



Mark Sapir · Iva Špakulová

## Almost all one-relator groups with at least three generators are residually finite

Received September 12, 2009 and in revised form March 31, 2010

**Abstract.** We prove that with probability tending to 1, a one-relator group with at least three generators and the relator of length  $n$  is residually finite, is a virtually residually (finite  $p$ -)group for all sufficiently large  $p$ , and is coherent. The proof uses both combinatorial group theory and non-trivial results about Brownian motions.

### 1. Introduction

Residual finiteness of one-relator groups has been one of the main topics in combinatorial group theory since the 1960s. The first non-residually finite examples were given in [2] (say, the Baumslag–Solitar groups  $BS(p, q) = \langle a, b \mid b^{-1}a^pb = b^q \rangle$  where  $p$  and  $q$  are different primes). Possibly the strongest “positive”, nonprobabilistic result so far is the result by D. Wise [16]: any one-relator group whose relator is a *positive word* satisfying the condition  $C'(1/6)$  is residually finite. The strongest “negative” result appeared recently in the paper by Baumslag, Miller and Troeger [1]: Let  $G = \langle a, b, \dots \mid r = 1 \rangle$  be a one-relator group with at least two generators and let  $G(r, w) = \langle a, b, \dots \mid r^{r^w} = r^2 \rangle$  where  $w$  is an element of a free group with free generators  $a, b, \dots$  that does not commute with  $r$ . Then [1, Theorem 1] asserts that the group  $G(r, w)$  is not residually finite. Note that the length of the relator of  $G(r, w)$  is at most a constant multiple of the length of the relator of  $G$ .

Another important problem about one-relator groups is whether every one-relator group is *coherent*, i.e. every finitely generated subgroup of it is finitely presented.

In this paper, we show that generically one-relator groups with at least three generators have both properties: they are residually finite and coherent. We also discuss the case of one-relator groups with two generators at the end of the paper.

We consider three natural models of choosing a random one-relator group.

**Model NR.** For every  $r \geq 0$ , consider the set  $S_r$  of all group words  $R$  of length  $r$  in a free group  $F_k = \langle x_1, \dots, x_k \rangle$ . On that set, we choose the uniform probability measure.

M. Sapir: Department of Mathematics, Vanderbilt University, Nashville, TN 37240, USA;  
e-mail: mark.sapir@vanderbilt.edu, url: www.math.vanderbilt.edu/~msapir

I. Špakulová: Department of Mathematics, Vanderbilt University, Nashville, TN 37240, USA;  
e-mail: iva.spakulova@gmail.com

By a *random one-relator group with  $k$  generators of complexity  $r$*  we mean the group with presentation  $\langle x_1, \dots, x_k \mid R = 1 \rangle$  where  $R$  is a random word from  $S_r$ .

**Model CR.** In this model, we consider the set  $CS_r$  of cyclically reduced words in  $F_k$  of length  $r$  and consider the uniform probability measure on that set. Then a *random one-relator group with  $k$  generators of complexity  $r$*  is a group  $\langle x_1, \dots, x_k \mid R = 1 \rangle$  where  $R$  is a random word from  $CS_r$ .

The third model is one considered, for example, in Kapovich–Schupp–Shpilrain [13].

**Model IC.** Consider the equivalence relation on the set of all cyclically reduced words from  $F_k$ : two words are equivalent if the corresponding one-relator groups are isomorphic. In this model we consider any set  $D_r$  of representatives of length exactly  $r$  of all equivalence classes of words containing words of length  $r$ . Consider the uniform probability measure on  $D_r$ . Then by a *random one-relator group with  $k$  generators of complexity  $r$*  we mean the group with presentation  $\langle x_1, \dots, x_k \mid R = 1 \rangle$  where  $R$  is a random word from  $D_r$ .

Now given any property  $P$  of groups and any of the three probabilistic models above, consider the probability  $p_r$  that a random  $k$ -generator one-relator group with relator of length  $r$  has property  $P$ . If  $p_r$  has a limit  $p$ , we say that a random  $k$ -generator one-relator group has this property with probability  $p$ .

We prove below (see Lemmas 3.1, 3.2) that if the limit of the probabilities  $p_r$  exists in the random model CR, then it coincides with the limit in the models NR and IC.

Recall that an *ascending HNN extension* of a free group (the mapping torus of a free group endomorphism) is a group of the form  $\text{HNN}_\phi(F_k) = \langle x_1, \dots, x_k, t \mid x_i^t = \phi(x_i), \dots, x_k^t = \phi(x_k) \rangle$  where  $\phi$  is an injective endomorphism of  $F_k$ .

Here is the main result of this paper.

**Theorem 1.1.** *A random  $k$ -generator one-relator group,  $k \geq 3$ , can be embedded into an ascending HNN extension of a finitely generated free group with probability 1. Therefore almost surely, such a group is residually finite, is a virtually residually (finite  $p$ -)group for every sufficiently large prime  $p$ , and is coherent.*

As an immediate corollary of Theorem 1.1, we deduce that one cannot replace a multiplicative constant in the result of Baumslag–Miller–Trager [1, Theorem 1] mentioned above by an additive constant: if  $n \geq 3$ , then there exists no map  $\phi: F_n \rightarrow F_n$  such that  $|\phi(R)| - |R|$  is bounded from above by some constant  $C$  and such that for every non-trivial  $R \in F_n$ , the group  $\langle F_n \mid \phi(R) = 1 \rangle$  is not residually finite. Indeed, it is easy to see that if such a map existed, the probability of a one-relator group with  $n$  generators to be residually finite would be bounded away from 1 as  $|R|$  tends to  $\infty$ .

Although the proof of Theorem 1.1 presented below is relatively short, it uses some strong results from different areas of mathematics that are very rarely employed together: geometric group theory, algebraic geometry and probability theory (Brownian motion). Namely, we use the result of Feighn and Handel [10] that ascending HNN extensions of free groups are coherent (that is mostly geometric group theory, more precisely discrete

Morse theory), results by Borisov and Sapir from [4], [5] that every ascending HNN extension of a free group is residually finite and even a virtually residually (finite  $p$ -)group for almost all primes  $p$  (that is essentially a result from algebraic geometry, namely the theory of quasi-fixed points of polynomial maps over finite fields), a result of Ol'shanskiĭ [14] about subgroups with congruence extension property of hyperbolic groups (this is geometric group theory), a result of Kapovich, Schupp and Shpilrain [13] about generic solvability of the isomorphism problem for one-relator groups (this result uses many parts of geometric group theory including the Arzhantseva–Ol'shanskiĭ method, boundaries of hyperbolic groups, etc.), and a result by Cranston, Hsu, and March [8] that the boundary of the convex hull of a Brownian trajectory is smooth almost surely in Wiener's measure (probability theory).

Note that almost all (with probability tending to 1) one-relator groups with a relator of size  $r \gg 1$  satisfy the small cancellation condition  $C'(1/6)$  and in fact  $C'(\lambda)$  for every fixed  $\lambda > 0$  [12]. Hence they are hyperbolic almost surely. It is still a major open question in group theory whether every hyperbolic group is residually finite. The positive answer would of course imply a part of Theorem 1.1. Most specialists believe, however (see, for example, Gromov's conjecture [12, 5.B]), that there are non-residually finite hyperbolic groups, and that, moreover, almost all hyperbolic groups and even most  $C'(\lambda)$ -groups are not residually finite. Constructing an example of a non-residually finite hyperbolic group is difficult because all non-elementary hyperbolic groups have “very many” quotients, including torsion and even, in most cases, bounded torsion, quotients [12, 14].

## 2. The theory of one-relator groups

### 2.1. The case of two generators

Let  $G = \langle a, b \mid R = 1 \rangle$  be a one-relator group,  $R$  is a cyclically reduced word in  $F_2 = \langle a, b \rangle$ . Consider a square lattice  $\Gamma$  in  $\mathbb{R}^2$ , the Cayley graph of  $\mathbb{Z}^2$ . We assume that horizontal edges are labeled by  $a$  and vertical edges are labeled by  $b$ . Let  $\psi : F_2 \rightarrow \mathbb{Z}^2$  be the abelianization map. Let  $w$  be the path in  $\Gamma$  starting at the origin  $(0, 0)$  and reading the word  $R$ . This  $w$  is called the *trace* of the relator  $R$ . Note that  $w$  can visit every vertex (edge) many times. Vertices (edges) visited only once are called *simple*. A line  $L$  in  $\mathbb{R}^2$  is said to be a *supporting* line of  $w$  if the path  $w$  lies on one side of  $L$  and has a common vertex with  $L$ .

**Theorem 2.1** (Brown [6, Theorem 4.4]). *Let  $G = \langle a, b \mid R = 1 \rangle$ , where  $R$  is a non-trivial cyclically reduced word in the free group on  $\{a, b\}$  and  $R \notin [F_2, F_2]$ . Let  $w$  be the trace of  $R$ , ending at a point  $(m, n)$ .*

*$G$  is an ascending HNN extension of a free group if and only if one of the two supporting lines of  $w$  parallel to the vector  $(m, n)$  intersects  $w$  in one simple vertex or one simple edge.*

2.2. *Embedding into 2-generated groups*

Let  $G = \langle x_1, \dots, x_k \mid R = 1 \rangle$ . If the sum of the exponents of  $x_i$  in  $R$  is 0, then we can apply the Magnus rewriting to  $R$ . It consists of

- removing all occurrences of  $x_i$  in  $R$ ;
- replacing every occurrence of a letter  $x_j$  in  $R$  by the letter  $x_{j,p}$  where  $p$  is the sum of exponents of the  $x_i$  in the prefix of  $R$  before that occurrence of  $x_j$ .

Let  $R'$  be the resulting word. The second indices  $p$  of letters in  $R'$  will be called the *Magnus  $x_i$ -indices*. We say that a certain Magnus index is *unique* if it occurs only once in  $R'$ .

We are going to use the following statement, which can be deduced from, say, a general result in [14] about hyperbolic groups.

**Lemma 2.2.** *Let  $w_1, \dots, w_k$  be words in the free group  $F_n$  satisfying  $C'(1/12)$ . Then the subgroup  $H = \langle w_1, \dots, w_k \rangle$  of  $F_n$  satisfies the congruence extension property, that is, for every normal subgroup  $N$  of  $H$ , the intersection of  $H$  and the normal closure  $N^{F_n}$  of  $N$  in  $F_n$  is  $N$ . In particular, the natural homomorphism  $H/N \rightarrow F_n/N^{F_n}$  is injective.*

Let  $\phi$  be the map  $F_k \rightarrow F_n$  (where  $F_n = \langle x_1, \dots, x_n \rangle$ ) given by  $x_i \mapsto w_i, i = 1, \dots, k$ , where  $w_1, \dots, w_k$  satisfy  $C'(1/12)$ . Lemma 2.2 immediately implies

**Lemma 2.3.** *The map  $\phi$  induces an injective homomorphism from  $G = \langle x_1, \dots, x_k \mid R = 1 \rangle$  to the one-relator  $n$ -generated group  $\langle x_1, \dots, x_n \mid \phi(R) = 1 \rangle$ .*

**Theorem 2.4.** *Consider a group  $G = \langle x_1, \dots, x_k \mid R = 1 \rangle$ , where  $R$  is a word in the free group on  $\{x_1, \dots, x_k\}$ ,  $k \geq 2$ . Assume the sum of the exponents of  $x_k$  in  $R$  is zero and the maximal Magnus  $x_k$ -index of  $x_1$  is unique. Then  $G$  can be embedded into an ascending HNN extension of a finitely generated free group.*

*Proof.* We may assume that the maximal Magnus  $x_k$ -index of  $x_1$  is greater than the one of  $x_i$ , for  $1 < i < k$ , otherwise apply the automorphism  $x_i \mapsto x_k^{-m} x_i x_k^m, x_j \mapsto x_j$  ( $j \neq i$ ) for  $m$  large enough.

Let  $n \gg 1$ . Consider the following words  $w_1, \dots, w_k \in F_2$ :

$$\begin{aligned} w_1 &= aba^2b \dots a^n ba^{n+1} ba^{-n-1} ba^{-n} b \dots a^{-2} ba^{-1} b, \\ w_i &= ab^i a^2 b^i \dots a^n b^i a^{-n} b^i \dots a^{-2} b^i a^{-1} b^i \quad \text{for } 1 < i < k, \\ w_k &= ab^k a^2 b^k \dots a^n b^k a^{-n} b^k \dots a^{-2} b^k. \end{aligned}$$

These words satisfy the following conditions:

- (1) For  $n$  large enough, these words and their cyclic shifts satisfy the small cancellation condition  $C'(1/12)$ . Indeed, the maximal length of a subword repeating twice as a prefix of cyclic shifts of  $w_i$  does not exceed  $2n + 3 + k$ , and the length of each  $w_i$  is at least  $n^2$ . For  $n$  large enough, we have  $(2n + 3 + k)/n^2 < 1/12$ .
- (2) The sum of the exponents of  $a$  in  $w_i, i < k$ , is 0, the sum of the exponents of  $a$  in  $w_k$  is 1.

- (3) The maximal Magnus  $a$ -index of  $b$  in  $w_1$  is  $(n + 1)(n + 2)/2$ , and this index is unique. The maximal Magnus  $a$ -indices of  $b$  in all other words are strictly smaller than the one in  $w_1$ .

By Lemma 2.3, the group  $G$  embeds into the 2-generated one-relator group with presentation  $\langle a, b \mid R(w_1, \dots, w_k) = 1 \rangle$ .

It remains to prove that  $R(w_1, \dots, w_k)$  satisfies the conditions of Lemma 2.1. Let  $R' = R(w_1, \dots, w_k)$ . Clearly the sum of the exponents of  $a$  in  $R'$  is zero. Every letter  $b$  with maximal Magnus  $a$ -index in  $R'$  comes from some occurrence of a word  $w_i$  substituted for the letter  $x_i$ . The sum of the exponents of  $a$  is nonzero only in  $w_k$ . Therefore the Magnus  $a$ -index of a letter  $b$  is the sum of the Magnus  $x_k$ -index of the letter  $x_i$  in  $R$ , for which it was substituted, and the Magnus  $a$ -index of  $b$  in  $w_i$ . The Magnus  $x_k$ -index in  $R$  is maximal for the letter  $x_1$  and the maximum is unique in  $R$ . The maximal Magnus  $a$ -index of  $b$  in  $w_1$  is also unique (and greater than in all other  $w_i$ 's). This gives the uniqueness of the maximal Magnus  $a$ -index in  $R'$ . Therefore there is a supporting line parallel to the  $b$ -axis that intersects the trace of  $R'$  in one simple edge corresponding to the letter  $b$  with the maximal Magnus  $a$ -index. Therefore by Lemma 2.1 the group  $\langle a, b \mid R' = 1 \rangle$  is an ascending HNN extension of a finitely generated free group.  $\square$

### 2.3. More than two generators and walks in $\mathbb{Z}^k$

In the case of more than two generators we generalize the notion of supporting line in the following way. Given a relator  $R$ , a nontrivial word in the free group on  $\{x_1, \dots, x_k\}$ , let  $w$  be its trace in the lattice  $\mathbb{Z}^k$ . For a letter  $t \in \{x_1, \dots, x_k\}$ , let  $w_t$  be a set of edges labeled by  $t$  in  $w$ . A vertex on  $w_t$  is called *simple* if it does not belong to two edges of  $w_t$ . In particular, if  $w$  contains two consecutive edges with labels  $t, t^{-1}$ , then the endpoints of those edges are not simple vertices.

**Definition 2.5.** A hyperplane  $P$  is a *supporting* hyperplane of  $w_t$  if the trace  $w_t$  lies on one side of  $P$  and has a common vertex with  $P$ . A hyperplane  $P$  is said to *touch*  $w$  if

- $P$  is parallel to the line containing the origin and the endpoint of  $w$ ,
- there is  $t \in \{x_1, \dots, x_k\}$  such that  $P$  is a supporting hyperplane of  $w_t$ ,
- the intersection of  $P$  and  $w_t$  consists of one simple vertex or one simple edge.

**Lemma 2.6.** Let  $G = \langle x_1, \dots, x_k \mid R = 1 \rangle$ , where  $R$  is a word in the free group on  $\{x_1, \dots, x_k\}$ ,  $k \geq 2$ . Let  $w$  be a trace of  $R$  in the lattice  $\mathbb{Z}^k$ . If there is a hyperplane  $P$  touching  $w$ , then  $G$  can be embedded into an ascending HNN extension of a free group.

*Proof.* We will embed  $G$  into a one-relator group on  $k + 1$  generators that satisfies the condition of Theorem 2.4.

If the normal vector of  $P$  has irrational entries, then there is a hyperplane  $P'$  whose normal vector has rational entries that also touches  $w$ . Thus we can assume that the normal vector of  $P$  pointing toward the half-space not containing  $w_t$  is  $(n(1), \dots, n(k))$  with integer entries.

Consider the following substitution  $\phi$ :

$$x_i \mapsto x_i z^{n(i)}, \quad i = 1, \dots, k.$$

Let  $H = \langle x_1, \dots, x_k, z \mid \phi(R) \rangle$ . Then  $G$  is embedded into  $H$  by  $\phi$ . Since the normal vector of  $P$  is orthogonal to the line connecting the origin and the endpoint of  $w$ , the sum of the exponents of  $z$  in  $\phi(R)$  is zero.

It remains to show that the maximal Magnus  $z$ -index of  $x_t$  in  $\phi(R)$  is unique.

We can assume that the edge in  $w_t$  intersecting  $P$  corresponds to the first letter of  $R$ . Assume that there is another letter  $x_t$  (at position  $j$ ) in  $\phi(R)$  with at least the same Magnus  $z$ -index as the first letter  $x_t$  in the word. Let  $m(i)$  be the total sum of the exponents of the letter  $x_i$  between those two occurrences of  $x_t$  (note that it is the same in  $R$  and in  $\phi(R)$ ). If the exponent of the first  $x_t$  is 1, then add 1 to  $m(t)$ . If the exponent of the other letter  $x_t$  (at position  $g$ ) is  $-1$ , then subtract 1 from  $m(t)$ . The Magnus  $z$ -index of the letter  $x_t$  at position  $j$  differs from the Magnus  $z$ -index of the first  $x_t$  by precisely  $m(1)n(1) + m(2)n(2) + \dots + m(k)n(k)$ .

Consider the edge corresponding to the first letter  $x_t$  and the edge of the letter  $x_t$  at position  $j$ . Connect their initial points in  $\mathbb{Z}^k$  by a vector (the vector connecting their terminal points is the same). It is easy to see that the coordinates of this vector are  $(m(1), \dots, m(k))$ . If the scalar product of this vector with the normal vector of  $P$  is non-negative, then one of the endpoints of the edge of the letter  $x_t$  at position  $g$  lies at  $P$  or on the other side than  $w_t$ . This is impossible, because  $P$  touches  $w$  (with respect to  $x_t$ ).  $\square$

**Remark 2.7.** Let  $R$  be a nonreduced word in  $\{x_1, \dots, x_n\}$ , and let  $R'$  be the cyclically reduced form of  $R$ . Let  $w, w'$  be the walks corresponding to  $R$  and  $R'$  respectively. If there exists a touching plane for  $w$ , then there exists a touching plane for  $w'$ . The proof easily proceeds by induction on the number of reductions.

Let  $w$  be the walk in  $\mathbb{Z}^k$  corresponding to  $R$ . Let  $\xi$  be the vector connecting the initial and the terminal points of  $w$ . Let  $t \in \{1, \dots, k\}$ . For every supporting plane  $P$  of  $w_t$  let  $P^+$  be the closed half-space of  $\mathbb{R}^k$  bounded by  $P$  and containing  $w_t$ . The intersection of all  $P^+$  is a convex polyhedron in  $\mathbb{R}^k$ . Let  $\Delta_0(t)$  denote the projection of the boundary  $\Delta(t)$  of that polyhedron onto the hyperplane orthogonal to  $\xi$ . Then  $\Delta(t)$  is the right cylinder with base  $\Delta_0(t)$ , i.e. the direct product  $\Delta_0(t) \times \mathbb{R}$ . A vertex of the random walk projected to a 0-cell of  $\Delta_0(t)$  is called a *corner*. For every vertex  $x$  that is a 0-cell of a  $\Delta_0(t)$ , the line  $x + \mathbb{R}\xi \subseteq \Delta(t)$  will be called the *support line* of  $w_t$ .

Lemma 2.6 immediately implies

**Lemma 2.8.** *If one of the support lines of  $w_t$  intersects  $w_t$  in a simple vertex or a simple edge, then  $G$  is embeddable into an ascending HNN extension of a free group.*

### 3. Random walks in $\mathbb{Z}^k$

#### 3.1. Preliminaries

Denote by  $P_n^{NR}$  the (uniform) measure on simple random walks of length  $n$  (not necessarily reduced) and by  $P_n^{NB}$  the uniform measure on nonbacktracking simple random walks

of length  $n$ . To model cyclically reduced words, we denote by  $P_n^{CR}$  the uniform measure on nonbacktracking simple random walks with last edge that is not the inverse of the first edge of the walk (note that asymptotically this happens with probability  $(2k - 1)/2k$ ). In all cases we can consider the sample space  $\Omega$  containing all walks of any finite length.

We say that an event  $A$  depends only on the cyclically reduced path of the random walk if  $w \in A$  if and only if  $w' \in A$ , where  $w'$  is the cyclically reduced path of  $w$ . An example of such an event is the event that a support line of the cyclically reduced path  $w'$  of a random walk  $w$  intersects  $w'$  in a simple vertex or a simple edge.

**Lemma 3.1.** *Let  $A$  be an event depending only on the cyclically reduced path of the random walk. Assume  $\lim_{n \rightarrow \infty} P_n^{CR}(A)$  exists. Then*

$$\lim_{n \rightarrow \infty} P_n^{NR}(A) = \lim_{n \rightarrow \infty} P_n^{CR}(A).$$

*Proof.* Let  $\lim_{n \rightarrow \infty} P_n^{CR}(A) = a$  and assume  $n_0$  is such that for all  $n > n_0$ ,

$$|P_n^{CR}(A) - a| < \epsilon.$$

If  $A$  depends only on the cyclically reduced path  $w'$  of a random walk  $w$ , then conditioning on the length of the cyclically reduced path  $|w'|$  we see that  $P_n^{NR}(A \mid |w'| = k) = P_k^{CR}(A)$ , provided  $P_n^{NR}(|w'| = k) > 0$ . Let  $n_1$  be such that for all  $n > n_1$ ,  $P_n^{NR}(|w'| < n_0) \leq \epsilon$ . Then

$$P_n^{NR}(A) = \sum_{k=0}^n P_n^{NR}(|w'| = k) P_k^{CR}(A),$$

and we can split the sum into two parts ( $k \leq n_0$  and  $k > n_0$ ) and obtain, for  $n > n_1$ ,

$$(1 - \epsilon)(a - \epsilon) < P_n^{NR}(A) < \epsilon + (a + \epsilon).$$

Therefore  $\lim_{n \rightarrow \infty} P_n^{NR}(A) = a$ . □

**Lemma 3.2.** *Assume  $\lim_{n \rightarrow \infty} P_n^{CR}(A)$  exists. Then the limit probability in the model IC exists as well and they are the same.*

*Proof.* By the result of Kapovich, Schupp and Shpilrain [13, Theorem C], there is a generic set  $Q$  of cyclically reduced words such that two one-relator groups with relators in  $Q$  are isomorphic if and only if their relators are obtained from each other by a relabeling automorphism and cyclic shift. This set is generic in the following sense:

$$\lim_{n \rightarrow \infty} \frac{|T_n \cap Q|}{|T_n|} = 1,$$

where  $T_n$  consists of all cyclically reduced words in  $F_k$  of length  $n$ .

Assume  $\lim_{n \rightarrow \infty} P_n^{CR}(A) = p$ . If we consider only words in  $Q$ , the limit clearly stays the same. Each word in  $D_r$  represents several words in  $Q \cup T_r$ . The maximum number of words which may be relators of pairwise isomorphic one-relator groups is the number of relabeling automorphisms ( $2^k k!$ ) times the number of cyclic shifts ( $r$ ). There are words

for which different cyclic shift are equal words after relabeling (these words are products of more than one copy of the same word in different alphabets), and there are words containing fewer than  $k$  letters, but the probability of obtaining such a word tends to 0 (exponentially) as the length grows. If we exclude such words from  $Q \cap T_r$ , then the set of groups given by relators in  $Q \cap T_r$  contains groups isomorphic to those with a relator from  $D_r$  (that is, the set from the model IC), each with the same multiplicity. Therefore the limit probability in the model IC is equal to  $p$  as in the model CR.  $\square$

Next we will need a modified version of Donsker's invariance principle. Denote by  $C$  the space of all continuous functions  $f : [0, 1] \rightarrow \mathbb{R}^k$  such that  $f(0) = 0$ , equipped with the sup norm.

**Theorem 3.3** (Donsker's Theorem modified). *Consider a piecewise linear function  $Y_n : [0, 1] \rightarrow \mathbb{R}^k$ , where the line segments connect the points  $Y_n(t) = S_{nt}/\sqrt{n}$  for  $t = 0, 1/n, 2/n, \dots, n/n = 1$ , where  $(S_n)$  has a distribution according to  $\mathbb{P}_n^{CR}$ . Then  $Y_n(t)$  converges in distribution to a Brownian motion as  $n \rightarrow \infty$ .*

*Proof.* First we prove that conditioning on the first step of a nonbacktracking random walk has asymptotically no influence on  $Y_n(t)$ , which allows us to switch between  $\mathbb{P}^{NB}$  and  $\mathbb{P}^{CR}$ . Next, we basically repeat the proof of Donsker's Theorem in [3, Theorem 10.1]. The Central Limit Theorem for nonbacktracking walks that we will use was proved in [15].

Let  $(R_n)$  be a nonbacktracking random walk. We cut the walk at time  $\ln(n)$ , splitting the walk into two (dependent) parts  $(R_{\ln(n)}^{(1)})$  and  $(R_{n-\ln(n)}^{(2)})$ . Define piecewise linear functions  $X(t)$  and  $Z(t)$  connecting the points  $X(t) = R_{nt}/\sqrt{n}$  and  $Z(t) = R_{nt}^{(2)}/\sqrt{n - \ln(n)}$  respectively. Clearly, the distance (in the sup norm) between  $X(t)$  and  $Z(t)$  goes to 0 as  $n \rightarrow \infty$ . Moreover the latter part of the walk  $(R_{n-\ln(n)}^{(2)})$  tends to be independent of the first step of  $(R_n)$  as  $n \rightarrow \infty$ . Therefore the piecewise linear functions obtained from walks with measures  $\mathbb{P}^{NB}$  and  $\mathbb{P}^{CR}$  have the same limit in distribution.

Next we show that the finite-dimensional distribution of  $Y_n(t)$  converges to the one of Brownian motion. By the result of Rivin [15, Theorem 5.1] the probability distribution of  $S_n/\sqrt{n}$  converges to a normal distribution on  $\mathbb{R}^k$ , whose mean is 0 and covariance matrix is diagonal, with entries

$$\sigma^2 = \frac{1}{\sqrt{2k-1}} \left[ 1 + \left( \frac{c+1}{c-1} \right)^{1/2} \right],$$

where  $c = k/\sqrt{2k-1}$ . By the previous paragraph this holds for  $R_n/\sqrt{n}$  as well.

Consider now the two-dimensional distribution that is the position at two time points,  $s < t$ . It is enough to show that  $S_{ns}/\sqrt{n}$  and  $(S_{nt} - S_{ns})/\sqrt{n}$  are asymptotically independent (the normal distribution of each of them was already established). The first step of  $(S_{nt} - S_{ns})$  is not independent of  $(S_{ns})$ , but asymptotically the distribution of  $(S_{nt} - S_{ns})/\sqrt{n}$  is independent of the first step. The convergence of finite-dimensional distributions for more time points can be proved in the same way.



It remains to show the tightness of the process. We refer to the proof in Billingsley [3, p. 69], and here we prove only the lemma needed. The claim is

$$\mathbb{P}_n^{NB} \left( \max_{i < n} |S_i| \geq \lambda \sigma \sqrt{n} \right) \leq \mathbb{P}_n^{NB} (|S_n| \geq (\lambda - \sqrt{2}) \sigma \sqrt{n}).$$

In order to prove this, we define events  $E_i = \{\max_{j < i} |S_j| < \lambda \sigma \sqrt{n} \leq |S_i|\}$ . Now we have:

$$\begin{aligned} \mathbb{P}_n^{NB} \left( \max_{i < n} |S_i| \geq \lambda \sigma \sqrt{n} \right) &\leq \mathbb{P}_n^{NB} (|S_n| \geq (\lambda - \sqrt{2}) \sigma \sqrt{n}) \\ &\quad + \sum_{i=1}^{n-1} \mathbb{P}_n^{NB} (E_i \cap \{|S_n| < (\lambda - \sqrt{2}) \sigma \sqrt{n}\}), \\ \mathbb{P}_n^{NB} (E_i \cap \{|S_n| < (\lambda - \sqrt{2}) \sigma \sqrt{n}\}) &\leq \mathbb{P}_n^{NB} (E_i \cap \{|S_n - S_i| \geq \sigma \sqrt{2n}\}) \\ &= \mathbb{P}_n^{NB} (E_i) \mathbb{P}_n^{NB} (|S_n - S_i| \geq \sigma \sqrt{2n}). \end{aligned}$$

The last equality follows from the fact that the length of  $S_n - S_i$  is independent of the walk up to time  $i$ . Now by Chebyshev's inequality  $\mathbb{P}_n^{NB} (|S_n - S_i| \geq \sigma \sqrt{2n}) \leq 1/2$ . The claim follows from

$$\begin{aligned} \sum_{i=1}^{n-1} \mathbb{P}_n^{NB} (E_i \cap \{|S_n| < (\lambda - \sqrt{2}) \sigma \sqrt{n}\}) &\leq \frac{1}{2} \sum_{i=1}^{n-1} \mathbb{P}_n^{NB} (E_i) \\ &\leq \frac{1}{2} \mathbb{P}_n^{NB} \left( \max_{i < n} |S_i| \geq \lambda \sigma \sqrt{n} \right). \quad \square \end{aligned}$$

### 3.2. Corners of random walk

Let  $(S_n)$  be a nonbacktracking random walk in  $\mathbb{Z}^k$  of length  $n$  with last edge that is not the inverse of the first edge (according to the measure  $\mathbb{P}_n^{CR}$ ). Recall that  $\Delta_0$  is the projection of the boundary of the convex hull of the random walk  $(S_n)$  onto the hyperplane orthogonal to  $\xi$ , the vector connecting the initial and the terminal points of the random walk. Denote by  $H_n$  the set of corners, which are the vertices of the random walk that project to the 0-cells of  $\Delta_0$ . We count the corners with their multiplicities.

**Lemma 3.4.** *Let  $(S_n)$  be a nonbacktracking random walk in  $\mathbb{Z}^k$  of length  $n$  with last edge that is not the inverse of the first edge (according to the measure  $\mathbb{P}_n^{CR}$ ). Let  $H_n$  be the set of its corners as defined above. Then for any integer  $m$ ,*

$$\mathbb{P}_n^{CR} (|H_n| < m) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

*Proof.* Consider a piecewise linear function  $X_n : [0, 1] \rightarrow \mathbb{R}^k$ , where the line segments connect the points  $X_n(t) = S_{nt}/\sqrt{n}$  for  $t = 0, 1/n, 2/n, \dots, n/n = 1$ . Recall that  $C$  is the space of all continuous functions  $f : [0, 1] \rightarrow \mathbb{R}^k$  such that  $f(0) = 0$ , equipped with the sup norm. By Theorem 3.3,  $X_n(t)$  converges in distribution to a Brownian motion as  $n \rightarrow \infty$ . Denote by  $A_m$  the subset of  $C$  such that  $f \in A_m$  if the convex hull of the

projection of  $f$  to the hyperplane orthogonal to  $f(1)$  is a  $(k - 1)$ -dimensional (convex) polytope with at most  $m$  0-cells. We will show that the set  $A_m$  is a closed subset of  $C$  in the sup norm and that the Wiener measure of  $A_m$  is zero. It follows from the weak convergence that  $P(X_n(t) \in A_m) \rightarrow 0$  as  $n \rightarrow \infty$ .

First we show that  $A_m$  is closed. Suppose  $f \notin A_m$  is the limit (in the sup norm) of  $f_n \in A_m$ . Let  $p_n$  (resp.  $p$ ) denote the projection on the hyperplane through the origin orthogonal to  $f_n(1)$  (resp.  $f(1)$ ). For any  $\epsilon > 0$  and for all but finitely many  $n$ , we have  $|p(f(t)) - p(p_n(f_n(t)))| < \epsilon$  for all  $t$ . If the convex hull of  $p_n(f_n)$  is a polytope with at most  $m$  0-cells, then the same holds for  $p(p_n(f_n))$ . Denote by  $B_n$  (resp.  $B$ ) the convex hull of  $p(p_n(f_n))$  (resp.  $p(f)$ ). Then for any  $\epsilon > 0$  the boundary of  $B$  is in the Hausdorff  $\epsilon$ -neighborhood of the boundary of  $B_n$  for all but finitely many  $n$ . We need to prove that if a convex body in  $\mathbb{R}^k$  is sufficiently close to some polytope with at most  $m$  0-cells, then the body itself is such a polytope. To prove that, enumerate the 0-cells of  $B_n$  somehow, say  $\{v_{n,1}, \dots, v_{n,m}\}$  (the last few 0-cells may coincide if the total number of 0-cells is smaller than  $m$ ). Choose one convergent subsequence  $\{v_{n_j(i),i}\}_{j=1,2,\dots}$  of  $\{v_{n,i}\}$  for each  $i$  in such a way that  $\{n_j(i+1) : j = 1, 2, \dots\}$  is a subset of  $\{n_j(i) : j = 1, 2, \dots\}$ . Let  $N^{(0)}$  be the sequence  $\{n_j(m) : j = 1, 2, \dots\}$ . There exists a subset  $N^{(1)}$  of  $N^{(0)}$  such that for every  $i, j \in \{1, \dots, m\}$  either  $v_{t,i}$  and  $v_{t,j}$  span a 1-cell in all  $B_t, t \in N^{(1)}$ , or they span a 1-cell in no  $B_t, t \in N^{(1)}$ . Proceeding by induction on the dimension of a cell, we can find an infinite subset  $N$  of natural numbers such that for every subset  $M \subset \{1, \dots, m\}$  either the vertices  $v_{t,i}, i \in M$ , span a cell in  $B_t$  for all  $t \in N$  or they span a cell in none of these  $B_t$ . For every  $M \subseteq \{1, \dots, m\}$  such that  $v_{t,i}, i \in M$ , span a cell  $F_M(t)$  of dimension  $j$  in all  $B_t, t \in N$ , the limit  $\lim_{t \in N} F_M(t)$  exists and is a Euclidean convex polytope of dimension  $j$  spanned by the 0-cells  $v_i, i \in M$ . Hence the convex hull  $B$  of  $p(f)$  is a convex polytope with at most  $m$  0-cells.

To see that  $A_m$  has measure 0 we introduce the following set  $D$  of continuous functions  $[0, 1] \rightarrow \mathbb{R}^k$ . A function  $f \in C$  is in  $D$  if the convex hull of its projection to some two-dimensional plane orthogonal to  $f(1)$  has a smooth boundary, i.e. it is a  $C^1$  curve in the plane. Clearly  $D \cap A_m$  is empty for all  $m$ . Let  $X_t$  be a standard Brownian motion in  $\mathbb{R}^k$ . Then  $X(t) - tX(1)$  is a Brownian bridge in  $\mathbb{R}^k$ . All projections of this Brownian bridge to  $\mathbb{R}^2$  are equivalent in distribution and give Brownian bridges in  $\mathbb{R}^2$ . To conclude that  $D$  has Wiener measure 1, it is enough to show that the convex hull of a planar Brownian bridge has a smooth boundary almost surely. For Brownian motions, that is proved in [8]. We are going to use almost the same argument.

Consider a Brownian bridge and pick any of its extreme points. Move the beginning of the time from 0 to this extreme point and rotate the plane so that the path is in the upper half-plane. The resulting process  $Y_t$  is a Brownian excursion, i.e. it stays in the upper half-plane and returns to the starting point. The same is true for  $Y_{1-t}$ . Then the transformation  $V_t := (1+t)Y_{t/(1+t)}$  is a Brownian meander (see for example [3, p. 68, Exercise 3]). Let  $V_t = (V_t(1), V_t(2))$ . By [7], for any  $c > 0$ , we have

$$P(\inf\{t : t > 0, |V_t(2)| \leq c|V_t(1)|\} = 0) = 1.$$

By reversing the transformation we obtain the same property for  $Y_t$  and  $Y_{1-t}$ . Now the claim follows using the argument from Theorem 1 in [8].  $\square$

We say that a random walk is *bad* if there is no 0-cell of  $\Delta_0$  such that only a single vertex is projected to it.

**Lemma 3.5.** *The probability that a  $k$ -dimensional nonbacktracking simple random walk (with last edge that is not the inverse of the first edge of the walk) is bad in the above sense tends to 0 for  $k > 2$ .*

*Proof.* Let  $(S_n)$  be a  $k$ -dimensional nonbacktracking simple random walk (with last edge that is not the inverse of the first edge of the walk),  $k > 2$ . The number of all cyclically reduced walks,  $|T_n|$ , equals asymptotically  $(2d - 1)^n$ . Let  $B_n$  be the set of all bad walks, i.e. for all 0-cells of  $\Delta_0$  we have at least two vertices projected to it.

Define a map  $\tau_i : B_n \rightarrow T_{n+4}$  that inserts a commutator at the  $i$ -th corner of the random walk in such a way that it produces a new corner. For example, if the corner is between the letters  $x_i$  and  $x_j$ , we can insert  $x_i x_j^{-1} x_i^{-1} x_j$  in between, so that the second vertex of these three new vertices projects outside of  $\Delta_0$  of the original walk. Note that the new walk is not bad anymore.

This map  $\tau_i$  is injective. Moreover the images of the same walk under  $\tau_i$  for different  $i$  are disjoint. The set of bad walks with more than  $K$  corners,  $U_K := \{w \mid w \in B_n, |H_n(w)| > K\}$ , is mapped by  $\tau_1, \dots, \tau_K$  into  $T_{n+4}$ . The union of their images,  $\bigcup_{i=1}^K \tau_i(U_K)$ , is of size  $K|U_K|$ .

For any integer  $K$  we can write

$$\begin{aligned} |B_n| &= |B_n \setminus U_K| + |U_K| \leq P(|H_n| < K) |T_n| + \frac{|T_{n+4}|}{K}, \\ \frac{|B_n|}{|T_n|} &\leq P(|H_n| < K) + \frac{1}{K} \frac{|T_{n+4}|}{|T_n|} \\ &\leq P(|H_n| < K) + \frac{(2d - 1)^4}{K} \rightarrow \frac{(2d - 1)^4}{K}. \end{aligned}$$

This holds for  $K$  arbitrarily large. The first summand tends to 0 as  $n \rightarrow \infty$ , by Lemma 3.5. This implies that the probability of a bad walk is less than any positive number.  $\square$

**Remark 3.6.** Using Lemma 3.1, the same result as in Lemma 3.5 follows for a simple random walk (we consider corners of its reduced form).

*Proof of Theorem 1.1.* Let  $G = \langle x_1, \dots, x_k \mid R = 1 \rangle$  be a random  $k$ -generator one-relator group,  $k > 2$ . If the trace of  $R$  is not bad in the above sense, then there is a hyperplane touching its cyclic reduction. Thus, by Lemma 3.5, there is a hyperplane touching  $R$  with probability tending to 1. By Lemma 2.6, this implies that the group can be embedded into an ascending HNN extension of a free group almost surely.  $\square$

#### 4. What if the number of generators is 2? Some open questions

Theorem 1.1 leaves the case of one-relator groups with two generators open. The reason is the following. In [4], Borisov and Sapir reported a result of computations saying that,

apparently, more than 94% of one-relator groups with two generators and a relator of size  $n \gg 1$  are ascending HNN extensions of free groups. Borisov and Sapir used the Monte-Carlo method for  $n \approx 10^6$ . Schupp and later Dunfield and Thurston [9] conducted similar experiments and came to the same conclusion. Thus in some sense we know (although it is not proved yet) that probably the majority of one-relator groups with two generators are residually finite and coherent. At the same time, Dunfield and Thurston noticed [9] that a 2-generated one-relator group is not almost surely an ascending HNN extension of a free group (that is, the probability that a 2-generated one-relator group with a relator of size  $n$  is an ascending HNN extension of a free group is bounded away from 1 as  $n \rightarrow \infty$ ).

Nevertheless the answer to the following question can be positive:

**Question 4.1.** Is it true that the probability that a 2-generated one-relator group is residually finite is 1?

Answering a question of M. Sapir, Fu and Virág proved (see [11]) that a 2-generated one-relator group is almost surely a very special HNN extension of a free group. That HNN extension is determined by three parameters: an integer  $k$  (the rank of the free group), an integer  $i$  between 1 and  $k$ , and a word  $w$  from the free group  $F_k$ . It is given by the following presentation:

$$H(k, i, w) = \langle a_1, \dots, a_k, t \mid ta_1t^{-1} = a_2, \dots, ta_{i-1}t^{-1} = a_i, ta_it^{-1} = w, \\ twt^{-1} = a_{i+1}, ta_{i+1}t^{-1} = a_{i+2}, \dots, ta_{k-1}t^{-1} = a_k \rangle.$$

**Question 4.2.** Is every group  $H(k, i, w)$  residually finite?

By [11], a positive answer to Question 4.2 implies a positive answer to Question 4.1.

*Acknowledgments.* The authors are grateful to Vadim Kaimanovich, Ilya Kapovich, Yuval Peres, Vladimir Shpilrain and Bálint Virág for their help.

This work was supported in part by the NSF grant DMS 0700811 and by a BSF (USA-Israeli) grant.

## References

- [1] Baumslag, G., Miller III, C. F., Troeger, D.: Reflections on the residual finiteness of one-relator groups. *Groups Geom. Dynam.* **1**, 209–219 (2007) Zbl 1141.20024 MR 2314044
- [2] Baumslag, G., Solitar, D.: Some two-generator one-relator non-Hopfian groups. *Bull. Amer. Math. Soc.* **68**, 199–201 (1962) Zbl 0108.02702 MR 0142635
- [3] Billingsley, P.: *Convergence of Probability Measures*. Wiley, New York (1968) Zbl 0172.21201 MR 0233396
- [4] Borisov, A., Sapir, M.: Polynomial maps over finite fields and residual finiteness of mapping tori of group endomorphisms. *Invent. Math.* **160**, 341–356 (2005) Zbl 1083.14023 MR 2138070
- [5] Borisov, A., Sapir, M.: Polynomial maps over  $p$ -adics and residual properties of mapping tori of group endomorphisms. *Int. Math. Res. Notices* **2009**, no. 16, 3002–3015 Zbl 1183.20031 MR 2533795

- 
- [6] Brown, K. S.: Trees, valuations, and the Bieri–Neumann–Strebel invariant. *Invent. Math.* **90**, 479–504 (1987) Zbl 0663.20033 MR 0914847
  - [7] Burdzy, K.: Brownian excursions from hyperplanes and smooth surfaces. *Trans. Amer. Math. Soc.* **295**, 35–57 (1986) Zbl 0601.60080 MR 0831187
  - [8] Cranston, M., Hsu, P., March, P.: Smoothness of the convex hull of planar Brownian motion. *Ann. Probab.* **17**, 144–150 (1989) Zbl 0678.60073 MR 0972777
  - [9] Dunfield, N. M., Thurston, D. P.: A random tunnel number one 3-manifold does not fiber over the circle. *Geom. Topol.* **10**, 2431–2499 (2006) Zbl 1139.57018 MR 2284062
  - [10] Feighn, M., Handel, M.: Mapping tori of free group automorphisms are coherent. *Ann. of Math. (2)* **149**, 1061–1077 (1999) Zbl 0938.20022 MR 1709311
  - [11] Fu, G. Y.: Random walks and random polynomials. PhD thesis, Univ. of Toronto (2007)
  - [12] Gromov, M.: Hyperbolic groups. In: *Essays in Group Theory*, S. M. Gersten (ed.), M.S.R.I. Publ. 8, Springer, Berlin, 75–263 (1987) Zbl 0634.20015 MR 0919829
  - [13] Kapovich, I., Schupp, P., Shpilrain, V.: Generic properties of Whitehead’s algorithm and isomorphism rigidity of random one-relator groups. *Pacific J. Math.* **223**, 113–140 (2006) Zbl 1149.20028 MR 2221020
  - [14] Ol’shanskii, A. Yu.: SQ-universality of hyperbolic groups. *Mat. Sb.* **186**, 119–132 (1995) (in Russian) Zbl 0864.20023 MR 1357360
  - [15] Rivin, I.: Growth in free groups (and other stories). arXiv:math/9911076v2 (1999)
  - [16] Wise, D. T.: The residual finiteness of positive one-relator groups. *Comment. Math. Helv.* **76**, 314–338 (2001) Zbl 1023.20011 MR 1839349