that $\int_G (\log \|g\|)^{1+\epsilon}\, \mathrm{d}\mu(g) < +\infty$ for some $\epsilon > 0$. Then, there is a constant $0 < \gamma < 1$ such that for any $a \in \mathbb{P}^1$ and every test function φ of class $\mathcal{C}^\beta$ on $\mathbb{P}^1$, with $0 < \beta \leq 1$, we have*

$$(2) \qquad \big|\langle (f_\mu^n)_* \delta_a - \nu, \varphi\rangle\big| \leq A_\beta \|\varphi\|_{\mathcal{C}^\beta} \gamma^{\beta n} \quad \textit{for every} \quad n \geq 0,$$

where $A_\beta > 0$ is a constant independent of n, a and φ.

Next, we give a new proof of the following known Central Limit Theorem for the random variables $\log \frac{\|g_n \cdots g_1 \cdot v\|}{\|v\|}$ where v is any non-zero vector in $\mathbb{C}^2$. The Lyapunov exponent is defined in Section 4.

Theorem 1.3 (Central Limit Theorem)**.** *Let μ be a probability measure on $G = \mathrm{PSL}_2(\mathbb{C})$. Assume that μ is non-elementary and has a finite second moment, i.e. $\int_G (\log \|g\|)^2\, \mathrm{d}\mu(g) < +\infty$. Let γ be the Lyapunov exponent of μ. Then there exists a number $\sigma > 0$ such that for any $v \in \mathbb{C}^2 \setminus \{0\}$.*

$$(3) \qquad \frac{1}{\sqrt{n}}\Big(\log \frac{\|g_n \cdots g_1 \cdot v\|}{\|v\|} - n\gamma\Big) \longrightarrow \mathcal{N}(0;\sigma^2) \quad \textit{in law,}$$

where $N(0;\sigma^2)$ is the centered normal distribution with variance σ^2.

Under an exponential moment condition, the above result is mainly due to Le Page [17] and was later refined by other authors (see for instance [15, 13]).

The question of whether this condition could be relaxed to an (optimal) second moment condition remained open until very recently, when Benoist-Quint gave an affirmative answer, [3]. Their most general result holds for probability measures on a semisimple connected linear real Lie group G whose support generates a Zariski dense semigroup. When $G = \mathrm{SL}_2(\mathbb{C})$ these probability measures are precisely the non-elementary ones and we recover Theorem 1.3 above. Our proof is independent of theirs and, in particular, does not rely on an a priori knowledge of the regularity of ν (although we also obtain such results later in Section 5). We expect that our method can be generalized to cover the general case of Benoist-Quint.

Our final result concerns the regularity of stationary measures. If μ is a non-elementary probability measure and ν is the associated stationary measure, the regularity of ν will depend on moment conditions on μ, see Definition 5.1. The statement of our main result (Theorem 5.6 below) and its proof rely on the theory of superpotentials introduced by Sibony and the first author [11]. We state here some more concrete consequences (see Corollaries 5.7 and 5.9) and refer to Section 5 for the general statements. Here $\mathbb{D}(a, r)$ denotes the disc of radius r and center a with respect to the standard metric on $\mathbb{P}^1$.

Theorem 1.4. *Let μ be a non-elementary probability measure on $G = \mathrm{PSL}_2(\mathbb{C})$.*

1. *If μ has a finite exponential moment, i.e. $\int_G \|g\|^p \, \mathrm{d}\mu(g) < +\infty$ for some $p > 0$, then there are constants $c, \alpha > 0$ such that $\nu(\mathbb{D}(a, r)) \leq cr^\alpha$ for every $a \in \mathbb{P}^1$ and $0 < r \leq 1$.*
2. *If μ has a finite first moment, i.e. $\int_G \log \|g\| \, \mathrm{d}\mu(g) < +\infty$ then there are constants $c, \alpha > 0$ such that $\nu(\mathbb{D}(a, r)) \leq c|\log r|^{-\alpha}$ for every $a \in \mathbb{P}^1$ and $0 < r \leq 1$.*

It is worth mentioning that many regularity results can be found in the literature. In particular, the first part of the above theorem is due to Guivarc'h [16], while the second part is related to a result of Benoist-Quint [3], who obtained a similar regularity estimate assuming a finite p^{th}-moment for some $p > 1$ (see also Remarks 5.8 and 5.10). Here they are obtained by completely different methods.

2. Action on Sobolev space and convergence to the stationary measure

This section is devoted to the proof of Theorem 1.1. We show that when μ has a finite first moment, i.e. when $\int_G \log \|g\| \, \mathrm{d}\mu(g) < +\infty$, the operator f_μ^*

acts continuously on the Sobolev space $W^{1,2}$. Later on, we prove that when μ is non-elementary this action has a spectral gap. As a consequence, we get an exponentially fast convergence of the Markov operator towards the stationary measure.

Consider the space

$$L^2_{(1,0)} := \{\phi : \phi \text{ is a } (1,0)\text{-form on } \mathbb{P}^1 \text{ with } L^2 \text{ coefficients}\}$$

equipped with the norm

$$\|\phi\|_{L^2} := \Big(\int_{\mathbb{P}^1} i\phi \wedge \overline{\phi}\Big)^{1/2}. \tag{4}$$

The space $L^2_{(0,1)}$ and the corresponding norm are defined analogously.

Proposition 2.1. *Let μ be a probability measure on $G = \mathrm{PSL}_2(\mathbb{C})$ and let f_μ be the associated generalized correspondence. Then the operator f_μ^*, which is well-defined on smooth $(1,0)$-forms, extends to a bounded linear operator $f_\mu^* : L^2_{(1,0)} \to L^2_{(1,0)}$ with norm bounded by 1. In other words, for ϕ in $L^2_{(1,0)}$ we have the inequality $\|f_\mu^*\phi\|_{L^2} \leq \|\phi\|_{L^2}$. Moreover, the equality holds if and only if $g_1^*\phi = g_2^*\phi$ for $\mu \otimes \mu$ almost every (g_1, g_2).*

Proof. The result follows from standard arguments from unitary representation theory, see Remark 2.10. We give the proof in our present setting since it will be useful in what follows.

By a direct computation we have

$$\begin{aligned} f_\mu^*(i\phi \wedge \overline{\phi}) - i f_\mu^*\phi \wedge \overline{f_\mu^*\phi} &= \\ &= \int_G g^*(i\phi \wedge \overline{\phi})\, \mathrm{d}\mu(g) - i\Big(\int_G g^*\phi\, \mathrm{d}\mu(g)\Big) \wedge \overline{\Big(\int_G g^*\phi\, \mathrm{d}\mu(g)\Big)} \\ &= \frac{1}{2}\int_G\int_G i(g_1^*\phi - g_2^*\phi) \wedge \overline{(g_1^*\phi - g_2^*\phi)}\, \mathrm{d}\mu(g_1)\mathrm{d}\mu(g_2). \end{aligned} \tag{5}$$

These identities are clear for smooth ϕ. We obtain the general case by the density of smooth forms in $L^2_{(1,0)}$.

Notice that the last integral in (5) defines a positive measure on $\mathbb{P}^1$. By integrating the left hand side of (5) over $\mathbb{P}^1$ and using the fact that the action of f_μ^* on measures preserves the total mass we get $\|\phi\|^2_{L^2} - \|f_\mu^*\phi\|^2_{L^2} \geq 0$. This is the desired inequality. From (5) it is also clear that $\|f^*\phi\|_{L^2} = \|\phi\|_{L^2}$ if and only if $g_1^*\phi = g_2^*\phi$ holds for all (g_1, g_2) outside a set of zero measure for $\mu \otimes \mu$. □

Consider now the Sobolev space $W^{1,2}$ of real valued L^1 functions on $\mathbb{P}^1$ with finite $\|\cdot\|_{W^{1,2}}$ norm, where

$$\|h\|_{W^{1,2}} := \Big|\int_{\mathbb{P}^1} h\,\omega_{\mathrm{FS}}\Big| + \|\partial h\|_{L^2}$$

and ω_{FS} stands for the Fubini-Study form on $\mathbb{P}^1$.

The following proposition was proved in [9].

Proposition 2.2. *Let U be a non-empty open subset of $\mathbb{P}^1$. Then the following norms on $W^{1,2}$ are equivalent to the norm $\|\cdot\|_{W^{1,2}}$.*

1. $\|h\|_1 := \|h\|_{L^1} + \|\partial h\|_{L^2}$
2. $\|h\|_2 := \|h\|_{L^2} + \|\partial h\|_{L^2}$
3. $\|h\|_3 := |\int_U h\,\omega_{FS}| + \|\partial h\|_{L^2}$
4. $\|h\|_4 := \int_U |h|\omega_{FS} + \|\partial h\|_{L^2}$.

Here and in what follows $\|g\| := \sup_{v\in\mathbb{C}^2\setminus\{0\}} \frac{\|g\cdot v\|}{\|v\|}$ will denote the operator norm of the matrix g. Notice that, for $g \in \mathrm{SL}_2(\mathbb{C})$, we have $\|g\| \geq 1$ and $\|g^{-1}\| = \|g\|$. This follows from Cartan's decomposition (see the proof of Lemma 2.4 below).

Proposition 2.3. *Let f_μ be the generalized correspondence associated with μ on $G = \mathrm{PSL}_2(\mathbb{C})$. Assume that $\int_G \log\|g\|\,\mathrm{d}\mu(g) < +\infty$. Then the Markov operator f_μ^* acting on smooth functions extends to a bounded linear operator from $W^{1,2}$ to itself.*

For the proof we need some preliminary results.

Lemma 2.4. *We have $g^*\omega_{FS} \leq \|g\|^4 \cdot \omega_{FS}$ and $g_*\omega_{FS} \leq \|g\|^4 \cdot \omega_{FS}$ for every $g \in \mathrm{PSL}_2(\mathbb{C})$.*

Proof. Using Cartan's decomposition we can write any element g in $\mathrm{PSL}_2(\mathbb{C})$ as $g = kak'$, where $k, k' \in \mathrm{SU}(2)$ and $a = \left(\begin{smallmatrix}\lambda & 0\\ 0 & \lambda^{-1}\end{smallmatrix}\right)$, for some $\lambda \geq 1$. We see that $\|g\| = \lambda$. Since $\mathrm{SU}(2)$ preserves ω_{FS} and $\|g\| = \|a\|$ we can assume that g is of the form $\left(\begin{smallmatrix}\lambda & 0\\ 0 & \lambda^{-1}\end{smallmatrix}\right)$, for some $\lambda \geq 1$.

In the standard affine coordinate of $\mathbb{P}^1$ we have $\omega_{\mathrm{FS}} = \frac{i}{2\pi}\frac{\mathrm{d}z\wedge\mathrm{d}\bar{z}}{(1+|z|^2)^2}$ and $g(z) = \lambda^2 z$, so

$$g^*\omega_{\mathrm{FS}} = \lambda^4\frac{i}{2\pi}\frac{\mathrm{d}z\wedge\mathrm{d}\bar{z}}{(1+\lambda^4|z|^2)^2} \leq \lambda^4\frac{i}{2\pi}\frac{\mathrm{d}z\wedge\mathrm{d}\bar{z}}{(1+|z|^2)^2} = \lambda^4\omega_{\mathrm{FS}} = \|g\|^4\omega_{\mathrm{FS}},$$

which proves the first inequality.

For the second inequality we apply the above argument for g^{-1} instead of g and use that $\|g^{-1}\| = \|g\|$. □

The following exponential estimate will be crucial for us. It will also be used in the proof of Theorem 1.2 in Section 3 and will be important in Section 5.

Proposition 2.5 (Moser-Trudinger estimate [20]). *Let $\mathcal{F}$ be a bounded family in $W^{1,2}$. Then there are constants $A > 0$ and $\alpha > 0$, depending on $\mathcal{F}$, such that*

$$\int_{\mathbb{P}^1} e^{\alpha\varphi^2}\omega_{FS} \leq A \quad \textit{for every} \quad \varphi \in \mathcal{F}.$$

Proof of Proposition 2.3. We need to show that $\|f_\mu^*\varphi\|_{W^{1,2}}$ is uniformly bounded in φ if $\|\varphi\|_{W^{1,2}} \leq 1$. For such φ we have, from Proposition 2.1, that $\|\partial f_\mu^*(\varphi)\|_{L^2} \leq 1$, so using Proposition 2.2 it remains to check that $\|f_\mu^*\varphi\|_{L^2}$ is uniformly bounded.

Let α and A be as in Proposition 2.5 for $\mathcal{F} = \{\varphi : \|\varphi\|_{W^{1,2}} \leq 1\}$. From Jensen's inequality and Lemma 2.4 we have

$$\begin{aligned}\exp\left(\int_{\mathbb{P}^1} \alpha(g^*\varphi)^2\omega_{\mathrm{FS}}\right) &\leq \int_{\mathbb{P}^1} e^{\alpha(g^*\varphi)^2}\omega_{\mathrm{FS}} = \int_{\mathbb{P}^1} e^{\alpha\varphi^2} g_*\omega_{\mathrm{FS}}\\ &\leq \|g\|^4 \int_{\mathbb{P}^1} e^{\alpha\varphi^2}\omega_{\mathrm{FS}} \leq A\|g\|^4.\end{aligned}$$

Taking the logarithm gives

$$\|g^*\varphi\|_{L^2} \leq A_2 + A_3 \log\|g\|, \tag{6}$$

for some constants $A_2, A_3 > 0$.

By Cauchy-Schwarz we have

$$\|f_\mu^*\varphi\|_{L^2} = \Big\| \int_G g^*\varphi \, \mathrm{d}\mu(g) \Big\|_{L^2} \leq \int_G \|g^*\varphi\|_{L^2} \, \mathrm{d}\mu(g).$$

Since $\int_G \log\|g\| \, \mathrm{d}\mu(g)$ is finite by assumption, it follows that $\|f_\mu^*\varphi\|_{L^2} \leq A_4$ for every $\varphi \in \mathcal{F}$ for some constant $A_4 > 0$. This finishes the proof. ☐

2.1. Non-elementary measures

Let μ be a probability measure on $\mathrm{PSL}_2(\mathbb{C})$. We will denote by T_μ the smallest closed sub-semigroup of $\mathrm{PSL}_2(\mathbb{C})$ containing the support of μ.

Definition 2.6. *Let R be a subset of $\mathrm{PSL}_2(\mathbb{C})$. We say that R is **elementary** if either R is conjugated to a subset of $\mathrm{PSU}(2)$ or if there is a finite subset of $\mathbb{P}^1$ which is invariant by every element of R. We say that a probability measure μ on $\mathrm{PSL}_2(\mathbb{C})$ is elementary if $\mathrm{supp}(\mu)$ is an elementary set.*

Remark 2.7. *(i) It is easy to see that R is elementary if and only if the closed semigroup generated by R is elementary. In particular μ is elementary if and only if T_μ is elementary.*

(ii) A subset R of $\mathrm{PSL}_2(\mathbb{C})$ is conjugated to a subset of $\mathrm{PSU}(2)$ if and only if the group generated by R is relatively compact. This follows from the fact that if the semigroup generated by R is relatively compact then there exists an R-invariant inner product on $\mathbb{C}^2$, obtained by averaging the standard inner product.

(iii) We can view $\mathbb{P}^1$ as the boundary of the 3-dimensional hyperbolic space $\mathbb{H}^3$. Then, any Möbius transformation of $\mathbb{P}^1$ extends to a homeomorphism of $\overline{\mathbb{H}^3} = \mathbb{H}^3 \cup \mathbb{P}^1$, called the Poincaré extension, that preserves the standard hyperbolic metric on $\mathbb{H}^3$. In this context, a subset R of $\mathrm{PSL}_2(\mathbb{C})$ is elementary if and only if it admits a finite orbit in $\overline{\mathbb{H}^3}$, see [1].

The following result is probably well-known. We include a proof for the convenience of the reader. Recall that an element g of $\mathrm{Aut}(\mathbb{P}^1)$ different from the identity is conjugated to either $z \mapsto z+1$ or $z \mapsto \lambda z$ for some $\lambda \in \mathbb{C}\setminus\{0,1\}$. In the former case, g is called *parabolic* and in the latter, g is called *elliptic* if $|\lambda| = 1$ or *loxodromic* if $|\lambda| \neq 1$, see also the appendix below.

Lemma 2.8. *Let $n \geq 1$. Then μ is non-elementary if and only if μ^{*n} is non elementary.*

Proof. Notice that $T_{\mu^{*n}} \subset T_\mu$, so if μ elementary then so is μ^{*n}.

Suppose now that μ is non-elementary and fix n. It follows from Lemma A.1 in the Appendix that T_μ contains a loxodromic element g_0. Then $T_{\mu^{*n}}$ contains a loxodromic element, namely, g_0^n. In particular $T_{\mu^{*n}}$ is non-compact and cannot be conjugated to subset of $\mathrm{PSU}(2)$.

To finish the proof we need to show that $\mathrm{supp}(\mu^{*n})$ leaves no finite set invariant. Suppose $F \subset \mathbb{P}^1$ is finite and invariant by $\mathrm{supp}(\mu^{*n})$. Then F is also invariant by $T_{\mu^{*n}}$. Since $\mathrm{id} \neq g_0^n \in T_{\mu^{*n}}$ is loxodromic, this implies that $F \subset \mathrm{Fix}(g_0)$. As μ is non-elementary, we can find another loxodromic element $g_1 \in T_\mu$ whose fix point set is disjoint from $\mathrm{Fix}(g_0)$ (see [1, Thm. 5.1.3]). Repeating the preceding argument for g_1 gives $F \subset \mathrm{Fix}(g_1)$. This implies that $F = \varnothing$, completing the proof. □

The main result of this section is the following.

Proposition 2.9. *Let μ be a non-elementary probability measure on $\mathrm{PSL}_2(\mathbb{C})$. Then there exists an $N \geq 1$ such that the norm of the operator $(f_\mu^N)^* : L^2_{(1,0)} \to L^2_{(1,0)}$ is strictly less than 1.*

Remark 2.10. *Proposition 2.9 can be interpreted using the language of unitary representations. Since the action of a given $g \in G = \mathrm{PSL}_2(\mathbb{C})$ preserves* *the total mass of positive measures, g acts as a unitary operator $\pi(g)$ on $L^2_{(1,0)}$. We define $\pi(\mu) := \int_G \pi(g)\,\mathrm{d}\mu(g)$. It is standard that $\|\pi(\mu)\| \le 1$ (cf. Proposition 2.1) and there are general theorems ensuring that, under some non-amenability hypothesis on μ, either the norm of $\pi(\mu)$ or its spectral radius is strictly less than one. See [2] and the references therein for a recent account on this kind of results. However, this is not the point of view we adopt here and our proof of Proposition 2.9 is independent. Moreover, when we consider matrices of arbitrary size, we have to deal with non-unitary representations. In dimension k we can also look at the unitary representation on the space of $(k,0)$-forms, but it is not clear to us how to explore that.*

Proof of Proposition 2.9. For $n \ge 1$ introduce

$$R^n := \{g_n \cdots g_2 g_1 : g_i \in \operatorname{supp}(\mu)\} \quad \text{and} \quad S^n := \{gh^{-1} : g, h \in R^n\}.$$

Notice that R^n is a dense subset of the support of the μ^{*n}.

By Proposition 2.1, $\|(f_\mu^n)^*\| \le 1$ for every $n \ge 1$. Suppose by contradiction that $\|(f_\mu^n)^*\| = 1$ for every $n \ge 1$. We will show that in this case μ must be elementary.

Since $\|f_\mu^*\| = 1$, there exists a sequence of $(1,0)$-forms $\{\phi_n\}_{n\ge 0}$ such that $\|\phi_n\|_{L^2} = 1$ and $\|f_\mu^*(\phi_n)\|_{L^2} \to 1$. By compactness, the sequence $\{i\phi_n \wedge \overline{\phi_n}\}_{n\ge0}$ of probability measures admits a subsequence, which we still denote by $\{i\phi_n \wedge \overline{\phi_n}\}_{n\ge0}$ for simplicity, that converges to a probability measure m.

By the proof of Proposition 2.1, the measures

$$\nu_n := \int_{G\times G} i(g_1^*\phi_n - g_2^*\phi_n) \wedge \overline{(g_1^*\phi_n - g_2^*\phi_n)}\,\mathrm{d}\mu(g_1)\mathrm{d}\mu(g_2)$$

tend to zero as $n \to \infty$. In particular,

$$\|\nu_n\| = \int_{G\times G} \|g_1^*\phi_n - g_2^*\phi_n\|^2_{L^2}\,\mathrm{d}\mu(g_1)\mathrm{d}\mu(g_2) \longrightarrow 0$$

as $n \to \infty$.

By Cauchy-Schwarz and the fact that $\|g^*\phi_n\|_{L^2} = \|\phi_n\|_{L^2} = 1$ for $g \in \mathrm{PSL}_2(\mathbb{C})$ we have

$$\begin{aligned}
\|g_1^*(i\phi_n \wedge \overline{\phi_n}) - g_2^*(i\phi_n \wedge \overline{\phi_n})\|_{L^1} = \\
&= \|ig_1^*\phi_n \wedge \overline{(g_1^*\phi_n - g_2^*\phi_n)} + i(g_1^*\phi_n - g_2^*\phi_n) \wedge \overline{g_2^*\phi_n}\|_{L^1} \\
&\le \|g_1^*\phi_n\|_{L^2}\|g_1^*\phi_n - g_2^*\phi_n\|_{L^2} + \|g_2^*\phi_n\|_{L^2}\|g_1^*\phi_n - g_2^*\phi_n\|_{L^2} \\
&= 2\|g_1^*\phi_n - g_2^*\phi_n\|_{L^2},
\end{aligned}$$

so

$$\int_{G\times G} \|g_1^*(i\phi_n \wedge \overline{\phi_n}) - g_2^*(i\phi_n \wedge \overline{\phi_n})\|_{L^1}^2 \,\mathrm{d}\mu(g_1)\mathrm{d}\mu(g_2) \longrightarrow 0$$

as $n \to \infty$.

By Lebesgue's dominated convergence theorem, it follows that

$$\int_{G\times G} \|g_1^*(m) - g_2^*(m)\|^2 \,\mathrm{d}\mu(g_1)\mathrm{d}\mu(g_2) = 0,$$

which implies that $g_1^*(m) = g_2^*(m)$ for $\mu \otimes \mu$ almost every (g_1, g_2).

Claim. $g_1^*(m) = g_2^*(m)$ for all $g_1, g_2 \in \mathrm{supp}(\mu)$.

Indeed, we know that $g_1^*(m) = g_2^*(m)$ holds for g_1 and g_2 on a set of full μ-measure. Now, such a set is dense in the support of μ for the standard distance on $\mathrm{PSL}_2(\mathbb{C})$ and $g \mapsto g^*m$ is continuous with respect to this distance. Hence $g_1^*(m) = g_2^*(m)$ for all $g_1, g_2 \in \mathrm{supp}(\mu)$ and the claim is proved.

The claim is equivalent to

$$(g_1 g_2^{-1})^* m = m \quad \text{for every} \quad g_1, g_2 \in \mathrm{supp}(\mu).$$

Which means m is invariant by $S := S^1$.

Since we also have $\|(f_\mu^n)^*\| = 1$ by assumption, we can replace f_μ by f_μ^n in the above proof and get, for each $n \geq 1$, a probability measure m_n invariant by S^n.

We can now finish the proof. After replacing f by f^{N_2}, R by R^{N_2} and S by S^{N_2} for some N_2 we may assume that S contains a non-elliptic element g_0 different from the identity. This is possible by Lemma A.1 from the appendix. By the above discussion, there exists a probability measure m_1 invariant by the pullback by every element of S. In particular $g_0^* m_1 = m_1$ and by iteration $(g_0^n)^* m_1 = m_1$ for every $n \geq 1$. Making $n \to \infty$ implies that $m_1 = \alpha_1 \delta_x + \beta_1 \delta_y$, with α_1, $\beta_1 \geq 0$ and $\alpha_1 + \beta_1 = 1$, where x and y are the fix points of g_0 (if g_0 is parabolic we set $x = y$). Notice that $S \subset S^n$ for every $n \geq 1$ so the measure m_n is also invariant by g_0. Hence $m_n = \alpha_n \delta_x + \beta_n \delta_y$ with $\alpha_n + \beta_n = 1$.

We will show now that μ is elementary. Let F^n be the largest finite S^n-invariant subset of $\mathbb{P}^1$. Notice that when $n \leq m$ we have $S^n \subset S^m$, hence $F^m \subset F^n$. We have $F^n \neq \varnothing$ for every $n \geq 1$, because S^n preserves the atomic measure m_n. Also, since F^1 is invariant by g_0 we have $F^1 \subset \{x, y\}$. We separate in a few cases.

Case 1. $F^1 = \{x\}$. In this case S fixes x, which means that there is a point p such that R maps p to x. As $\varnothing \neq F^2 \subset F^1$ we have $F^2 = \{x\}$, so S^2 also fixes x. Hence there is a point q such that R^2 maps q to x. For $g \in R$ we have that $g^2 \in R^2$, so $g \cdot p = x = g^2 \cdot q$. Hence $g \cdot q = p$ for every $g \in R$. This implies that p is S-invariant, so we must have $p = x$. We conclude that $g \cdot x = x$ for every $g \in R = \mathrm{supp}(\mu)$, so μ is elementary.

Case 2. $F^1 = \{x, y\}$ and $F^2 = \{x\}$ or $\{y\}$. In that case we can replace f_μ by f_μ^2 and repeat the argument of Case 1, see also Lemma 2.8.

Case 3. $F^1 = \{x, y\}$ and $F^2 = \{x, y\}$. If $x = y$ we fall in Case 1, so we may assume $x \neq y$. In this case the set $\{x, y\}$ is S-invariant, which means that there are points p, q such that R maps $\{p, q\}$ to $\{x, y\}$. Analogously, $\{x, y\}$ is S^2-invariant so there are points r, s such that R^2 maps $\{r, s\}$ to $\{x, y\}$. For $g \in R$ we have that $g^2 \in R^2$, so $g\{p, q\} = \{x, y\} = g^2\{r, s\}$. Hence $\{p, q\} = g\{r, s\}$ for every $g \in R$, which implies that $\{p, q\}$ is S-invariant. By the maximality of F^1 we get $\{x, y\} = \{p, q\}$. Hence $R = \mathrm{supp}(\mu)$ maps $\{x, y\}$ to itself, so μ is elementary.

Summing up, we have shown that if $\|(f_\mu^n)^*\| = 1$ for every $n \geq 1$ then μ must be elementary, thus completing the proof. ☐

Once we know that, up to taking iterates, $f_\mu^* : L^2_{(1,0)} \to L^2_{(1,0)}$ has norm less than one, we will have that $f_\mu^* : W^{1,2} \to W^{1,2}$ has a spectral gap. It is then well known how to use this to produce a stationary measure. This is the content of the next result.

Theorem 2.11. *Let μ be a non-elementary probability measure on $G = \mathrm{PSL}_2(\mathbb{C})$. Assume that μ has a finite first moment, i.e. $\int_G \log\|g\| \mathrm{d}\mu(g) < +\infty$. Then μ admits a stationary measure ν that can be extended to a continuous linear functional on $W^{1,2}$ with the following properties*

1. *There are constants $A > 0$ and $0 < \lambda < 1$ such that*
$$\left\|(f_\mu^n)^* h - \langle \nu, h\rangle\right\|_{W^{1,2}} \leq A\|\partial h\|_{L^2}\lambda^n, \text{ for every } n \geq 0 \text{ and } h \in W^{1,2}.$$
2. *$|\langle \nu, h\rangle| \leq A'\|h\|_{W^{1,2}}$ for some constant $A' > 0$ independent of h.*

In particular, ν has no mass on polar subsets of $\mathbb{P}^1$.

Proof. By Proposition 2.9 we may assume, after replacing f_μ by f_μ^N for some $N \geq 1$, that the norm of f_μ^* acting on $L^2_{(1,0)}$ is less than one. Let $0 < \lambda < 1$ be its value.

For $h \in W^{1,2}$, let

$$c_0 := \int_X h\omega_{\mathrm{FS}} \qquad \text{and} \qquad h_0 := h - c_0$$

and define inductively

$$c_n := \int_X (f_\mu^* h_{n-1})\omega_{\rm FS} \qquad \text{and} \qquad h_n := f_\mu^* h_{n-1} - c_n.$$

Then

$$(f_\mu^n)^* h = h_n + c_n + c_{n-1} + \cdots + c_1 + c_0. \tag{7}$$

By Proposition 2.3, we have $h_n \in W^{1,2}$ for all n. Notice that $\langle \omega_{\rm FS}, h_n\rangle = 0$, so by Poincaré-Sobolev inequality we have $\|h_n\|_{L^2} \le A_1 \|\partial h_n\|_{L^2}$ for some constant $A_1 > 0$. We also have $\partial h_n = f_\mu^*(\partial h_{n-1})$ for every n. Then

$$\|h_n\|_{L^2} \le A_1 \|\partial h_n\|_{L^2} = A_1 \|(f_\mu^n)^*(\partial h)\|_{L^2} \le A_1 \lambda^n \|\partial h\|_{L^2}.$$

By Propositions 2.2 and 2.3, there is a constant $A_2 > 0$ such that $\|f^*\varphi\|_{L^2} \le A_2 \|\varphi\|_{W^{1,2}}$ for every $\varphi \in W^{1,2}$. Hence, we have

$$\begin{aligned} |c_n| &= \Big| \int_X (f_\mu^* h_{n-1})\omega_{\rm FS}\Big| \le \|f_\mu^* h_{n-1}\|_{L^2} \le A_2 \|h_{n-1}\|_{W^{1,2}} \\ &= A_2 \|\partial h_{n-1}\|_{L^2} = A_2 \|(f_\mu^{n-1})^*(\partial h)\|_{L^2} \le A_2 \lambda^{n-1} \|\partial h\|_{L^2}. \end{aligned}$$

Set $c_h := \sum_{k=0}^\infty c_k$ and define the linear functional ν by

$$\langle \nu, h\rangle := c_h \quad \text{for } h \in W^{1,2}.$$

Clearly, this constant is finite and satisfies the second estimate in the theorem for a suitable constant $A' > 0$. We also have $\langle \nu, \mathbf{1}\rangle = 1$ according to (7) and if h is smooth and non-negative we have $\langle \nu, h\rangle = \lim_{n\to\infty} \int_{\mathbb{P}^1} (f_\mu^n)^* h \, \omega_{\rm FS} \ge 0$. So, by Riesz Representation Theorem, ν defines a probability measure on $\mathbb{P}^1$.

We have from (7) that

$$\|(f_\mu^n)^* h - \langle \nu, h\rangle\|_{L^2} = \|(f_\mu^n)^* h - c_h\|_{L^2} = \Big\| h_n - \sum_{k=n+1}^\infty c_k \Big\|_{L^2}$$

$$\le \|h_n\|_{L^2} + \sum_{k=n+1}^\infty |c_k| \le A_1 \lambda^n \|\partial h\|_{L^2} + \sum_{k=n+1}^\infty A_2 \lambda^{k-1} \|\partial h\|_{L^2} \le A_3 \|\partial h\|_{L^2} \lambda^n$$

for some constant $A_3 > 0$. On the other hand, by Proposition 2.2 and the definition of λ, we obtain

$$\|(f_\mu^n)^* h - c_h\|_{W^{1,2}} \lesssim \|(f_\mu^n)^* h - c_h\|_{L^2} + \|(f_\mu^n)^*(\partial h)\|_{L^2}$$

$$\leq \|(f_\mu^n)^*h - c_h\|_{L^2} + \lambda^n \|\partial h\|_{L^2}$$
$$\leq A_4 \lambda^n \|\partial h\|_{L^2}$$

for some constant $A_4 > 0$. Thus, we get the first estimate in the theorem for a suitable constant $A > 0$.

In order to show that ν is stationary we notice that, from the obtained estimate, we have $(f_\mu^n)^*h \to \langle \nu, h\rangle$ in $W^{1,2}$ for every $h \in W^{1,2}$. In particular, if φ is a smooth test function, then

$$\langle \nu, \varphi\rangle = \lim_{n\to\infty} (f_\mu^{n+1})^*\varphi = \lim_{n\to\infty} (f_\mu^n)^* f_\mu^* \varphi = \langle \nu, f_\mu^*\varphi\rangle = \langle (f_\mu)_*\nu, \varphi\rangle,$$

showing that $(f_\mu)_*\nu = \nu$, that is, ν is stationary.

We now prove the last statement. If $E \subset \mathbb{P}^1$ is a polar set then, by definition, there is a quasi-subharmonic function u on $\mathbb{P}^1$ such that $E \subseteq \{u = -\infty\}$. We may assume that $u \leq -1$ and u is the limit of a decreasing sequence of smooth negative functions u_n with $dd^c u_n \geq -\omega_{\text{FS}}$. Then $h := -\log(-u)$ belongs to $W^{1,2}$ and is the decreasing limit of the sequence $h_n := -\log(-u_n)$ which is bounded in $W^{1,2}$, see [10] and [23, Ex.1]. The function h is defined everywhere and is bounded from above, so $\langle \nu, h\rangle$ coincides with the integral of h with respect to μ. The fact $h = -\infty$ on E and that $\langle \nu, h\rangle$ is finite imply that $\nu(E) = 0$. The proof is now complete. □

Remark 2.12. *As mentioned in the Introduction, it is well known since Furstenberg that a non-elementary measure admits a unique stationary measure. Hence, the measure ν in the above theorem is necessarily the unique μ-stationary measure and our result says that the iterates of the Markov operator f_μ^* converge exponentially fast with respect to the Sobolev norm to the operator $\varphi \mapsto \langle \nu, \varphi\rangle \mathbf{1}$. The uniqueness of the stationary measure also follows from Theorem 1.2.*

The proof of Theorem 1.1 follows immediately from Theorem 2.11 and Remark 2.12. The following consequence of Theorem 2.11 will be used later.

Corollary 2.13. *Let μ and ν be as in Theorem 2.11. Then $\|\varphi\|_\nu := |\langle \nu, \varphi\rangle| + \|\partial\varphi\|_{L^2}$ defines a norm on $W^{1,2}$ which is equivalent to $\|\cdot\|_{W^{1,2}}$.*

Proof. Clearly $\|\cdot\|_\nu \lesssim \|\cdot\|_{W^{1,2}}$ by Theorem 2.11. We now prove the reverse inequality. Let $\varphi \in W^{1,2}$ and define $m(\varphi) := \int \varphi\, \omega_{\text{FS}}$. Then $\|\varphi\|_{W^{1,2}} = |m(\varphi)| + \|\partial\varphi\|_{L^2}$. By Theorem 2.11, we have

$$|\langle \nu, \varphi - m(\varphi)\rangle| \lesssim \|\varphi - m(\varphi)\|_{W^{1,2}} = \|\partial\varphi\|_{L^2}.$$

Hence

$$|m(\varphi)| = |\langle \nu, m(\varphi)\rangle| \leq |\langle \nu, \varphi - m(\varphi)\rangle| + |\langle \nu, \varphi\rangle| \lesssim \|\varphi\|_\nu.$$

This gives $\|\varphi\|_{W^{1,2}} \lesssim \|\varphi\|_\nu$ and completes the proof. □

3. Equidistribution of points

This section is devoted to the proof of Theorem 1.2.

We will need the following consequence of Proposition 2.5. A proof can be found in [9]. In what follows, we say that a real valued function u on $\mathbb{P}^1$ is (M,γ)-*Hölder continuous* if $|u(x) - u(y)| \leq M \operatorname{dist}(x,y)^\gamma$ for every $x, y \in \mathbb{P}^1$. When $\gamma = 1$ we say that u is M-*Lipschitz*.

Lemma 3.1. *Let $\mathcal{F}$ be a bounded subset of $W^{1,2}$. There is a constant $A = A(\mathcal{F}) > 0$ (independent of M and γ) such that if $\varphi \in \mathcal{F}$ is (M,γ)-Hölder continuous for some constants $M \geq 1$ and $0 < \gamma \leq 1$, then*

$$\|\varphi\|_\infty \leq A\gamma^{-1}(1 + \log M).$$

Proof of Theorem 1.2. By the Theory of Interpolation between Banach spaces it is enough to prove the result for $\beta = 1$, see [22]. We can normalize φ so that $\|\varphi\|_{\mathcal{C}^1} \leq 1$ and $\langle \nu, \varphi\rangle = 0$.

Let $\varphi_n := (f_\mu^n)^*\varphi$. Since

$$\langle (f_\mu^n)_*\delta_a, \varphi\rangle = \langle \delta_a, (f_\mu^n)^*\varphi\rangle = \varphi_n(a)$$

we need to show that $\|\varphi_n\|_\infty \leq A\gamma^n$ for some constants $A > 0$ and $0 < \gamma < 1$.

Let λ_0 be the norm of f_μ^* acting on $L^2_{(1,0)}$. By Proposition 2.9, after replacing μ by μ^{*N} for some $N \geq 1$ if necessary, we may assume that $0 < \lambda_0 < 1$. Let $C_n := e^{\delta_0^n}$ where $\delta_0 > 1$ is a constant such that $1 < \delta_0^{1+\epsilon} < \frac{1}{\lambda_0}$. Set

$$\mathcal{A}^{(n)} := \{(g_1, \ldots, g_n) \in G^n : \|g_n \cdots g_1\| \leq C_n\}$$

and

$$\mathcal{B}^{(n)} := \{(g_1, \ldots, g_n) \in G^n : \|g_n \cdots g_1\| > C_n\}.$$

We can then write $\varphi_n = \varphi_n^{(1)} + \varphi_n^{(2)}$, where

$$\varphi_n^{(1)}(x) := \int_{\mathcal{A}^{(n)}} \varphi(g_n \cdots g_1 \cdot x)\, \mathrm{d}\mu^n(g_1, \ldots, g_n)$$

and

$$\varphi_n^{(2)}(x) := \int_{\mathcal{B}^{(n)}} \varphi(g_n \cdots g_1 \cdot x)\, \mathrm{d}\mu^n(g_1, \ldots, g_n).$$

We will show separately that $\|\varphi_n^{(1)}\|_\infty$ and $\|\varphi_n^{(2)}\|_\infty$ are bounded by $A\gamma^n$ for some constants $A > 0$ and $0 < \gamma < 1$.

We start by estimating $\varphi_n^{(2)}$. Let $M_n = \int (\log \|g_n \cdots g_1\|)^{1+\epsilon}\, \mathrm{d}\mu^n(g_1, \ldots, g_n)$ be the $(1+\epsilon)$-moment of μ^{*n}. By assumption M_1 is finite. We also have that $M_n \leq n^{1+\epsilon} M_1$ by the sub-additivity of $\log \|g\|$. This implies that

$$\mu^{\otimes n}(\mathcal{B}^{(n)}) \leq \frac{M_1 \cdot n^{1+\epsilon}}{(\log C_n)^{1+\epsilon}} = \frac{M_1 \cdot n^{1+\epsilon}}{\delta_0^{n(1+\epsilon)}}. \tag{8}$$

Since $\|\varphi\|_\infty \leq 1$, the definition of $\varphi_n^{(2)}$ implies that $\|\varphi_n^{(2)}\|_\infty \leq M_1 n^{1+\epsilon} \delta_0^{-n(1+\epsilon)}$, which is bounded by $A_2 \gamma_2^n$ for some constants $A_2 > 0$ and $0 < \gamma_2 < 1$.

In order to estimate $\varphi_n^{(1)}$ choose a constant δ such that $\delta_0 < \delta < \delta_0^{1+\epsilon}$ and set $\widehat{\varphi}_n = \delta^n \varphi_n$ and $\widehat{\varphi}_n^{(j)} = \delta^n \varphi_n^{(j)}$, $j = 1, 2$. We have $\widehat{\varphi}_n = \widehat{\varphi}_n^{(1)} + \widehat{\varphi}_n^{(2)}$.

Claim. $\widehat{\varphi}_n$, $\widehat{\varphi}_n^{(1)}$ and $\widehat{\varphi}_n^{(2)}$ belong to a bounded family in $W^{1,2}$.

Proof. We will prove that $\widehat{\varphi}_n$ and $\widehat{\varphi}_n^{(2)}$ belong to a bounded family. Then the result for $\widehat{\varphi}_n^{(1)}$ will follow because $\widehat{\varphi}_n^{(1)} = \widehat{\varphi}_n - \widehat{\varphi}_n^{(2)}$.

By the invariance of ν we have that $\langle \nu, \widehat{\varphi}_n \rangle = \delta^n \langle \nu, \varphi_n \rangle = \delta^n \langle \nu, \varphi \rangle = 0$. We also have that

$$\|\partial \widehat{\varphi}_n\|_{L^2} = \delta^n \|\partial \varphi_n\|_{L^2} = \delta^n \|(f_\mu^n)^* \partial \varphi\|_{L^2} \lesssim \delta^n \lambda_0^n \|\partial \varphi\|_{L^2} \leq (\delta_0^{1+\epsilon})^n \lambda_0^n \|\partial \varphi\|_{L^2}$$

is bounded uniformly in n since $\delta < \delta_0^{1+\epsilon} < \frac{1}{\lambda_0}$. Therefore $\widehat{\varphi}_n$ is a bounded family in $W^{1,2}$ for $n \geq 1$.

We now prove that $\widehat{\varphi}_n^{(2)}$ belong to a bounded family. Using (8) and the definition of $\varphi_n^{(2)}$ we have that

$$|\langle \omega_{\mathrm{FS}}, \varphi_n^{(2)} \rangle| \leq \|\varphi_n^{(2)}\|_\infty \leq M_1 n^{1+\epsilon} \delta_0^{-n(1+\epsilon)}$$

and, by Cauchy-Schwarz inequality and Proposition 2.1

$$\|\partial \varphi_n^{(2)}\|_{L^2} \leq \|\partial \varphi\|_{L^2} M_1 n^{1+\epsilon} \delta_0^{-n(1+\epsilon)} \leq M_1 n^{1+\epsilon} \delta_0^{-n(1+\epsilon)}.$$

Hence $\|\widehat{\varphi}_n^{(2)}\|_{W^{1,2}} \lesssim M_1 n^{1+\epsilon} \delta^n \delta_0^{-n(1+\epsilon)}$. Since $1 < \delta < \delta_0^{1+\epsilon}$, the last quantity is bounded uniformly in n. This proves the claim. □

We can now finish the proof of the theorem. Notice that $\varphi_n^{(1)}$ is $A_0 C_n^2$-Lipschitz for some universal constant $A_0 > 0$. This is not difficult to check using Cartan's decomposition as in Lemma 2.4. Therefore $\widehat{\varphi}_n^{(1)}$ is $A_0 \delta^n C_n^2$-Lipschitz. By Lemma 3.1 and the above claim we get

$$\|\widehat{\varphi}_n^{(1)}\|_\infty \leq B\big(1 + \log(A_0 \delta^n C_n^2)\big) = B'\big(1 + n \log \delta + 2\delta_0^n\big)$$

for some constants $B, B' > 0$, giving

$$\|\varphi_n^{(1)}\|_\infty \leq B' \delta^{-n}\big(1 + n \log(\lambda_0^{-1}) + 2\delta_0^n\big).$$

Since $1 < \delta_0 < \delta$ we get $\|\varphi_n^{(1)}\|_\infty \leq A_1 \gamma_1^n$ for some constants $A_1 > 0$ and $0 < \gamma_1 < 1$.

Taking $A = 2\max\{A_1, A_2\}$ and $\gamma = \max\{\gamma_1, \gamma_2\}$ gives $\|\varphi_n\|_\infty \leq A\gamma^n$, finishing the proof. $\square$

4. Central Limit Theorem

This section is devoted to the proof of Theorem 1.3. We begin by recalling some basic notions, see [6] for more details.

Let μ be a probability measure on $G = \mathrm{PSL}_2(\mathbb{C})$ satisfying the first moment condition $\int \log \|g\| \, \mathrm{d}\mu(g) < +\infty$. Then, the (upper) **Lyapunov exponent** of μ is defined as

(9)
$$\gamma := \lim_{n \to \infty} \frac{1}{n} \mathbf{E}(\log \|g_n \cdots g_1\|) = \lim_{n \to \infty} \frac{1}{n} \int \log \|g_n \cdots g_1\| \, \mathrm{d}\mu(g_1) \cdots \mathrm{d}\mu(g_n).$$

It follows from Kingman's subadditive ergodic theorem that

$$\gamma = \lim_{n \to \infty} \frac{1}{n} \log \|g_n \cdots g_1\| \quad \text{almost surely.}$$

We also have the following formula, due to Furstenberg (cf. [6, III.3.4]):

$$\gamma = \int_{\mathbb{P}^1} \int_G \log \frac{\|g \cdot v\|}{\|v\|} \mathrm{d}\mu(g) \mathrm{d}\nu(x), \quad x = [v]. \tag{10}$$

Here and in what follows v will denote a non-zero vector in $\mathbb{C}^2$ and $x = [v]$ will be the corresponding point in $\mathbb{P}^1$. We will call v a *lift* of x. Notice that the quantity $\frac{\|g \cdot v\|}{\|v\|}$ is independent of the choice of lift.

For the proof of Theorem 1.3, we will apply the method of Gordin-Liverani. Recall their theorem.

Theorem 4.1 (Gordin-Liverani, [14, 18])**.** *Let $(X, \mathfrak{m})$ be a probability space and let $F : X \to X$. Assume that $\mathfrak{m}$ is F-invariant and ergodic. Let $F^* : \phi \mapsto \phi \circ F$ be the pullback operator acting on $L^2(\mathfrak{m})$ and denote by $\Lambda : L^2(\mathfrak{m}) \to L^2(\mathfrak{m})$ its adjoint.*

Let $\widetilde{\varphi} \in L^2(\mathfrak{m})$ be such that $\langle \mathfrak{m}, \widetilde{\varphi} \rangle = 0$ and assume $\widetilde{\varphi}$ is not a coboundary, that is, not of the form $\widetilde{\varphi} = \psi \circ F - \psi$ for some $\psi \in L^2(\mathfrak{m})$. If

$$\sum_{n \geq 0} \|\Lambda^n \widetilde{\varphi}\|^2_{L^2(\mathfrak{m})} < +\infty,$$

then the sequence of random variables $Z_n := \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \widetilde{\varphi} \circ F^j$ converges in distribution to a Gaussian random variable of mean 0 and variance $\sigma > 0$, where

$$\sigma^2 = -\langle \mathfrak{m}, (\widetilde{\varphi})^2 \rangle + 2 \sum_{n \geq 0} \langle \mathfrak{m}, \widetilde{\varphi} \cdot (\widetilde{\varphi} \circ F^n) \rangle.$$

Our approach is to first apply the above theorem to a certain dynamical system on $X = G^{\mathbb{N}^*} \times \mathbb{P}^1$ and the observable $\widetilde{\varphi}(\mathbf{g}, x) = \log \frac{\|g_1^{-1} v\|}{\|v\|} + \gamma$, where $\mathbf{g} = (g_1, g_2, \ldots) \in G^{\mathbb{N}^*}$. After that, we will translate the corresponding CLT to the CLT for the random variables $Y_n^v = \log \frac{\|g_n \cdots g_1 \cdot v\|}{\|v\|}$.

Let μ be a non-elementary probability measure on $G = \mathrm{PSL}_2(\mathbb{C})$ and denote by ν the unique μ-stationary measure on $\mathbb{P}^1$. We have the following fundamental result, see [6, Prop. II.3.3].

Proposition 4.2 (Furstenberg)**.** *For almost every sequence $\mathbf{g} = (g_1, g_2, \ldots)$ there exists a point $Z(\mathbf{g}) \in \mathbb{P}^1$ such that*

$$\lim_{n \to \infty} (g_1 \cdots g_n)_* \nu = \delta_{Z(\mathbf{g})}.$$

Furthermore the distribution of $Z(\mathbf{g})$ is ν, that is,

$$\int_{G^{\mathbb{N}^*}} \delta_{Z(\mathbf{g})} \, \mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}) = \nu. \tag{11}$$

An alternative way of phrasing the above result is to say that there exists a map $Z : G^{\mathbb{N}^*} \to \mathbb{P}^1$ defined $\mu^{\mathbb{N}^*}$-almost everywhere such that $Z_* \mu^{\mathbb{N}^*} = \nu$.

Let $X := G^{\mathbb{N}^*} \times \mathbb{P}^1$. Consider the shift map

$$T : G^{\mathbb{N}^*} \to G^{\mathbb{N}^*}, \quad T((g_1, g_2, \ldots)) = (g_2, g_3, \ldots)$$

and the fibered product

$$F : X \to X, \quad F(\mathbf{g}, x) = (T\mathbf{g}, g_1^{-1} \cdot x).$$

It follows from Proposition 4.2 that

$$g_1^{-1}Z(\mathbf{g}) = Z(T\mathbf{g}) \quad \text{and} \quad g_1 Z(T\mathbf{g}) = Z(\mathbf{g}). \tag{12}$$

In particular, F maps $(\mathbf{g}, Z(\mathbf{g}))$ to $(T\mathbf{g}, Z(T\mathbf{g}))$. Define a probability measure $\mathfrak{m}$ on X by

$$\mathfrak{m} := \int_{G^{\mathbb{N}^*}} \delta_{\mathbf{g}} \otimes \delta_{Z(\mathbf{g})} \, \mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}). \tag{13}$$

Lemma 4.3. *The measure $\mathfrak{m}$ is F-invariant and ergodic.*

Proof. The result is well known. The invariance of $\mathfrak{m}$ follows from (12) and a direct computation. The ergodicity of $\mathfrak{m}$ comes from the ergodicity of $\mu^{\mathbb{N}^*}$. See also [4, p. 33] for a more general statement. □

In what follows we identify functions on $\mathbb{P}^1$ with functions on $G^{\mathbb{N}^*} \times \mathbb{P}^1$ that depend only on the $\mathbb{P}^1$ variable. Similarly, we identify functions on $G \times \mathbb{P}^1$ with functions on $G^{\mathbb{N}^*} \times \mathbb{P}^1$ that depend only on the $\mathbb{P}^1$ variable and the first entry g_1 of the sequence $\mathbf{g} = (g_1, g_2, \ldots)$. Recall that Λ is the adjoint of F^* acting on $L^2(\mathfrak{m})$.

Lemma 4.4. *Let ψ be a function on $G \times \mathbb{P}^1$ viewed as a function on $G^{\mathbb{N}^*} \times \mathbb{P}^1$. Assume that $\psi \in L^2(\mathfrak{m})$. Then $\Lambda\psi(\mathbf{g}, x)$ depends only on x and is given by*

$$\Lambda\psi(x) = \int_G \psi(h, h \cdot x) \, \mathrm{d}\mu(h). \tag{14}$$

In particular, if ψ is a function on $\mathbb{P}^1$ viewed as a function on $G^{\mathbb{N}^*} \times \mathbb{P}^1$ we have $\Lambda\psi = f_\mu^*\psi$.

Proof. Let ψ be as above. Using (12) we have, for any $\phi \in L^2(\mathfrak{m})$,

$$\begin{aligned}
\langle \phi, \Lambda\psi \rangle_{L^2(\mathfrak{m})} = \langle F^*\phi, \psi \rangle_{L^2(\mathfrak{m})} &= \int_{G^{\mathbb{N}^*}} \phi(T\mathbf{g}, g_1^{-1}Z(\mathbf{g})) \, \psi(g_1, Z(\mathbf{g})) \mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}) \\
&= \int_{G^{\mathbb{N}^*}} \phi(T\mathbf{g}, Z(T\mathbf{g})) \psi(g_1, g_1 Z(T\mathbf{g})) \mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}) \\
&= \int_{G^{\mathbb{N}^*}} \phi(T\mathbf{g}, Z(T\mathbf{g})) \left(\int_G \psi(g_1, g_1 Z(T\mathbf{g}) \mathrm{d}\mu(g_1) \right) \mathrm{d}\mu^{\mathbb{N}^*}(T\mathbf{g}) \\
&= \int_{G^{\mathbb{N}^*}} \phi(\mathbf{g}', Z(\mathbf{g}')) \left(\int_G \psi(g_1, g_1 Z(\mathbf{g}') \mathrm{d}\mu(g_1) \right) \mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}') \\
&= \Big\langle \phi, \int_G \psi(g_1, g_1 \cdot x) \mathrm{d}\mu(g_1) \Big\rangle_{L^2(\mathfrak{m})},
\end{aligned}$$

where on the second to last step we used the change of coordinates $\mathbf{g}' = T\mathbf{g}$ and the fact that $\mu^{\mathbb{N}^*}$ is T-invariant. Since $\phi \in L^2(\mathfrak{m})$ is arbitrary, this proves (14). The final statement is straightforward. This proves the lemma. □

Lemma 4.5. *If $\psi \in L^2(\mathfrak{m})$ depends only on the $\mathbb{P}^1$ variable then $\langle \mathfrak{m}, \psi\rangle = \langle \nu, \psi\rangle$. In particular, for such ψ we have $\|\psi\|_{L^p(\mathfrak{m})} = \|\psi\|_{L^p(\nu)}$ for $p = 1$ or 2.*

Proof. From the definition of $\mathfrak{m}$ and the fact that $Z_*\mu^{\mathbb{N}^*} = \nu$ (cf. eq. (11)), it follows that

$$\int_{G^{\mathbb{N}^*}\times\mathbb{P}^1} \psi(x)\,\mathrm{d}\mathfrak{m}(\mathbf{g}, x) = \int_{G^{\mathbb{N}^*}} \psi(Z(\mathbf{g}))\,\mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}) = \int_{\mathbb{P}^1} \psi(x)\,\mathrm{d}\nu(x).$$

This gives us the first assertion. Similar identities for $|\psi|$ and $|\psi|^2$ give the second assertion. □

Consider now the function

$$\varphi : G^{\mathbb{N}^*} \times \mathbb{P}^1 \to \mathbb{R}, \quad \varphi(\mathbf{g}, x) = \log \frac{\|g_1^{-1}\cdot v\|}{\|v\|}, \quad x = [v]. \tag{15}$$

Notice that $\varphi \circ F^j(\mathbf{g}, x) = \varphi(T^j\mathbf{g}, g_j^{-1}\cdots g_1^{-1}\cdot x) = \log \frac{\|g_{j+1}^{-1}g_j^{-1}\cdots g_1^{-1}\cdot v\|}{\|g_j^{-1}\cdots g_1^{-1}\cdot v\|}$. So, the associated Birkhoff sum is

$$\sum_{j=0}^{n-1} \varphi \circ F^j(\mathbf{g}, x) = \log \frac{\|g_n^{-1}\cdots g_1^{-1}\cdot v\|}{\|v\|}, \quad x = [v]. \tag{16}$$

Lemma 4.6. *We have $\langle \mathfrak{m}, \varphi\rangle = -\gamma$.*

Proof. Let $Z(\mathbf{g}) \in \mathbb{P}^1$ be as in Proposition 4.2. Let $W(\mathbf{g}) \in \mathbb{C}^2 \setminus \{0\}$ be a lift of $Z(\mathbf{g})$. By (12) we can chose W so that $g_1^{-1}W(\mathbf{g}) = W(T\mathbf{g})$. Using the definition of $\mathfrak{m}$, the fact that $\mu^{\mathbb{N}^*}$ is invariant by T and equations (11) and (10) we get

$$\begin{aligned}
\langle \mathfrak{m}, \varphi\rangle &= \int_{G^{\mathbb{N}^*}} \varphi(Z(\mathbf{g}))\,\mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}) = \int_{G^{\mathbb{N}^*}} \log\frac{\|g_1^{-1}W(\mathbf{g})\|}{\|W(\mathbf{g})\|}\,\mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}) \\
&= \int_{G^{\mathbb{N}^*}} \log\frac{\|W(T\mathbf{g})\|}{\|g_1 W(T\mathbf{g})\|}\,\mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}) = \int_G\int_{G^{\mathbb{N}^*}} \log\frac{\|W(\mathbf{g}')\|}{\|g_1 W(\mathbf{g}')\|}\,\mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}')\,\mathrm{d}\mu(g_1) \\
&= \int_G\int_{\mathbb{P}^1} \log\frac{\|v\|}{\|g_1\cdot v\|}\mathrm{d}\nu(x)\mathrm{d}\mu(g_1) = -\gamma.
\end{aligned}$$

The lemma follows. □

Proposition 4.7. *Let φ be the function in (15). Then $\widetilde{\varphi} = \varphi - \langle \mathfrak{m}, \varphi \rangle$ belongs to $L^2(\mathfrak{m})$ and satisfies Gordin's condition. Namely,*

$$\sum_{n\geq 0} \|\Lambda^n \widetilde{\varphi}\|^2_{L^2(\mathfrak{m})} < +\infty.$$

Proof. Let us first check that $\widetilde{\varphi} \in L^2(\mathfrak{m})$. Let W be a lift of Z as in the proof of Lemma 4.6. We have

$$\begin{aligned}
\langle \mathfrak{m}, |\varphi|^2 \rangle &= \int_{G^{\mathbb{N}^*}\times\mathbb{P}^1} \left(\log \frac{\|g_1^{-1}\cdot v\|}{\|v\|} \right)^2 \mathrm{d}\mathfrak{m}(\mathbf{g}, x) \\
&= \int_{G^{\mathbb{N}^*}} \left(\log \frac{\|g_1^{-1}\cdot W(\mathbf{g})\|}{\|W(\mathbf{g})\|} \right)^2 \mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}) \\
&\leq \int_G \sup_{x\in\mathbb{P}^1, [v]=x} \left(\log \frac{\|g_1^{-1}\cdot v\|}{\|v\|} \right)^2 \mathrm{d}\mu(g_1) \\
&= \int_G (\log \|g_1\|)^2 \,\mathrm{d}\mu(g_1) < +\infty,
\end{aligned}$$

where we have used that $\|g\| = \|g^{-1}\|$ for $g \in \mathrm{SL}_2(\mathbb{C})$ and the assumption that μ has a finite second moment. So $\varphi \in L^2(\mathfrak{m})$, which implies that $\widetilde{\varphi} \in L^2(\mathfrak{m})$ as claimed.

Let us now prove Gordin's estimate. We begin by noticing that $\widetilde{\varphi}(\mathbf{g}, x)$ depends only on the first entry of $\mathbf{g}$, so we may apply Lemma 4.4. Then

$$\psi(x) := \Lambda\widetilde{\varphi}(x) = \int_G \log \frac{\|v\|}{\|g\cdot v\|} \mathrm{d}\mu(g) + \gamma, \quad x = [v]$$

depends only on the $\mathbb{P}^1$ variable and

$$\Lambda^n \widetilde{\varphi} = \Lambda^{n-1}\psi = (f_\mu^*)^{n-1}\psi.$$

We claim that $\psi \in W^{1,2}$. In order to see that, define $\theta_g(x) := \log \frac{\|g\cdot v\|}{\|v\|}$, $x = [v]$. Then $\psi(x) = \int_G -\theta_g(x)\mathrm{d}\mu(g) + \gamma$. Now, each θ_g is a smooth function and we have from Lemma A.6 in the appendix that $\|\theta_g\|_{W^{1,2}} \lesssim 1 + \log \|g\|$. Then

$$\|\psi\|_{W^{1,2}} \leq \int_G \|\theta_g\|_{W^{1,2}} \mathrm{d}\mu(g) + \gamma \lesssim \int_G (1 + \log \|g\|)\mathrm{d}\mu(g) + \gamma < +\infty,$$

showing that $\psi \in W^{1,2}$.

Now, from Lemma 4.5 and the invariance of $\mathfrak{m}$ we get

$$\langle \nu, \psi \rangle = \langle \mathfrak{m}, \psi \rangle = \langle \mathfrak{m}, \Lambda \widetilde{\varphi} \rangle = \langle \mathfrak{m}, \widetilde{\varphi} \rangle = 0.$$

This can also be checked directly using (10) and the expression of ψ.

From Theorem 2.11 we have that $(f_\mu^*)^{n-1}\psi$ converges to $\langle \nu, \psi \rangle = 0$ in $W^{1,2}$ exponentially fast. Since, also by Theorem 2.11, ν acts continuously on $W^{1,2}$ and $\| |h| \|_{W^{1,2}} \lesssim \|h\|_{W^{1,2}}$ for $h \in W^{1,2}$ (cf. [10, Prop. 4.1]) we get

$$\|\Lambda^n \widetilde{\varphi}\|_{L^1(\nu)} = \|(f_\mu^*)^{n-1}\psi\|_{L^1(\nu)} \lesssim \lambda^n$$

for some constant $0 < \lambda < 1$.

Observe now that, using Lemma 4.4

$$\begin{aligned}\Lambda^n \varphi(x) &= (f_\mu^*)^{n-1} \int_G \varphi(g_1, g_1 \cdot x) \mathrm{d}\mu(g_1) = (f_\mu^*)^{n-1} \int_G \log \frac{\|v\|}{\|g_1 \cdot v\|} \mathrm{d}\mu(g_1) \\ &= \int_{G^n} \log \frac{\|g_2 g_3 \cdots g_n \cdot v\|}{\|g_1 g_2 \cdots g_n \cdot v\|} \mathrm{d}\mu(g_1) \cdots \mathrm{d}\mu(g_n).\end{aligned}$$

Hence $\|\Lambda^n \varphi\|_\infty \leq \int_G \log \|g_1\| \mathrm{d}\mu(g_1)$ for every $n \geq 1$. In particular, there is a constant C such that $\|\Lambda^n \widetilde{\varphi}\|_\infty \leq C$ for every $n \geq 1$.

By interpolating between the spaces $L^\infty(\nu) \subset L^2(\nu) \subset L^1(\nu)$ we conclude that

$$\|\Lambda^n \widetilde{\varphi}\|_{L^2(\mathfrak{m})} = \|\Lambda^n \widetilde{\varphi}\|_{L^2(\nu)} \lesssim \|\Lambda^n \widetilde{\varphi}\|^{1/2}_{L^\infty(\nu)} \|\Lambda^n \widetilde{\varphi}\|^{1/2}_{L^1(\nu)} \lesssim \lambda^{n/2},$$

which gives $\sum_{n \geq 0} \|\Lambda^n \widetilde{\varphi}\|^2_{L^2(\mathfrak{m})} < +\infty$. The proof is complete. □

Lemma 4.8. *The function $\widetilde{\varphi} = \varphi - \langle \mathfrak{m}, \varphi \rangle$ is not a coboundary.*

Proof. Assume by contradiction that $\widetilde{\varphi} = \psi \circ F - \psi$ for some $\psi \in L^2(\mathfrak{m})$. Then $\widetilde{\varphi} \circ F^j = \psi \circ F^{j+1} - \psi \circ F^j$ for $j \geq 0$ and

$$\sum_{j=0}^{n-1} \widetilde{\varphi} \circ F^j = \psi \circ F^n - \psi. \tag{17}$$

The L^2 norm of the right-hand side of (17) is bounded by $\|\psi \circ F^n\|_{L^2(\mathfrak{m})} + \|\psi\|_{L^2(\mathfrak{m})} = 2\|\psi\|_{L^2(\mathfrak{m})}$. In particular, this quantity is bounded independently of n.

We will now show that the L^2 norm of the left-hand side of (17) is unbounded as n goes to infinity. This contradiction will end the proof.

From (16) we have that

$$\sum_{j=0}^{n-1} \widetilde{\varphi} \circ F^j(\mathbf{g}, x) = \log \frac{\|g_n^{-1} \cdots g_1^{-1} \cdot v\|}{\|v\|} + n\gamma, \quad x = [v].$$

Let $W(\mathbf{g})$ be as in the proof of Lemma 4.6. Then,

$$\begin{aligned}
\Big\| \sum_{j=0}^{n-1} \widetilde{\varphi} \circ F^j \Big\|^2_{L^2(\mathfrak{m})} &= \int \Big(\log \frac{\|g_n^{-1} \cdots g_1^{-1} \cdot W(\mathbf{g})\|}{\|W(\mathbf{g})\|} + n\gamma \Big)^2 \, \mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}) \\
&= \int \Big(\log \frac{\|W(T^n\mathbf{g})\|}{\|g_n \cdots g_1 W(T^n \mathbf{g})\|} + n\gamma \Big)^2 \, \mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}) \\
&= \int \Big(-\log \frac{\|g_n \cdots g_1 \cdot W(\mathbf{g}')\|}{\|W(\mathbf{g}')\|} + n\gamma \Big)^2 \, \mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}') \, \mathrm{d}\mu(g_1) \cdots \mathrm{d}\mu(g_n) \\
&= \int \Big(-\log \frac{\|g_n \cdots g_1 \cdot v\|}{\|v\|} + n\gamma \Big)^2 \, \mathrm{d}\nu(x) \, \mathrm{d}\mu(g_1) \cdots \mathrm{d}\mu(g_n).
\end{aligned} \tag{18}$$

Let ζ_n be the random variable $\log \frac{\|g_n \cdots g_1 \cdot v\|}{\|v\|} - n\gamma$ on $G^{\mathbb{N}^*} \times \mathbb{P}^1$, where $x = [v]$ has law ν and the g_i have law μ. Notice that the last integral in (18) is the variance of ζ_n. Hence, in order to prove that (18) is unbounded it is enough to show that the sequence of the distributions of ζ_n is not tight (that is, not relatively compact in the space of probability measures on $\mathbb{R}$, see [5]).

It follows from [6, V.8.5 and V.8.6] that for every fixed $v \in \mathbb{C}^2 \setminus \{0\}$ and any $c > 0$ we have

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \mathbf{P}_x \Big(\Big| \log \frac{\|g_k \cdots g_1 \cdot v\|}{\|v\|} - k\gamma \Big| > c \Big) = 1,$$

where $\mathbf{P}_x$ denotes the probability with respect to $\mu^{\mathbb{N}^*} \otimes \delta_x$. Using Fubini's Theorem and Lebesgue's Dominated Convergence Theorem we get

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \mathbf{P} \Big(\Big| \log \frac{\|g_k \cdots g_1 \cdot v\|}{\|v\|} - k\gamma \Big| > c \Big) = 1,$$

where $\mathbf{P}$ denotes the probability with respect to $\mu^{\mathbb{N}^*} \otimes \nu$. This implies that the sequence of the distributions of ζ_n is not tight, thus finishing the proof. □

We will need the next proposition that shows that for most sequences $\mathbf{g}$ the quantities $\|g_1 \ldots g_n\|$ and $\frac{\|g_1 \cdots g_n \cdot v\|}{\|v\|}$ are comparable for any given $v \in \mathbb{C}^2 \setminus \{0\}$. See [6, III.3.2] and [4, Rmk. 4.26].

Proposition 4.9. *Let μ be a non-elementary probability measure on* $\mathrm{PSL}_2(\mathbb{C})$*. Then for any $\varepsilon > 0$ there exists a $\delta > 0$ such that, for every non-zero $v \in \mathbb{C}^2$*

$$(19)\quad \mu^{\mathbb{N}^*}\left\{\mathbf{g} = (g_1, g_2, \dots) : \delta \leq \frac{\|g_n \cdots g_1 \cdot v\|}{\|g_n \cdots g_1\| \|v\|} \leq 1 \quad \textit{for all } n \geq 1\right\} \geq 1 - \varepsilon.$$

Proof of Theorem 1.3. Consider the dynamical system $F : X \to X$, the measure $\mathfrak{m}$ on X and the function $\widetilde{\varphi}$ introduced above. By Proposition 4.7, Lemmas 4.6 and 4.8 we can apply Gordin-Liverani's Theorem to $\widetilde{\varphi}$ and F. This gives that the sequence of random variables $Z_n = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \widetilde{\varphi} \circ F^j$ converges in distribution to a Gaussian random variable of mean zero and variance $\sigma > 0$.

Let $Y_n := \frac{1}{\sqrt{n}}(\log \frac{\|g_n \cdots g_1 \cdot v\|}{\|v\|} - n\gamma)$ be random variables on $G^{\mathbb{N}^*} \times \mathbb{P}^1$, where $x = [v]$ has law ν and the g_i have law μ. We claim that Z_n and $-Y_n$ have the same distribution. Since the Gaussian law is symmetric around the origin, the convergence of Z_n to the normal distribution will give the convergence of Y_n to the same distribution. In order to do so, we compare the characteristic functions $\chi_{Z_n}(t)$ and $\chi_{Y_n}(t)$ of Z_n and Y_n.

From (16) and Lemma 4.6 we have $Z_n = \frac{1}{\sqrt{n}}(\log \frac{\|g_n^{-1} \cdots g_1^{-1} \cdot v\|}{\|v\|} + n\gamma)$. Then

$$\begin{aligned}
\chi_{Z_n}(t) = \mathbf{E}(e^{itZ_n}) &= \int e^{\frac{it}{\sqrt{n}}(\log \frac{\|g_n^{-1} \cdots g_1^{-1} \cdot v\|}{\|v\|} + n\gamma)} \mathrm{d}\mathfrak{m}(\mathbf{g}, x) \\
&= \int e^{\frac{it}{\sqrt{n}}(\log \frac{\|g_n^{-1} \cdots g_1^{-1} \cdot W(\mathbf{g})\|}{\|W(\mathbf{g})\|} + n\gamma)} \mathrm{d}\mu^{\mathbb{N}^*}(\mathbf{g}) \\
&= \int e^{\frac{it}{\sqrt{n}}(-\log \frac{\|g_n \cdots g_1 \cdot v\|}{\|v\|} + n\gamma)} \mathrm{d}\nu(x)\, \mathrm{d}\mu(g_1) \cdots \mathrm{d}\mu(g_n),
\end{aligned}$$

where in the last step we used the same argument as in (18).

On the other hand

$$\begin{aligned}
\chi_{-Y_n}(t) = \mathbf{E}(e^{-itY_n}) &= \int e^{-itY_n} \mathrm{d}\nu(x)\, \mathrm{d}\mu(g_1) \cdots \mathrm{d}\mu(g_n) \\
&= \int e^{\frac{it}{\sqrt{n}}(-\log \frac{\|g_n \cdots g_1 \cdot v\|}{\|v\|} + n\gamma)} \mathrm{d}\nu(x)\, \mathrm{d}\mu(g_1) \cdots \mathrm{d}\mu(g_n) = \chi_{Z_n}(t).
\end{aligned}$$

As the characteristic function of a random variable determines its distribution, we conclude that Z_n and $-Y_n$ have the same distribution.

By the above remarks, the sequence of random variables Y_n on $G^{\mathbb{N}^*} \times \mathbb{P}^1$ converges in law to $\mathcal{N}(0; \sigma^2)$. From Proposition 4.9 we conclude that for any nonzero $v \in \mathbb{C}^2$ the sequence of random variables $Y_n^v := \frac{1}{\sqrt{n}}\big(\log \frac{\|g_n \cdots g_1 \cdot v\|}{\|v\|} - n\gamma\big)$ on $G^{\mathbb{N}^*}$ converges in law to $\mathcal{N}(0; \sigma^2)$. The proof is now complete. □

5. Regularity of the stationary measure

We now study the regularity of stationary measures. Throughout this section μ will be a non-elementary probability measure on $G = \mathrm{PSL}_2(\mathbb{C})$ such that $\int_G \log \|g\| \, \mathrm{d}\mu(g) < +\infty$ and ν will denote the unique μ-stationary measure. We will also replace μ by μ^{*N} for some $N \geq 1$ when necessary and assume that the norm of f_μ^* acting on $L^2_{(1,0)}$ is strictly less than one (cf. Proposition 2.9). Notice that μ and μ^{*N} have the same stationary measure.

As we will see, the regularity of ν will depend on the moments of μ. We'll need the following notion.

Definition 5.1. *Let $\chi : [0, +\infty) \to [0, +\infty)$ be a non-negative function. The χ-moment of a probability measure μ on $G = \mathrm{PSL}_2(\mathbb{C})$ is the number*

$$\int_G \chi(\log \|g\|) \, \mathrm{d}\mu(g) \in \mathbb{R}_{\geq 0} \cup \{+\infty\}.$$

If the above integral is finite, we say that μ satisfies the χ-moment condition, or equivalently, that μ has a finite χ-moment.

In particular, if $\chi(s) = s^p$ (resp. $\chi(s) = e^{ps}$) for some $p > 0$ we say that μ satisfies the p^{th}-moment condition (resp. an exponential moment condition).

Recall that we are assuming that μ has a finite first moment. In particular, if $\chi(s) \lesssim s$ for s large, then μ satisfies the χ-moment condition. Hence, we'll often assume $\chi(s) \gtrsim s$ for s large. It is also natural to consider χ convex and increasing. In that case, it follows from the sub-additivity of $\log \|g\|$ that if μ has a finite χ-moment then μ^{*n} has a finite χ_n-moment, where $\chi_n(s) := \chi(\frac{1}{n}s)$ for $n \geq 1$. In particular, if μ has a finite p^{th} moment or a finite exponential moment then the same is true for μ^{*n}.

We now introduce a notion of regularity for probability measures following the theory of super-potentials, [11].

Consider the unit ball in $W^{1,2}$

$$\mathbb{B} := \{\varphi \in W^{1,2} : \|\varphi\|_{W^{1,2}} \leq 1\}.$$

Let $\|\cdot\|$ be an auxiliary norm on $W^{1,2}$ and denote by dist the distance induced by $\|\cdot\|$. We will be interested in norms that are weaker than $\|\cdot\|_{W^{1,2}}$.

Definition 5.2. *Let m be a probability measure on $\mathbb{P}^1$. We say that m has a Hölder continuous super-potential with respect to $W^{1,2}$ and* dist *if the restriction of m to $\mathbb{B}$ is a Hölder continuous function with respect to* dist.

The functional on $W^{1,2}$ defined by m is a kind of superpotential of m (compare with [11]). Notice that the above notion doesn't change if we replace $\mathbb{B}$ by any bounded open subset of $W^{1,2}$. In particular, we can replace $\mathbb{B}$ by the unit ball of $W^{1,2}$ with respect to any norm on $W^{1,2}$ that is equivalent to $\|\cdot\|_{W^{1,2}}$.

It will be convenient to work with the following norm and corresponding ball:

$$\|\varphi\|_\nu := |\langle \nu, \varphi\rangle| + \|\partial\varphi\|_{L^2} \quad \text{and} \quad \mathbb{B}_\nu := \{\varphi \in W^{1,2} : \|\varphi\|_\nu \leq 1\}.$$

It follows from Corollary 2.13 that $\|\cdot\|_\nu$ is equivalent to $\|\cdot\|_{W^{1,2}}$. For later use, define also

$$\mathbb{B}^0_\nu := \{\varphi \in \mathbb{B}_\nu : \langle \nu, \varphi\rangle = 0\}$$

and

$$\Lambda := f^*_\mu : W^{1,2} \longrightarrow W^{1,2}.$$

Since ν is stationary, $\mathbb{B}_\nu$ and $\mathbb{B}^0_\nu$ are invariant by Λ. Moreover, $\frac{1}{2}(\Lambda - \mathrm{id})$ maps $\mathbb{B}_\nu$ to $\mathbb{B}^0_\nu$ and by Proposition 2.9 there is a constant $\delta > 1$ such that $\mathbb{B}^0_\nu$ is invariant by $\widetilde{\Lambda} := \delta\Lambda$.

Denote by $\mathbb{D}(a,r)$ the disc of center a and of radius r in $\mathbb{P}^1$. Fix $0 < \epsilon_0 < \frac{1}{2}$. For $0 < \epsilon < \epsilon_0$, $0 < r \leq 1$ and $a \in \mathbb{P}^1$, set

$$u^\epsilon_{a,r}(z) := \max\Big(-\log \frac{\mathrm{dist}_{\mathbb{P}^1}(z,a)}{2r}, 0\Big)^{\frac{1}{2}-\epsilon}.$$

Then $u^\epsilon_{a,r}$ belongs to $W^{1,2}$ and it is supported by $\mathbb{D}(a,2r)$. One can also check that the $u^\epsilon_{a,r}$ belong to a bounded subset of $W^{1,2}$, see [23, Ex. 2].

Define

$$\mathcal{V}_\epsilon(r) := \max_{a\in\mathbb{P}^1} \|u^\epsilon_{a,r}\| < +\infty.$$

Proposition 5.3. *Assume that m has a Hölder continuous super-potential with respect to $W^{1,2}$ and* dist. *Then there are constants $c > 0$ and $\alpha > 0$ independent of ϵ such that for $0 < r \leq 1$*

$$m(\mathbb{D}(a,r)) \leq c\,\mathcal{V}_\epsilon(r)^\alpha.$$

Proof. Notice that $u^\epsilon_{a,r} \geq d$ on $\mathbb{D}(a,r)$ where $d := \sqrt{\log 2}$. Since m has a Hölder continuous super-potential, we have

$$m(\mathbb{D}(a,r)) \leq d^{-1}|\langle m, u^\epsilon_{a,r}\rangle| = d^{-1}|m(u^\epsilon_{a,r}) - m(0)| \leq c\|u^\epsilon_{a,r} - 0\|^\alpha \leq c\,\mathcal{V}_\epsilon(r)^\alpha,$$

for some positive constants c and α. This ends the proof. □

Proposition 5.4. *Let μ be a non-elementary probability measure on $G = \mathrm{PSL}_2(\mathbb{C})$ having a finite first moment. Assume that $\Lambda = f_\mu^* : W^{1,2} \to W^{1,2}$ is bounded with respect to the norm $\|\cdot\|$ with $\|\cdot\|_{L^1} \leq c\|\cdot\|$ and $\|\cdot\| \leq c\|\cdot\|_{W^{1,2}}$ for some constant $c > 0$. Then ν has a Hölder continuous super-potential with respect to $W^{1,2}$ and the distance* dist.

We will need the following lemma.

Lemma 5.5. *Let K be a metric space. Let $A \geq 1$ be a constant and let $F_n : K \to K$ be a sequence of Lipschitz maps on K such that $\|F_n\|_{\mathrm{Lip}} \leq A^n$ for every n. Then for any bounded Hölder continuous function $\vartheta : K \to \mathbb{C}$ and any $\delta > 1$, the function*

$$\sum_{n\geq 0} \delta^{-n} (\vartheta \circ F_n)$$

is also Hölder continuous. If furthermore K has finite diameter, then the assumption on the boundedness of ϑ is superfluous.

Proof. In the particular case where $F_n = F^n$ for some Lipschitz map F this is Lemma 1.19 in [12]. It can be easily checked that the proof given there extends to the present setting. □

Proof of Proposition 5.4. We apply Lemma 5.5 to $K := \mathbb{B}_\nu$, $F_n := \widetilde{\Lambda}^n \circ \big(\frac{1}{2}(\Lambda - \mathrm{id})\big)$ and ϑ the restriction of ω_{FS} to $\mathbb{B}_\nu$. Recall that $\widetilde{\Lambda} = \delta\Lambda$ for some $\delta > 1$ and that both $\widetilde{\Lambda}$ and $\frac{1}{2}(\Lambda - \mathrm{id})$ preserve $\mathbb{B}_\nu$.

Since $\|\cdot\|_{L^1} \lesssim \|\cdot\|$ by hypothesis, we have $|\vartheta(\varphi)| \leq \|\varphi\|_{L^1} \lesssim \|\varphi\|$ for $\varphi \in \mathbb{B}_\nu$. Hence ϑ is a Lipschitz function on K. Moreover, since $\Lambda = f_\mu^*$ is bounded with respect to $\|\cdot\|$ by assumption, the maps $\widetilde{\Lambda}$ and $\frac{1}{2}(\Lambda - \mathrm{id})$ are also Lipschitz on K. So we have $\|F_n\|_{\mathrm{Lip}} \leq A^n$ for some constant $A \geq 1$. Notice also that the assumption $\|\cdot\| \lesssim \|\cdot\|_{W^{1,2}}$ implies that $\mathbb{B}_\nu$ has finite diameter with respect to dist.

We now have, for $\varphi \in K$

$$\begin{aligned} 2\,\delta^{-n}\, \vartheta \circ F_n(\varphi) &= \vartheta \circ \Lambda^n \circ (\Lambda - \mathrm{id})(\varphi) = \langle \omega_{\mathrm{FS}}, \Lambda^{n+1}(\varphi) - \Lambda^n(\varphi) \rangle \\ &= \langle (f_\mu^{n+1})_* \omega_{\mathrm{FS}} - (f_\mu^n)_* \omega_{\mathrm{FS}}, \varphi \rangle. \end{aligned}$$

It follows from Theorem 2.11 that $\lim_{n\to\infty} (f_\mu^n)_*(\omega_{\mathrm{FS}}) = \nu$. Therefore,

$$2 \sum_{n\geq 0} \delta^{-n} (\vartheta \circ F_n) = -\omega_{\mathrm{FS}} + \lim_{n\to\infty} (f_\mu^n)_*(\omega_{\mathrm{FS}}) = -\omega_{\mathrm{FS}} + \nu$$

By Lemma 5.5 we get that $-\omega_{\mathrm{FS}}+\nu$ defines a Hölder continuous function on $\mathbb{B}_\nu$. It follows that ν defines a Hölder continuous function on $\mathbb{B}_\nu$. The proof is complete. $\square$

We now apply the above results for some choices of the norm $\|\cdot\|$. Consider a Young's function $\Phi : [0,\infty) \to [0,\infty)$, that is, a convex increasing function such that

$$\lim_{t\to 0}\frac{\Phi(t)}{t}=0 \quad \text{and} \quad \lim_{t\to\infty}\frac{\Phi(t)}{t}=\infty.$$

We also assume that $e^{-t^2}\Phi(t)$ is bounded. Consider the Luxemburg norm (or gauge norm)

$$\|\varphi\|_\Phi := \inf\Big\{A\in[0,\infty) : \int_{\mathbb{P}^1}\Phi\Big(\frac{|\varphi|}{A}\Big)\omega_{\mathrm{FS}}\leq 1\Big\}$$

and the associated Birnbaum-Orlicz space $L_\Phi(\mathbb{P}^1)$ consisting of measurable functions on $\mathbb{P}^1$ having finite $\|\cdot\|_\Phi$ norm, see [21].

The distance associated with this norm is denoted by dist_Φ. Since we are assuming that $e^{-t^2}\Phi(t)$ is bounded we have, by Moser-Trudinger's estimate (Proposition 2.5), that $\|\cdot\|_\Phi \lesssim \|\cdot\|_{W^{1,2}}$ and $W^{1,2}\subset L_\Phi(\mathbb{P}^1)$.

Define also the function $\eta_\Phi : \mathbb{R}_{\geq 0}\to[0,+\infty]$ by

$$\eta_\Phi(s) := \sup_{\varphi\in W^{1,2}\setminus\{0\}}\frac{\|\varphi\|_{e^s\Phi}}{\|\varphi\|_\Phi}.$$

Theorem 5.6. *Let μ be a non-elementary probability measure on $G=\mathrm{PSL}_2(\mathbb{C})$ having a finite χ-moment and let ν be the associated stationary measure. Assume that $\eta_\Phi(4s)\lesssim\chi(s)+1$ for $s\geq 0$. Then ν has a Hölder continuous super-potential with respect to $W^{1,2}$ and the distance* dist_Φ.

Proof. It is well-known that $\|\cdot\|_{L^1}\lesssim\|\cdot\|_\Phi$ (see [21]) and we have seen that $\|\cdot\|_\Phi\lesssim\|\cdot\|_{W^{1,2}}$. So by Proposition 5.4, it is enough to check that $\Lambda : W^{1,2}\to W^{1,2}$ is bounded with respect to the norm $\|\cdot\|_\Phi$. We have for $\varphi\in W^{1,2}$

$$\|\Lambda\varphi\|_\Phi=\Big\|\int_G g^*\varphi\,\mathrm{d}\mu(g)\Big\|_\Phi\leq\int_G\|g^*\varphi\|_\Phi\,\mathrm{d}\mu(g).$$

Since, by assumption, μ has a finite χ-moment, it is enough to show that

$$\|g^*\varphi\|_\Phi\lesssim(\chi(\log\|g\|)+1)\|\varphi\|_\Phi \quad \text{for every } g\in G.$$

Set $s := \log\|g\|$. Recall from Lemma 2.4 that $g_*\omega_{\mathrm{FS}} \leq \|g\|^4\omega_{\mathrm{FS}}$. Then, for any $A > 0$, we have

$$\int_{\mathbb{P}^1} \Phi\Big(\frac{|g^*\varphi|}{A}\Big)\omega_{\mathrm{FS}} = \int_{\mathbb{P}^1} \Phi\Big(\frac{|\varphi|}{A}\Big) g_*\omega_{\mathrm{FS}}$$
$$\leq \int_{\mathbb{P}^1} \|g\|^4 \Phi\Big(\frac{|\varphi|}{A}\Big)\omega_{\mathrm{FS}} = \int_{\mathbb{P}^1} e^{4s}\Phi\Big(\frac{|\varphi|}{A}\Big)\omega_{\mathrm{FS}}.$$

Hence

$$\|g^*\varphi\|_\Phi \leq \|\varphi\|_{e^{4s}\Phi} \leq \eta_\Phi(4s)\|\varphi\|_\Phi \lesssim (\chi(s)+1)\|\varphi\|_\Phi = (\chi(\log\|g\|)+1)\|\varphi\|_\Phi.$$

The theorem follows. □

We can now use Theorem 5.6 to obtain explicit regularity properties of ν in terms of the moments of μ. The idea is the following: assuming that μ has a finite χ-moment, find a suitable Young's function Φ so that $\eta_\Phi(4s) \lesssim \chi(s)+1$. Theorem 5.6 will then give that ν has a Hölder continuous super-potential with respect to dist_Φ. Together with Proposition 5.3 this will give an estimate for the mass of ν on small discs.

The following corollaries illustrate two extremal cases where our method applies. The same idea can be extended to other moment conditions on μ.

Corollary 5.7. *Let μ be a non-elementary measure on $\mathrm{PSL}_2(\mathbb{C})$ with finite exponential moment and let ν be the associated stationary measure. Then there is a number $q \in [1,\infty)$ such that ν has a Hölder continuous super-potential with respect to $W^{1,2}$ and the L^q-norm. In particular, there are constants $\theta > 0, A > 0, c > 0$ and $\alpha > 0$ such that*

$$\int_{\mathbb{P}^1} e^{\theta|\varphi|^2}\,\mathrm{d}\nu \leq A \quad \textit{and} \quad \nu(\mathbb{D}(a,r)) \leq cr^\alpha$$

for every $\varphi \in W^{1,2}$ with $\|\varphi\|_{W^{1,2}} \leq 1$, $a \in \mathbb{P}^1$ and $0 < r \leq 1$.

Proof. Fix a number q large enough and choose $\Phi(t) = t^q$. It can be easily seen that $\|\cdot\|_\Phi$ is the L^q-norm and that $\eta_\Phi(s) = e^{s/q}$. By assumption, μ has a finite χ-moment where $\chi(s) = e^{ps}$ for some $p > 0$. Since q is large, we have $\eta_\Phi(4s) \lesssim \chi(s)$. By Theorem 5.6, ν has a Hölder continuous super-potential with respect to $W^{1,2}$ and the norm L^q.

Let $\varphi \in W^{1,2}$ such that $\|\varphi\|_{W^{1,2}} \leq 1$. For $N \geq 1$, define $\varphi_N := \min(|\varphi|, N)$. Then φ_N belongs to a bounded subset of $W^{1,2}$ (cf. [10, Prop. 4.1]). Define also $\psi_N := \varphi_{N+1} - \varphi_N$. Notice that $0 \le \psi_N \le 1$, $\psi_N \equiv 0$ on $\{|\varphi| \le N\}$, and $\psi_N \equiv 1$ on $\{|\varphi| \ge N+1\}$. Therefore

$$\nu\{N \le |\varphi| \le N+1\} \le \nu(\psi_{N-1}) \lesssim \|\psi_{N-1}\|_{L^q}^{\beta} \lesssim \text{area}\{|\varphi| \ge N-1\}^{\beta/q},$$

where $\beta > 0$ is the Hölder exponent of the functional defined by ν and the area is with respect to ω_{FS}.

From Proposition 2.5 it follows that $\text{area}\{|\varphi| \ge N-1\} \lesssim e^{-\alpha' N^2}$ for some $\alpha' > 0$, so

$$\nu\{N \le |\varphi| \le N+1\} \lesssim e^{-\alpha'' N^2} \text{ for some } \alpha'' > 0. \tag{20}$$

Now, for $\theta > 0$ small enough, the first estimate in the corollary follows after cutting the integral $\int_{\mathbb{P}^1} e^{\theta|\varphi|^2} \mathrm{d}\nu$ along the subsets $\{N \le |\varphi| \le N+1\}$ and using (20).

It is not difficult to see that for our choice of Φ we have $\mathcal{V}_\epsilon(r) \lesssim r^\gamma$ for every $0 < \gamma < 2/q$. Then, the second estimate in the corollary follows by applying Proposition 5.3. □

Remark 5.8. *A measure m satisfying $m(\mathbb{D}(a,r)) \le cr^\alpha$ for some constants $c, \alpha > 0$ is often called Hölder regular. The Hölder regularity of ν under an exponential moment condition is an old result due to Guivarc'h, see [6, VI.4] and [16].*

Corollary 5.9. *Let μ be a non-elementary measure on $\mathrm{PSL}_2(\mathbb{C})$ with finite first moment and let ν be the associated stationary measure. Then there are constants $c > 0$ and $\alpha > 0$ such that*

$$\nu(\mathbb{D}(a,r)) \le c|\log r|^{-\alpha}$$

for every $a \in \mathbb{P}^1$ and $0 < r \le 1$.

Proof. As above, we will apply Theorem 5.6 for a suitable function Φ. Observe that the function $t \mapsto e^{-t^{-3}}$ is convex and increasing on some interval $[0, t_0]$ in $\mathbb{R}_{\ge 0}$. We extend it to a convex increasing function Φ on $\mathbb{R}_{\ge 0}$ such that $\Phi(t) = e^{t^2}$ for t large enough.

Claim 1. We have $\eta_\Phi(4s) \lesssim s + 1 = \chi(s) + 1$ for $s \ge 0$.

It is enough to prove that $\eta_\Phi(4s) \le ks$ for some constant $k > 0$ and s large enough. Let $\varphi \in W^{1,2}$ be such that $\int_{\mathbb{P}^1} \Phi(|\varphi|)\omega_{\text{FS}} = 1$. We need to show that

$$\int_{\mathbb{P}^1} \Phi\Big(\frac{|\varphi|}{k's}\Big)\omega_{\text{FS}} \le e^{-4s} \tag{21}$$

for some constant $k' > 0$ and s large enough.

Observe that $\Phi(t) \le e^{t^2} + \text{const}$ on $\mathbb{R}_{\ge 0}$. We have

$$\int_{|\varphi|>2\sqrt{s}} \Phi\Big(\frac{|\varphi|}{s}\Big)\omega_{\mathrm{FS}} \lesssim \int_{|\varphi|>2\sqrt{s}} e^{|\varphi|^2/s^2}\omega_{\mathrm{FS}}$$
$$\le e^{-s}\int_{|\varphi|>2\sqrt{s}} e^{|\varphi|^2}\omega_{\mathrm{FS}} \le e^{-s}\int_{\mathbb{P}^1}\Phi(|\varphi|)\omega_{\mathrm{FS}} = e^{-s}.$$

On the other hand, we have

$$\int_{|\varphi|\le 2\sqrt{s}} \Phi\Big(\frac{|\varphi|}{s}\Big)\omega_{\mathrm{FS}} = \int_{|\varphi|\le 2\sqrt{s}} e^{-(|\varphi|/s)^{-3}}\omega_{\mathrm{FS}} \le \int_{\mathbb{P}^1} e^{-s}\omega_{\mathrm{FS}} = e^{-s}.$$

This gives (21) for $k' = 5$, ending the proof of the claim.

By Theorem 5.6 and the claim, ν has a Hölder continuous super-potential with respect to $W^{1,2}$ and the distance dist_Φ.

To finish the proof, we now need to estimate the function $\mathcal{V}_\epsilon(r)$ appearing in Proposition 5.3. It is enough to consider a fixed value of ϵ. Take $\epsilon := 1/4$ and set $u := u_{a,\epsilon}$.

Claim 2. We have $\|u\|_\Phi \le |\log r|^{-1/8}$ for r small enough.

Set $A := |\log r|^{-1/8}$. By the definition of $\|\cdot\|_\Phi$, we need to check that

$$\int_{\mathbb{P}^1}\Phi\Big(\frac{|u|}{A}\Big)\omega_{\mathrm{FS}} < 1.$$

In order to simplify the notation, assume that $a = 0$ and denote by $|z|$ the distance between z and 0. Then $u = |\log(|z|/2r)|^{1/4}$ on $|z| < 2r$ and zero elsewhere. Observe that $|u| > A$ if and only if $|z| < 2re^{-|\log r|^{-1/2}}$. Moreover, we have $|u| \le |\log|z||^{1/4}$. Thus, using that ω_{FS} is comparable with $i dz \wedge d\overline{z}$ near 0, we have for $s := -\log|z|$ and r small

$$\int_{|u|>A}\Phi\Big(\frac{|u|}{A}\Big)\omega_{\mathrm{FS}} = \int_{|z|<2re^{-|\log r|^{-1/2}}}\Phi\Big(\frac{|u|}{A}\Big)\omega_{\mathrm{FS}} \lesssim \int_{|z|<3r} e^{|u|^2/A^2}\omega_{\mathrm{FS}}$$
$$\lesssim \int_{|\log r|-3}^{\infty} e^{A^{-2}s^{1/2}}e^{-2s}ds$$
$$\lesssim \int_{|\log r|-3}^{\infty} e^{-2s+2s^{3/4}}ds \lesssim \int_{|\log r|-3}^{\infty} e^{-s}ds = O(r).$$

Recall that $\Phi(0)=0$ and that u is supported by $\mathbb{D}(a,2r)$. Then

$$\int_{|u|\leq A}\Phi\Big(\frac{|u|}{A}\Big)\omega_{\rm FS}\lesssim \operatorname{area}(\mathbb{D}(a,2r))=O(r^2).$$

The claim follows.

The last claim gives that $\mathcal{V}_\epsilon(r)\lesssim|\log r|^{-\gamma}$ for $\epsilon=1/4$, r small and a suitable constant $\gamma>0$. The corollary then follows from Proposition 5.3. □

Remark 5.10. *A similar type of regularity under a finite p^{th} moment condition was obtained by Benoist-Quint in [3]. This is a crucial ingredient in their proof of the Central Limit Theorem. We note that the exponents appearing in this section can be made explicit. We chose not do so in order to keep the paper less technical.*

Appendix A. Elementary sets and auxiliary lemmas

We present in this appendix some results used in the text. A number of them are probably known to experts.

Let us first recall the classification of elements of $\operatorname{Aut}(\mathbb{P}^1)$. In this appendix we shall denote by the same symbol g an element of $\operatorname{Aut}(\mathbb{P}^1)$, its corresponding matrix in $\mathrm{SL}_2(\mathbb{C})$ and its class in $\mathrm{PSL}_2(\mathbb{C})$. This should not cause any confusion.

Recall that an element g of $\operatorname{Aut}(\mathbb{P}^1)$ different from the identity is conjugated to either $z\mapsto z+1$ or $z\mapsto\lambda z$ for some $\lambda\in\mathbb{C}\setminus\{0,1\}$. In the former case, g is called *parabolic* and in the latter, g is called *elliptic* if $|\lambda|=1$ or *loxodromic* if $|\lambda|\neq 1$. A parabolic automorphism has a single fixed point that attracts every point of $\mathbb{P}^1$. An elliptic automorphism has two different neutral fixed points and a loxodromic automorphism g admits two fixed points a and b such that $g^n(z)\to a$ and $g^{-n}(z)\to b$ as n tends to infinity, for any $z\in\mathbb{P}^1\setminus\{a,b\}$. In terms of the trace of the corresponding matrices, $g\neq\mathrm{Id}$ is parabolic if $\mathrm{Tr}^2 g=4$, elliptic if $\mathrm{Tr}^2 g\in[0,4)$ and loxodromic if $\mathrm{Tr}^2 g\notin[0,4]$.

Now let R be a subset of $\mathrm{PSL}_2(\mathbb{C})$. For $n\geq 1$ denote

$$R^n:=\{g_n\cdots g_1: g_i\in R\}\quad\text{ and }\quad S^n:=\{gh^{-1}: g,h\in R^n\}.$$

Recall that R is non-elementary if its support does not preserve a finite subset of $\mathbb{P}^1$ and if the semi-group generated by R is not relatively compact, see Definition 2.6 and Remark 2.7.

Lemma A.1. *Let R be a non-elementary subset of* $\mathrm{PSL}_2(\mathbb{C})$*. Then there exist integers $N_1 \geq 1$ and $N_2 \geq 1$ such that R^{N_1} contains a loxodromic element and S^{N_2} contains a non-elliptic element.*

Proof. The first assertion is well known. First, we extend the action of $\mathrm{PSL}_2(\mathbb{C})$ to the 3-dimensional hyperbolic space $\mathbb{H}^3$ (see Remark 2.7-(iii)). Then, the results from [7, Chapter 6] imply that the semi-group generated by R contains a loxodromic element g_0. This gives the first assertion.

We now prove the second assertion. Since R^{N_1} is non-elementary, we can find another element h_0 in R^{N_1} whose fixed point set is different from that of g_0. If $|\mathrm{Fix}(g_0) \cap \mathrm{Fix}(h_0)| = 1$, Lemma A.2 below implies that $g_0 h_0 g_0^{-1} h_0^{-1} \in S^{2N_1}$ is parabolic. If $\mathrm{Fix}(g_0) \cap \mathrm{Fix}(h_0) = \varnothing$, Lemmas A.3, A.4 and A.5 below show that there is an $N_3 \geq 1$ such that $g_0^{N_3}(h_0^{-1})^{N_3} \in S^{N_1 N_3}$ is loxodromic. This proves the second assertion and concludes the proof of the lemma. □

Lemma A.2. *If g, h are two non-trivial elements in* $\mathrm{PSL}_2(\mathbb{C})$*, g has 2 fixed points on $\mathbb{P}^1$ and $|\mathrm{Fix}(g) \cap \mathrm{Fix}(h)| = 1$, then $ghg^{-1}h^{-1}$ is parabolic.*

Proof. See [19, p. 12]. □

Lemma A.3. *Let $g, h \in \mathrm{PSL}_2(\mathbb{C})$. If g is loxodromic and h is elliptic then there is an $N \geq 1$ such that $g^N h^N$ is loxodromic.*

Proof. We can assume that the fixed points of g are 0 and ∞ and $g = \left(\begin{smallmatrix} t & 0 \\ 0 & t^{-1} \end{smallmatrix}\right)$, where $|t| > 1$. Since h is elliptic the set $\{h^n : n \geq 1\}$ is relatively compact in $\mathrm{PSL}_2(\mathbb{C})$. Hence, there exists a subsequence h^{n_k}, converging to some elliptic $r = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}_2(\mathbb{C})$. After replacing h^{n_k} by h^{2n_k} and r by r^2 if necessary we may assume that $a \neq 0$. Denoting by a_n, b_n, c_n, d_n the entries of h^n we have that $a_{n_k} \to a$ and $d_{n_k} \to d$. Then $|\mathrm{Tr}^2(g^{n_k} h^{n_k})| = |t^{n_k} a_{n_k} + t^{-n_k} d_{n_k}|^2 \to \infty$. If we choose N so that $|\mathrm{Tr}^2(g^N h^N)| > 4$ then $g^N h^N$ is loxodromic. □

Lemma A.4. *Let $g, h \in \mathrm{PSL}_2(\mathbb{C})$. If g is loxodromic and h is parabolic then there exists an $N \geq 1$ such that $g^N h^N$ is loxodromic.*

Proof. We can write $g = A\left(\begin{smallmatrix} t & 0 \\ 0 & 1/t \end{smallmatrix}\right)A^{-1}, h = B\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)B^{-1}$, where $|t| > 1$ and $A, B \in \mathrm{PSL}_2(\mathbb{C})$. Define a_i, b_i, c_i, d_i by $A^{-1}B = \left(\begin{smallmatrix} a_1 & b_1 \\ c_1 & d_1 \end{smallmatrix}\right)$ and $B^{-1}A = \left(\begin{smallmatrix} a_2 & b_2 \\ c_2 & d_2 \end{smallmatrix}\right)$. Then

$$\begin{aligned} \mathrm{Tr}(g^n h^n) &= \mathrm{Tr}\left(A \begin{pmatrix} t^n & 0 \\ 0 & 1/t^n \end{pmatrix} A^{-1} B \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} B^{-1} \right) \\ &= \mathrm{Tr}\left(\left(\begin{pmatrix} t^n & 0 \\ 0 & 1/t^n \end{pmatrix} A^{-1} B \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} B^{-1} A \right) \right. \end{aligned}$$

$$= \mathrm{Tr}\left(\begin{pmatrix} t^n & 0 \\ 0 & 1/t^n \end{pmatrix}\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}\right)$$
$$= a_1c_2nt^n + a_1a_2t^n + b_1c_2t^n + c_1b_2/t^n + c_1d_2n/t^n + d_1d_2/t^n$$
$$= a_1c_2nt^n + t^n + c_1d_2n/t^n + 1/t^n,$$

which shows that $|\mathrm{Tr}^2(g^nh^n)|$ is unbounded as $n \to \infty$. Hence g^nh^n is loxodromic for n large enough. □

Lemma A.5. *If g, h are both loxodromic and* $\mathrm{Fix}(g) \cap \mathrm{Fix}(h) = \varnothing$, *then g^Nh^N is loxodromic for some $N \geq 1$.*

Proof. Write $g = A\left(\begin{smallmatrix} t & 0 \\ 0 & 1/t \end{smallmatrix}\right)A^{-1}, h = B\left(\begin{smallmatrix} s & 0 \\ 0 & 1/s \end{smallmatrix}\right)B^{-1}$, where $|t| > 1, |s| > 1$ and $A, B \in \mathrm{PSL}_2(\mathbb{C})$. Define a_i, b_i, c_i, d_i by $A^{-1}B = \left(\begin{smallmatrix} a_1 & b_1 \\ c_1 & d_1 \end{smallmatrix}\right)$ and $B^{-1}A = \left(\begin{smallmatrix} a_2 & b_2 \\ c_2 & d_2 \end{smallmatrix}\right)$. Then

$$\mathrm{Tr}(g^nh^n) = \mathrm{Tr}\left(A\begin{pmatrix} t^n & 0 \\ 0 & 1/t^n \end{pmatrix}A^{-1}B\begin{pmatrix} s^n & 0 \\ 0 & 1/s^n \end{pmatrix}B^{-1}\right)$$
$$= \mathrm{Tr}\left(\begin{pmatrix} t^n & 0 \\ 0 & 1/t^n \end{pmatrix}A^{-1}B\begin{pmatrix} s^n & 0 \\ 0 & 1/s^n \end{pmatrix}B^{-1}A\right)$$
$$= \mathrm{Tr}\left(\begin{pmatrix} t^n & 0 \\ 0 & 1/t^n \end{pmatrix}\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}\begin{pmatrix} s^n & 0 \\ 0 & 1/s^n \end{pmatrix}\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}\right)$$
$$= a_1a_2t^ns^n + b_1c_2t^n/s^n + c_1b_2s^n/t^n + d_1d_2/(t^ns^n).$$

Suppose for contradiction that for every $n \geq 1$, g^nh^n is not loxodromic. Then $|\mathrm{Tr}(g^nh^n)|$ is a bounded sequence. It follows that $a_1a_2 = 0$. Without loss of generality, assume $a_1 = 0$. We get $A^{-1}B = \left(\begin{smallmatrix} 0 & b_1 \\ -1/b_1 & d_1 \end{smallmatrix}\right)$ and $B^{-1}A = (A^{-1}B)^{-1} = \left(\begin{smallmatrix} d_1 & -b_1 \\ 1/b_1 & 0 \end{smallmatrix}\right)$, so

$$h = B\begin{pmatrix} s & 0 \\ 0 & 1/s \end{pmatrix}B^{-1} = A\begin{pmatrix} 0 & b_1 \\ -1/b_1 & d_1 \end{pmatrix}\begin{pmatrix} s & 0 \\ 0 & 1/s \end{pmatrix}\begin{pmatrix} d_1 & -b_1 \\ 1/b_1 & 0 \end{pmatrix}A^{-1}$$
$$= A\begin{pmatrix} * & 0 \\ * & * \end{pmatrix}A^{-1}.$$

This implies that h and g have the same fixed point $A([0:1])$, contradicting the hypothesis $\mathrm{Fix}(g) \cap \mathrm{Fix}(h) = \varnothing$. This proves the lemma. □

For $g \in \mathrm{PSL}_2(\mathbb{C})$ let $\theta_g(x) := \log \frac{\|g\cdot v\|}{\|v\|}$, $x = [v]$. Then θ_g is a smooth function on $\mathbb{P}^1$ and $\|\theta_g\|_\infty = \log\|g\|$. The following estimate was used in Section 4.

Lemma A.6. *We have* $\|\theta_g\|_{W^{1,2}} \lesssim 1 + \log \|g\|$.

Proof. Since $\|\theta_g\|_\infty = \log \|g\|$ it follows that $\|\theta_g\|_{L^1} \leq \log \|g\|$. So, from Proposition 2.2 we only need to estimate $\|\partial\theta_g\|_{L^2}$.

Set $\omega_g := i\partial\theta_g \wedge \overline{\partial\theta_g}$ so that $\|\partial\theta_g\|_{L^2}^2 = \int_{\mathbb{P}^1} \omega_g$. By Cartan's decomposition we can write $g = k'ak$ where $k, k' \in \mathrm{SU}(2)$ and $a \in \mathrm{SL}_2(\mathbb{C})$ is diagonal with positive eigenvalues. Since k' and k preserve the euclidean norm we have

$$\theta_g(x) = \log \frac{\|k'ak \cdot v\|}{\|v\|} = \log \frac{\|ak \cdot v\|}{\|v\|} = \log \frac{\|ak \cdot v\|}{\|k \cdot v\|} = \theta_a(k \cdot x),$$

that is $\theta_g = k^*\theta_a$. Hence $\omega_g = k^*\omega_a$ and since $\mathrm{SU}(2)$ is compact we have that $\omega_g \sim \omega_a$. This allows us to assume that g is of the form $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{smallmatrix}\right)$, for some $\lambda \geq 1$.

Let $z = [z : 1]$ be the standard affine coordinate in $\mathbb{P}^1 \setminus \{\infty\}$. In this coordinate we have $g(z) = \lambda^2 z$, so

$$\theta_g(z) = \frac{1}{2} \log \frac{\lambda^4|z|^2 + 1}{|z|^2 + 1} = \frac{1}{2} \log(\lambda^4|z|^2 + 1) - \frac{1}{2} \log(|z|^2 + 1).$$

Hence

$$\omega_g = i\partial\theta_g \wedge \overline{\partial\theta_g} = \frac{(\lambda^4 - 1)^2|z|^2}{4(\lambda^4|z|^2 + 1)^2(|z|^2 + 1)^2} idz \wedge d\overline{z}.$$

Then, by Lemma A.7 below we get $\int_{\mathbb{P}^1} \omega_g \lesssim \log \lambda^4 = 4 \log \|g\|$. Hence $\|\partial\theta_g\|_{L^2} \lesssim (\log \|g\|)^{1/2}$. This, together with the above estimate for $\|\theta_g\|_{L^1}$, implies the lemma. □

Lemma A.7. *Let* $\beta > 1$ *and denote by* z *the standard affine coordinate in* $\mathbb{C} \subset \mathbb{P}^1$. *Then*

$$\int_{\mathbb{C}} \frac{(\beta - 1)^2|z|^2}{(\beta|z|^2 + 1)^2(|z|^2 + 1)^2} idz \wedge d\overline{z} \leq 2\pi \frac{\beta - 1}{\beta + 1} \log \beta.$$

Proof. Multiplying the integral on left hand side by $\frac{\beta+1}{\beta-1}$ gives

$$\begin{aligned} &\int \frac{(\beta^2 - 1)|z|^2}{(\beta|z|^2 + 1)^2(|z|^2 + 1)^2} idz \wedge d\overline{z} \leq \int \frac{(\beta^2 - 1)|z|^2}{(\beta^2|z|^4 + 1)(|z|^4 + 1)} idz \wedge d\overline{z} \\ &= \iint \frac{(\beta^2 - 1)r^2}{(\beta^2 r^4 + 1)(r^4 + 1)} 2r\, dr\, d\theta = \iint \frac{\beta^2 - 1}{2(\beta^2 t + 1)(t + 1)}\, dt\, d\theta \end{aligned}$$

$$= \pi \int \Big(\frac{\beta^2}{\beta^2 t + 1} - \frac{1}{t+1}\Big)\, dt = \pi \left[\log \frac{\beta^2 t + 1}{t+1}\right]_0^\infty = 2\pi \log \beta,$$

giving the desired inequality. □

Acknowledgements

We thank the referee for valuable comments on the first version of this manuscript. This paper was partially written during the visit of the first author to the Institute of Mathematical Sciences and Department of Mathematics in Chinese University of Hong Kong. He would like to thank these organisations and Prof. Conan Leung for their very warm hospitality. This work was supported by the NUS grants C-146-000-047-001, AcRF Tier 1 R-146-000-248-114 and R-146-000-259-114.

References

[1] Alan F. Beardon. *The geometry of discrete groups*, volume 91 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1983.

[2] Bachir Bekka and Yves Guivarc'h. A spectral gap property for random walks under unitary representations. *Geom. Dedicata*, **118**:141–155, 2006.

[3] Yves Benoist and Jean-François Quint. Central limit theorem for linear groups. *Ann. Probab.*, **44**(2):1308–1340, 2016.

[4] Yves Benoist and Jean-François Quint. *Random walks on reductive groups*, volume 62 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer, Cham, 2016.

[5] Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[6] Philippe Bougerol and Jean Lacroix. *Products of random matrices with applications to Schrödinger operators*, volume 8 of *Progress in Probability and Statistics*. Birkhäuser Boston, Inc., Boston, MA, 1985.

[7] Tushar Das, David Simmons, and Mariusz Urbański. *Geometry and dynamics in Gromov hyperbolic metric spaces*, volume 218 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2017. With an emphasis on non-proper settings.

[8] Jean-Pierre Demailly. *Complex Analytic and Differential Geometry.* http://www-fourier.ujf-grenoble.fr/~demailly/.

[9] Tien-Cuong Dinh, Lucas Kaufmann, and Hao Wu. Dynamics of holomorphic correspondences on Riemann surfaces. *Internat. J. Math.*, **31**(5):2050036, 21, 2020.

[10] Tien-Cuong Dinh and Nessim Sibony. Decay of correlations and the central limit theorem for meromorphic maps. *Comm. Pure Appl. Math.*, **59**(5):754–768, 2006.

[11] Tien-Cuong Dinh and Nessim Sibony. Super-potentials of positive closed currents, intersection theory and dynamics. *Acta Math.*, **203**(1):1–82, 2009.

[12] Tien-Cuong Dinh and Nessim Sibony. Dynamics in several complex variables: endomorphisms of projective spaces and polynomial-like mappings. In *Holomorphic dynamical systems*, volume 1998 of *Lecture Notes in Math.*, pages 165–294. Springer, Berlin, 2010.

[13] I. Ya. Goldsheid and G. A. Margulis. Lyapunov exponents of a product of random matrices. *Uspekhi Mat. Nauk*, **44**(5(269)):13–60, 1989.

[14] M. I. Gordin. The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR*, **188**:739–741, 1969.

[15] Y. Guivarc'h and A. Raugi. Frontière de Furstenberg, propriétés de contraction et théorèmes de convergence. *Z. Wahrsch. Verw. Gebiete*, **69**(2):187–242, 1985.

[16] Yves Guivarc'h. Produits de matrices aléatoires et applications aux propriétés géométriques des sous-groupes du groupe linéaire. *Ergodic Theory Dynam. Systems*, **10**(3):483–512, 1990.

[17] Émile Le Page. Théorèmes limites pour les produits de matrices aléatoires. In *Probability measures on groups (Oberwolfach, 1981)*, volume 928 of *Lecture Notes in Math.*, pages 258–303. Springer, Berlin-New York, 1982.

[18] Carlangelo Liverani. Central limit theorem for deterministic systems. In *International Conference on Dynamical Systems (Montevideo, 1995)*, volume 362 of *Pitman Res. Notes Math. Ser.*, pages 56–75. Longman, Harlow, 1996.

[19] Bernard Maskit. *Kleinian groups, volume 287 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 1988.

[20] Jürgen Moser. A sharp form of an inequality by N. Trudinger. *Indiana Univ. Math. J.*, **20**:1077–1092, 1970/71.

[21] M. M. Rao and Z. D. Ren. *Theory of Orlicz spaces*, volume 146 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 1991.

[22] Hans Triebel. *Interpolation theory, function spaces, differential operators*, volume 18 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam-New York, 1978.

[23] Gabriel Vigny. Dirichlet-like space and capacity in complex analysis in several variables. *J. Funct. Anal.*, **252**(1):247–277, 2007.

Tien-Cuong Dinh
Department of Mathematics
National University of Singapore
10, Lower Kent Ridge Road
Singapore 119076
E-mail: matdtc@nus.edu.sg

Lucas Kaufmann
Department of Mathematics
National University of Singapore
10, Lower Kent Ridge Road
Singapore 119076
Current address:
Center for Complex Geometry
Institute for Basic Science
55 Expo-ro, Yuseong-gu
Daejeon 34126
South Korea
E-mail: lucaskaufmann@ibs.re.kr; lucaskaufmann.math@gmail.com

Hao Wu
Department of Mathematics
National University of Singapore
10, Lower Kent Ridge Road
Singapore 119076
E-mail: matwu@nus.edu.sg