

GROWTH SERIES FOR THE COMMUTATOR SUBGROUP

MARK POLLICOTT AND RICHARD SHARP

(Communicated by Ronald M. Solomon)

Dedicated to William Parry on the occasion of his 60th birthday

ABSTRACT. In this paper we study growth series associated to commutator subgroups of surface groups. We conclude that these growth series are not rational.

0. INTRODUCTION

In the last few years there has been considerable interest in the growth series for the word length of surface groups G relative to a finite symmetric set of generators G_0 . If for $h \in G$ we let $|h|$ denote the word length of h (i.e. the smallest number of generators from G_0 required in writing h) and write $N(n) = \#\{h \in G : |h| = n\}$, then the growth series is given by

$$f(t) = \sum_{n=1}^{\infty} N(n)t^n.$$

Such growth series have been studied by Cannon [1], Floyd and Plotnick [2], Parry [5] and others. It is known that the function $f(t)$ is rational and, moreover, the number of elements with word length equal to n has purely exponential asymptotic growth, i.e., $N(n) \sim C\lambda^n$ for some $C > 0$ and $\lambda > 1$.

Gromov has studied a modified growth series which counts conjugacy classes for surface groups (and more general groups), and has shown that it is rational [4, §8.5].

The purpose of this note is to study a growth series for those elements $h \in G$ which lie in the commutator subgroup $[G, G] \triangleleft G$. We shall conclude that the results for this series are very different to those for $f(t)$ above.

We shall concentrate on the following important example. Let G be the fundamental group of a compact orientable surface of genus $g \geq 2$ with the following presentation:

$$G = \left\langle a_1, \dots, a_g, b_1, \dots, b_g : \prod_{i=1}^g [a_i, b_i] = 1 \right\rangle,$$

Received by the editors April 1, 1994.

1991 *Mathematics Subject Classification*. Primary 20F05, 58F15.

The first author was supported by The Royal Society through a University Research Fellowship.

The second author was supported by the UK SERC under grant number GR/G51930.

so that $G_0 = \{a_1^{\pm 1}, \dots, a_g^{\pm 1}, b_1^{\pm 1}, \dots, b_g^{\pm 1}\}$. We denote $M(n) = \#\{h \in [G, G] : |h| = n\}$ and then introduce the commutator growth series

$$f_0(t) = \sum_{n=1}^{\infty} M(n)t^n.$$

Our main result is the following.

Theorem. (i) *There exists $C > 0$ such that $M(2n) \sim C \frac{\lambda^{2n}}{n^g}$ as $n \rightarrow +\infty$; and*

(ii) *$M(2n + 1) = 0$ for all $n \geq 1$.*

In particular, the growth series $f_0(t)$ is not rational (and has an essential singularity at $t = \frac{1}{\lambda}$).

1. THE ESSENTIAL SINGULARITY OF $f_0(t)$

In this short section we shall show that the asymptotic formula for $M(n)$ in the theorem immediately implies the final conclusion of the theorem that $f_0(t)$ must have an essential singularity at $t = \frac{1}{\lambda}$.

By part (i) of the theorem we can choose a constant $c > 1$ such that $\frac{1}{c} \frac{\lambda^{2n}}{n^g} \leq M(2n) \leq c \frac{\lambda^{2n}}{n^g}$ for $n \geq 1$. If we denote $k_g(t) := \sum_{n=1}^{\infty} n^g M(n)t^n$, then the series expansion $-\log(1 - z) = \sum_{n=1}^{\infty} \frac{z^n}{n}$ allows us to bound

$$\begin{aligned} \frac{1}{c} (-\log(1 - (t\lambda)^2)) &= \frac{1}{c} \sum_{n=1}^{\infty} \frac{\lambda^{2n}}{n} t^{2n} \\ &\leq k_{g-1}(t) \\ &\leq c \sum_{n=1}^{\infty} \frac{\lambda^{2n}}{n} t^{2n} \\ &= c (-\log(1 - (t\lambda)^2)). \end{aligned}$$

In particular, we see that $k_{g-1}(t)$ has a (non-algebraic) essential singularity at $t = \frac{1}{\lambda}$.

We observe that $k_{g-1}(t) = (t \frac{d}{dt}) k_{g-2}(t)$, and by induction

$$k_{g-1}(t) = \left(t \frac{d}{dt}\right)^{g-1} k_0(t).$$

It follows that $k_0(t)$ also has a (non-algebraic) essential singularity at $t = \frac{1}{\lambda}$, since otherwise it would contradict the corresponding conclusion for $k_{g-1}(t)$ above. Finally, we observe that $f_0(t) = k_0(t)$, and so our claim is proved.

2. AUTOMATA AND THE Δ INVARIANT

We call an *automaton* for the group G (with the standard presentation) the following object:

- (a) a directed graph (V, E) with vertex set V (including a distinguished vertex $*$) and edge set E (with no edges terminating at $*$); and
- (b) a labelling map $\rho : E \rightarrow G_0$

such that

- (i) there is a bijection between finite paths in the graph starting at the vertex $*$ and passing through the consecutive edges e_1, \dots, e_n , say, and elements $h \in G$ given by the correspondence $h = \rho(e_1) \dots \rho(e_n)$; and
- (ii) the word length $|h|$ is equal to the path length n .

The existence of such an automaton can be found in [7] (and in a more general context in [1]).

It will be helpful to consider the realization of the group G as a Fuchsian group, i.e., a discrete group of isometries of the Poincaré disc, and the coding of its limit set (as a subshift of finite type) which can be found in [7].

Let A denote the incidence matrix of the graph G , i.e., A is a square matrix indexed by V with $A(i, j)$ denoting the number of edges from i to j . An important property of the matrix A is that the submatrix B formed by restricting A to the entries indexed by $V - \{*\}$ is aperiodic. By the Perron-Frobenius theorem, B has a simple positive maximal eigenvalue λ and the other eigenvalues are all contained in the set $\{z \in \mathbb{C} : |z| < \lambda\}$. Furthermore, it is easy to see that the eigenvalues of the matrix A are precisely 0 and the eigenvalues of B .

The existence of the automaton means that the counting function $N(n)$ can be written in the form

$$N(n) = \sum_{j \in V} A^n(*, j).$$

In particular this implies that the growth series $f(t)$ is a rational function and its smallest pole is $\frac{1}{\lambda}$.

Remark. The referee pointed out to us that the asymptotic formula in Theorem 1 has an interpretation similar in spirit to the theory of random walks on the group \mathbb{Z}^{2g} . Define a map $\psi : G \rightarrow \mathbb{Z}^{2g}$ by setting $\psi(a_i^{\pm 1}) = \pm e_i$ and $\psi(b_i^{\pm 1}) = \pm e_{g+i}$, where $e_i \in \mathbb{Z}^{2g}$ is the vector having a one in the i -th position and zeros elsewhere. If one considers all paths of length n , say, on the graph (V, E) , then we can take the image under $\psi : G \rightarrow \mathbb{Z}^{2g}$ of the corresponding group elements $h \in G$. The evolution of the images $\psi(h)$ from the case $|h| = n$ to the case $|h| = n + 1$ is reminiscent of a random walk on \mathbb{Z}^{2g} (although many of the usual trajectories in the standard random walk will be absent due to the restrictions imposed by (V, E)). Our result could be reformulated in terms of counting the number of such trajectories for which $\psi(h)$ “returns” to $0 \in \mathbb{Z}^{2g}$. If one considers the “standard” random walk on \mathbb{Z}^{2g} , where the transition probabilities are given by $P(n, m) = 1/4g$ if $n - m = \pm e_i$, for some $1 \leq i \leq 2g$, and zero otherwise, then it is a standard result that the probability of returning to 0 in $2n$ steps is comparable to n^{-g} [8].

We can define a homomorphism $\alpha : G \rightarrow \mathbb{Z}^{2g}$ where:

- (a) $\alpha_i(h) \in \mathbb{Z}$ ($i = 1, \dots, g$) is the number of occurrences of a_i minus the number of occurrences of a_i^{-1} in the edge labelling of the associated path; and
- (b) $\alpha_i(h) \in \mathbb{Z}$ ($i = g + 1, \dots, 2g$) is the number of occurrences of b_i minus the number of occurrences of b_i^{-1} .

Note that $h \in [G, G]$ if and only if $\alpha(h) = 0$.

For each $t \in \mathbb{T}^{2g}$ we define a matrix A_t by $A_t(i, j) = A(i, j)e^{2\pi it \cdot \alpha(i, j)}$ for $i, j \in V$. If we recall the simple orthogonality identity

$$\int_{\mathbb{T}^{2g}} e^{2\pi it \cdot y} dt = \begin{cases} 1 & \text{for } y = 0, \\ 0 & \text{for } y \in \mathbb{Z}^{2g} - \{0\}, \end{cases}$$

then we see that we can write

$$M(n) = \sum_{|h|=n} \int_{\mathbb{T}^{2g}} e^{2\pi i t \cdot \alpha(h)} dt = \sum_{j \in V} \int_{\mathbb{T}^{2g}} A_t^n(*, j) dt.$$

Similarly, we can define a matrix B_t by $B_t(i, j) = B(i, j)e^{2\pi i t \cdot \alpha(\rho(i, j))}$. We shall denote the eigenvalues of B_t by $\lambda_j(t)$, $j = 1, \dots, k$ (where $k = |V| - 1$), arranged so that $|\lambda_1(t)| \geq |\lambda_2(t)| \geq \dots \geq |\lambda_k(t)|$. We observe that the eigenvalues of A_t are precisely given by $\{\lambda_1(t), \dots, \lambda_k(t), 0\}$.

We require the following standard result on matrices.

Wielandt's theorem [3, p.57]. *The eigenvalues of the matrix B_t must satisfy the inequality $|\lambda_j(t)| \leq \lambda$ for $j = 1, \dots, k$. Furthermore, $\lambda_1(t) = e^{2\pi i \theta} \lambda$ if and only if $B_t = e^{2\pi i \theta} DBD^{-1}$, where D is a diagonal matrix whose diagonal entries are complex numbers of modulus unity. In particular, if $|\lambda_1(t)| = \lambda$, then $\lambda_1(t)$ is simple.*

Assume that for $t_1 \in \mathbb{T}^{2g}$ we have that A_{t_1} has a maximal eigenvalue of modulus λ . We shall now show that there are only finitely many such t_1 . To do this, we shall introduce the following important object, familiar in the coding theory of subshifts of finite type. We let $\Delta \subset \mathbb{Z}^{2g}$ denote the subgroup defined by

$$\Delta = \{ \alpha(\rho(i_0, i_1) \dots \rho(i_{n-1}, i_n)) - \alpha(\rho(j_0, j_1) \dots \rho(j_{n-1}, j_n)) : \\ i_k, j_k \in V - \{*\}, \quad k = 0, \dots, n, \quad i_0 = i_n, j_0 = j_n, n \in \mathbb{N} \}.$$

(The fact that Δ is closed under addition is an easy consequence of the aperiodicity of B .)

We claim that $\Delta = S$, where we define

$$S := \{ (n_1, \dots, n_{2g}) : n_1 + \dots + n_{2g} \in 2\mathbb{Z} \}.$$

First we shall show that $\Delta \supset S$. Observe that it suffices to show that Δ contains the points

$$\{ (0, \dots, 0, n_i, n_{i+1}, 0, \dots, 0) : i = 0, \dots, 2g - 1, \text{ with } n_i + n_{i+1} \in 2\mathbb{Z} \}.$$

We want to appeal to some features of an explicit automaton for this presentation of the group. We recall the construction of Series (cf. [7]) in which a fundamental domain for the surface of genus g in the Poincaré disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ is the exterior to $4g$ isometric circles. The endpoints of these arcs on the unit circle divide the circle into $8g$ arcs, which we label in order as I_1, I_2, \dots, I_{8g} . We can assume that the indexing has been chosen so that the *even* indices correspond to arcs interior to a single isometric circle and the *odd* indices correspond to arcs interior to two isometric circles. We associate to the even indices i the element $g_i \in G_0$ with the associated isometric circle, and to the odd indices i the element $g_i \in G_0$ associated to the furthest clockwise of the two isometric circles.

The associated automaton can now be expressed as follows. We take the vertices to be the pairs (i, j) for which $I_j \subset g_i(I_i)$, and we allow a directed edge from (i, j) to (k, l) (labelled by $g_j \in G_0$) if $I_k \subset g_j(I_j)$. The property that we need to observe

is that for i even, since $I_i \subset g_i(I_i)$ we have both that (i, i) is a vertex and there is an edge from this vertex to itself labelled by g_i .

Returning to our consideration of Δ , let $a_1, a_2 \in G_0$. Then since $\alpha(a_1) = (1, 0, \dots, 0)$ and $\alpha(a_2) = (0, 1, 0, \dots, 0)$ we see by the above observations that $\alpha(a_1) - \alpha(a_2) = (1, -1, 0, \dots, 0) \in \Delta$. Similarly, since $\alpha(a_1) = (1, 0, \dots, 0)$ and $\alpha(a_1^{-1}) = (-1, 0, \dots, 0)$ $(2, 0, \dots, 0) \in \Delta$. This is enough to show that Δ contains the points $(n_1, n_2, 0, \dots, 0)$, with $n_1 + n_2 \in 2\mathbb{Z}$. The other cases are similar, and so we see that $S \subset \Delta$.

To see the reverse inclusion, we need only observe that

$$\Delta \subset \left\{ (x_1 - y_1, \dots, x_{2g} - y_{2g}) : \begin{array}{l} x_i, y_i \in \mathbb{Z}, i = 1, \dots, 2g, \text{ and } \sum_{i=1}^{2g} |x_i| = \sum_{i=1}^{2g} |y_i| \end{array} \right\} \subset S.$$

We are now in a position to conclude that there are only two values of $t \in \mathbb{T}^{2g}$ such that A_t has an eigenvalue of modulus λ . By Proposition 3 of [6], the set of such t can be identified with $\Delta^\perp = \widehat{\mathbb{Z}^{2g}/\Delta}$, and this has cardinality 2.

3. THE ASYMPTOTIC FORMULA

From the last section we conclude that for the two values $t = 0$ or $t = t_1$ we have that A_t has an eigenvalue $\lambda_1(t)$ of maximal modulus with $|\lambda_1(t)| = \lambda$. We choose a (small) neighbourhood U_0 of $0 \in \mathbb{T}^{2g}$ and then write $U_1 = U_0 + t_1$.

Using analytic perturbation theory for the spectrum of A_t and Wielandt's theorem, we can assume that for U_0 sufficiently small we can choose $\epsilon > 0$ so that $|\lambda_1(t)| \geq \lambda - \epsilon$ and $|\lambda_i(t)| \leq \lambda - \epsilon$, for $i = 2, \dots, k$ and $t \in U_0 \cup U_1$, and so that $|\lambda_i(t)| \leq \lambda - \epsilon$ for $i = 1, \dots, k$ and $t \in \mathbb{T}^{2g} - (U_0 \cup U_1)$. We associate to $\lambda_1(t)$ a left eigenvector v_t and a right eigenvector w_t for A_t (both normalized to have unit length).

Lemma 1. For $t \in U_1$, $\lambda(t) = -\lambda(t - t_1)$, $v_t = v_{t-t_1}$, $w_t = w_{t-t_1}$.

Proof. The result becomes apparent upon the simple observation that $A_t^2 = A_{t-t_1}^2$, cf. Lemma 3 of [6].

These observations allow us to write

$$\begin{aligned} (3.1) \quad M(n) &= \sum_{|h|=n} \int_{\mathbb{T}^{2g}} e^{2\pi i t \cdot \alpha(h)} \\ &= \sum_{j \in V} \int_{\mathbb{T}^{2g}} A_t^n(*, j) dt \\ &= \int_{\mathbb{T}^{2g}} \xi A_t^n \mathbf{1} dt \end{aligned}$$

where $\xi = (1, 0, 0, \dots, 0)$ with the 1 occurring in position $*$ and $\mathbf{1} = (1, 1, \dots, 1)^T$.

For $t \in U_0 \cup U_1$ we shall write $\xi = a_t v_t + v'_t$ and $\mathbf{1} = b_t w_t + w'_t$, where $v'_t \in (\mathbb{R}v_t)^\perp$ and $w'_t \in (\mathbb{R}w_t)^\perp$. Note that by analytic perturbation theory the quantities

a_t, b_t, v_t, w_t vary analytically in t . Substituting these identities into (3.1), we have

$$\begin{aligned}
 (3.2) \quad M(n) &= \int_{U_0} ((a_t b_t) v_t w_t (\lambda_1(t))^n + v'_t A_t^n w'_t) dt \\
 &\quad + \int_{U_1} ((a_t b_t) v_t w_t (\lambda_1(t))^n + v'_t A_t^n w'_t) dt \\
 &\quad + \int_{\mathbb{T}^{2g} - (U_0 \cup U_1)} \xi A_t^n \mathbf{1} dt \\
 &= \int_{U_0} (1 + (-1)^n) (\lambda_1(t))^n a_t b_t v_t w_t dt + O((\lambda - \epsilon)^n)
 \end{aligned}$$

by the estimates following Lemma 1.

To complete the asymptotic estimate for $M(n)$ we need the following result.

Proposition 1. *There exists a constant $C > 0$ such that*

$$\int_{U_0} (\lambda_1(t))^n a_t b_t v_t w_t dt = C \lambda^n \frac{\Gamma(n+1)}{\Gamma(n+g+1)} \left(1 + O\left(\frac{1}{n^{1/2}}\right) \right).$$

Proof. Noting that $a_t b_t v_t w_t = a_0 b_0 v_0 w_0 + O(|t|)$, this will follow from the calculations in [6] once we have established the results on the derivatives of $\lambda_1(t)$ given in Lemma 4 of [6]. These results are an immediate consequence of the following lemma.

Lemma 2. *The function $\alpha \circ \rho : E \rightarrow \mathbb{Z}^{2g}$ satisfies*

$$\sum_{i,j \in V} \frac{v_0(j)(\alpha \circ \rho)(i,j) A(i,j)}{\lambda} = 0$$

Proof of Lemma 2. Formulated in terms of subshifts of finite type, this summation is merely $\int (\alpha \circ \rho) d\mu$, where μ is the measure of maximal entropy for the subshift associated to the transition matrix A [6]. It then becomes clear that we can write $\int (\alpha \circ \rho) d\mu = \lim_{n \rightarrow +\infty} \frac{1}{N(n)} \sum_{|h|=n} \frac{\alpha(h)}{n}$. However, noting that $\alpha(h) = -\alpha(h^{-1})$ and $|h| = |h^{-1}|$, we see that each of these summations vanishes, giving the required result.

Substituting the estimate in Proposition 1 into the identity (3.2) gives that

$$M(n) = (1 - (-1)^n) C \lambda^n \frac{\Gamma(n+1)}{\Gamma(n+g+1)} \left(1 + O\left(\frac{1}{n^{1/2}}\right) \right)$$

Since by Stirling’s formula $\frac{\Gamma(n+1)}{\Gamma(n+g+1)} \sim \frac{e^{-(g+\frac{1}{2})}}{n^g}$, the asymptotic formula stated in part (i) of the theorem follows.

To complete the proof of the theorem we only have to show part (ii). This follows easily on noting that any commutator $[a, b]$, $a, b \in G$ must have even word length.

REFERENCES

1. J. Cannon, *The combinatorial structure of co-compact discrete hyperbolic groups*, *Geom. Dedicata* **16** (1984), 123-148. MR **86j**:20032
2. W. Floyd and S. Plotnick, *Growth functions on Fuchsian groups and the Euler characteristic*, *Invent. math.* **88** (1987), 1-29. MR **88m**:22023
3. F. R. Gantmacher, *The Theory of Matrices, Vol. II*, Chelsea, New York, 1974. MR **21**:6372c (earlier ed.)
4. M. Gromov, *Hyperbolic groups*, MSRI publications, Vol. 8 (S. Gersten, ed.), Springer, Berlin, 1987, pp. 75-263. MR **89e**:20070
5. W. Parry, *Counterexamples involving growth series and the Euler characteristic*, *Proc. Amer. Math. Soc.* **102** (1988), 49-51. MR **89c**:20051
6. M. Pollicott and R. Sharp, *Rates of recurrence for \mathbb{Z}^q and \mathbb{R}^q extensions of subshifts of finite type*, *J. London Math. Soc.* **49** (1994), 401-416. MR **95f**:54037
7. C. Series, *Symbolic dynamics for geodesic flows*, *Acta Mathematica* **146** (1981), 103-128. MR **82f**:58071
8. N. Varopoulos, L. Saloff-Coste and T. Coulhon, *Analysis and Geometry on Groups*, Cambridge University Press, Cambridge, 1992. MR **95f**:43008

DEPARTMENT OF MATHEMATICS, WARWICK UNIVERSITY, COVENTRY, CV4 7AL, ENGLAND

Current address: Department of Mathematics, University of Manchester, Oxford Road, Manchester M13 9PL, England

SCHOOL OF MATHEMATICAL SCIENCES, QUEEN MARY AND WESTFIELD COLLEGE, MILE END ROAD, LONDON, E1 4NS, ENGLAND

Current address: Department of Mathematics, University of Manchester, Oxford Road, Manchester M13 9PL, England