Combinatorics, Probability and Computing (2015) 24, 838–856. © Cambridge University Press 2015
doi:10.1017/S0963548314000819

A Numerical Lower Bound for the Spectral Radius of Random Walks on Surface Groups

S. GOUEZEL

IRMAR, CNRS UMR 6625, Université de Rennes 1, 35042 Rennes, France
(e-mail: sebastien.gouezel@univ-rennes1.fr)

Received 15 October 2013; revised 6 June 2014; first published online 29 January 2015

Estimating numerically the spectral radius of a random walk on a non-amenable graph is complicated, since the cardinality of balls grows exponentially fast with the radius. We propose an algorithm to get a bound from below for this spectral radius in Cayley graphs with finitely many cone types (including for instance hyperbolic groups). In the genus 2 surface group, it improves by an order of magnitude the previous best bound, due to Bartholdi.

2010 *Mathematics subject classification*: Primary 60B15
Secondary 20F67, 05C81

1. Main algorithm

Let Γ be a countable group, generated by a finite symmetric set S of cardinality $|S|$. The simple random walk $X_0, X_1, \dots$ on Γ is defined by $X_0 = e$, the identity of Γ, and $X_{n+1} = X_n s$ with probability $1/|S|$ for any $s \in S$. A crucial numerical parameter of this random walk is its *spectral radius* $\rho = \lim \mathbb{P}(X_{2n} = e)^{1/2n}$. Equivalently, let W_n denote the number of words of length n in the generators that represent e in Γ; then $\mathbb{P}(X_n = e) = W_n/|S|^n$, so that $\rho = \lim W_{2n}^{1/2n}/|S|$. It is equivalent to study the spectral radius or the *cogrowth* $\lim W_{2n}^{1/2n}$.

The spectral radius is at most 1, and $\rho = 1$ if and only if Γ is amenable. In the free group with d generators, the generating function $\sum W_n z^n$ can be computed explicitly (it is algebraic), and the exact value of the spectral radius follows: $\rho = \sqrt{2d-1}/d$. Since words that reduce to the identity in the free group also reduce to the identity in any group with the same number of generators, one infers that in any group Γ, $\rho \geqslant 2\sqrt{|S|-1}/|S|$. Moreover, equality holds if and only if the Cayley graph of Γ is a tree [6].

In general, there are no explicit formulas for ρ, and even giving precise numerical estimates is a delicate question. In this short note, we will describe an algorithm giving bounds from below on ρ in some classes of groups, particularly for the fundamental group

Γ_g of a compact surface of genus $g \geqslant 2$, given by its usual presentation

$$\Gamma_g = \langle a_1, \dots, a_g, b_1, \dots, b_g \mid [a_1, b_1] \cdots [a_g, b_g] = e \rangle. \tag{1.1}$$

Since there are 4 generators in Γ_2, the above trivial bound obtained by comparison to the free group gives $\rho \geqslant 0.661437$. Our main estimate is the following result.

Theorem 1.1. *In the surface group Γ_2, one has $\rho \geqslant 0.662772$.*

This improves on the previously best known result, due to Bartholdi [1], giving $\rho \geqslant \rho_{\mathrm{Bar}} = 0.662421$.[1] Bartholdi's method is to study a specific class of paths from the identity to itself (called cactus trees), for which he can compute the generating function. The radius of convergence of this generating function is a lower bound for ρ.

The best known upper bound for ρ in Γ_2 is $\rho \leqslant \rho_{\mathrm{Nag}} = 0.662816$, due to Nagnibeda [7]. Non-rigorous numerical estimates[2] suggest that $\rho = 0.662812\dots$, so the upper bound is still sharper than our lower bound, although our lower bound is an order of magnitude better than the bound of [1]: indeed, $\rho_{\mathrm{Nag}} - \rho_{\mathrm{Bar}} \sim 4.10^{-4}$, while our estimate ρ from Theorem 1.1 satisfies $\rho_{\mathrm{Nag}} - \rho \sim 4.10^{-5}$.

Nagnibeda's upper bound does not rely on a counting argument for closed paths, but on another spectral interpretation of ρ. Indeed, ρ is also the spectral radius of the Markov operator Q on $\ell^2(\Gamma)$ corresponding to the random walk, *i.e.*, the convolution with the probability measure μ which is uniformly distributed on S (see for instance [10, Corollary 10.2]). It is also the norm of this operator, since it is symmetric. Nagnibeda gets the above upper bound by using a lemma of Gabber about norms of convolution operators on graphs and the precise geometry of Γ_2.

Our approach to get Theorem 1.1 is very similar to Nagnibeda's. To bound from below the norm of the convolution operator Q, it is sufficient to exhibit one function u (which ought to be close to a hypothetical eigenfunction for the element ρ of the spectrum of Q) for which $\|Qu\|/\|u\|$ is large. This is exactly what we will do.

[1] Bartholdi claims that $\rho \geqslant 0.662418$, but implementing his algorithm in multiprecision, one in fact gets the better bound $\rho \geqslant 0.662421$.

[2] I obtained this estimate as follows. One can count exactly the number W_n of words of length n representing the identity in the group, for reasonable n, say up to $n = 24$ – for the record, $W_{24} = 4214946994935248$ – giving the first values of the sequence $p_n = \mathbb{P}(X_{2n} = e)$. We know rigorously from [5] that $p_n \sim C\rho^{2n}/n^{3/2}$ when $n \to \infty$. Define $q_n = \log(n^{3/2} p_n)/(2n)$; it follows that $q_n \to \log \rho$. In the free group, where p_n is known very explicitly, the sequence q_n has a further expansion in powers of $1/n$. Assuming that the same holds in the surface group, we get $q_n = \log \rho + \sum_{k=1}^{K} a_k/n^k + o(1/n^K)$, where the a_k are unknown. Using the known value of q_{24}, this gives an estimate for $\log \rho$, with an error of the order of $1/24$, which is very bad. However, it is possible to accelerate the convergence of sequences having an asymptotic expansion in powers of $1/n$: there are explicit recipes (for instance Richardson extrapolation or Wynn's rho algorithm) taking such a sequence, and giving a new sequence converging to the same limit, with an expansion in powers of $1/n$, but starting at $1/n^2$. Iterating this process, one can eliminate the first few terms, and get a speed of convergence $O(1/n^L)$ for any L (but one needs to know enough terms of the initial sequence). Applying this process to our sequence q_n, one gets the claimed estimate for ρ. To make this rigorous, one would need to know that an asymptotic expansion of q_n exists, with explicit bounds on the a_k and on the $o(1/n^K)$ term. This seems completely out of reach.

For any $\alpha < \rho$, the function $u_\alpha = \sum_{n=0}^{\infty} \alpha^n Q^n \delta_e$ is in ℓ^2, and $\|Qu_\alpha\|/\|u_\alpha\|$ converges to ρ when α tends to ρ. Unfortunately, u_α is not explicit enough. To find estimates, one should instead find an ansatz for the function u, depending on finitely many parameters, and then optimize over these parameters.

A first strategy would be the following: take a very large ball B_n in the Cayley graph, and compute the function u supported in this ball such that $\|Qu\|/\|u\|$ is largest. This gives a lower bound ρ_n on ρ, and ρ_n converges to ρ when n tends to infinity. However, this strategy is computationally not efficient at all: one would need to take a very large n to obtain good estimates (since most mass of u_α is supported close to infinity if α is close to ρ), and the cardinality of B_n grows exponentially with n. On the other hand, it can be implemented in any finitely presented group for which the word problem is solvable (see for instance [3] for examples in Baumslag–Solitar and Thompson groups). We will use a more efficient method, but one which requires more assumptions on the group: it should have finitely many cone types.

To illustrate our method of construction of u, let us describe it quickly in the case of the free group $\mathbb{F}_d$ with d generators. The sphere $\mathbb{S}^n$ of radius $n \geqslant 1$ has cardinality $2d(2d-1)^{n-1}$. Fix some $\alpha < 1/\sqrt{2d-1}$, and define a function u_α by $u_\alpha(x) = \alpha^n$ for $x \in \mathbb{S}^n$, $n \geqslant 1$. This function belongs to $\ell^2(\mathbb{F}_d)$. We write $x \sim y$ if x and y are neighbours in the Cayley graph of Γ, and $x \to y$ if $x \sim y$ and $d(e,y) = d(e,x) + 1$. Then

$$\langle Qu_\alpha, u_\alpha \rangle = \frac{1}{2d} \sum_{x \sim y} u_\alpha(x) u_\alpha(y) = \frac{1}{d} \sum_{x \to y} u_\alpha(x) u_\alpha(y) = \frac{1}{d} \sum_{n=1}^{\infty} (2d-1)\alpha^{2n+1} |\mathbb{S}^n|,$$

since a point in $\mathbb{S}^n$ has $2d-1$ successors in $\mathbb{S}^{n+1}$. Since

$$\langle u_\alpha, u_\alpha \rangle = \sum_{n=1}^{\infty} \alpha^{2n} |\mathbb{S}^n|,$$

we get

$$\langle Qu_\alpha, u_\alpha \rangle = \alpha \frac{2d-1}{d} \langle u_\alpha, u_\alpha \rangle.$$

Hence,

$$\rho = \|Q\| \geqslant \alpha \frac{2d-1}{d}.$$

Letting α tend to $1/\sqrt{2d-1}$, we finally obtain $\rho \geqslant \sqrt{2d-1}/d$, which is the true value of the spectral radius.

In the free group, it is natural to take a function u that is constant of the sphere $\mathbb{S}^n$ of radius n, since all the points in such a sphere are equivalent: the automorphisms of the Cayley graph of $\mathbb{F}_d$ fixing the identity act transitively on $\mathbb{S}^n$. In more general groups, for instance surface groups, this is not the case. Intuitively, we would like to take a function that decays exponentially as above, but with different values on different equivalence classes under the automorphism group. However, this automorphism group is finite in the case of surface groups, so instead of true equivalence classes (which are finite), we will consider larger classes, of points that 'locally behave in the same way', and we will construct functions that are constants on such classes of points (leaving only finitely many parameters which one can optimize using a computer).

This intuition is made precise with the notion of *type* of the elements of the group (as in [7]). Let Γ be a countable group generated by a finite symmetric set S. Assume that there are no cycles of odd length, so that any edge can be oriented from the closer point from e to the farther point. Let $\mathcal{S}(x)$ be the set of successors of x, *i.e.*, the points y which are neighbours of x with $|y| = |x| + 1$, where we write $|x| = d(e, x)$.

Definition 1.2. Let T be a finite set, let t be a function from Γ to T and let M be a square matrix indexed by T. We say that (T, t, M) is a *type system* for (Γ, S) if, for all i and j in T, for all but finitely many $x \in \Gamma$ with $t(x) = j$, one has

$$\mathrm{Card}\{y \in \mathcal{S}(x) \ : \ t(y) = i\} = M_{ij}.$$

We will often simply say that t is a type system, since it determines T and M.

In other words, if one knows the type of a point x, then one knows the number of successors of each type, thanks to the matrix M. For instance, in $\mathbb{F}_d$, one can use one single type, with $M_{11} = 2d - 1$: every point but the identity has $2d - 1$ successors.

Using a type system, we will be able to find a lower bound for the spectral radius of the simple random walk. While the argument works in general, it is more convenient to formulate using an additional assumption, which is satisfied for surface groups.

Definition 1.3. A type system (T, t, M) is Perron–Frobenius if the matrix M is Perron–Frobenius, *i.e.*, some power M^n has only positive entries.

The algorithm to estimate the spectral radius follows.

Theorem 1.4. *Let (Γ, S) be a countable group with a finite symmetric generating set, whose Cayley graph has no cycle of odd length. Let (T, t, M) be a Perron–Frobenius type system for (Γ, S).*

Define a new matrix $\tilde{M}$ by $\tilde{M}_{ij} = M_{ij}/p_i$, where p_i is the number of predecessors of a point of type i (it is given by $p_j = |S| - \sum_i M_{ij}$). Since it is Perron–Frobenius, its dominating eigenvalue e^v is simple. Let $(A_1, \dots, A_k)$ be a corresponding eigenvector, with positive entries, let D be the diagonal matrix with entries A_i, and let $M' = D^{-1/2} M D^{1/2}$. Define

$$\lambda = \max_{|q|=1} \langle M' q, q \rangle. \tag{1.2}$$

Then

$$\rho \geqslant \frac{2e^{-v/2}\lambda}{|S|}. \tag{1.3}$$

Proof. Let $s_n(i) = \mathrm{Card}\{x \in \mathbb{S}^n \ :\ t(x) = i\}$. By definition of a type system, if n is large enough (say $n \geqslant n_0$),

$$p_i s_{n+1}(i) = \sum_{y \in \mathbb{S}^n} \mathrm{Card}\{x \in \mathcal{S}(y)\ :\ t(x) = i\} = \sum_j M_{ij} s_n(j).$$

This shows that $s_{n+1} = \tilde{M} s_n$. Therefore, the cardinality of $\mathbb{S}^n$ grows like ce^{nv} for some $c > 0$. Moreover, $s_n(i) = c' A_i e^{nv} + O(e^{n(v-\varepsilon)})$ for some $\varepsilon > 0$.

Take some parameters $b_1, \dots, b_k > 0$ to be chosen later, and let $\alpha < e^{-v/2}$. We define a function u_α by $u_\alpha(x) = \alpha^n b_i$ if $x \in \mathbb{S}^n$ and $t(x) = i$ with $n \geqslant n_0$. For $n < n_0$, let $u_\alpha(x) = 0$. We have when α tends to $e^{-v/2}$

$$\begin{aligned}
\langle Q u_\alpha, u_\alpha \rangle &= \frac{1}{|S|} \sum_{x \sim y} u_\alpha(x) u_\alpha(y) = \frac{2}{|S|} \sum_x \sum_{y \in \mathcal{S}(x)} u_\alpha(x) u_\alpha(y) \\
&= \frac{2}{|S|} \sum_{n \geqslant n_0} \sum_{i,j} s_n(j) b_j \alpha^n M_{ij} b_i \alpha^{n+1} \\
&= \frac{2\alpha}{|S|} \sum_{n \geqslant n_0} \sum_{i,j} c' A_j e^{nv} b_j \alpha^{2n} M_{ij} b_i + O(1) \\
&= \frac{2\alpha}{|S|} \sum_{i,j} c' A_j b_j M_{ij} b_i / (1 - \alpha^2 e^v) + O(1).
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\langle u_\alpha, u_\alpha \rangle &= \sum_{n \geqslant n_0} \sum_i s_n(i) b_i^2 \alpha^{2n} = \sum_{n \geqslant n_0} \sum_i c' A_i e^{nv} b_i^2 \alpha^{2n} + O(1) \\
&= \sum_i c' A_i b_i^2 / (1 - \alpha^2 e^v) + O(1).
\end{aligned}$$

We have $\rho \geqslant \langle Q u_\alpha, u_\alpha \rangle / \langle u_\alpha, u_\alpha \rangle$. Comparing the above two equations and letting α tend to $e^{-v/2}$, we get

$$\rho \geqslant \frac{2e^{-v/2}}{|S|} \frac{\sum_{i,j} A_j b_j M_{ij} b_i}{\sum_i A_i b_i^2}.$$

To conclude, we need to optimize in b_i. Writing b_i as $A_i^{-1/2} c_i$, this lower bound becomes

$$\frac{2e^{-v/2}}{|S|} \frac{\sum A_j^{1/2} c_j M_{ij} A_i^{-1/2} c_i}{\sum c_i^2}.$$

The maximum of the last factor is the maximum on the unit sphere of the quadratic form with matrix $M' = D^{-1/2} M D^{1/2}$. This proves (1.3). □

Remark 1.5. It follows from the formula (1.2) that λ is the maximum on the unit sphere of $\langle M'' q, q \rangle$, where M'' is the symmetric matrix $(M' + M'^*)/2$. Since any symmetric matrix is diagonal in some orthogonal basis, it also follows that λ is the maximal eigenvalue of M'', *i.e.*, its spectral radius. Hence, it is easy to compute using standard algorithms.

The formula given by Theorem 1.4 depends not only on the geometry of the group, but also on the choice of a type system: in a given group (with a given system of generators), there may be several type systems, giving different estimates. We will take advantage of this fact for surface groups in Section 3: applying Theorem 1.4 with the canonical type system for surface groups, constructed by Cannon, we obtain in (3.1) an estimate for the spectral radius which is weaker than the estimate of Theorem 1.1. This stronger estimate is proved by applying Theorem 1.4 to a different type system, constructed as a refinement of the canonical type system.

This dependence on the choice of a type system should be contrasted with the upper bound of Nagnibeda in [7]. Indeed, it is shown in [8] that this upper bound, computed using a type system, has a purely geometric interpretation (it is the spectral radius of a random walk on the tree of geodesics of the group), which does not depend on the choice of the type system. In particular, the refined type system we use to prove Theorem 1.1 cannot improve the upper bound of Nagnibeda.

2. Geometric interpretation

In this section, we describe a geometric interpretation of Theorem 1.4, similar to Nagnibeda's interpretation in [8] of the bound she obtained in [7].

We first recall Nagnibeda's construction. Consider a group Γ with a finite system of generators S, whose Cayley graph has no cycle of odd length. Let X be its tree of geodesics, *i.e.*, the graph whose vertices are the finite geodesics in Γ originating from the identity e, and where one puts an edge from a geodesic with length n to its extensions with length $n+1$. There is a canonical projection π_X from X to Γ, taking a geodesic to its endpoint. One can think of X as obtained from Γ by unfolding the loops based at e.

Consider the random walk in X whose transitions are as follows. From x, one goes to any of its successors with probability $1/|S|$, and to its unique predecessor with probability $p_x/|S|$ where p_x is $|S|$ minus the number of successors of x (it is the number of predecessors in Γ of the projection $\pi_X(x)$). This random walk on X does *not* project to the simple random walk on Γ, since it does not follow loops in Γ (the projected random walk is not Markov in general). The transition probabilities coincide when going towards infinity, but not when going back towards the identity. One expects that the probability of coming back to the identity is higher in X than in Γ, thanks to the following heuristic: since the process in X is less random when coming back toward the identity, once the walk is in a subset where it comes back often to the identity, it cannot easily escape from this subset, and therefore returns even more.

To illustrate this heuristic, suppose that two points x and x' in X have successors y and y' in X with $p_y = p_{y'} = 2$, and consider a new random walk in which y and y' are identified (this is what the projection π_X does, all over the place), so that from this new point one can either jump back to x or to x' with probability $1/|S|$. Let u_n and u'_n be the probabilities in X of being at time n at x and x'. For the sake of argument, we will assume some form of symmetry, *i.e.*, u_n and u'_n are also the probabilities of reaching e at time n starting respectively from x or x'. In X, one can form paths from e to itself of length $2n+2$ by jumping to x in time n, then to y, then back to x, and then from x to e.

This happens with probability

$$u_n \cdot \frac{1}{|S|} \cdot \frac{2}{|S|} \cdot u_n.$$

One can do the same with x', giving an overall probability

$$a = \frac{2}{|S|^2}(u_n^2 + {u'_n}^2).$$

On the other hand, if y and y' are identified, then from this new point one can either jump back to x or to x'. The corresponding probability of coming back to e at time $2n+2$ following such paths is therefore

$$b = \frac{1}{|S|^2}(u_n + u'_n)^2.$$

As $2(v^2 + w^2) \geqslant (v+w)^2$, we have $a \geqslant b$, *i.e.*, the probability of returning to e using corresponding paths is bigger in X than in the random walk where y and y' are identified. This explains our heuristic that more randomness in the choice of predecessors in the graph creates a mixing effect that decreases the spectral radius.

In [7] and [8], Nagnibeda justifies this heuristic rigorously as follows. Consider a group Γ and a generating system S such that the Cayley graph of Γ with respect to S has no cycle of odd length, and finitely many cone types. By applying a spectral lemma of Gabber, she gets an upper bound ρ (given by a minimax formula, complicated to estimate in general) for the spectral radius ρ_Γ of the simple random walk on (Γ, S). Since the tree of geodesics X also has finitely many cone types, she is able to compute exactly the spectral radius ρ_X of the random walk in X. It turns out that this is exactly ρ. Hence, $\rho_\Gamma \leqslant \rho_X$, the interest of this formula being that ρ_X can be easily computed (it is algebraic as the tree X has finitely many cone types: see [9]). Note, however, that this bound does not come from a direct argument using the projection $\pi_X : X \to \Gamma$, but rather from two separate computations in X and in Γ.

We now turn to a similar geometric interpretation of the lower bound given in Theorem 1.4. We are looking for a natural random walk, related to the original random walk on Γ, whose spectral radius can be computed exactly and coincides with the lower bound given in (1.3). Following the above heuristic, this random walk should have more randomness than the original random walk regarding the choice of predecessors, to decrease the probability of returning to the origin.

We use the setting and notations of Theorem 1.4. In particular, Γ is a group with a type system (T, t, M), and $(A_1, \dots, A_k)$ and $\tilde{M}$ are defined in the statement of this theorem. We define a random walk as follows. It is a walk on the space $Y = \mathbb{Z} \times T$ (where T is the space of types), whose transition probabilities are given by

$$p((n,j) \to (n+1,i)) = M_{ij}/|S|, \quad p((n,j) \to (n-1,i)) = e^{-v}\frac{A_i M_{ji}}{A_j|S|}. \tag{2.1}$$

The spectral radius of this random walk is by definition $\rho_Y = \lim \mathbb{P}(X_{2n} = e)^{1/2n}$ (this is *not* a spectral definition). Note that this random walk admits a quasi-transitive $\mathbb{Z}$-action (*i.e.*, Y is endowed with a free action of $\mathbb{Z}$, with finite quotient, and the transition probabilities are invariant under $\mathbb{Z}$). Such random walks are well studied: see for instance [10, Section 8.B].

Theorem 2.1. *With the notations of Theorem 1.4, the random walk on Y has spectral radius $\rho_Y = 2e^{-v/2}\lambda/|S|$.*

Hence, the result of Theorem 1.4 reads $\rho_\Gamma \geqslant \rho_Y$, and ρ_Y is easy to compute.

Let us first explain why this random walk is natural, and related to the random walk on Γ. Starting from a point $x \in \Gamma$, of type j and length n, the original random walk goes to any of its successors with probability $1/|S|$. In particular, it reaches points of type i and length $n+1$ with probability $M_{ij}/|S|$, just like the probability given in (2.1). On the other hand, it goes to any of its predecessors with probability $1/|S|$, but the types of these predecessors depend on x, not only on j. A random walk which is simpler to estimate may be constructed by randomizing the predecessors: from x, one chooses to go to any point of length $n-1$ and type i, provided that there is an edge from type i to type j, *i.e.*, $M_{ji} > 0$. The probability of going from x to such points should be given by the average number of predecessors of type i to a point of type j. Writing

$$s_n(i) = \operatorname{Card}\{x \in \mathbb{S}^n \ : \ t(x) = i\},$$

this quantity is

$$\frac{\sum_{|x|=n,t(x)=j} \sum_{|y|=n-1,t(y)=i} 1(x \in \mathcal{S}(y))}{\sum_{|x|=n,t(x)=j} 1} = \frac{s_{n-1}(i)M_{ji}}{s_n(j)}. \tag{2.2}$$

As $s_n(i) \sim c' A_i e^{nv}$ for some $c' > 0$ (see the proof of Theorem 1.4), the quantity in (2.2) converges when $n \to \infty$ to $e^{-v}A_i M_{ji}/A_j$, giving the transition probability (2.1) in the limit. Note that, in this randomized random walk, all the points of the same length and the same type are equivalent. Hence, we may identify them, to get a smaller space and a simpler random walk. This is precisely our random walk on Y.

Thus, the random walk on Y is obtained by starting from the random walk on Γ, randomizing the choice of predecessors, going in the asymptotic regime $n \to \infty$, and identifying the points on the sphere that are equivalent. One can define a projection map $\pi_Y : \Gamma \to Y$ by $\pi_Y(x) = (|x|, t(x))$, under which the two random walks correspond in a loose sense (going towards infinity, the transition probabilities are the same, but coming back towards the identity they differ, just as for the projection π_X in Nagnibeda's construction, with more randomness in Y than in Γ).

Proof of Theorem 2.1. We define two matrices P^+ and P^- giving the transition probabilities of the walk on Y to the right and to the left, respectively, *i.e.*,

$$P^+_{ij} = M_{ij}/|S|, \quad P^-_{ij} = e^{-v}\frac{A_i M_{ji}}{A_j |S|}.$$

In other words, $P^+ = M/|S|$ and $P^- = e^{-v} D M^* D^{-1}/|S|$, where D is the diagonal matrix with entries A_i and M^* is the transpose of M.

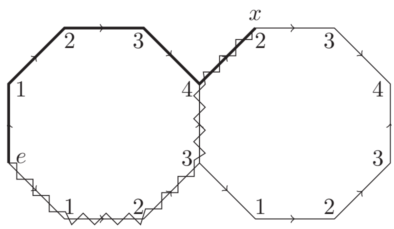

Figure 1. Two examples of geodesics from e to x.

Although this is clear from the geometric construction, let us first check algebraically that the transition probabilities in (2.1) indeed define probabilities, *i.e.*, $\sum_i P^+_{ij} + P^-_{ij} = 1$ for all j. Let p_j be the number of predecessors of a point of type j in Γ. By definition of A, the matrix $\tilde{M}_{ji} = M_{ji}/p_j$ satisfies $\tilde{M}A = e^v A$. Hence,

$$\sum_i M_{ji} A_i = p_j \sum_i \tilde{M}_{ji} A_i = p_j e^v A_j.$$

Therefore,

$$\sum_i M_{ij} + \sum_i e^{-v} \frac{A_i M_{ji}}{A_j} = |S| - p_j + e^{-v} \frac{p_j e^v A_j}{A_j} = |S|,$$

proving that (2.1) defines transition probabilities.

While one can give a pedestrian proof of the equality $\rho_Y = 2e^{-v/2}\lambda/|S|$, it is more efficient to use results available in the literature. Define a function φ on $\mathbb{R}$ by $\varphi(c) = \rho(e^c P^+ + e^{-c} P^-)$ (where this quantity is the spectral radius of a *bona fide* finite-dimensional matrix). It is proved in [10, Proposition 8.20 and Theorem 8.23] that φ is convex, that it tends to infinity at $\pm\infty$, and that its minimum is precisely ρ_Y. Since the spectral radius is invariant under transposition and conjugation, we have

$$\begin{aligned} |S|\varphi(c) &= \rho(e^c M + e^{-c} e^{-v} D M^* D^{-1}) = \rho(e^c M^* + e^{-c-v} D^{-1} M D) \\ &= \rho(e^c D M^* D^{-1} + e^{-c-v} M) = |S|\varphi(-c-v). \end{aligned}$$

Hence, the function $c \mapsto \varphi(c)$ is symmetric around $c = -v/2$. As it is convex, it attains its minimum at $c = -v/2$. Therefore,

$$\rho_Y = \varphi(-v/2) = \frac{e^{-v/2}}{|S|} \rho(M + D M^* D^{-1}) = \frac{e^{-v/2}}{|S|} \rho(D^{-1/2} M D^{1/2} + D^{1/2} M^* D^{-1/2}).$$

By Remark 1.5, the last term is equal to 2λ. □

3. Application to surface groups

3.1. Cannon's types

Consider a countable group with the word distance coming from a finite generating set S. The *cone* of a point x is the set of points y for which there is a geodesic from e to y going through x. The *cone type* of x is the set $\{x^{-1}y\}$, for y in the cone of x. Note that knowing the cone type of a point determines the number of its successors, and the number of its successors having any given cone type. Cannon proved that, in any hyperbolic group, there are finitely many cone types. Therefore, such a group admits a type system in the sense of Definition 1.2. This is in particular the case of the surface groups Γ_g. However, the number of cone types is too large, and it is more convenient for practical purposes to reduce them using symmetries. We obtain Cannon's canonical types for the surface groups, described in [2] or [4] as follows.

The hyperbolic plane can be tessellated by regular $4g$-gons, with $4g$ of them around each vertex. The Cayley graph of Γ_g (with its usual presentation (1.1)) is dual to this (self-dual) tessellation, and is therefore isomorphic to it. Define the type of a point $x \in \Gamma_g$ as the maximal length along the last $4g$-gon of a geodesic starting from e and ending at x. Beware that one really has to take the maximum: for instance, in Figure 1, the thick geodesic from e to x shares only one edge with the last octagon, while the wiggly one shares two edges. Hence, the type of x is 2.

The type can also be described combinatorially as follows. Write $x = c_1 \cdots c_{|x|}$ as a product of minimal length in the generators $a_1, \dots, b_g$, look at the length n of its longest common suffix with a fundamental relator (*i.e.*, a cyclic permutation of the basic relation $[a_1, b_1] \cdots [a_g, b_g]$ in (1.1) or its inverse: $c_{|x|-n+1} \cdots c_{|x|}$ should be a subword of the basic relation or its inverse, up to cyclic permutation), and take the maximum of all such n over all ways to write $x = c_1 \cdots c_{|x|}$. It is obvious that the geometric and combinatorial descriptions are equivalent; we will mostly rely on the geometric one.

The type of a group element x can be at most $2g$ (otherwise, taking the same path but going the other way around the last $4g$-gon, one would get a strictly shorter path, contradicting the fact that the initial path is geodesic), and it is 0 only for the identity. Points x of type $i < 2g$ have only one predecessor, and $4g-1$ successors. Among them, two are followers of x on the $4g$-gon to its left and to its right, while the other ones correspond to newly created $4g$-gons (whose closest point to e is x). It follows that those points have $4g-3$ successors of type 1, one of type 2 and one of type $i+1$. Points of type $2g$ are special, since they have two predecessors (one can reach them with a geodesic either from the left or from the right of a single $4g$-gon). They have two successors of type 2, corresponding to the extremal outgoing edges of x (they extend the two $4g$-gons adjacent to both incoming edges to x), and the $4g-4$ remaining successors are on newly created $4g$-gons, and are of type 1. See Figure 2 for an illustration in genus 2 (of course, additional octagons should be drawn around all outgoing edges, but since this is notoriously difficult to do in a Euclidean drawing, we have to rely on the reader's imagination).

Keeping only the types from 1 to $2g$ (since type 0 only happens for the identity, while Definition 1.2 allows us to discard finitely many points), we obtain a type system for Γ_g with $T = \{1, \dots, 2g\}$, where the matrix M has been described in the previous paragraph. For instance, in genus 2,

$$M = \begin{pmatrix} 5 & 5 & 5 & 4 \\ 2 & 1 & 1 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

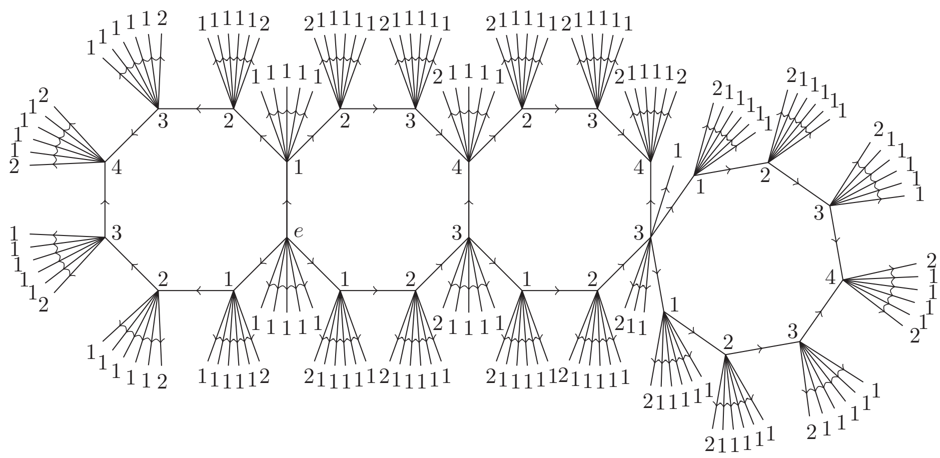

Figure 2. Types of the points in (part of) the Cayley graph of Γ_2.

One can now apply the algorithm of Theorem 1.4 to this matrix to bound the spectral radius of the simple random walk from below. All points but points of type 4 have one predecessor, so the matrix $\tilde{M}$ is

$$\tilde{M} = \begin{pmatrix} 5 & 5 & 5 & 4 \\ 2 & 1 & 1 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \end{pmatrix}.$$

The dominating eigenvalue of this matrix is $e^v = 6.979835\ldots$ (this is also the growth of the group), while the corresponding eigenvector is

$$A = (0.715987\ldots, 0.246211\ldots, 0.035274\ldots, 0.002526\ldots).$$

We obtain that the matrix M'' of Remark 1.5 is

$$M'' = \begin{pmatrix} 5 & 3.171316\ldots & 0.554905\ldots & 0.118814\ldots \\ 3.171316\ldots & 1 & 1.510223\ldots & 0.101307\ldots \\ 0.554905\ldots & 1.510223\ldots & 0 & 1.868132\ldots \\ 0.118814\ldots & 0.101307\ldots & 1.868132\ldots & 0 \end{pmatrix},$$

with dominating eigenvalue $\lambda = 7.000902\ldots$. Finally,

$$\rho \geqslant \frac{2e^{-v/2}\lambda}{8} = 0.662477\ldots. \tag{3.1}$$

This is already slightly better than Bartholdi's estimate $\rho \geqslant 0.662421$, but much weaker than the estimate $\rho \geqslant 0.662772$ that we claimed in Theorem 1.1 (to be compared with the 'true' value $\rho \sim 0.662812$).

In the next sections, we will explain how to get better estimates by using different type systems, that distinguish between more points (but, of course, give rise to larger matrices M and therefore to more computer-intensive computations).

Remark 3.1. Using cone types instead of Cannon's canonical types does not give rise to better estimates for ρ (although the number of types is much larger). Indeed, if some type system t' is obtained from some type system t by quotienting by some symmetries of t, then the dominating eigenvector of $\tilde{M}(t)$, being unique, is invariant by those symmetries and reduces to the dominating eigenvector of $\tilde{M}(t')$. It follows that the dominating eigenvector of $M''(t)$ is also invariant by those symmetries, and that the dominating eigenvalue of $M''(t)$ is the same as that of $M''(t')$. Hence, the estimates on ρ given by Theorem 1.1 for t and t' are the same.

3.2. Suffix types

There are many ways to define new type systems in surface groups, that separate more points. If a type system is finer than another one, then the estimate on the spectral radius coming from Theorem 1.4 is better, but the matrix involved in the computation is larger. To get manageable estimates, we should find the right balance.

In this subsection, we describe a very simple extension of Cannon's canonical type systems in surface groups, which we call suffix types. Given a point $x \in \Gamma_g$, there can be several geodesics from e to x. Consider the longest ending that is common to all these geodesics, say $x_{n-k+1}, \dots, x_n$ (with $x_n = x$), and define the suffix type of x to be

$$t_{\mathrm{suff}}(x) = (t(x_n), t(x_{n-1}), \dots, t(x_{n-k+1})),$$

where t is the canonical type of Cannon.

For any x, $t_{\mathrm{suff}}(x)$ is easy to compute inductively.

- If $t(x) = 0$, *i.e.*, $x = e$, then $t_{\mathrm{suff}}(x) = (0)$.
- If x is of type $2g$, it has two predecessors, so the common ending to all geodesics ending at x is simply x, and $t_{\mathrm{suff}}(x) = (2g)$.
- If $t(x) \in \{1, \dots, 2g-1\}$, then x has a unique predecessor z. The common ending to all geodesics ending at x is the common ending to all geodesics ending at z, followed by x. Hence, $t_{\mathrm{suff}}(x) = (t(x), t_{\mathrm{suff}}(z))$.

It also follows from this description that, if one knows $t_{\mathrm{suff}}(x)$, it is easy to determine $t_{\mathrm{suff}}(y)$ for any successor y of x: if $t(y) = 2g$, then $t_{\mathrm{suff}}(y) = (2g)$, otherwise x is the only predecessor of y and $t_{\mathrm{suff}}(y) = (t(y), t_{\mathrm{suff}}(x))$.

We have shown that t_{suff} shares most properties of type systems as described in Definition 1.2, except that it does not take its values in a finite set. To ensure this additional property, one should truncate the suffix type. For instance, one can fix some maximal length k, and define the k-truncated suffix type $t_{\mathrm{suff}}^{(k)}(x)$ by keeping only the first k elements of $t_{\mathrm{suff}}(x)$ if its length is $> k$.

The following proposition is obvious from the previous discussion.

Proposition 3.2. *For any $k \geqslant 1$, the k-truncated suffix type system $t_{\mathrm{suff}}^{(k)}$ is a (Perron–Frobenius) type system in the sense of Definition 1.2.*

The matrix size increases with k, but the estimates on the spectral radius following from Theorem 1.4 get better. For instance, in Γ_2, for $k = 5$, the matrix size is 148, and we get $\rho \geqslant 0.662694$.

A drawback of this truncation process is that it truncates uniformly, independently of the likeliness of the type, while it should be more efficient to extend mostly those types that are more likely to happen. This intuition leads to another truncation process: fix a system of weights $w = (w_0, \dots, w_{2g}) \in [0, +\infty)^{2g+1}$, a threshold k, and truncate a suffix type $(t_0, t_1, \dots)$ at the smallest n such that $t_0 + \cdots + t_n > k$. This gives another type system denoted by $t_{\mathrm{suff}}^{(k,w)}$ ($t_{\mathrm{suff}}^{(k)}$ corresponds to the weights $w = (1, \dots, 1)$ and the threshold $k-1$). Define for instance a weight system $\bar{w}$ by $\bar{w}_0 = 1$ and $\bar{w}_i = i$ for $i \geqslant 1$: the corresponding type system $t_{\mathrm{suff}}^{(k,\bar{w})}$ truncates more quickly the suffix types involving a lot of large types, that happen less often in the group. Hence, it should give a smaller matrix than the naive truncation only according to length, while retaining a comparatively good estimate for the spectral radius.

This intuition is correct. For instance, in Γ_2, using $t_{\mathrm{suff}}^{(k,\bar{w})}$ with $k = 6$, one gets a matrix with size 109 and an estimate $\rho \geqslant 0.662697$: the matrix is smaller than for the naive truncation $t_{\mathrm{suff}}^{(5)}$, while the estimate on the spectral radius is better.

We can now push the computations, to a larger matrix size. Using the weight $\bar{w}$ and the truncation threshold $k = 25$ in Γ_2, one obtains a type system where the matrix is of size 2 774 629, and the following estimate on the spectral radius.

Proposition 3.3. *In Γ_2, one has $\rho \geqslant 0.662757$.*

This is definitely better than (3.1), but not yet as good as Theorem 1.1.

Here follow a few comments on the practical implementation. There are three main steps in the algorithm of Theorem 1.4.

(1) Compute the matrix M corresponding to the type system.
(2) Find the eigenvector A, to define the matrix M'.
(3) Find the maximal expansion rate of M'.

Computing the matrix of the type system is a matter of simple combinatorics: we explained above all the transitions from one suffix type to the next ones. The resulting matrix M is very sparse: each type has at most $2g$ successors. However, it is extremely large, so that finding the eigenvector A and then the maximal expansion rate of M' might seem computationally expensive. This is not the case, as we now explain.

Let $A^{(0)}$ be the eigenvector for the original Cannon type, so that

$$\mathrm{Card}\{x \in \mathbb{S}^n \ : \ t(x) = i\} \sim A_i^{(0)} e^{nv}.$$

Further, let $M^{(0)}$ be the matrix for the original Cannon type. Given a new type $\bar{i} = (i_0, \dots, i_m)$, the entry $A_{\bar{i}}$ of the eigenvector A for the new type $t_{\mathrm{suff}}^{(w,k)}$ is such that

$$\mathrm{Card}\{x \in \mathbb{S}^n \ : \ t_{\mathrm{suff}}^{(w,k)}(x) = \bar{i}\} \sim A_{\bar{i}} e^{nv}.$$

Such a point x can be obtained uniquely by starting from a point $y \in \mathbb{S}^{n-m}$ with type i_m, and then taking successors respectively of type $i_{m-1}, \dots, i_0$. Hence,

$$\begin{aligned} \mathrm{Card}\{x \in \mathbb{S}^n \ : \ t_{\mathrm{suff}}^{(w,k)}(x) = (i_0, \dots, i_m)\} \\ = M_{i_0 i_1}^{(0)} \cdots M_{i_{m-1} i_m}^{(0)} \, \mathrm{Card}\{y \in \mathbb{S}^{n-m} \ : \ t(y) = i_m\}. \end{aligned} \tag{3.2}$$

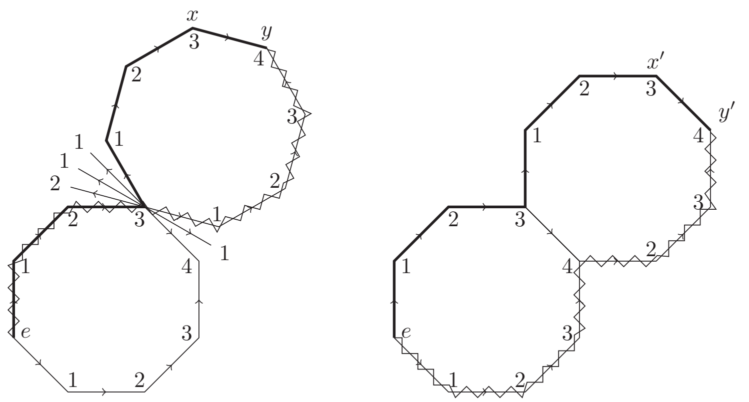

Figure 3. The points x and x' have the same essential type, contrary to y and y'.

It follows that the new eigenvector is given by

$$A_{\bar{i}} = M^{(0)}_{i_0 i_1} \cdots M^{(0)}_{i_{m-1} i_m} A^{(0)}_{i_m} e^{-mv}.$$

This shows that A is very easy to compute.

By Remark 1.5, to determine the maximal expansion rate λ of M', it suffices the find the maximal eigenvalue of $M'' = (M' + M'^*)/2$. This matrix is real, symmetric, with non-negative coefficients, and it is Perron–Frobenius (*i.e.*, it has one single maximal eigenvalue). It follows that, for any vector v with positive coefficients, $\lambda = \lim \|M''^n v\| / \|M''^{n-1} v\|$ (and moreover this sequence is non-decreasing: see for instance [10, Corollary 10.2]). Hence, one can readily estimate λ from below, by starting from a fixed vector v and iterating M''. Again, there is no issue of instability or complexity.

3.3. Essential types

To improve the suffix types, to separate even more points, one can for instance replace the canonical Cannon types with the true cone types. However, the matrix size increases so quickly that this is not usable in practice. Moreover, this does not solve the main problem of suffix types: they do not separate at all points with Cannon type $2g$, although such points are clearly not always equivalent. In this subsection we introduce a new type system that can separate such points, which we call the *essential type*.

The basic idea (which will not work directly) is to memorize not only the common ending of all geodesics ending at a point x, but all the parts that are common to such geodesics: *i.e.*, the sequence $F_{\mathrm{ess}}(x) = (x_0 = e, x_1, \dots, x_n = x)$ (with $n = |x|$) where $x_i = *$ if there are two geodesics from e to x that differ at position i, and x_i is the point that is common to all those geodesics at position i otherwise. We then associate with x the sequence

$$t_{\mathrm{ess}}(x) = (t(x_n), t(x_{n-1}), \dots, t(x_0)),$$

where $t(x_i)$ is the Cannon type of x_i if $x_i \neq *$, and $t(*) = *$.

The problem with this notion is that $t_{\mathrm{ess}}(x)$ does not determine $t_{\mathrm{ess}}(y)$ for y a successor of x: in Figure 3, the points x and x' have the same essential type $(3, 2, 1, 3, 2, 1, 0)$, while their

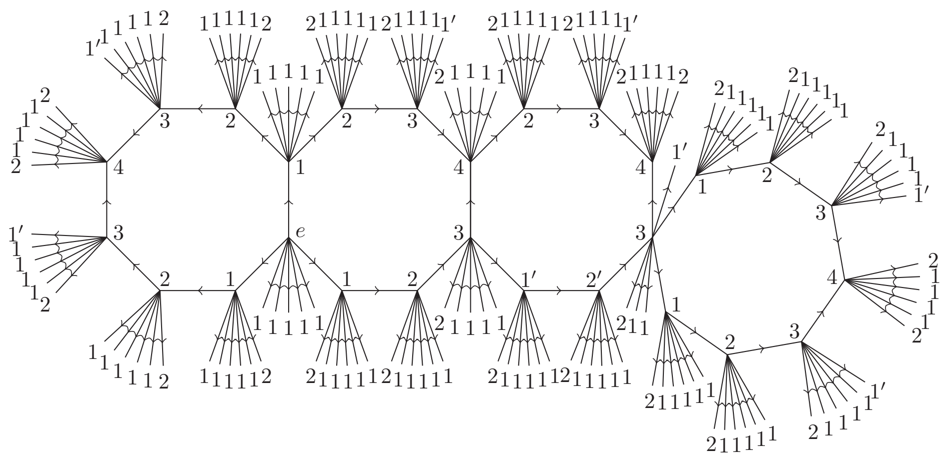

Figure 4. Modified Cannon types of the points in (part of) the Cayley graph of Γ_2.

successors y and y' have respective essential types $(4,*,*,*,3,2,1,0)$ and $(4,*,*,*,*,*,*,0)$ (this follows from the fact that the thick paths and wiggly paths are geodesics). This shows that, as we have defined it, t_{ess} cannot be used to define a type system.

This problem can be solved if we do not use the Cannon types directly in the definition of t_{ess}, but a slightly refined notion, the Cannon modified type, taking values in $\{1,\dots,2g,1',2'\}$. The modified type of a point is the same as its Cannon type, except for some points of Cannon types 1 and 2, which have modified types $1'$ and $2'$ respectively. Considering any point y of type $2g-1$, it has a unique successor z of type $2g$, and a unique successor x of type 1 that is on the same $4g$-gon as z. We say that x is of modified type $1'$. Moreover, x has a unique successor of type 2 that is also on the same $4g$-gon as z; we say that it is of modified type $2'$. See Figure 4 for an example in genus 2. By definition, the modified Cannon type t' is also a type system. The transition matrix is the same as for the usual Cannon type, except for the following.

- A point of type $2g-1$ has one successor of type $1'$, one successor of type $2g$, one successor of type 2, and $4g-4$ successors of type 1.
- A point of type $1'$ has one successor of type 2, one successor of type $2'$, and $4g-3$ successors of type 1.
- A point of type $2'$ has one successor of type 2, one successor of type 3, and $4g-3$ successors of type 1.

We define $t_{\text{ess}}(x) = (t'(x_n),\dots,t'(x_0))$, where $(x_0,\dots,x_n) = F_{\text{ess}}(x)$ and $t'(*) = *$.

Proposition 3.4. *The essential type $t_{\text{ess}}(x)$ of a point x determines the essential type of its successors.*

Proof. We argue by induction on $|x| = d(x,e)$.

Consider a point x, and one of its successors y. If the type of y is not $2g$, then x is the unique predecessor of y, and $t_{\text{ess}}(y) = (t'(y), t_{\text{ess}}(x))$. Assume now that $t'(y) = 2g$ (so that

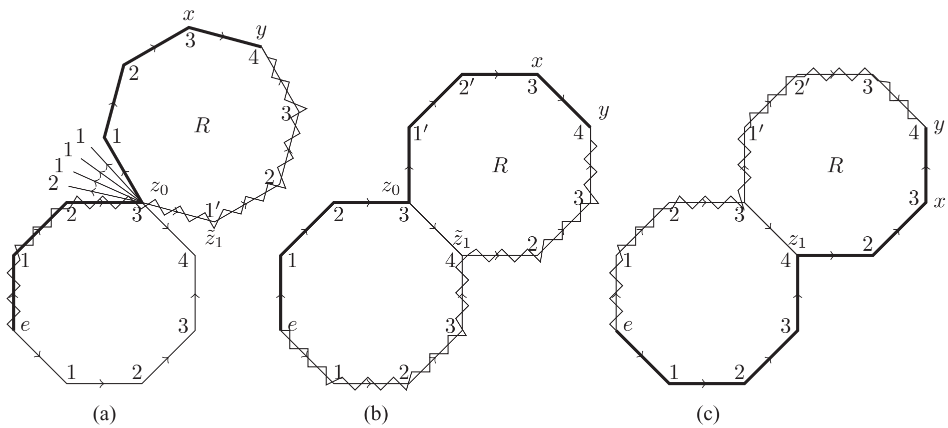

Figure 5. Determining the essential type of y from that of x.

$t'(x) = 2g - 1$). Let $z_{2g-1} = x$, and define inductively z_i as the unique predecessor of z_{i+1} for $i \geqslant 1$, so that $F(x) = (e, \ldots, z_1, z_2, \ldots, z_{2g-1} = x)$. Those points are on a common $4g$-gon R, and $t(z_i) = i$ for $i \geqslant 2$, while $t(z_1)$ can be anything. In the same way, let $\tilde{z}_{2g-1}, \ldots, \tilde{z}_1$ be the successive pre-images of y going around R in the other direction. They also satisfy $t(\tilde{z}_i) = i$ for $i \geqslant 2$.

If $t(z_1) \neq 2g$, then z_1 has a unique pre-image z_0, which also belongs to R. Moreover, z_0 is the unique closest point to e on R. The path $P = (z_0, z_1, \ldots, z_{2g-1}, y)$ is a geodesic path going around R in one direction, and $\tilde{P} = (z_0, \tilde{z}_1, \ldots, \tilde{z}_{2g-1}, y)$ is also geodesic and goes around R in the other direction. If, along $\tilde{P}$, all points different from z_0, y have type $< 2g$ (so that they have a unique pre-image), then any geodesic from e to y has to follow either P or $\tilde{P}$, so that $t_{\text{ess}}(y) = (2g, *, \ldots, *, t_{\text{ess}}(z_0))$, with $2g - 1$ ambiguous points. See Figure 5(a) for an illustration.

The only way to have a point of type $2g$ along $\tilde{P}$ is if $t(z_0) = 2g - 1$ and $t(\tilde{z}_1) = 2g$, since $t(\tilde{z}_i) = i$ for $i \geqslant 2$: see Figure 5(b). By definition of the modified type, this happens exactly when $t'(z_1) = 1'$ and $t'(z_2) = 2'$. In this case, a geodesic from e to y can either go through z_0 and then follow P, or go through $\tilde{z}_1$ and then follow $(\tilde{z}_2, \ldots, \tilde{z}_{2g-1}, y)$. In the first case, if one truncates the geodesic when it reaches z_0 and then adds the edge $[z_0\tilde{z}_1]$, one gets a geodesic from e to $\tilde{z}_1$. It follows that the essential type of y is $(2g, *, \ldots, *, \widehat{t_{\text{ess}}}(\tilde{z}_1))$, where $\widehat{t_{\text{ess}}}(\tilde{z}_1)$ is the essential type of $\tilde{z}_1$ minus its first entry (*i.e.*, the type $2g$ of $\tilde{z}_1$), and where there are $2g - 1$ stars. Since $\tilde{z}_1$ is a successor of z_0, the induction hypothesis ensures that its essential type can be determined from that of z_0, which is given by that of x. When $t(z_1) \neq 2g$, we have shown that in all situations $t_{\text{ess}}(x)$ determines $t_{\text{ess}}(y)$ in an algorithmic way.

Assume now that $t(z_1) = 2g$: see Figure 5(c). This case is very similar to the previous one. To reach y, one has to reach z_1 or its pre-image on R, and then reach y by going around R in one direction or the other. It follows that $t_{\text{ess}}(y) = (2g, *, \ldots, *, \widehat{t_{\text{ess}}}(z_1))$ as above. Since $t_{\text{ess}}(z_1)$ is obtained by removing the last $2g - 2$ entries of $t_{\text{ess}}(x)$, it follows again that $t_{\text{ess}}(x)$ determines $t_{\text{ess}}(y)$. $\square$

For $k > 0$, let $t_{\text{ess}}^{(k)}(x)$ be the truncated essential type, obtained by keeping the first k entries of the essential type of x. The above proof also shows that $t_{\text{ess}}^{(k)}(x)$ determines $t_{\text{ess}}^{(k+1)}(y)$ (and therefore $t_{\text{ess}}^{(k)}(y)$) for any successor y of x. In the same way, if one considers a truncation according to a weight $w = (w_0, w_1, \dots, w_{2g}, w_{1'}, w_{2'}, w_*)$ and a threshold k, then $t_{\text{ess}}^{(k,w)}(x)$ determines $t_{\text{ess}}^{(k,w)}(y)$ for any successor y of x, if the weight w_* is maximal among all weights. This last requirement is necessary for the following reason: the essential type of y can be obtained from that of x by replacing some entries with stars; if this could decrease the weight of the resulting sequence, one might need to look further to determine $t_{\text{ess}}^{(k,w)}(y)$, and $t_{\text{ess}}^{(k,w)}(x)$ might not be sufficient.

Under these conditions, it follows that $t_{\text{ess}}^{(k)}$ and $t_{\text{ess}}^{(k,w)}$ are (Perron–Frobenius) type systems, in the sense of Definition 1.2. Hence, we can use Theorem 1.4 to estimate the spectral radius of the corresponding random walk. Again, it turns out that it is more efficient to truncate using weights than length.

In genus 2, taking the weights $w = (1, 2, 3, 4, 1, 2, 4)$ and the threshold $k = 25$, we obtain a matrix of size 8 999 902. The corresponding bound on the spectral radius is $\rho \geqslant 0.662772$, proving Theorem 1.1. Those bounds were obtained on a personal computer with a memory of 12GB (memory is indeed the main limiting factor, since one should store all truncated essential types to create the matrix M). With more memory, one would get better estimates, but it is unlikely that those estimates converge to the true spectral radius when k tends to infinity: to recover it, it is probably necessary to distinguish even more points, for instance by using Cannon's cone types instead of the canonical types (but this would become totally impractical).

In higher genus, here are the bounds we obtain.

Theorem 3.5. *In genus* 3*, using* $t_{\text{ess}}^{(k,w)}$ *with* $w = (1, 2, 3, 4, 5, 6, 1, 2, 6)$ *and* $k = 25$*, we get a matrix of size* 7 307 293 *and the estimate* $\rho \geqslant 0.5527735593$.

In genus 4*, using* $t_{\text{suff}}^{(k,w)}$ *with* $w = (1, 2, 3, 4, 5, 6, 7, 8)$ *and* $k = 24$*, we get a matrix of size* 4 120 495 *and the estimate* $\rho \geqslant 0.48412292068$.

When the genus increases, the groups look more and more like free groups. This means that the spectral radius is very close to that of the random walk on a tree with the corresponding number of generators (*i.e.*, $\sqrt{4g-1}/(2g)$ in general, specializing to 0.55277079 in genus 3 and 0.48412291827 in genus 4), and to get a significant improvement one needs to take very large matrices. The path counting arguments of Bartholdi [1], on the other hand, are more and more precise when the groups looks more and more like a free group: in genus 3, he gets $\rho \geqslant 0.5527735401$, which is just slightly worse than our estimate, while requiring considerably less computer power. In genus 4, he gets $\rho \geqslant 0.48412292074$, which is already better than our estimate, and the situation is certainly the same in higher genus.

In genus 3, the upper bound of Nagnibeda [7] is $\rho \leqslant 0.552792$, while our lower bound (or Bartholdi's) is much closer to the naive lower bound coming from the free group. It is unclear which one is closer to the real value of the spectral radius.

For the practical implementation, as in the end of Section 3.2, it is important to know the asymptotics of the number of points on $\mathbb{S}^n$ having a given truncated essential type. We illustrate how to compute such asymptotics in three significant examples, which can be combined to handle the general case.

(1) Assume first that the type $(i_0, \dots, i_m)$ does not contain any ambiguous letter, *i.e.*, $i_\ell \neq *$ for all ℓ. In this case, the formula (3.2) still holds.

(2) Assume now that the type is of the form $(i, *, \dots, *, j)$ for some types $i, j \neq *$, and some number N of stars. Let x be a point with the above truncated essential type, on a sphere $\mathbb{S}^n$, and let y be the point of modified type j, on the sphere $\mathbb{S}^{n-N-1}$, such that any geodesic from e to x goes through y. Since the type $2g$ is the only one to have several predecessors, necessarily $i = 2g$. On the other hand, j can be any type in $\{1, 1', 2, 2', 3, \dots, 2g\}$.

The discussion in the proof of Proposition 3.4 (see in particular Figure 5(b,c)) implies that $N = (2g-1)m$, for some integer m: a geodesic from e to x goes through y, and then it follows m $4g$-gons.

Let us first study the case $m = 1$. Consider a point y of type j, and a $4g$-gon R based at y (*i.e.*, y is the closest point to e on this $4g$-gon). There are $4g-2$ such R if $j \neq 2g$ and $4g-3$ if $j = 2g$ (since a point of type $2g$ has two incoming edges). On R, consider the point x that is the farthest from e: it has type $2g$, and one can reach it from y by going around R in one direction or the other. It follows that

$$\begin{aligned}\operatorname{Card}\{x \in \mathbb{S}^n \ :\ t^{(w,k)}_{\text{suff}}(x) = (2g, *, \dots, *, j) \text{ with } N = 2g-1 \text{ stars}\}\\ = a_j \operatorname{Card}\{y \in \mathbb{S}^{n-N-1} \ :\ t'(y) = j\},\end{aligned}$$

where $a_j = 4g-3$ if $j \in \{2g-1, 2g\}$, and $a_j = 4g-2$ otherwise. The case of a point y of type $j = 2g-1$ is special since, on the $4g$-gon R containing the successors of y of types $1'$ and $2'$, one of the geodesic paths from y to the opposite point x goes through a vertex of type $2g$, giving rise to further ambiguities. Hence, this $4g$-gon should be discarded from the above counting, leaving only $2g-3$ suitable $4g$-gons.

In the case of a general $m \geqslant 1$, the number of points x corresponding to a given point y of type j may be obtained first by choosing a suitable $4g$-gon R_1 based at y (giving a_j choices). Further ambiguities can only be obtained by choosing one of the two predecessors (of type $2g-1$) of the point that is opposite to y on R_1, and then following the $4g$-gon R_2 based at this point and containing its successors of type $1'$ and $2'$. This gives two choices for R_2, then two more choices for the next $4g$-gon R_3, and so on. In the end, we obtain

$$\begin{aligned}\operatorname{Card}\{x \in \mathbb{S}^n \ :\ t^{(w,k)}_{\text{suff}}(x) = (2g, *, \dots, *, j) \text{ with } N = (2g-1)m \text{ stars}\}\\ = a_j 2^{m-1} \operatorname{Card}\{y \in \mathbb{S}^{n-N-1} \ :\ t'(y) = j\}.\end{aligned}$$

(3) Finally, assume that the type is $(i, *, \dots, *)$ with some number N of stars (the situation is very similar to the previous one). Necessarily, as above, $i = 2g$. Write $N = (2g-1)m + k$ for some $k \in \{1, \dots, 2g-1\}$.

For $m = 0$, a point of type $2g$ has two predecessors, so there are always ambiguities regarding his first $2g-1$ ancestors. Hence, the set of points we are considering is simply the set of points of type $2g$ on the sphere $\mathbb{S}^n$, and there is nothing to do.

For $m = 1$, there can be further ambiguities only if x has an ancestor x' of generation $2g-1$ that has type $2g$, and x is at the tip of the $4g$-gon R based at one of the two predecessors of x' (of type $2g-1$), and containing x'. There are two choices for R.

Proceeding inductively in the case of a general m, we get

$$\begin{aligned}\operatorname{Card}\{x \in \mathbb{S}^n \ : \ t_{\text{suff}}^{(w,k)}(x) = (2g, *, \dots, *) \text{ with } N = (2g-1)m + k \text{ stars}\}\\ = 2^m \operatorname{Card}\{y \in \mathbb{S}^{n-(2g-1)m} \ : \ t(y) = 2g\}.\end{aligned}$$

A general truncated essential type is the concatenation of successive types of the form just described in the examples. We can thus count the number of points with a given type just by combining the above formulas.

Acknowledgements

We thank the anonymous referee for pointing out the reference [8], and suggesting that there might be a possible geometric interpretation to the lower bound in Theorem 1.4, as in [8]. This led to Section 2.

References

[1] Bartholdi, L. (2004) Cactus trees and lower bounds on the spectral radius of vertex-transitive graphs. In *Random Walks and Geometry*, de Gruyter, pp. 349–361.
[2] Cannon, J. W. (1983) The growth of the closed surface groups and the compact hyperbolic Coxeter groups. Preprint.
[3] Elder, M., Rechnitzer, A. and Wong, T. (2012) On the cogrowth of Thompson's group *F*. *Groups Complex. Cryptol.* **4** 301–320.
[4] Floyd, W. J. and Plotnick, S. P. (1987) Growth functions on Fuchsian groups and the Euler characteristic. *Inventio Math.* **88** 1–29.
[5] Gouëzel, S. and Lalley, S. P. (2013) Random walks on co-compact Fuchsian groups. *Ann. Sci. Éc. Norm. Supér.* (4) **46** 129–173.
[6] Kesten, H. (1959) Symmetric random walks on groups. *Trans. Amer. Math. Soc.* **92** 336–354.
[7] Nagnibeda, T. (1997) An upper bound for the spectral radius of a random walk on surface groups. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* (POMI) **240** 154–165, 293–294.
[8] Nagnibeda, T. (2004) Random walks, spectral radii, and Ramanujan graphs. In *Random Walks and Geometry*, de Gruyter, pp. 487–500.
[9] Nagnibeda, T. and Woess, W. (2002) Random walks on trees with finitely many cone types. *J. Theoret. Probab.* **15** 383–422.
[10] Woess, W. (2000) *Random Walks on Infinite Graphs and Groups*, Vol. 138 of *Cambridge Tracts in Mathematics*, Cambridge University Press.