

# ORBIT COUNTING IN CONJUGACY CLASSES FOR FREE GROUPS ACTING ON TREES

GEORGE KENISON AND RICHARD SHARP

ABSTRACT. In this paper we study the action of the fundamental group of a finite metric graph on its universal covering tree. We assume the graph is finite, connected and the degree of each vertex is at least three. Further, we assume an irrationality condition on the edge lengths. We obtain an asymptotic for the number of elements in a fixed conjugacy class for which the associated displacement of a given base vertex in the universal covering tree is at most  $T$ . Under a mild extra assumption we also obtain a polynomial error term.

## 1. INTRODUCTION

Let  $G$  be a finite connected graph. We always assume that each vertex of  $G$  has degree at least 3, in which case the fundamental group of  $G$  is a free group  $F$  on  $k \geq 2$  generators. We make  $G$  into a metric graph by assigning to each edge  $e$  a positive real length  $l(e)$ . The length of a path in  $G$  is given by the sum of the lengths of its edges. We assume the set of closed geodesics in  $G$  (i.e. closed paths without backtracking) has lengths not contained in a discrete subgroup of  $\mathbb{R}$ .

The universal cover of  $G$  is an infinite tree  $\mathcal{T}$  and the metric on  $G$  lifts to a metric  $d_{\mathcal{T}}$  on  $\mathcal{T}$ . We consider each edge in  $\mathcal{T}$  as an isometric copy of a real interval. Then the ball of radius  $T$  centred at  $o \in \mathcal{T}$  is the set

$$B(o, T) = \{y \in \mathcal{T} : d_{\mathcal{T}}(o, y) < T\}.$$

The *volume* of  $B(o, T)$  is the sum of the metric edge lengths in  $B(o, T)$ . Let  $h > 0$  denote the *volume entropy* of  $\mathcal{T}$  given by

$$\lim_{T \rightarrow \infty} \frac{1}{T} \log \text{vol}(B(o, T)).$$

We note the volume entropy is independent of the choice of  $o \in \mathcal{T}$ .

Let  $x \in F$  and fix a base vertex  $o \in \mathcal{T}$ . We define  $L : F \rightarrow \mathbb{R}^+$  by  $L(x) = d_{\mathcal{T}}(o, ox)$ . Guillopé [5] showed that

$$\#\{x \in F : L(x) \leq T\} \sim ce^{hT} \quad \text{as } T \rightarrow \infty,$$

for some  $c > 0$ . Here  $f(T) \sim g(T)$  means that  $f(T)/g(T) \rightarrow 1$  as  $T \rightarrow \infty$ .

Let  $\mathfrak{C}$  be a non-trivial conjugacy class in  $F$ . Then  $\mathfrak{C}$  is infinite and it is interesting to study the restriction of the above counting problem to this conjugacy class, i.e. to study the asymptotic behaviour of

$$N_{\mathfrak{C}}(T) := \#\{x \in \mathfrak{C} : L(x) \leq T\}.$$

The following is our main result.

**Theorem 1.1.** *Suppose that  $G$  is a finite connected metric graph such that the degree of each vertex is at least 3 and the set of lengths of closed geodesics in  $G$  is not contained in a discrete subgroup of  $\mathbb{R}$ . Let  $\mathfrak{C}$  be a non-trivial conjugacy class in  $F$ . Then, for some constant  $C > 0$ , depending on  $\mathfrak{C}$ ,*

$$N_{\mathfrak{C}}(T) \sim Ce^{hT/2}, \quad \text{as } T \rightarrow \infty.$$

We can also obtain a polynomial error term in our approximation to  $N_{\mathfrak{C}}(T)$  subject to a mild additional condition on the lengths of closed geodesics (see Theorem 5.2).

In the case of a co-compact group of isometries of the hyperbolic plane, an analogue of Theorem 1.1 was obtained by Huber [6] in the 1960s. If  $\Gamma$  is a co-compact Fuchsian group and  $\mathfrak{C}$  is a non-trivial conjugacy class, he showed that

$$\#\{g \in \mathfrak{C} : d_{\mathbb{H}^2}(o, og) \leq T\} \sim Ce^{T/2}, \quad \text{as } T \rightarrow \infty,$$

for some  $C > 0$  depending on  $\Gamma$  and  $\mathfrak{C}$  (while the unrestricted counting function  $\#\{g \in \Gamma : d_{\mathbb{H}^2}(o, og) \leq T\}$  is asymptotic to a constant times  $e^T$ ). Very recently, Parkkonen and Paulin [9] have studied the same problem in higher dimensions and variable curvature, obtaining many results. In particular, they have shown that for the fundamental group of a compact negatively curved manifold acting on its universal cover, the conjugacy counting function has exponential growth rate equal to  $h/2$ , where  $h$  is the topological entropy of the geodesic flow. They have an ergodic-geometric approach using, in particular, the mixing properties of the Bowen–Margulis measure.

More closely related to our situation, suppose that  $G$  is a  $(q + 1)$ -regular graph (i.e. each vertex has degree  $q + 1$ ) with each edge given length 1. Then Douma [3] showed that

$$N_{\mathfrak{C}}(n) \sim Cq^{\lfloor (n-l(\mathfrak{C}))/2 \rfloor}, \quad \text{as } n \rightarrow \infty$$

(for  $n \in \mathbb{Z}^+$ ), for some  $C > 0$ , where  $l(\mathfrak{C})$  is the length of the closed geodesic in the conjugacy class  $\mathfrak{C}$ .

Since the first version of this paper was written, we learned that Broise-Alamichel, Parkkonen and Paulin also have results for graphs and metric graphs which include Theorem 1.1 and the result of Douma. They also consider the more general situation of graphs of groups in the sense of Bass–Serre theory. We understand that an account of this work is in preparation [1].

In contrast to the ergodic-geometric approach of Parkkonen and Paulin in [9] or the use of spectral theory of the graph Laplacian in [3] (which is inspired by Huber’s original spectral approach [6]), we use a method based on a symbolic coding of the group  $F$  in terms of a subshift of finite type. We may then study a generating function via the spectra of a family of matrices. In the next section, we set out the background we shall need, discussing shifts of finite type and some concepts from ergodic theory. In section 3, we describe how the lengths on the graph may be encoded in terms of a function on our subshift and use this function to define a family of matrices and sketch a proof of Guillopé’s result given above. In section 4 we introduce a generating function appropriate to our problem and carry out an analysis which leads to the proof of Theorem 1.1. In the final section we discuss error terms.

We are grateful to the referee for a very careful reading of our paper. Their comments have considerably improved the exposition.

## 2. PRELIMINARIES

A (finite, connected) graph  $G = (V, E)$  consists of a finite collection of vertices  $V$  and edges  $E$ . Let  $E^o$  denote the oriented edge set of the graph  $G$ . For each  $e \in E^o$  we indicate the edge with reversed orientation by  $\bar{e}$ . A *path* is a sequence of consecutive oriented edges  $e_0, \dots, e_{n-1}$ ; and we call a path *non-backtracking* if, in addition,  $e_{i+1} \neq \bar{e}_i$  for  $i = 0, \dots, n-1$ . Path  $e_0, \dots, e_{n-1}$  is said to be *closed* if the terminal vertex of  $e_{n-1}$  and the initial vertex of  $e_0$  are the same. We say a closed path  $e_0, \dots, e_{n-1}$  is a *closed geodesic* if the path is non-backtracking and  $e_{n-1} \neq \bar{e}_0$ , i.e. each path given by a cyclic permutation of  $e_0, \dots, e_{n-1}$  is non-backtracking.

The condition that each vertex has degree at least 3 ensures that the fundamental group of  $G$  is a free group  $F$  on  $k \geq 2$  generators and that the universal cover is an infinite tree  $\mathcal{T}$ . We put a metric on  $G$  by assigning a positive length to each edge and this lifts to a metric on  $\mathcal{T}$ .

Fix a generating set  $\mathcal{A} = \{a_1, \dots, a_k\}$ ,  $k \geq 2$ , for the free group  $F$  and write  $\mathcal{A}^{-1} = \{a_1^{-1}, \dots, a_k^{-1}\}$ . We shall say  $x_0 \cdots x_{n-1}$ , with each  $x_i \in \mathcal{A} \cup \mathcal{A}^{-1}$ , is a *reduced word* if  $x_{i+1} \neq x_i^{-1}$  for  $i = 0, \dots, n-1$  and a *cyclically reduced word* if, in addition,  $x_{n-1} \neq x_0^{-1}$ . Each non-identity element of  $F$  has a unique representation as a reduced word and the *word length*  $|x|$  of  $x$  is the number of terms in its reduced word representation. We regard the identity element 1 as a reduced word of length zero. For each  $m \geq 0$ , let  $W_m$  denote the set of reduced words of length at most  $m$  and let  $W'_m$  denote the set of reduced words of length exactly  $m$ . We let  $W^* = \bigcup_{n=0}^{\infty} W'_n$  denote the set of all finite reduced words.

Let  $\mathfrak{C}$  be a non-trivial conjugacy class in  $F$ . Then  $\mathfrak{C}$  contains a cyclically reduced word  $x_0 \cdots x_{n-1} \in W^* \setminus \{1\}$ . The only other cyclically reduced words in this conjugacy class are obtained by cyclic permutation and the elements of  $\mathfrak{C}$  represented by non-cyclically reduced words have word length greater than  $n$ .

Recall the function  $L : F \rightarrow \mathbb{R}^+$  given by  $L(x) = d_{\mathcal{T}}(o, ox)$ . We will use the following lemma from [15].

**Lemma 2.1.** *There exists  $N \in \mathbb{N}$  such that if  $n \geq N$  and  $x_0 \cdots x_{n-1}$  is a reduced word then*

$$L(x_0 x_1 \cdots x_{n-1}) - L(x_1 \cdots x_{n-1}) = L(x_0 x_1 \cdots x_{N-1}) - L(x_1 \cdots x_{N-1}).$$

It is useful to consider infinite sequences and dynamics on them. In particular, it is convenient to associate closed geodesics in the graph to periodic orbits of the shift map. We can define a function on both finite and infinite reduced words which will encode the lengths  $L(x)$  and which will also give the lengths of closed geodesics by summing the function around the corresponding periodic orbits. Introducing the shift map on infinite sequences also has the advantage of allowing us to use thermodynamic concepts from ergodic theory, for example pressure and equilibrium states defined below, and standard results about differentiating pressure. Let  $\Sigma$  denote the set of infinite reduced words in  $\mathcal{A} \cup \mathcal{A}^{-1}$ :

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

We will need to study the dynamical system on  $\Sigma$  given by the shift map  $\sigma : \Sigma \rightarrow \Sigma$ , defined by  $\sigma((x_n)_{n=0}^{\infty}) = (x_{n+1})_{n=0}^{\infty}$ . This is a subshift of finite type and we will refer to [11] for the general theory of these systems.

Writing  $x = (x_n)_{n=0}^{\infty}, y = (y_n)_{n=0}^{\infty} \in \Sigma$ , we endow  $\Sigma$  with the metric defined by  $d(x, x) = 0$  and, for  $x \neq y$ ,  $d(x, y) = (1/2)^k$ , where  $k = \min\{n \in \mathbb{N} : x_n \neq y_n\}$ . This makes  $\Sigma$  into a Cantor set and makes the map  $\sigma$  continuous. Furthermore,  $\sigma$  is topologically mixing (i.e. if  $U, V \subset \Sigma$  are non-empty open sets then  $\sigma^{-n}U \cap V \neq \emptyset$  for all sufficiently large  $n$ ).

A function  $f : \Sigma \rightarrow \mathbb{R}$  is said to be *locally constant* if there exists an  $N \in \mathbb{N}$  such that for any two elements  $x, y \in \Sigma$  with  $x_n = y_n$  for every  $0 \leq n \leq N$  we have  $f(x) = f(y)$ . If  $f$  is locally constant then it is *Hölder continuous* for every positive exponent (i.e. where  $f$  is Hölder continuous of exponent  $\alpha > 0$  if there exists a positive constant  $\kappa$  such that  $|f(x) - f(y)| \leq \kappa d(x, y)^\alpha$ , for every  $x, y \in \Sigma$ ). We will use the notation

$$f^n = f + f \circ \sigma + \cdots + f \circ \sigma^{n-1}.$$

Two Hölder continuous functions  $f, g : \Sigma \rightarrow \mathbb{R}$  are said to be *cohomologous* if  $f = g + u \circ \sigma - u$  for some continuous function  $u : \Sigma \rightarrow \mathbb{R}$ . Clearly, two cohomologous functions have the same integral with respect to each  $\sigma$ -invariant measure.

Let  $\mathcal{M}$  denote the set of  $\sigma$ -invariant probability measures on  $\Sigma$ . We denote by  $h(\mu)$  the measure theoretic entropy of  $\sigma$  with respect to  $\mu$ . For a continuous function  $f : \Sigma \rightarrow \mathbb{R}$ , we define its *pressure*  $P(f)$  by

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

We say that  $m \in \mathcal{M}$  is an *equilibrium state* for  $f$  if the supremum is attained at  $m$ . When  $f$  is a Hölder continuous function the Variational Principle ([11], Theorem 3.5) tells us the equilibrium state is unique.

**Remark 2.2.** Another subshift of finite type naturally associated to the graph is obtained by taking infinite paths, i.e. infinite sequences of oriented edges with the restriction that  $e'$  can follow  $e$  only if  $e$  terminates at the initial vertex of  $e'$ . The advantage of our approach is that it makes it easier to systematically enumerate the elements of a given conjugacy class. On the other hand, it requires more work to represent the edge lengths and we introduce a function that does this below.

We will also need to consider the spectra of non-negative matrices and complex matrices. We say that a non-negative (square) matrix  $A$  is *irreducible* if, for each pair of indices  $(i, j)$ , there exists  $n \geq 1$  such that  $A^n(i, j) > 0$  and that  $A$  is *aperiodic* if there exists  $n \geq 1$  such that, for each pair of indices  $(i, j)$ ,  $A^n(i, j) > 0$ . We will use the following two standard results.

**Theorem 2.3** (Perron–Frobenius Theorem [4]). *Suppose that  $A$  is an aperiodic matrix with non-negative entries. Then  $A$  has a simple and positive eigenvalue  $\beta$  such that  $\beta$  is strictly greater in modulus than all the remaining eigenvalues of  $A$ . The left and right eigenvectors associated to the eigenvalue  $\beta$  have strictly positive entries. Moreover,  $\beta$  is the only eigenvalue of  $A$  that has an eigenvector whose entries are all non-negative.*

**Theorem 2.4** (Wielandt’s Theorem [4]). *Suppose that  $A$  is a square matrix with complex entries and let  $|A|$  be the matrix whose entries are given by  $|A|(s, s') = |A(s, s')|$ . Suppose further that  $|A|$  is aperiodic and let  $\beta$  be its maximal eigenvalue guaranteed by the Perron–Frobenius Theorem. Then the moduli of the eigenvalues of  $A$  are bounded above by  $\beta$ . Moreover,  $A$  has an eigenvalue of the form  $\beta e^{i\theta}$  (with  $\theta \in [0, 2\pi)$ ) if and only if  $A = e^{i\theta} D |A| D^{-1}$  where  $D$  is a diagonal matrix whose entries along the main diagonal all have modulus one.*

### 3. LENGTH FUNCTIONS, MATRICES AND SPECTRA

We will prove Theorem 1.1 by studying the analytic properties of a generating function  $\sum_{x \in \mathfrak{C}} e^{-sL(x)}$ . More precisely, we will show that the generating function is analytic in the half-plane  $\operatorname{Re}(s) > h/2$ , has a simple pole at  $s = h/2$  and, crucially, apart from this pole has an analytic extension to a neighbourhood of  $\operatorname{Re}(s) = h/2$ . To do this, we will show that the generating function can essentially be written in terms of a family of weighted matrices and their eigenvectors. The required analytic properties will then follow from results about the spectra of these matrices. In turn, the key spectral property (see Lemma 3.7 below) is a consequence of the hypothesis the lengths of closed geodesics in our metric graph do not lie in a discrete subgroup of the real numbers.

In this section, we will set up the machinery required to study the generating function. We first introduce a function defined on (finite and infinite) reduced words which encodes information about lengths on the graph. (This is similar to the constructions in [14] and [15].) We will then introduce our weighted matrices and establish some of their properties.

**Definition 3.1.** We define a function  $r : \Sigma \rightarrow \mathbb{R}$  by

$$r((x_i)_{i=0}^{\infty}) = L(x_0 x_1 \cdots x_{N-1}) - L(x_1 \cdots x_{N-1}).$$

We also define  $r : W^* \rightarrow \mathbb{R}$  by  $r(1) = 0$  and, for  $n \geq 1$ ,

$$r(x_0 \cdots x_{n-1}) = L(x_0 \cdots x_{n-1}) - L(x_1 \cdots x_{n-1}).$$

Note that, by Lemma 2.1, if  $n \geq N$  then

$$r(x_0 \cdots x_{n-1}) = L(x_0 \cdots x_{N-1}) - L(x_1 \cdots x_{N-1}).$$

We may also extend the definition of the shift map  $\sigma$  to finite reduced words by defining  $\sigma : W^* \rightarrow W^*$  by

$$\sigma(x_0 x_1 \cdots x_{n-1}) = (x_1 \cdots x_{n-1})$$

and  $\sigma 1 = 1$ . We shall continue to write  $r^n = r + r \circ \sigma + \cdots + r \circ \sigma^{n-1}$ . The next lemma is immediate from the definition of  $r$ .

**Lemma 3.2.** *For any finite reduced word  $x_0 \cdots x_{n-1}$ , we have  $L(x_0 \cdots x_{n-1}) = r^n(x_0 \cdots x_{n-1})$ .*

The following lemma connects closed geodesics in the graph  $G$  to periodic points in the subshift of finite type  $\sigma : \Sigma \rightarrow \Sigma$ . The fact that the sums of  $r$  over periodic points do not lie in a discrete subgroup will be crucial in establishing that our generating function has no non-real poles on its abscissa of convergence.

**Lemma 3.3.** *Let  $\gamma$  be the unique closed geodesic corresponding to the periodic orbit  $\{x, \sigma x, \dots, \sigma^{n-1} x\}$  ( $\sigma^n x = x$ ) with  $x = (x_i)_{i=0}^{\infty} \in \Sigma$ . Then  $r^n(x) = l(\gamma)$ . In particular,  $\{r^n(x) : \sigma^n x = x, n \geq 1\}$  is not contained in a discrete subgroup of  $\mathbb{R}$ .*

*Proof.* Let  $(x_0 \cdots x_{n-1})^m$  denote the  $m$ -fold concatenation of the cyclically reduced word  $x_0 \cdots x_{n-1}$ . By the definition of  $r$  and Lemma 3.2, we have

$$|r^{mn}(x) - L((x_0 \cdots x_{n-1})^m)| \leq 2N \|r\|_{\infty}$$

and so, since  $r^{mn}(x) = mr^n(x)$ , we deduce

$$r^n(x) = \lim_{m \rightarrow \infty} \frac{1}{m} L((x_0 \cdots x_{n-1})^m).$$

Now consider the closed geodesic  $\gamma$  in  $G$ . This lifts to a geodesic path in  $\mathcal{T}$ , from some vertex  $p$  to  $pg$ , where  $g \in F$  is conjugate to  $x_0 \cdots x_{n-1}$ . For each  $m \geq 1$ , we have  $ml(\gamma) = d_{\mathcal{T}}(p, pg^m)$ . For any vertex  $q \in \mathcal{T}$ , the triangle inequality gives

$$d_{\mathcal{T}}(p, pg^m) - 2d_{\mathcal{T}}(p, q) \leq d_{\mathcal{T}}(q, qg^m) \leq d_{\mathcal{T}}(p, pg^m) + 2d_{\mathcal{T}}(p, q),$$

so that

$$l(\gamma) = \lim_{m \rightarrow \infty} \frac{1}{m} d_{\mathcal{T}}(q, qg^m).$$

If  $x_0 \cdots x_{n-1} = w^{-1}gw$  then putting  $q = ow^{-1}$  gives

$$\begin{aligned} l(\gamma) &= \lim_{m \rightarrow \infty} \frac{1}{m} d_{\mathcal{T}}(ow^{-1}, ow^{-1}g^m) = \lim_{m \rightarrow \infty} \frac{1}{m} d_{\mathcal{T}}(o, o(x_0 \cdots x_{n-1})^m) \\ &= \lim_{m \rightarrow \infty} \frac{1}{m} L((x_0 \cdots x_{n-1})^m). \end{aligned}$$

This completes the proof.  $\square$

Now we turn to the definition of the matrices we use to analyse the generating function.

We begin by defining an unweighted transition matrix  $A$ , whose rows and columns are indexed by  $W_{N-1}$ , the set of reduced words of length at most  $N-1$ , where the number  $N \geq 2$  is given by Lemma 2.1. The entries of  $A$  are defined as follows. We have  $A(x, y) = 1$  if there exists  $n \leq N-2$  and there exists  $x_0, x_1, \dots, x_{n-2} \in \mathcal{A} \cup \mathcal{A}^{-1}$  such that  $x$  has the reduced word representation  $x = x_0 x_1 \cdots x_{n-2}$

and  $y$  has the reduced word representation  $y = x_1 \cdots x_{n-2}$ , or if there exists  $x_0, x_1, \dots, x_{N-2}, y_{N-2} \in \mathcal{A} \cup \mathcal{A}^{-1}$  such that  $x$  has the reduced word representation  $x = x_0 x_1 \cdots x_{N-2}$  and  $y$  has the reduced word representation  $y = x_1 \cdots x_{N-2} y_{N-2}$ . We have  $A(x, y) = 0$  in all other cases.

We next define the numbers we shall use to define weighted matrices compatible with  $A$ .

**Definition 3.4.** For each pair  $(x, y) \in W_{N-1} \times W_{N-1}$  with  $A(x, y) = 1$ , we define a number  $R(x, y)$  by

$$R(x, y) = r(x_0 x_1 \cdots x_{N-2} y_{N-2}) = L(x_0 x_1 \cdots x_{N-2} y_{N-2}) - L(x_1 \cdots x_{N-2} y_{N-2}),$$

in the case where  $x = x_0 x_1 \cdots x_{N-2}$  and  $y = x_1 x_2 \cdots x_{N-2} y_{N-2}$ , and

$$R(x, y) = r(x_0 x_1 \cdots x_{n-2}) = L(x_0 x_1 \cdots x_{n-2}) - L(x_1 \cdots x_{n-2}),$$

in the case where  $x = x_0 x_1 \cdots x_{n-2}$  and  $y = x_1 x_2 \cdots x_{n-2}$ , with  $n \leq N$ .

We now introduce the family of weighted matrices with which we encode edge lengths. The matrices  $A_s$ , with  $s \in \mathbb{C}$ , have rows and columns indexed by  $W_{N-1}$  with entries

$$A_s(x, y) = \begin{cases} 0 & \text{if } A(x, y) = 0, \\ e^{-sR(x, y)} & \text{otherwise.} \end{cases}$$

We will be interested in the spectral properties of  $A_s$ . If  $s \in \mathbb{R}$  then  $A_s$  has non-negative entries but it is not aperiodic or even irreducible and we cannot apply the Perron–Frobenius Theorem directly. Similarly, we cannot apply Wielandt’s Theorem directly to  $A_s$  when  $s \in \mathbb{C}$ . Instead, we will consider the submatrix of  $A_s$  with rows and columns indexed by  $W'_{N-1}$ , which we will denote by  $B_s$ . One can easily see that  $A_s$  and  $B_s$  have the same non-zero spectrum, though  $A_s$  has additional zero eigenvalues.

When  $s \in \mathbb{R}$ , the matrix  $B_s$  is aperiodic and so, by the Perron–Frobenius Theorem, has a simple and positive eigenvalue  $\beta(s)$ , which is strictly greater in modulus than all of the other eigenvalues of  $B_s$ . Let  $\tilde{u}(s)$  and  $\tilde{v}(s)$  denote the left and right eigenvectors of  $B_s$  corresponding to  $\beta(s)$ , normalised so that  $\tilde{u}(s)$  is a probability vector and  $\tilde{u}(s) \cdot \tilde{v}(s) = 1$ .

We have the following lemma.

**Lemma 3.5.** *For  $s \in \mathbb{R}$ , suppose that  $\beta(s)$  is the maximal eigenvalue of the matrix  $B_s$ . Then  $\beta(s)$  is real analytic and is related to the integral of  $r : \Sigma \rightarrow \mathbb{R}$  by the formula*

$$\beta'(s) = -\beta(s) \int r d\mu_{-sr},$$

where  $\mu_{-sr}$  is the equilibrium state for  $-sr$ . Furthermore,  $\int r d\mu_{-sr} > 0$ .

*Proof.* The analyticity of an isolated simple eigenvalue is standard. Furthermore,  $\beta(s) = \exp P(-sr)$ , where  $P$  is the pressure function for the shift  $\sigma : \Sigma \rightarrow \Sigma$ . The formula for the derivative is then given in [11], for example. For any  $n \geq 1$ ,  $r$  is cohomologous to  $n^{-1}r^n$  and, for  $n$  sufficiently large,  $r^n$  is strictly positive. These observations prove positivity of the integral.  $\square$

**Corollary 3.6.** *There exists a unique positive real number  $h$  such that  $\beta(h) = 1$ .*

*Proof.* By comparing with the trace of  $B_s^n$ , one can easily check that  $\beta(0) > 1$  and that  $\beta(s) < 1$  for sufficiently large  $s$ . By Lemma 3.5,  $\beta(s)$  is strictly decreasing, so the required number  $h$  exists and is unique.  $\square$

We will also need to consider  $B_s$  for  $s \in \mathbb{C}$ . In particular, we have the following result.

**Lemma 3.7.** *For  $t \neq 0$ , the matrix  $B_{h+it}$  does not have 1 as an eigenvalue.*

*Proof.* Lemma 3.3 tells us  $\{r^n(x): \sigma^n x = x, n \geq 1\}$  is not contained in a discrete subgroup of  $\mathbb{R}$ . It then follows from [10] that the result holds.  $\square$

For  $s \in \mathbb{R}$ , let  $u(s)$  and  $v(s)$  denote, respectively, the left and right eigenvectors of  $A_s$  associated to the eigenvalue  $\beta(s)$ . We choose to normalise  $u(s)$  so that each entry is strictly positive and scale the entries so that  $(u(s))_1 = 1$ . (Note that here the subscript 1 denotes the entry index by the identity element in  $F$ , i.e. the word of length zero.) We normalise  $v(s)$  so that the entries  $(v(s))_y$  for  $y \in W'_{N-1}$  are strictly positive, whilst all other entries are 0. We then scale the entries of  $v(s)$  so that  $u(s) \cdot v(s) = 1$ .

We conclude this section by explaining why the number  $h$ , from Corollary 3.6, that gives  $\beta(h) = 1$  is equal to the volume entropy. We will do this by sketching a proof of Guillopé's theorem. This will also serve as a prelude to the analysis in the next section, where the arguments will be given in more detail.

To evaluate the asymptotic behaviour of  $\#\{x \in F: L(x) \leq T\}$  we first establish analytic properties of the complex generating function

$$\eta(s) = \sum_{n=1}^{\infty} \sum_{|x|=n} e^{-sL(x)}.$$

Letting  $\mathbf{1} = (1, \dots, 1)$  and  $\mathbf{e}_x$  denote the standard unit vector associated to  $x \in W_{N-1}$ , we can rewrite  $\eta(s)$  in terms of the matrices  $A_s$ :

$$\eta(s) = \sum_{n=1}^{\infty} \mathbf{1} A_s^n \mathbf{e}_1.$$

This converges absolutely for  $\beta(\operatorname{Re}(s)) < 1$ , i.e. for  $\operatorname{Re}(s) > h$ .

Since  $\beta(s)$  is a simple eigenvalue, it varies analytically and we can show that  $\eta(s) = c_0/(s-h) + \phi_1(s)$  where  $\phi_1(s)$  is an analytic function in a neighbourhood of  $s = h$  and  $c_0 = -(v(s) \cdot \mathbf{1})/\beta'(h) > 0$ . Furthermore, using Lemma 3.7, we may show that  $\eta(s)$  has no further poles on the line  $\operatorname{Re}(s) = h$ .

The generating function  $\eta(s)$  is related to the counting function  $N(T) := \#\{x \in F: L(x) \leq T\}$  by the following Stieltjes integral:

$$\eta(s) = \int_0^{\infty} e^{-sT} dN(T).$$

This enables us to apply the Ikehara–Wiener Tauberian Theorem.

**Theorem 3.8** (Ikehara–Wiener Tauberian Theorem ([11], Theorem 6.7)). *Suppose that the function  $\eta(s) = \int_0^{\infty} e^{-sT} dN(T)$  is analytic for  $\operatorname{Re}(s) > h$ , has a simple pole at  $s = h$ , and the function  $\eta(s) - c_0/(s-h)$  has an analytic extension to a neighbourhood of  $\operatorname{Re}(s) \geq h$ . Then*

$$N(T) \sim \frac{c_0 e^{hT}}{h}, \quad \text{as } T \rightarrow \infty.$$

As a consequence, we recover Guillopé's result that

$$\#\{x \in F: L(x) \leq T\} \sim c e^{hT} \quad \text{as } T \rightarrow \infty,$$

for some  $c > 0$ , and hence that the constant  $h$  defined by  $\beta(h) = 1$  is the volume entropy.

#### 4. A COMPLEX GENERATING FUNCTION

Suppose that  $\mathfrak{C}$  is a non-trivial conjugacy class of the free group  $F$ . Let  $\mathfrak{C}_n$  denote the set of  $x \in \mathfrak{C}$  such that  $|x| = n$  and suppose that

$$k = \min_{x \in \mathfrak{C}} |x|.$$

The set of  $g \in \mathfrak{C}_k$  are precisely those elements in  $\mathfrak{C}$  whose reduced word representations are cyclically reduced. In fact, if  $g = g_1 \cdots g_k \in \mathfrak{C}_k$  then all the remaining elements of  $\mathfrak{C}_k$  are given by cyclic permutations of the letters in  $g_1 \cdots g_k$ .

The reduced word representation of each  $x \in \mathfrak{C}$  takes the form

$$x = w^{-1}gw = w_m^{-1} \cdots w_1^{-1}g_1 \cdots g_k w_1 \cdots w_m$$

with  $g = g_1 \cdots g_k \in \mathfrak{C}_k$ ; and  $w = w_1 \cdots w_m \in W'_m$  subject to the restriction  $w_1 \neq g_1, g_k^{-1}$  in order that no pairwise cancellation occurs when concatenating  $w^{-1}$ ,  $g$  and  $w$ . For a given  $g = g_1 \cdots g_k$  we say that  $w \in W'_m(g)$  if  $w \in W'_m$  and  $w_1 \neq g_1, g_k^{-1}$ . Clearly  $\mathfrak{C}_n$  is non-empty if and only if  $n = k + 2m$  for  $m \in \mathbb{N}$ .

For  $x \in \mathfrak{C}$  with sufficiently long word length, the following lemma gives a useful decomposition of  $L(x)$ .

**Lemma 4.1.** *Suppose that  $x \in \mathfrak{C}$  such that  $x = w^{-1}gw$  here  $g \in \mathfrak{C}_k$  and  $w \in W'_m(g)$  with  $m \geq N - 1$ . Let  $y = w_1 \cdots w_{N-1}$  then*

$$L(x) = L(w^{-1}gw) = L(y^{-1}gy) - 2L(y) + 2L(w).$$

In order to study the counting problem in the conjugacy class  $\mathfrak{C}$ , we define a generating function

$$\eta_{\mathfrak{C}}(s) = \sum_{x \in \mathfrak{C}} e^{-sL(x)}.$$

We use the reduced word representation of elements in  $\mathfrak{C}$  to rewrite  $\eta_{\mathfrak{C}}(s)$  as follows:

$$\begin{aligned} \eta_{\mathfrak{C}}(s) &= \sum_{n=1}^{\infty} \sum_{x \in \mathfrak{C}_n} e^{-sL(x)} \\ &= \sum_{g \in \mathfrak{C}_k} \sum_{y \in W'_{N-1}(g)} e^{-s(L(y^{-1}gy) - 2L(y))} \sum_{m=N-1}^{\infty} \sum_{\substack{w \in W'_m \\ y=w_1 \cdots w_{N-1}}} e^{-2sL(w)} + \phi(s) \end{aligned}$$

where  $\phi(s)$  is an entire function. We may write  $\eta_{\mathfrak{C}}(s)$  in terms of  $A_{2s}$  by using

$$\sum_{m=N-1}^{\infty} \sum_{\substack{w \in W'_m \\ y=w_1 \cdots w_{N-1}}} e^{-2sL(w)} = \sum_{m=N-1}^{\infty} \mathbf{e}_y A_{2s}^m \mathbf{e}_1.$$

We will need the following classical result from linear algebra.

**Lemma 4.2.** *Let  $M$  be a  $d \times d$  matrix with real entries. Suppose that  $M$  has a simple eigenvalue  $\beta$ , and that  $u$  and  $v$  are the associated left and right eigenvectors, normalised so that  $u \cdot v = 1$ . Then  $w \in \mathbb{R}^d$  can be written  $w = (u \cdot w)v + \bar{v}$ , where  $\bar{v}$  is in the span of generalised right eigenvectors of  $M$  not associated to  $\beta$ .*

*Proof.* Let  $\{v\} \cup \mathcal{S}$  be a Jordan basis for  $M$ . Then  $\mathbb{R}^d = \mathbb{R}v \oplus \text{span}(\mathcal{S})$  and  $u$  is orthogonal to each element of  $\mathcal{S}$ . The result follows.  $\square$

**Proposition 4.3.** *The generating function  $\eta_{\mathfrak{C}}(s)$  is analytic for  $\text{Re}(s) > h/2$ , has a simple pole at  $s = h/2$  with positive residue and, apart from this, has an analytic extension to a neighbourhood of  $\text{Re}(s) \geq h/2$ .*

*Proof.* For  $\sigma \in \mathbb{R}$ ,  $A_{2\sigma}$  has spectral radius  $\beta(2\sigma)$ . Since this is strictly decreasing and  $\beta(h) = 1$ , it is clear that  $\sum_{m=N-1}^{\infty} \mathbf{e}_y^T A_{2\sigma}^m \mathbf{e}_1$  converges for  $\sigma > h/2$  and hence that, for  $s \in \mathbb{C}$ ,  $\eta_{\mathfrak{C}}(s)$  is analytic for  $\text{Re}(s) > h/2$ .

We now consider the analyticity of  $\eta(s)$  for  $s$  in a neighbourhood of  $h/2 + it \in \mathbb{C}$ , for an arbitrary  $t \in \mathbb{R}$ . We will let  $\text{spr}(M)$  denote the spectral radius of a matrix  $M$ . By Wielandt's Theorem (Theorem 2.4), either

- (1)  $\text{spr}(A_{h+2it}) = \text{spr}(B_h) = 1$ , in which case  $A_{h+2it}$  has a simple eigenvalue  $\beta(h + 2it)$  with  $|\beta(h + 2it)| = \beta(h) = 1$  and such that the remaining eigenvalues are strictly smaller in modulus; or



(2)  $\text{spr}(A_{h+2it}) < \text{spr}(B_h) = 1$ .

When (1) holds, standard eigenvalue perturbation theory gives that the simple eigenvalue  $\beta(2s)$  persists and is analytic for  $s$  in a neighbourhood of  $h + it$ , as are the corresponding left and right eigenvectors  $u(2s)$  and  $v(2s)$  [8]. By Lemma 4.2 and recalling that  $(u(2s))_1 = 1$ , we have  $\mathbf{e}_1 = v(2s) + \bar{v}(2s)$ , where  $\bar{v}(2s)$  is a vector in the subspace spanned by the generalised eigenvectors associated to the non-maximal eigenvalues of  $A_{2s}$ . Thus we have

$$\begin{aligned} \sum_{m=N-1}^{\infty} \mathbf{e}_y^T A_{2s}^m \mathbf{e}_1 &= \sum_{m=N-1}^{\infty} (\mathbf{e}_y \cdot v(2s)) \beta(2s)^m + \sum_{m=N-1}^{\infty} \mathbf{e}_y^T A_{2s}^m \bar{v}(2s) \\ &= \frac{(\mathbf{e}_y \cdot v(2s)) \beta(2s)^{N-1}}{1 - \beta(2s)} + \phi_2(s), \end{aligned}$$

where  $\phi_2(s)$  is analytic in a neighbourhood of  $h/2 + it$ . Therefore,  $\eta_{\mathfrak{C}}(s)$  is analytic in a neighbourhood of  $h/2 + it$  unless  $\beta(h + 2it) = 1$ . Lemma 3.7 tells us that this only occurs when  $t = 0$ . When (2) holds, we immediately obtain that  $\sum_{m=N-1}^{\infty} \mathbf{e}_y^T A_{2s}^m \mathbf{e}_1$  converges and hence that  $\eta_{\mathfrak{C}}(s)$  converges to an analytic function for  $s$  in a neighbourhood of  $h/2 + it$ .

For  $s$  in a neighbourhood of  $h/2$ , we have that, modulo an analytic function,

$$\eta_{\mathfrak{C}}(s) = \sum_{g \in \mathfrak{C}_k} \sum_{y \in W'_{N-1}(g)} e^{-s(L(y^{-1}gy) - 2L(y))} \frac{(\mathbf{e}_y \cdot v(2s)) \beta(2s)^{N-1}}{1 - \beta(2s)}.$$

From the analyticity of  $\beta$  and the fact that  $\beta(h) = 1$ , we obtain that, in a neighbourhood of  $h/2$ ,

$$\eta_{\mathfrak{C}}(s) = \frac{c}{s - h/2} + \phi_3(s),$$

where  $\phi_3(s)$  is analytic and  $c > 0$ . The latter holds because  $\mathbf{e}_y \cdot v(h) > 0$ , for each  $y \in \bigcup_{g \in \mathfrak{C}_k} W'_{N-1}(g)$ , and, by Lemma 3.5,  $-1/(2\beta'(h)) > 0$ .

Combining the above observations, we have that  $\eta_{\mathfrak{C}}(s)$  is analytic for  $\text{Re}(s) > h/2$  and, apart from a simple pole at  $s = h/2$ , has an analytic extension to a neighbourhood of  $\text{Re}(s) \geq h/2$ . Furthermore, the residue at the simple pole is positive.  $\square$

Since  $\eta_{\mathfrak{C}}(s)$  is related to the counting function  $N_{\mathfrak{C}}(T) = \#\{x \in \mathfrak{C} : d_{\mathcal{T}}(o, ox) \leq T\}$  via the Stieltjes integral

$$\eta_{\mathfrak{C}}(s) = \int_0^{\infty} e^{-sT} dN_{\mathfrak{C}}(T),$$

we may apply the Ikehara–Wiener Tauberian Theorem (Theorem 3.8) to conclude that

$$N_{\mathfrak{C}}(T) \sim C \frac{e^{hT/2}}{h/2},$$

for some  $C > 0$ . This completes the proof of Theorem 1.1.

## 5. ERROR TERMS

In this final section we discuss the error terms which may appear when estimating  $N_{\mathfrak{C}}(T)$ . We first note the following, which may be deduced from the analysis above and the arguments in [12] (Propositions 6 and 7).

**Proposition 5.1.** *There is never an exponential error term in Theorem 1.1, i.e. for no  $\epsilon > 0$  do we have  $N_{\mathfrak{C}}(T) = Ce^{hT/2} + O(e^{(h-\epsilon)T/2})$ .*

A more interesting problem is to ask when there is a polynomial error term for  $N_{\mathfrak{C}}(T)$ . Recall that an irrational number  $\alpha$  is said to be *Diophantine* if there exist  $c > 0$  and  $\beta > 1$  such that  $|q\alpha - p| \geq cq^{-\beta}$  for all  $p, q \in \mathbb{Z}$ ,  $q > 0$ . We have the following.

**Theorem 5.2.** *Suppose that  $G$  is a finite connected metric graph such that the degree of each vertex is at least 3. Suppose also that  $G$  contains two closed geodesics  $\gamma$  and  $\gamma'$  such that  $l(\gamma)/l(\gamma')$  is Diophantine. Then there exists  $\delta > 0$  such that*

$$N_{\mathfrak{C}}(T) = Ce^{hT/2} + O(e^{hT/2}T^{-\delta}).$$

Note that, since the lengths of the closed geodesics in  $G$  are not contained in a discrete subgroup of  $\mathbb{R}$  if and only if there are two closed geodesics the ratio of whose lengths is irrational, the hypothesis of Theorem 5.2 is strictly stronger than that of Theorem 1.1.

The crucial new ingredient is to obtain a bound on powers of the matrix  $A_s$ , for  $\operatorname{Re}(s)$  close to  $h$  and  $\operatorname{Im}(s)$  away from zero. To do this, we use the work of Dolgopyat [2] on transfer operators, where it is a key ingredient in studying the mixing rate of flows (cf. also [13]). Since  $r : \Sigma \rightarrow \mathbb{R}$  is locally constant, the transfer operator  $L_{-sr}$ , defined pointwise by

$$(L_{-sr}\psi)(x) = \sum_{\sigma y=x} e^{-sr(y)}\psi(y),$$

acts on the space of complex valued Hölder continuous functions with exponent  $\alpha$ , for any  $\alpha > 0$ . For definiteness, we will consider the action on the Banach space of Lipschitz functions,  $L_{-sr} : C_{\text{Lip}}(\Sigma, \mathbb{C}) \rightarrow C_{\text{Lip}}(\Sigma, \mathbb{C})$ , with the norm  $\|f\| = \|f\|_{\infty} + |f|_{\text{Lip}}$ , where

$$|f|_{\text{Lip}} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)}.$$

Let  $V_{N-1} \subset C_{\text{Lip}}(\Sigma, \mathbb{C})$  denote the finite dimensional subspace consisting of locally constant functions depending on the first  $N-1$  coordinates. This has dimension  $D := \#W'_{N-1}$ . Then  $L_{-sr} : V_{N-1} \rightarrow V_{N-1}$  and  $\|L_{-sr}|_{V_{N-1}}\| \leq \|L_{-sr}\|$ . Furthermore, the restriction of  $L_{-sr}$  to  $V_{N-1}$  can be identified with the action of  $B_s$  on  $\mathbb{C}^D$ . We can therefore use the results of [2] on bounding the norm of iterates of transfer operators to give a bound on the norm of the iterates of  $B_s$  and hence  $A_s$ .

For the remainder of this section, it will be convenient to write  $s = \varsigma + it$ .

**Proposition 5.3.** *Under the hypotheses of Theorem 5.2, there exist constants  $C_1, C_2 > 0$ ,  $\tau > 0$  and  $t_1 \geq 1$  such that when  $|t| \geq t_1$  and  $m \geq 1$ ,*

$$\|A_{\varsigma+it}^{2\nu m}\| \leq C_1 |t| \beta(\varsigma)^{2\nu m} (1 - |t|^{-\tau})^{m-1},$$

where  $\nu = \lfloor C_2 \log |t| \rfloor$ .

We will also use the elementary inequality (cf. Proposition 2.1 of [11]):

**Lemma 5.4.** *There exists a constant  $D_1 > 0$ , independent of  $t$  and uniform in  $\varsigma$ , such that, for all  $n > 0$ , we have*

$$\|A_{\varsigma+it}^n\| \leq \beta(\varsigma)^n (D_1 |t| + 2^{-n}).$$

*Proof.* Let  $\phi$  denote the strictly positive eigenfunction of the operator  $L_{-sr}$  associated to the eigenvalue  $\beta(\varsigma)$ , guaranteed by the Ruelle-Perron-Frobenius Theorem for this operator (see [11], Theorem 2.2). Let

$$g = -sr + \log \phi - \log \phi \circ \sigma - \log \beta(\varsigma).$$

Then  $L_g 1 = 1$  and it follows from Proposition 2.1 of [11] that

$$|L_{g-it}^n \psi|_{\text{Lip}} \leq D_0 |t| \|\psi\|_{\infty} + 2^{-n} |\psi|_{\text{Lip}},$$

for some  $D_0 > 0$  independent of  $t$  and uniform in  $\varsigma$ , and

$$\|L_{g-it\tau}^n \psi\|_\infty \leq \|\psi\|_\infty.$$

(The only difference from the statement given in [11] is the appearance of the term  $|t|$  and the uniformity of  $D_0$  but this follows from an inspection of the proof.) In particular, we have

$$\|L_{g-it\tau}^n \psi\| \leq (D_1 |t| + 2^{-n}) \|\psi\|.$$

for some  $D_1 > 0$  independent of  $t$  and uniform in  $\sigma$ . Since

$$L_{-(\varsigma+it)r} = \beta(\varsigma) \Delta_\phi L_{g-it\tau} \Delta_\phi^{-1},$$

where  $\Delta_\phi$  is the multiplication operator  $\Delta_\phi(\psi) = \phi\psi$ , the result follows.  $\square$

Combining Proposition 5.3 and Lemma 5.4 gives the following bound on  $\|A_s^n\|$  for all  $n \in \mathbb{N}$ .

**Proposition 5.5.** *Let  $n = 2\nu m + l$  where  $m = \lfloor \frac{n}{2\nu} \rfloor$  and  $l \in \{0, \dots, 2\nu - 1\}$ , then, for  $|t| \geq t_1$*

$$\|A_{\varsigma+it}^n\| \leq C_3 |t|^2 \beta(\varsigma)^n (1 - |t|^{-\tau})^{m-1},$$

for some  $C_3 > 0$  independent of  $t$  and uniform in  $\varsigma$ .

We use this to study the analyticity of  $\eta_{\mathfrak{C}}(s)$  (using a simpler version of the arguments in [13]).

**Proposition 5.6.** *There exist constants  $\rho > 0$  and  $t_2 \geq t_1$  such that  $\eta_{\mathfrak{C}}(s)$  has an analytic extension to the region*

$$\mathcal{R}(\rho) = \{\varsigma + it \in \mathbb{C} : 2\varsigma > h - 1/|2t|^\rho, |t| \geq t_2\},$$

where it satisfies the bound  $|\eta_{\mathfrak{C}}(\varsigma + it)| = O(|t|^{2+\rho})$ .

*Proof.* Suppose that  $\varsigma + it \in \mathbb{C}$  satisfies  $2\varsigma > h - |2t|^{-\rho}$  and  $|2t| \geq t_1$ . Take  $\rho > \tau$  (with  $\tau$  the constant from Proposition 5.3). Since

$$\beta(2\varsigma) = 1 + \beta'(h)(2\varsigma - h) + O((2\varsigma - h)^2)$$

with  $\beta'(h) < 0$ , for  $|t| \geq t_2$ , where  $t_2 \geq t_1$ , we have  $(1 - |2t|^{-\tau})^{1/2\nu} < \beta(2\varsigma)^{-1}$ .

We can estimate  $|\eta_{\mathfrak{C}}(s)| \leq M \sum_{m=1}^{\infty} |A_{2\varsigma}^m \mathbf{e}_1|$ , for some  $M > 0$ . Thus, for  $\varsigma + it \in \mathcal{R}(\rho)$ , we have, by Proposition 5.5,

$$\begin{aligned} \sum_{m=1}^{\infty} |A_{2(\varsigma+it)}^m \mathbf{e}_1| &\leq \sum_{m=1}^{\infty} C_3 |2t|^2 \beta(2\varsigma)^m (1 - |2t|^{-\tau})^{\lfloor m/2\nu \rfloor - 1} \\ &\leq \frac{C_3 |2t|^2 \beta(2\varsigma)}{(1 - |2t|^{-\tau})^{2-2/\nu} (1 - \beta(2\varsigma) (1 - |2t|^{-\tau})^{1/2\nu})} \\ &\leq \frac{C_4 |2t|^2}{1 - \beta(2\varsigma) (1 - |2t|^{-\tau})^{1/2\nu}} = O(|t|^{2+\rho}), \end{aligned}$$

which shows that  $\eta_{\mathfrak{C}}(s)$  is analytic in the desired region and gives the bound.  $\square$

Let  $\xi_{\mathfrak{C}}(s)$  be the normalised generating function given by

$$\xi_{\mathfrak{C}}(s) = \eta_{\mathfrak{C}}(sh/2) = \sum_{x \in \mathfrak{C}} e^{-shL(x)/2}.$$

We immediately deduce that  $\xi_{\mathfrak{C}}(s)$  is analytic in the half-plane  $\operatorname{Re}(s) > 1$ , has an analytic extension to a neighbourhood of  $\operatorname{Re}(s) \geq 1$  apart from the simple pole at  $s = 1$ , which has positive residue, and, furthermore, that there exist positive constants  $\rho$  and  $t_3 = 2t_2/h$  such that  $\xi_{\mathfrak{C}}(s)$  has an analytic extension to

$$\mathcal{R}_\xi(\rho) = \left\{ \varsigma + it : \varsigma > 1 - \frac{1}{h^{\rho+1}|t|^\rho}, |t| > t_3 \right\};$$
 and

where it satisfies  $|\xi_{\mathfrak{C}}(s)| = O(|t|^{2+\rho})$ .

Let us introduce a normalised counting function

$$\psi_0(T) = \sum_{e^{hL(x)/2} \leq T} 1.$$

Adapting the arguments of [13], we will establish an error term for  $\psi_0(T)$ , from which Theorem 5.2 will follow since  $\psi_0(e^{hT/2}) = N_{\mathfrak{C}}(T)$ . We introduce the following family of auxiliary functions. Let  $\psi_1(T) = \int_1^T \psi_0(u) du$  and continue inductively so that

$$\psi_k(T) = \int_1^T \psi_{k-1}(u) du = \frac{1}{k!} \sum_{e^{hL(x)/2} \leq T} (T - e^{hL(x)/2})^k.$$

We use the following identity ([7], Theorem B, page 31) to connect the functions  $\xi_{\mathfrak{C}}(s)$  and  $\psi_k(T)$ . If  $k$  is a positive integer and  $d > 0$  we have

$$\frac{1}{2\pi i} \int_{d-i\infty}^{d+i\infty} \frac{y^s}{s(s+1)\cdots(s+k)} ds = \begin{cases} 0 & \text{if } 0 < y < 1, \\ \frac{1}{k!} (1 - 1/y)^k & \text{if } y \geq 1. \end{cases}$$

This gives us the following.

**Lemma 5.7.** *For  $d > 1$  we may write*

$$\psi_k(T) = \frac{1}{2\pi i} \int_{d-i\infty}^{d+i\infty} \frac{\xi_{\mathfrak{C}}(s) T^{s+k}}{s(s+1)\cdots(s+k)} ds.$$

We briefly outline the method to approximate the integral for  $\psi_k(T)$ . First, compare the integral for  $\psi_k(T)$  to the truncated integral on the line segment  $[d - iR, d + iR]$ , where  $R = (\log T)^\varepsilon$  and  $0 < \varepsilon < 1/\rho$ . Let us choose  $d = 1 + 1/\log T$  and then since  $\xi_{\mathfrak{C}}(d) = O((d-1)^{-1})$ , we deduce

$$\left| \psi_k(T) - \frac{1}{2\pi i} \int_{d-iR}^{d+iR} \frac{\xi_{\mathfrak{C}}(s) T^{s+k}}{s(s+1)\cdots(s+k)} ds \right| = O\left(\frac{(\log T) T^{k+1}}{R^{k+1}}\right) = O\left(\frac{T}{(\log T)^{k\varepsilon}}\right).$$

We evaluate the truncated integral using Cauchy's Residue Theorem. Consider a closed contour  $\Gamma \cup [d - iR, d + iR]$ . Here  $\Gamma$  is the union of the line segments  $[d + iR, c(R) + iR]$ ,  $[c(R) - iR, d - iR]$  and  $[c(R) + iR, c(R) - iR]$ , where

$$c(R) = 1 - \frac{1}{2h\rho+1 R^\rho},$$

so that  $\Gamma$  lies in  $\mathcal{R}_{\xi}(\rho)$ . Note that  $\Gamma \cup [d - iR, d + iR]$  encloses the simple pole of  $\xi_{\mathfrak{C}}(s)$  at  $s = 1$ . Cauchy's Residue Theorem gives that, for some  $c > 0$ ,

$$\frac{1}{2\pi i} \int_{d-iR}^{d+iR} \frac{\xi_{\mathfrak{C}}(s) T^{s+k}}{s(s+1)\cdots(s+k)} ds = \frac{cT^{k+1}}{(k+1)!} + \frac{1}{2\pi i} \int_{\Gamma} \frac{\xi_{\mathfrak{C}}(s) T^{s+k}}{s(s+1)\cdots(s+k)} ds.$$

We consider the contribution made by each of the line segments in  $\Gamma$ . First, integrating over the interval  $[d + iR, c(R) + iR]$ , we have

$$\begin{aligned} \left| \int_{d+iR}^{c(R)+iR} \frac{\xi_{\mathfrak{C}}(s) T^{s+k}}{s(s+1)\cdots(s+k)} ds \right| &\leq \frac{T^{d+k}}{R^{k+1}} \left| \int_{d+iR}^{c(R)+iR} \xi_{\mathfrak{C}}(s) ds \right| \\ &= O\left(\frac{T^{d+k}}{(\log T)^{(k-\rho-1)\varepsilon}}\right) \end{aligned}$$

and similarly, for  $[c(R) - iR, d - iR]$ , we have

$$\left| \int_{c(R)-iR}^{d-iR} \frac{\xi_{\mathfrak{C}}(s) T^{s+k}}{s(s+1)\cdots(s+k)} ds \right| = O\left(\frac{T^{d+k}}{(\log T)^{(k-\rho-1)\varepsilon}}\right).$$

We estimate the modulus of the integral along  $[c(R) + iR, c(R) - iR]$  by

$$\begin{aligned} T^{c(R)+k} \left| \int_{c(R)+iR}^{c(R)-iR} \frac{\xi_{\mathfrak{c}}(s)}{s(s+1)\cdots(s+k)} ds \right| &= O \left( T^{c(R)+k} \int_1^R t^{1+\rho-k} dt \right) \\ &= O(T^{c(R)+k} R^{2+\rho-k}), \end{aligned}$$

which means for any positive  $\gamma$  we have

$$T^{c(R)+k} R^{2+\rho-k} = T^{k+1} e^{-\frac{\log T}{2h\rho+1(\log T)^{\varepsilon\rho}} (\log T)^{(2+\rho-k)\varepsilon}} = O(T^{k+1} (\log T)^{-\gamma}).$$

Together the integral estimates give us an error term for  $\psi_k(T)$ :

$$\psi_k(T) = c' T^{k+1} + O \left( \frac{T^{k+1}}{(\log T)^{(k-\rho-1)\varepsilon}} \right)$$

where  $c' > 0$ . Then repeatedly applying the inequality

$$\psi_{j-1}(T - \Delta T) \Delta T \leq \psi_j(T) - \psi_j(T - \Delta T) \leq \psi_{j-1}(T) \Delta T,$$

where

$$\Delta T = T(\log T)^{-(k-\rho-1)2^{j-k-1}\varepsilon},$$

we obtain

$$\psi_0(T) = CT + O(T(\log T)^{-\delta}),$$

where  $C, \delta > 0$ . Thus we have the error term  $N_{\mathfrak{c}}(T) = Ce^{hT/2} + O(e^{hT/2} T^{-\delta})$ , completing the proof of Theorem 5.2.

#### REFERENCES

- [1] A. Broise-Alamichel, J. Parkkonen and F. Paulin, *Counting paths in graphs*, in preparation.
- [2] D. Dolgopyat, *Prevalence of rapid mixing in hyperbolic flows*, Ergodic Theory Dynam. Systems 18, 1097-1114, 1998.
- [3] F. Douma, *A lattice point problem on the regular tree*, Discrete Math. 311, 276-281, 2011.
- [4] F. Gantmacher, *The Theory of Matrices, Volume II*, AMS Chelsea Publishing, Providence RI, 2000.
- [5] L. Guillopé, *Entropies et spectres*, Osaka J. Math. 31, 247-289, 1994.
- [6] H. Huber, *Zur analytischen Theorie hyperbolischer Raumformen und Bewegungsgruppen II*, Math. Ann. 142, 385-398, 1961.
- [7] A. Ingham, *The distribution of prime numbers*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1990.
- [8] T. Kato, *Perturbation theory for linear operators*, Reprint of the 1980 edition, Classics in Mathematics, Springer-Verlag, Berlin, 1995.
- [9] J. Parkkonen and F. Paulin, *On the hyperbolic orbital counting problem in conjugacy classes*, Math. Z. 279, 1175-1196, 2015.
- [10] W. Parry, *An analogue of the prime number theorem for closed orbits of shifts of finite type and their suspensions*, Israel J. Math. 45, 41-52, 1983.
- [11] W. Parry and M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérisque 187-188, Société Mathématique de France, 1990.
- [12] M. Pollicott, *Meromorphic extensions of generalised zeta functions*, Invent. math. 85, 147-164, 1986.
- [13] M. Pollicott and R. Sharp, *Error terms for closed orbits of hyperbolic flows*, Ergodic Theory Dynam. Systems 21, 545-562, 2001.
- [14] R. Sharp, *Distortion and entropy for automorphisms of free groups*, Discrete Contin. Dynam. Systems 26, 347-363, 2009.
- [15] R. Sharp, *Comparing length functions on free groups*, in "Spectrum and Dynamics", CRM Conference Proceedings & Lecture Notes 52, 185-207, 2010.

MATHEMATICS INSTITUTE, UNIVERSITY OF WARWICK, COVENTRY CV4 7AL, U.K.  
*E-mail address:* G.Kenison@warwick.ac.uk

MATHEMATICS INSTITUTE, UNIVERSITY OF WARWICK, COVENTRY CV4 7AL, U.K.  
*E-mail address:* R.J.Sharp@warwick.ac.uk