

Statistics in conjugacy classes in free groups

George Kenison¹ · Richard Sharp² 

Received: 17 August 2017 / Accepted: 4 February 2018 / Published online: 9 February 2018
© The Author(s) 2018. This article is an open access publication

Abstract In this paper, we establish statistical results for a convex co-compact action of a free group on a $CAT(-1)$ space where we restrict to a non-trivial conjugacy class in the group. In particular, we obtain a central limit theorem where the variance is twice the variance that appears when we do not make this restriction.

Keywords Free group · Conjugacy class · Convex co-compact · Central limit theorem · Subshift of finite type · Measure of maximal entropy · Thermodynamic formalism · Transfer operators · Generating functions

Mathematics Subject Classification 20F65 · 20F67 · 20F69 · 37C30 · 37D35 · 60B10

1 Introduction and results

Let Γ be a free group on $p \geq 2$ generators acting convex co-compactly on a $CAT(-1)$ space (X, d) (i.e the quotient of the intersection of X and the convex hull of the limit set of Γ is compact). There has been considerable work in trying to understand the statistics of such an action. For example, the following result (a particular case of the Švarc–Milnor lemma) is well-known. Fix a free generating set $\mathcal{A} = \{a_1, \dots, a_p\}$ and let $|\cdot|$ denote word length on Γ with respect to \mathcal{A} . Then, for an arbitrary base point $o \in X$, there exist constants $C_1, C_2 > 0$ such that

$$C_1|x| \leq d(o, xo) \leq C_2|x| \tag{1.1}$$

✉ Richard Sharp
r.j.sharp@warwick.ac.uk

George Kenison
george.kenison@bristol.ac.uk

¹ School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK

² Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK

for all $x \in \Gamma$. Thus $|x|$ and $d(o, xo)$ are comparable quantities and it is natural to ask if more precise estimates hold, at least typically or on average.

One such result is the following. Write $\Gamma_n := \{x \in \Gamma : |x| = n\}$. Then the averages

$$\frac{1}{\#\Gamma_n} \sum_{x \in \Gamma_n} \frac{d(o, xo)}{n} \tag{1.2}$$

converge to some $\lambda > 0$, as $n \rightarrow \infty$ [14, 15], where the positivity follows immediately from the lower bound in (1.1). (See Remark 1.3(i) below for a further discussion.) Furthermore, subject to a mild non-degeneracy condition, namely that the set $\{d(o, xo) - \lambda|x| : x \in \Gamma\}$ is not bounded, the distribution of $(d(o, xo) - \lambda n)/\sqrt{n}$ with respect to the normalised counting measure on Γ_n converges to a normal distribution $N(0, \sigma^2)$, as $n \rightarrow \infty$, for some finite $\sigma^2 > 0$.

In this paper, we shall consider the corresponding questions when we restrict our group elements to a non-trivial conjugacy class. Let \mathfrak{C} be a non-trivial conjugacy class in Γ and let $k = \min\{|x| : x \in \mathfrak{C}\}$. Let $\mathfrak{C}_n = \{x \in \mathfrak{C} : |x| = n\}$ and note that \mathfrak{C}_n is non-empty if and only if $n = k + 2m$, $m \in \mathbb{Z}^+$.

Theorem 1.1 *We have*

$$\lim_{m \rightarrow \infty} \frac{1}{\#\mathfrak{C}_{k+2m}} \sum_{x \in \mathfrak{C}_{k+2m}} \frac{d(o, xo)}{k + 2m} = \lambda.$$

Subject to an additional condition, we also have a central limit theorem.

Theorem 1.2 *Suppose that the set $\{d(o, xo) - \lambda|x| : x \in \Gamma\}$ is not bounded. Then the distribution of $(d(o, xo) - \lambda(k + 2m))/\sqrt{k + 2m}$ with respect to normalised counting measure on \mathfrak{C}_{k+2m} converges to a normal distribution $N(0, 2\sigma^2)$, as $n \rightarrow \infty$, i.e.*

$$\begin{aligned} & \lim_{m \rightarrow \infty} \frac{1}{\#\mathfrak{C}_{k+2m}} \# \left\{ x \in \mathfrak{C}_{k+2m} : \frac{d(o, xo) - \lambda(k + 2m)}{\sqrt{k + 2m}} < y \right\} \\ &= \frac{1}{2\sqrt{\pi}\sigma} \int_{-\infty}^y e^{-t^2/4\sigma^2} dt. \end{aligned}$$

A noteworthy feature of this result is that the variance is twice the variance that appears in the unrestricted case. Theorems 1.1 and 1.2 will follow from more general results proved below.

Remark 1.3 (i) The existence of the limit

$$\lim_{n \rightarrow \infty} \frac{1}{\#\Gamma_n} \sum_{x \in \Gamma_n} \frac{d(o, xo)}{n}$$

follows from Proposition 8 of [14]. The results in [14] are proved for co-compact groups of isometries of real hyperbolic space but go over to co-compact groups of isometries of CAT(−1) spaces by the arguments of [15]. Some explanation may be in order here. The paper [15] is written in the context of compact manifolds (possibly with boundary) with variable negative curvature. In our situation, X corresponds to the universal cover of M and Γ to the fundamental group, acting as isometries on X . Given a point $p \in M$ and a non-identity element $x \in \Gamma$ (thought of as $\pi_1(M, p)$), the number $l(x)$ is defined to be the length of the shortest geodesic arc from p to itself in the homotopy class determined by x . This can be reinterpreted as the number $d(o, xo)$, where o is a lift of p to X , returning us to our

original setting. Although the results of [15] are stated for manifolds of negative curvature, the arguments used there, in particular the key Lemma 1, only require that X be a CAT(−1) space. A consequence of this lemma is that $d(o, xo)$ can be written as the Birkhoff sum of a Hölder continuous function on an associated subshift of finite type (Proposition 3 of [15]); this shows that $d(o, xo)$ satisfies the assumption (A1) in the next section. (Of course, the assumption (A2) below is trivially satisfied.)

The existence of a limit in (1.2) continues to hold if Γ is a word hyperbolic group following an observation of Calegari and Fujiwara [1], using a result of Coornaert [3].

(ii) The number $\lambda > 0$ may also be characterised in the following way. Let Σ be the space of infinite reduced words on $\mathcal{A} \cup \mathcal{A}^{-1}$ and let μ_0 be the measure of maximal entropy for the shift map $\sigma : \Sigma \rightarrow \Sigma$ —these objects are defined in Sect. 2. Then, for μ_0 -a.e. $(x_i)_{i=0}^\infty \in \Sigma$,

$$\lim_{n \rightarrow \infty} \frac{d(o, x_0x_1 \cdots x_{n-1}o)}{n} = \lambda.$$

This follows from the representation of $d(o, xo)$ as a Birkhoff sum of a Hölder continuous function on $\Sigma \cup \Gamma$ and the ergodic theorem. (See, for example, Lemma 4.4 and Corollary 4.5 of [16].)

(iii) The fact that the variance in Theorem 1.2 is independent of the choice of conjugacy class is a consequence of the hypothesis that $\{d(o, xo) - \lambda|x| : x \in \Gamma\}$ is unbounded, which is a condition on the behaviour of the displacement function $d(o, xo)$ over the whole group Γ . (The same may be said of the assumption (A3) in the next section.)

(iv) It is interesting to have examples where the above hypothesis that $S = \{d(o, xo) - \lambda|x| : x \in \Gamma\}$ is unbounded holds. The hypothesis may be reformulated as follows. For $x \in \Gamma$, define homogeneous length functions associated to $d(o, xo)$ and $|x|$:

$$\ell(x) := \lim_{n \rightarrow \infty} \frac{d(o, x^n o)}{n} \quad \text{and} \quad \|x\| := \lim_{n \rightarrow \infty} \frac{|x^n|}{n}.$$

Then $\ell(x)$ and $\|x\|$ are positive and depend only on the conjugacy class of x , so we may write $\ell(\mathcal{C})$ and $\|\mathcal{C}\|$. Furthermore, $\ell(\mathcal{C})$ is the length of the closed geodesic on the quotient $\Gamma \backslash X$ in the free homotopy class determined by \mathcal{C} . If S were bounded then we would have $\ell(\mathcal{C}) = \lambda\|\mathcal{C}\|$ for all non-trivial conjugacy classes \mathcal{C} . In particular, the length spectrum of $\Gamma \backslash X$, i.e. the set of lengths of closed geodesics, would be contained in the set $\lambda\mathbb{Z}$. However, it is known that the length spectrum is not contained in a discrete subgroup of the reals when X is the real hyperbolic space \mathbb{H}^k , $k \geq 2$ or when X is a simply connected surface of pinched variable negative curvature [4], so the hypothesis holds in these cases. More generally, though the hypothesis may fail in particular cases, it will typically hold. For example, if X is a metric tree with quotient metric graph $\Gamma \backslash X$ then to ensure the hypothesis is satisfied, one only requires that $\Gamma \backslash X$ has two closed paths whose lengths have irrational ratio.

(v) The above results still hold if $d(o, xo)$ is replaced by a Hölder length function $L(x)$ as defined in [7].

We end the introduction by outlining the contents of the paper. In Sect. 2 we discuss the relationship between free groups and subshifts of finite type and state more general versions of Theorems 1.1 and 1.2. In Sect. 3 we introduce the transfer operators that we use for our analysis and discuss some of their properties. In Sect. 4 we introduce a generating

function $\eta_{\mathcal{C}}(z, s)$ related to the conjugacy class \mathcal{C} , where z and s are complex variables. In the geometric setting considered above, this generating function takes the form

$$\eta_{\mathcal{C}}(z, s) = \sum_{m=0}^{\infty} z^{k+2m} \sum_{x \in \mathcal{C}_{k+2m}} e^{sd(o,xo)}.$$

In particular, the variable z is associated to the word length and the variable s to the geometric length (or to a more general weighting below). This generating function is perhaps the main new innovation of the paper, though its analysis is inspired by work on a somewhat similar function in [9]. This allows us to prove our first main result. We conclude the paper in Sect. 5 by proving a central limit theorem over a non-trivial conjugacy class. The results in this paper form part of the first author’s Ph.D. thesis at the University of Warwick.

2 Free groups and subshifts

As above, let Γ be a free group with free generating set $\mathcal{A} = \{a_1, \dots, a_p\}$, $p \geq 2$. Write $\mathcal{A}^{-1} = \{a_1^{-1}, \dots, a_p^{-1}\}$. A word $x_0 \dots x_{n-1}$, with letters $x_k \in \mathcal{A} \cup \mathcal{A}^{-1}$, is said to be *reduced* if $x_{k+1} \neq x_k^{-1}$ for each $k \in \{0, \dots, n - 2\}$ and *cyclically reduced* if, in addition, $x_0 \neq x_{n-1}^{-1}$. Every non-identity element $x \in \Gamma$ has a unique representation as a reduced word $x = x_0x_1 \dots x_{n-1}$ and we define the *word length* $|x|$ of x , by $|x| = n$. We associate to the identity element the empty word and set $|1| = 0$. Let $\Gamma_n = \{x \in \Gamma : |x| = n\}$.

Let \mathcal{C} be a non-trivial conjugacy class in Γ and let $k = \inf\{|x| : x \in \mathcal{C}\} > 0$. The set of elements with shortest word length in the conjugacy class is precisely the set of elements with cyclically reduced word representations. In fact, if $g = g_1 \dots g_k \in \mathcal{C}$ is cyclically reduced then all cyclically reduced words in \mathcal{C} are given by cyclic permutations of the letters in $g_1 \dots g_k$. Let $\mathcal{C}_n = \{x \in \mathcal{C} : |x| = n\}$ and note that \mathcal{C}_n is non-empty if and only if $n = k + 2m$. If $x \in \mathcal{C}_{k+2m}$ then its reduced word representation is of the form $w_m^{-1} \dots w_1^{-1} g_1 \dots g_k w_1 \dots w_m$, for some cyclically reduced $g = g_1 \dots g_k \in \mathcal{C}_k$ and $w = w_1 \dots w_m \in \Gamma_m$ with $w_1 \neq g_1, g_k^{-1}$. Hence it is convenient to introduce the notation $\Gamma_m(g) = \{w \in \Gamma_m : w_1 \neq g_1, g_k^{-1}\}$. A simple calculation shows that the number of elements in \mathcal{C}_{k+2m} is given by $\#\mathcal{C}_{k+2m} = (2p - 2)(2p - 1)^{m-1} \#\mathcal{C}_k$.

We associate to the free group Γ a dynamical system called a *subshift of finite type*. This subshift of finite type is formed from the space of infinite reduced words (with the obvious definition) adjoined to the elements of Γ together with the dynamics given by the action of the shift map. It will be convenient to describe this space by means of a transition matrix. Define a $p \times p$ matrix A , with rows and columns indexed by $\mathcal{A} \cup \mathcal{A}^{-1}$, by $A(a, b) = 0$ if $b = a^{-1}$ and $A(a, b) = 1$ otherwise. We then define

$$\Sigma = \left\{ (x_n)_{n=0}^{\infty} \in (\mathcal{A} \cup \mathcal{A}^{-1})^{\mathbb{Z}^+} : A(x_n, x_{n+1}) = 1, \forall n \in \mathbb{Z}^+ \right\}.$$

The shift map $\sigma : \Sigma \rightarrow \Sigma$ is defined by $(\sigma(x_n)_{n=0}^{\infty}) = (x_{n+1})_{n=0}^{\infty}$. We give $\mathcal{A} \cup \mathcal{A}^{-1}$ the discrete topology, $(\mathcal{A} \cup \mathcal{A}^{-1})^{\mathbb{Z}^+}$ the product topology and Σ the subspace topology; then σ is continuous. Since the matrix A is *aperiodic* (i.e. there exists $n \geq 1$ such that for each pair of indices (s, t) , $A^n(s, t) > 0$), $\sigma : \Sigma \rightarrow \Sigma$ is *mixing* (i.e. for every pair of non-empty open sets $U, V \subset \Sigma$ there is an $n \in \mathbb{Z}^+$ such that $\sigma^{-k}U \cap V \neq \emptyset$ for $k \geq n$).

We augment Σ by defining $\Sigma^* = \Sigma \cup \Gamma$, where the elements of Γ are identified with finite reduced words in the obvious way. The shift map naturally extends to a map $\sigma : \Sigma^* \rightarrow \Sigma^*$,

where, for the finite reduced word $x_0x_1 \cdots x_{n-1} \in \Gamma$, we set $\sigma(x_0x_1 \cdots x_{n-1}) = x_1 \cdots x_{n-1}$; and for the empty word $\sigma 1 = 1$. It is sometimes useful to think of an element of Γ as an infinite sequence ending in an infinite string of 1s.

We endow Σ^* with the following metric, consistent with the topology on Σ . Fix $0 < \theta < 1$ then let $d_\theta(x, y) = 0$ and, for $x \neq y$, let $d_\theta(x, y) = \theta^k$, where $k = \min\{n \in \mathbb{Z}^+ : x_n \neq y_n\}$. For a finite word $x = x_0x_1 \cdots x_{m-1} \in \Gamma_m$ we take $x_n = 1$ (the empty symbol) for each $n \geq m$. Then $\sigma : \Sigma^* \rightarrow \Sigma^*$ is continuous and Γ is a dense subset of Σ^* .

We will write \mathcal{M} for the set of σ -invariant Borel probability measures on Σ . For $\nu \in \mathcal{M}$, we write $h(\nu)$ for its entropy. We define the pressure of a continuous function $f : \Sigma \rightarrow \mathbb{R}$ by

$$P(f) := \sup_{\nu \in \mathcal{M}} \left(h(\nu) + \int f \, d\nu \right).$$

If f is Hölder continuous then the supremum is attained at a unique $\mu_f \in \mathcal{M}$, called the equilibrium state of f . (If $f : \Sigma^* \rightarrow \mathbb{R}$ then we write $P(f) := P(f|_\Sigma$.) The equilibrium state of zero μ_0 is also called the measure of maximal entropy and $P(0)$ is equal to the topological entropy h of $\sigma : \Sigma \rightarrow \Sigma$. It is easy to calculate that $h = \log(2p - 1)$ (the logarithm of the largest eigenvalue of A) and that μ_0 is characterised by

$$\mu_0([w]) = (2p)^{-1} (2p - 1)^{-(n-1)},$$

where, for a reduced word $w = w_0w_1 \cdots w_{n-1} \in \Gamma_n$, $[w]$ is the associated cylinder set $[w] \subset \Sigma^*$ by $[w] = \{(x_j)_{j=0}^\infty : x_j = w_j, j = 0, \dots, n - 1\}$. (Technically, this defines μ_0 as a measure on Σ^* with support equal to Σ .)

Two Hölder continuous functions $f, g : \Sigma^* \rightarrow \mathbb{R}$ are *cohomologous* if there exists a continuous function $u : \Sigma^* \rightarrow \mathbb{R}$ such that $f = g + u \circ \sigma - u$. Two Hölder continuous functions have the same equilibrium state if and only if they differ by the sum of a coboundary and a constant. A function $f : \Sigma^* \rightarrow \mathbb{R}$ is *locally constant* if there exists $n \geq 1$ such that for all pairs $x, y \in \Sigma$ with $x_k = y_k$ for $0 \leq k \leq n$, $f(x) = f(y)$. Locally constant functions are automatically Hölder continuous for any choice of Hölder exponent. For a function $f : \Sigma^* \rightarrow \mathbb{R}$ we denote by $f^n(x)$ the Birkhoff sum

$$f^n(x) := f(x) + f(\sigma x) + \cdots + f(\sigma^{n-1}x).$$

We have the following result [12, 18].

Proposition 2.1 *If $f : \Sigma \rightarrow \mathbb{R}$ is Hölder continuous then, for $t \in \mathbb{R}$, $t \mapsto P(tf)$ is real analytic,*

$$\left. \frac{dP(tf)}{dt} \right|_{t=0} = \int f \, d\mu_0$$

and

$$\left. \frac{d^2P(tf)}{dt^2} \right|_{t=0} = \sigma_f^2 := \lim_{n \rightarrow \infty} \frac{1}{n} \int \left(f^n(x) - n \int f \, d\mu_0 \right)^2 \, d\mu_0.$$

Furthermore, $\sigma_f^2 = 0$ if and only if f is cohomologous to a constant.

For convenience, in the work that follows we shall interchangeably refer to elements $x \in \Gamma$ and the associated element of the sequence space $x \in \Sigma^*$. We now state the technical result from which Theorem 1.1 follows. We consider functions $F : \Gamma \rightarrow \mathbb{R}$ which satisfy the following two assumptions.

- (A1) There exists a Hölder continuous function $f : \Sigma^* \rightarrow \mathbb{R}$ so that $F(x) = f^n(x)$ for each $x \in \Gamma_n$ with $n \geq 0$, and
- (A2) $F(x) = F(x^{-1})$.

We will prove the following.

Theorem 2.2 *Suppose that $F : \Gamma \rightarrow \mathbb{R}$ satisfies assumptions (A1) and (A2). There exists $\bar{F} \in \mathbb{R}$ such that*

$$\lim_{m \rightarrow \infty} \frac{1}{\#\mathcal{C}_{k+2m}} \sum_{x \in \mathcal{C}_{k+2m}} \frac{F(x)}{k + 2m} = \bar{F}.$$

Furthermore, $\bar{F} = \int f \, d\mu_0$.

We remark that, without the restriction to a conjugacy class, the analogous result

$$\lim_{n \rightarrow \infty} \frac{1}{\#\Gamma_n} \sum_{x \in \Gamma_n} \frac{F(x)}{n} = \bar{F}$$

holds subject only to (A1). This follows from the analysis in [14] or from a large deviations argument following the ideas of Kifer [10] as employed in [13].

We also establish a central limit theorem for the group elements in Γ restricted to a non-trivial conjugacy class. In addition to assumptions (A1) and (A2), we require a third assumption.

- (A3) $F(\cdot) - \bar{F}|\cdot|$ is unbounded as a function from Γ to \mathbb{R} .

Lemma 2.3 *Let F and f be as in (A1). Then $F(\cdot) - \bar{F}|\cdot|$ is bounded if and only if $f|_\Sigma$ is cohomologous to a constant.*

Proof For simplicity, we will write $f|_\Sigma = f$. If $F(\cdot) - \bar{F}|\cdot|$ is bounded then $\{f^n(x) - n \int f \, d\mu_0 : x \in \Gamma_n, n \geq 1\}$ is a bounded set. Since f is Hölder continuous, this implies that

$$\left\{ f^n(x) - n \int f \, d\mu_0 : x \in \Sigma, n \geq 1 \right\}$$

is also bounded. In particular, $(f^n - n \int f \, d\mu_0)^2/n$ converges uniformly to zero and it is easy to deduce that $\sigma_f^2 = 0$. Therefore, by Proposition 2.1, f is cohomologous to a constant.

On the other hand, if f is cohomologous to a constant then, again by Hölder continuity, $\{F(x) - \bar{F}|x| : x \in \Gamma\} = \{f^n(x) - n \int f \, d\mu_0 : x \in \Gamma_n, n \geq 1\}$ is bounded. □

It is a well-known result that if $f : \Sigma \rightarrow \mathbb{R}$ is not cohomologous to a constant then the process $f \circ \sigma^n, n \geq 1$, satisfies a central limit theorem with respect to μ_0 with variance σ_f^2 , i.e., that $(f^n - n \int f \, d\mu_0) / \sqrt{n}$ converges in distribution to a normal random variable with mean zero and variance $\sigma_f^2 > 0$ or, explicitly, that for $a \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} \mu_0 \left\{ x \in \Sigma : \left(f^n(x) - n \int f \, d\mu_0 \right) / \sqrt{n} \leq a \right\} = \frac{1}{\sqrt{2\pi\sigma_f^2}} \int_{-\infty}^a e^{-u^2/2\sigma_f^2} \, du$$

[2, 18]. Furthermore, analogues of this hold for the periodic points of $\sigma : \Sigma \rightarrow \Sigma$ [2] and, by adapting the proof, for pre-images of a given point. This gives a central limit theorem for F over Γ_n (without the assumption (A2)). Particular cases of this have appeared in articles

by Rivin [17] for homomorphisms, and Horsham and Sharp [7] (see also [6]) for quasi-morphisms. Calegari and Fujiwara [1] prove a central limit theorem for quasimorphisms on Gromov hyperbolic groups, but have more restrictions on the regularity of the quasimorphism. Restricting to a non-trivial conjugacy class, we have the following theorem.

Theorem 2.4 *Suppose that $F : \Gamma \rightarrow \mathbb{R}$ satisfies assumptions (A1), (A2) and (A3). Then the sequence*

$$\frac{1}{\#\mathfrak{C}_{k+2m}} \# \left\{ x \in \mathfrak{C}_{k+2m} : (F(x) - (k + 2m)\bar{F})/\sqrt{k + 2m} \leq a \right\}$$

converges to the distribution function of a normal random variable with mean 0 and positive variance $2\sigma_f^2$.

We note the limiting distribution function is independent of the choice of non-trivial conjugacy class. Further, it is interesting that the variance in Theorem 2.4 is twice the variance when we do not restrict elements $x \in \Gamma$ to a non-trivial conjugacy class.

Proof of Theorems 1.1 and 1.2 As in the introduction, let the free group Γ act convex co-compactly on a CAT(−1) space (X, d) . Then it was shown in [15] that $F(x) := d(o, x_o)$ satisfies (A1). (In fact, the result in [15] is stated when X is a simply connected manifold with bounded negative curvatures but the proof only requires the CAT(−1) property.) Assumption (A2) is clearly satisfied. Therefore, Theorem 1.1 follows from Theorem 2.2. Furthermore, the additional assumption on $d(o, x_o)$ in Theorem 1.2 matches (A3) and so Theorem 1.2 also follows. □

3 Transfer operators

In this section we recall results from the theory of transfer operators that will be used to deduce Theorems 2.2 and 2.4. Let $\mathcal{F}_\theta(\Sigma, \mathbb{C})$ denote the space of d_θ -Lipschitz functions $f : \Sigma \rightarrow \mathbb{C}$. This is a Banach space with respect to the norm $\| \cdot \|_\theta = \| \cdot \|_\infty + | \cdot |_\theta$, where

$$|f|_\theta := \sup_{x \neq y} \frac{|f(x) - f(y)|}{d_\theta(x, y)}.$$

Any Hölder continuous function becomes Lipschitz by changing the choice of θ (i.e. if f has Hölder exponent α with respect to d_θ then $f \in \mathcal{F}_{\theta^\alpha}(\Sigma, \mathbb{C})$), so there is no loss of generality in restricting to these spaces. Given $g \in \mathcal{F}_\theta(\Sigma, \mathbb{C})$, the *transfer operator* $L_g : \mathcal{F}_\theta(\Sigma, \mathbb{C}) \rightarrow \mathcal{F}_\theta(\Sigma, \mathbb{C})$ is defined pointwise by

$$L_g \omega(x) = \sum_{\sigma y=x} e^{g(y)} \omega(y).$$

We have the following standard result [12, 18].

Proposition 3.1 (Ruelle–Perron–Frobenius Theorem) *Suppose that $g \in \mathcal{F}_\theta(\Sigma, \mathbb{C})$ is real-valued. Then $L_g : \mathcal{F}_\theta(\Sigma, \mathbb{C}) \rightarrow \mathcal{F}_\theta(\Sigma, \mathbb{C})$ has a simple eigenvalue equal to $e^{P(g)}$, associated strictly positive eigenfunction ψ and eigenmeasure ν (i.e. $L_g \psi = e^{P(g)} \psi$ and $L_g^* \nu = e^{P(g)} \nu$), normalised so that ν is a probability measure and $\int \psi \, d\nu = 1$. Furthermore, the rest of the spectrum of L_g is contained in a disk of radius strictly smaller than $e^{P(g)}$.*

The equilibrium state μ_g is given by $d\mu_g = \psi dv$. We say that g is *normalised* if $L_g 1 = 1$ (which in particular implies $P(g) = 0$). If we replace g by $g' = g - P(g) + u - u \circ \sigma$ where $u = \log \psi$ then g' is normalised and g and g' have the same equilibrium state.

Suppose that $f, g \in \mathcal{F}_\theta(\Sigma, \mathbb{C})$ are real-valued functions. We consider small perturbations of the operator L_g of the form L_{g+sf} for values of $s \in \mathbb{C}$ in a neighbourhood of the origin. Since $e^{P(g)}$ is a simple isolated eigenvalue of L_g , for small perturbations of s close to the origin this eigenvalue persists so that the operator L_{g+sf} has a simple eigenvalue $\beta(s)$ and corresponding eigenfunction ψ_s that vary analytically with s and satisfy $\beta(0) = e^{P(g)}$ and $\psi_0 = \psi$ [8]. Furthermore, by the upper semi-continuity of the spectral radius, there exists $\varepsilon > 0$ such that, for s close to the origin, the remainder of the spectrum of L_{g+sf} lies in a disk of radius $e^{P(g)-\varepsilon}$. We extend the definition of pressure by setting $e^{P(g+sf)} = \beta(s)$.

We find it useful to consider $\sigma : \Sigma^* \rightarrow \Sigma^*$ as a subshift of finite type and will use the previous notation and concepts introduced for Σ in this setting. We modify the definition of the transfer operator $L_{sf} : \mathcal{F}_\theta(\Sigma^*, \mathbb{C}) \rightarrow \mathcal{F}_\theta(\Sigma^*, \mathbb{C})$ as follows:

$$L_{sf}\omega(x) = \sum_{\substack{\sigma y=x \\ y \neq 1}} e^{sf(y)}\omega(y).$$

Here 1 denotes the identity element in Γ , considered as an infinite word $(1, 1, \dots)$. We note the transfer operator we use differs from the usual definition by excluding the preimage $y = 1$ from the summation over the set $\{y \in \Sigma^* : \sigma y = x\}$; however, the definition of this transfer operator agrees with our previous definition for each $x \neq 1$. Following Lemma 2 of [14], $L_{sf} : \mathcal{F}_\theta(\Sigma^*, \mathbb{C}) \rightarrow \mathcal{F}_\theta(\Sigma^*, \mathbb{C})$ has the same isolated eigenvalues as $L_{sf} : \mathcal{F}_\theta(\Sigma \cup \{1\}, \mathbb{C}) \rightarrow \mathcal{F}_\theta(\Sigma \cup \{1\}, \mathbb{C})$. Since the modified definition of L_{sf} excludes the eigenvalue $e^{sf(1)}$ associated to the eigenfunction $\chi_{\{1\}}$ (the indicator function of the set $\{1\}$), $L_{sf} : \mathcal{F}_\theta(\Sigma^*, \mathbb{C}) \rightarrow \mathcal{F}_\theta(\Sigma^*, \mathbb{C})$ therefore has the same isolated eigenvalues as $L_{sf} : \mathcal{F}_\theta(\Sigma, \mathbb{C}) \rightarrow \mathcal{F}_\theta(\Sigma, \mathbb{C})$. Furthermore, again by Lemma 2 of [14], $L_{sf} : \mathcal{F}_\theta(\Sigma^*, \mathbb{C}) \rightarrow \mathcal{F}_\theta(\Sigma^*, \mathbb{C})$ is quasi-compact with essential spectral radius at most $\theta e^{P(\text{Re}(s)f)}$, and so it suffices to consider the spectral theory of L_{sf} on $\mathcal{F}_\theta(\Sigma, \mathbb{C})$.

4 Proof of Theorem 2.2

In this section, we will prove Theorem 2.2. We introduce a generating function $\eta_{\mathcal{C}}(z, s)$ on two complex variables given by

$$\eta_{\mathcal{C}}(z, s) = \sum_{m=0}^{\infty} z^{k+2m} \sum_{x \in \mathcal{C}_{k+2m}} e^{sF(x)} = \sum_{m=0}^{\infty} z^{k+2m} \sum_{g \in \mathcal{C}_k} \sum_{w \in \Gamma_m(g)} e^{sf^{k+2m}(w^{-1}gw)}$$

(wherever the series converges). We prove the theorem by studying the asymptotic behaviour, as $m \rightarrow \infty$, of the coefficient of z^{k+2m} in the power series

$$\frac{\partial}{\partial s} \eta_{\mathcal{C}}(z, s) \Big|_{s=0} = \sum_{m=0}^{\infty} z^{k+2m} \sum_{x \in \mathcal{C}_{k+2m}} F(x).$$

We will find the following bound useful in the proof of Theorem 2.2.

Lemma 4.1 *Suppose that $f \in \mathcal{F}_\theta(\Sigma^*, \mathbb{C})$, $g \in \mathcal{C}_k$ and $w \in \Gamma_m(g)$ then there exists a constant $K > 0$, independent of m , such that*

$$|f^{k+2m}(w^{-1}gw) - f^m(w) - f^k(g) - f^m(w^{-1})| \leq K.$$

Proof We have $f^{k+2m}(w^{-1}gw) = f^m(w^{-1}gw) + f^k(gw) + f^m(w)$. Thus

$$\begin{aligned} &|f^{k+2m}(w^{-1}gw) - f^m(w) - f^k(g) - f^m(w^{-1})| \\ &\leq |f^m(w^{-1}gw) - f^m(w^{-1})| + |f^k(gw) - f^k(g)| \leq \frac{2|f|_{\theta\theta}}{1-\theta} \end{aligned}$$

and we are done. □

By Lemma 4.1,

$$\exp(s f^{k+2m}(w^{-1}gw)) = \exp(s(f^m(w) + f^k(g) + f^m(w^{-1}))) + s\kappa_w + \xi_w(s)$$

where $\kappa_w = f^{k+2m}(w^{-1}gw) - f^m(w) - f^k(g) - f^m(w^{-1})$ is uniformly bounded for $w \in \Gamma$ (by Lemma 4.1) and $\xi_w(s) = s^2\zeta_w(s)$, with $\zeta_w(s)$ an entire function. By this approximation and assumption (A2), we have

$$\eta_{\mathcal{C}}(z, s) = \sum_{m=0}^{\infty} z^{k+2m} \sum_{g \in \mathcal{C}_k} \sum_{w \in \Gamma_m(g)} e^{s(f^k(g)+2f^m(w))} + \delta(z, s)$$

where

$$\delta(z, s) = \sum_{m=0}^{\infty} z^{k+2m} \sum_{g \in \mathcal{C}_k} \sum_{w \in \Gamma_m(g)} (s\kappa_w + \xi_w(s)).$$

Let $\chi_g : \Sigma^* \rightarrow \mathbb{R}$ be the locally constant function given by

$$\chi_g((w_n)_{n=0}^{\infty}) = \begin{cases} 0 & \text{if } w_0 = g_1, g_k^{-1}, \text{ and} \\ 1 & \text{otherwise.} \end{cases}$$

We introduced the function χ_g in order to write $\eta_{\mathcal{C}}(z, s)$ in terms of the transfer operator. We have

$$\begin{aligned} \eta_{\mathcal{C}}(z, s) &= \sum_{g \in \mathcal{C}_k} e^{s f^k(g)} \sum_{m=0}^{\infty} z^{k+2m} \sum_{w \in \Gamma_m} e^{2s f^m(w)} \chi_g(w) + \delta(z, s), \\ &= \sum_{g \in \mathcal{C}_k} e^{s f^k(g)} \sum_{m=0}^{\infty} z^{k+2m} (L_{2sf}^m \chi_g)(1) + \delta(z, s). \end{aligned}$$

Thus the power series $\sum_{m=0}^{\infty} z^{k+2m} \sum_{x \in \mathcal{C}_{k+2m}} F(x)$ can be written in terms of the transfer operator since

$$\begin{aligned} \left. \frac{\partial}{\partial s} \eta_{\mathcal{C}}(z, s) \right|_{s=0} &= \sum_{g \in \mathcal{C}_k} \sum_{m=0}^{\infty} \left. \frac{\partial}{\partial s} z^{k+2m} (L_{2sf}^m \chi_g)(1) \right|_{s=0} \\ &+ \sum_{g \in \mathcal{C}_k} f^k(g) \sum_{m=0}^{\infty} z^{k+2m} (L_0^m \chi_g)(1) + \left. \frac{\partial}{\partial s} \delta(z, s) \right|_{s=0}. \end{aligned}$$

We analyse the growth of the coefficients of the power series in the following sequence of lemmas.

Lemma 4.2 *The coefficient of z^{k+2m} in the power series $\sum_{m=0}^{\infty} z^{k+2m} (L_0^m \chi_g)(1)$ grow with order $O(e^{mh})$.*

The coefficient in the next lemma grows with the same order.

Lemma 4.3 *The coefficient of z^{k+2m} in the power series $\frac{\partial}{\partial s} \delta(z, s)|_{s=0}$ grow with order $O(e^{mh})$.*

Proof Since, for each $w \in \Gamma$, $\xi'_w(0) = 0$,

$$\frac{\partial}{\partial s} \delta(z, s) \Big|_{s=0} = \sum_{m=0}^{\infty} z^{k+2m} \sum_{g \in \mathfrak{C}_k} \sum_{w \in \Gamma_m(g)} \kappa_w.$$

For each $w \in \Gamma$ we have $|\kappa_w| \leq K$. Thus the coefficient of z^{k+2m} is bounded in modulus by

$$\sum_{g \in \mathfrak{C}_k} \sum_{w \in \Gamma_m(g)} K = K \#\mathfrak{C}_{k+2m} = K(2p-2)(2p-1)^{m-1} \#\mathfrak{C}_k = O(e^{mh}),$$

from which the lemma follows. □

We decompose the transfer operator L_{sf} into the projection R_s associated to the eigenspace associated to the eigenvalue $e^{P(sf)}$ and $Q_s = L_{sf} - e^{P(sf)} R_s$. For $s \in \mathbb{C}$ in a neighbourhood of $s = 0$, the operators R_s and Q_s are analytic. We use this operator decomposition to obtain the estimates in the next two lemmas.

Lemma 4.4 *The coefficient of z^{k+2m} in the power series*

$$\frac{\partial}{\partial s} \sum_{g \in \mathfrak{C}_k} \sum_{m=0}^{\infty} z^{k+2m} Q_{2s}^m \chi_g(1) \Big|_{s=0}$$

grow with order $O(e^{m(h-\varepsilon)})$, for some $\varepsilon > 0$.

Proof Suppose that $s \in \mathbb{C}$ such that $0 \leq |s| < \delta_1$ then, as discussed in Sect. 3, if δ_1 is sufficiently small each perturbed operator L_{2sf} has a simple maximal eigenvalue $e^{P(2sf)}$. Moreover, for $|s| < \delta_1$, there exists $\varepsilon_1(\delta_1) > 0$ such that

$$\limsup_{m \rightarrow \infty} \|Q_{2s}^m\|^{1/m} \leq e^{h-\varepsilon_1}.$$

We consider the analyticity of the series

$$\sum_{g \in \mathfrak{C}_k} \sum_{m=0}^{\infty} z^{k+2m} Q_{2s}^m \chi_g(1).$$

Suppose that we fix $z \in \mathbb{C}$ such that $|z| < e^{-h+\varepsilon_1}$, then the series converges for each $s \in \mathbb{C}$ with $|s| < \delta_1$. Meanwhile, given $s \in \mathbb{C}$ such that $|s| < \delta_1$, the series converges for each $z \in \mathbb{C}$ with $|z| < e^{-h+\varepsilon_1}$. Thus, by Hartogs’ theorem [11, Theorem 1.2.5], the series converges to an analytic function in the polydisk $\{s \in \mathbb{C} : |s| < \delta_1\} \times \{z \in \mathbb{C} : |z| < e^{-h+\varepsilon_1}\}$. Thus the power series

$$\frac{\partial}{\partial s} \sum_{g \in \mathfrak{C}_k} \sum_{m=0}^{\infty} z^{k+2m} Q_{2s}^m \chi_g(1) \Big|_{s=0}$$

is analytic for $|z| < e^{-h+\varepsilon_1}$ and so we estimate the coefficients of the power series by $O(e^{m(h-\varepsilon)})$ with $0 < \varepsilon < \varepsilon_1$. □

There is one power series left to study.

Lemma 4.5 Let $P'(0)$ denote the derivative of the function $P(sf)$ evaluated at $s = 0$. The coefficient of z^{k+2m} in the power series

$$\left. \frac{\partial}{\partial s} \sum_{m=0}^{\infty} z^{k+2m} e^{mP(2sf)} R_{2s} \chi_g(1) \right|_{s=0}$$

is $2me^{mh} P'(0) R_0 \chi_g(1) + e^{mh} \frac{d}{ds} R_{2s} \chi_g(1) \Big|_{s=0}$.

Proof We have

$$\begin{aligned} & \left. \frac{\partial}{\partial s} \sum_{m=0}^{\infty} z^{k+2m} e^{mP(2sf)} R_{2s} \chi_g(1) \right|_{s=0} \\ &= \sum_{m=0}^{\infty} z^{k+2m} 2me^{mh} P'(0) R_0 \chi_g(1) + \sum_{m=0}^{\infty} z^{k+2m} e^{mh} \frac{d}{ds} R_{2s} \chi_g(1) \Big|_{s=0}, \end{aligned}$$

from which the result follows. □

Combining the above lemmas, we find that the coefficient of z^{k+2m} in $\frac{\partial}{\partial s} \eta_{\mathfrak{C}}(z, s) \Big|_{s=0}$ satisfies the estimate

$$\sum_{g \in \mathfrak{C}_k} 2me^{mh} P'(0) R_0 \chi_g(1) + O(e^{mh}).$$

Returning to Theorem 2.2 we now have

$$\frac{1}{\#\mathfrak{C}_{k+2m}} \sum_{x \in \mathfrak{C}_{k+2m}} \frac{F(x)}{k+2m} = \frac{2m}{k+2m} P'(0) \frac{e^{mh}}{\#\mathfrak{C}_{k+2m}} \sum_{g \in \mathfrak{C}_k} R_0 \chi_g(1) + O\left(\frac{1}{m}\right).$$

Thus we have

$$\lim_{m \rightarrow \infty} e^{-mh} \sum_{x \in \mathfrak{C}_{k+2m}} \frac{F(x)}{k+2m} = \int f d\mu_0 \sum_{g \in \mathfrak{C}_k} R_0 \chi_g(1).$$

If we substitute $f : \Sigma^* \rightarrow \mathbb{R}$ given by $f(x) = 1$ for each $x \in \Sigma^*$ into the preceding limit we obtain

$$\lim_{m \rightarrow \infty} \frac{\#\mathfrak{C}_{k+2m}}{e^{mh}} = \sum_{g \in \mathfrak{C}_k} R_0 \chi_g(1).$$

Hence we have the desired result,

$$\lim_{m \rightarrow \infty} \frac{1}{\#\mathfrak{C}_{k+2m}} \sum_{x \in \mathfrak{C}_{k+2m}} \frac{F(x)}{k+2m} = \int f d\mu_0.$$

5 Proof of Theorem 2.4

In this section we will prove Theorem 2.4. By Levy’s Continuity Theorem (cf. [5, Theorem 2, Chapter XV §3]), the theorem will follow if we show that the characteristic functions

$$\varphi_m(t) = \frac{1}{\#\mathfrak{C}_{k+2m}} \sum_{x \in \mathfrak{C}_{k+2m}} e^{it(F(x) - (k+2m)\bar{F})/\sqrt{k+2m}}.$$

converge pointwise to $e^{-\sigma_f^2 t^2}$, the characteristic function of the normal distribution with mean zero and variance $2\sigma_f^2$.

Suppose that F satisfies (A1), (A2) and (A3). By replacing F with $F - \overline{F} | \cdot |$ (which still satisfies the three assumptions) or, equivalently, f with $f - \int f d\mu_0$, we may assume without loss of generality that $\int f d\mu_0 = 0$. This reduction does not change the variance. We may then write

$$\varphi_m(t) = \frac{1}{\#\mathcal{C}_{k+2m}} \sum_{x \in \mathcal{C}_{k+2m}} e^{itf^{k+2m}(x)/\sqrt{k+2m}}.$$

We recall the approximation, which we obtain from Lemma 4.1,

$$\exp(sf^{k+2m}(w^{-1}gw)) = \exp\left(s(2f^m(w) + f^k(g))\right) + s\kappa_w + \xi_w(s),$$

where $\kappa_w = f^{k+2m}(w^{-1}gw) - 2f^m(w) - f^k(g)$ is uniformly bounded for $w \in \Gamma$ and $\xi_w(s)$ is an entire function such that $\xi_w(0) = 0$. Using the above approximation, we write $\varphi_m(t)$ as the sum of a leading term and an error term:

$$\frac{1}{\#\mathcal{C}_{k+2m}} \sum_{g \in \mathcal{C}_k} e^{\tau f^k(g)/2} \sum_{w \in \Gamma_m(g)} e^{\tau f^m(w)} + \rho_m(t),$$

where $\tau = 2it/\sqrt{k+2m}$ and the error term $\rho_m(t)$ is given by

$$\rho_m(t) = \frac{1}{\#\mathcal{C}_{k+2m}} \sum_{g \in \mathcal{C}_k} \sum_{w \in \Gamma_m(g)} \frac{it\kappa_w}{\sqrt{k+2m}} + \xi_w\left(\frac{it}{\sqrt{k+2m}}\right).$$

Since the bound on κ_w is uniform and $\xi_w(0) = 0$, we find that $\rho_m(t) \rightarrow 0$ as $m \rightarrow \infty$. We rewrite the leading term using the transfer operator as

$$\frac{1}{\#\mathcal{C}_{k+2m}} \sum_{g \in \mathcal{C}_k} e^{\tau f^k(g)/2} L_{\tau f}^m \chi_g(1).$$

For sufficiently large m , the simple maximal eigenvalue $e^{P(\tau f)}$ of the perturbed operator $L_{\tau f}$ persists and also plays a crucial role in determining the limit of $\varphi_m(t)$ as $m \rightarrow \infty$. Before we establish the limit, we first analyse the pressure function and establish a preliminary limit for $e^{m(P(\tau f)-h)}$ as $m \rightarrow \infty$.

Recall that the pressure function $P(sf)$ (defined as the principal branch of the logarithm of $e^{P(sf)}$) is analytic in a neighbourhood of $s = 0$ and that $P'(0) = \int f d\mu_0 = 0$. By analyticity we can choose $\delta > 0$ such that if $|s| < \delta$ then

$$P(2sf) = h + 2\sigma_f^2 s^2 + s^3 \vartheta(s),$$

for some function $\vartheta(s)$ that is analytic in a neighbourhood of $s = 0$. For sufficiently large m , with $\tau = 2it/\sqrt{k+2m}$ as before, we have

$$(k+2m)P(\tau f) = (k+2m)h - 2\sigma_f^2 t^2 - \frac{4it^3 \vartheta(\tau)}{3\sqrt{k+2m}}$$

and so

$$\frac{e^{(k+2m)P(\tau f)}}{e^{(k+2m)h}} = e^{-2\sigma_f^2 t^2} \exp\left\{-\frac{4it^3 \vartheta(\tau)}{3\sqrt{k+2m}}\right\},$$

from which the next proposition and corollary follow.

Proposition 5.1 *We have the following limit*

$$\lim_{m \rightarrow \infty} \frac{e^{(k+2m)P(\tau f)}}{e^{(k+2m)h}} = e^{-2\sigma_f^2 t^2}.$$

Corollary 5.2 *We have the limit*

$$\lim_{m \rightarrow \infty} \frac{e^{mP(\tau)}}{e^{mh}} = e^{-\sigma_f^2 t^2}.$$

We use the notation $\beta(\tau) = e^{P(\tau f)}$ and $\beta(0) = e^h$ in the proof of Proposition 5.3.

Proposition 5.3 *The limit of $\varphi_m(t)$ as $m \rightarrow \infty$ is $e^{-\sigma_f^2 t^2}$.*

Proof Written in terms of the transfer operator and a null sequence $(\rho_m(t))_{m=0}^\infty$, the function $\varphi_m(t)$ is equal to

$$\frac{1}{\#\mathfrak{C}_{k+2m}} \sum_{g \in \mathfrak{C}_k} e^{\tau f^k(g)/2} L_{\tau f}^m \chi_g(1) + \rho_m(t).$$

We recall the decomposition of the transfer operator into $L_{sf} = \beta(s)R_s + Q_s$. For sufficiently large m , the leading term is given by

$$\frac{\beta(\tau)^m}{\#\mathfrak{C}_{k+2m}} \sum_{g \in \mathfrak{C}_k} e^{\tau f^k(g)/2} R_\tau \chi_g(1) + \frac{1}{\#\mathfrak{C}_{k+2m}} \sum_{g \in \mathfrak{C}_k} e^{\tau f^k(g)/2} Q_\tau^m \chi_g(1).$$

Since the spectral radius of Q_τ is strictly less than $|\beta(\tau)|$, we find $\|\beta(\tau)^{-m} Q_\tau^m\| = O(\kappa^m)$ for some $\kappa \in (0, 1)$ and so we have

$$\frac{1}{\#\mathfrak{C}_{k+2m}} \sum_{g \in \mathfrak{C}_k} e^{\tau f^k(g)/2} Q_\tau^m \chi_g(1) = O\left(\frac{\beta(\tau)^m}{\beta(0)^m} \kappa^m\right).$$

By Corollary 5.2 we have $\lim_{m \rightarrow \infty} \beta(\tau)^m / \beta(0)^m = e^{-\sigma_f^2 t^2}$ and so

$$\lim_{m \rightarrow \infty} \frac{1}{\#\mathfrak{C}_{k+2m}} \sum_{g \in \mathfrak{C}_k} e^{\tau f^k(g)/2} Q_\tau^m \chi_g(1) = 0.$$

We now turn our attention to the asymptotics for the term

$$\frac{\beta(\tau)^m}{\#\mathfrak{C}_{k+2m}} \sum_{g \in \mathfrak{C}_k} e^{\tau f^k(g)/2} R_\tau \chi_g(1).$$

In order to approximate this term, we first write the projection R_τ in terms of R_0 . Since the projection is analytic for τ in a neighbourhood of 0 we have, for sufficiently large m , $e^{\tau f^k(g)/2} R_\tau \chi_g(1) = R_0 \chi_g(1) + O(t/\sqrt{k+2m})$. We recall that $\#\mathfrak{C}_{k+2m} = (\beta(0) - 1)\beta(0)^{m-1}\#\mathfrak{C}_k$ and so

$$\frac{\beta(\tau)^m}{\#\mathfrak{C}_{k+2m}} \sum_{g \in \mathfrak{C}_k} e^{\tau f^k(g)/2} R_\tau \chi_g(1) = \frac{\beta(\tau)^m}{\#\mathfrak{C}_{k+2m}} \sum_{g \in \mathfrak{C}_k} R_0 \chi_g(1) + O\left(\frac{\beta(\tau)^m t}{\beta(0)^m \sqrt{k+2m}}\right).$$

We recall the limit

$$\lim_{m \rightarrow \infty} \frac{\#\mathfrak{C}_{k+2m}}{\beta(0)^m} = \sum_{g \in \mathfrak{C}_k} R_0 \chi_g(1)$$

and so, together with the above approximation, we find the limit of $\varphi_m(t)$ as $m \rightarrow \infty$ is given by

$$\lim_{m \rightarrow \infty} \frac{\beta(\tau)^m}{\#\mathfrak{C}_{k+2m}} \sum_{g \in \mathfrak{C}_k} e^{\tau f^k(g)/2} R_\tau \chi_g(1) = \lim_{m \rightarrow \infty} \frac{\beta(\tau)^m}{\beta(0)^m} = e^{-\sigma_f^2 t^2},$$

which is the desired result. □

Acknowledgements George Kenison was funded by the Engineering and Physical Sciences Research Council (DTA Award Number 1359001).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Calegari, D., Fujiwara, K.: Combable functions, quasimorphisms, and the central limit theorem. *Ergod. Theory Dyn. Syst.* **30**, 1343–1369 (2010)
2. Coelho, Z., Parry, W.: Central limit asymptotics for shifts of finite type. *Israel J. Math.* **69**(2), 235–249 (1990)
3. Coornaert, M.: Mesures de Patterson-Sullivan sur le bord d'un espace hyperbolique au sens de Gromov. *Pac. J. Math.* **159**, 241–270 (1993)
4. Dal'bo, F.: Remarques sur le spectre des longueurs d'une surface et comptages. *Bol. Soc. Bras. Mat.* **30**, 199–221 (1999)
5. Feller, W.: *An Introduction to Probability Theory and Its Applications*, vol. II, 2nd edn. Wiley, New York (1971)
6. Horsham, M.: Central limit theorems for quasi-morphisms of surface groups. Ph.D. thesis, Manchester (2008)
7. Horsham, M., Sharp, R.: Lengths, quasi-morphisms and statistics for free groups. In: Kotani M, Naito H, Tate T (eds) *Spectral Analysis in Geometry and Number Theory*, volume 484 of *Contemporary Mathematics*, pp. 219–237. American Mathematical Society, Providence, RI (2009)
8. Kato, T.: *Perturbation Theory for Linear Operators*. *Classics in Mathematics*. Springer, Berlin (1995). **(Reprint of the 1980 edition)**
9. Kenison, G., Sharp, R.: Orbit counting in conjugacy classes for free groups acting on trees. *J. Topol. Anal.* **9**, 631–647 (2017)
10. Kifer, Y.: Large deviations in dynamical systems and stochastic processes. *Trans. Am. Math. Soc.* **321**, 505–524 (1990)
11. Krantz, S.: *Function theory of several complex variables*. AMS Chelsea Publishing, Providence (2001). **(Reprint of the 1992 edition)**
12. Parry, W., Pollicott, M.: Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque* **187–188** (1990)
13. Pollicott, M., Sharp, R.: Large deviations and the distribution of pre-images of rational maps. *Commun. Math. Phys.* **181**, 733–739 (1996)
14. Pollicott, M., Sharp, R.: Comparison theorems and orbit counting in hyperbolic geometry. *Trans. Am. Math. Soc.* **350**(2), 473–499 (1998)
15. Pollicott, M., Sharp, R.: Poincaré series and comparison theorems for variable negative curvature. In: Turaev V, Vershik A (eds) *Topology, Ergodic Theory, Real Algebraic Geometry*, volume 202 of *American Mathematical Society Translation Series 2*, pp. 229–240. American Mathematical Society, Providence (2001)
16. Pollicott, M., Sharp, R.: Statistics of matrix products in hyperbolic geometry. In: Kolyada S, Manin Y, Möller M, Moree P, Ward T (eds) *Dynamical Numbers: Interplay Between Dynamical Systems and Number Theory*, *Contemporary Mathematics*, vol. 532, pp. 213–230 (2011)
17. Rivin, I.: Growth in free groups (and other stories)—twelve years later. *Ill. J. Math.* **54**(1), 327–370 (2010)
18. Ruelle, D.: *Thermodynamic Formalism*, volume 5 of *Encyclopedia of Mathematics and Its Applications*. Addison-Wesley, Reading (1978)