Cogrowth and spectral gap of generic groups

Yann Ollivier

29th November 2004

Abstract

We prove that that for all ε, having cogrowth exponent at most $1/2+\varepsilon$ (in base $2m-1$ with m the number of generators) is a generic property of groups in the density model of random groups. This generalizes a theorem of Grigorchuk and Champetier. More generally we show that the cogrowth of a random quotient of a torsion-free hyperbolic group stays close to that of this group.

This proves in particular that the spectral gap of a generic group is as large as it can be.

Cogrowth of generic groups. The spectral gap of an infinite group (with respect to a given set of generators) is a quantity controlling the speed of convergence of the simple random walk on the group (see [K]); up to parity problems it is equal to the first eigenvalue of the discrete Laplacian. By a formula of Grigorchuk (Theorem 4.1 of [Gri], see also section 1.1 below) this quantity can also be expressed combinatorially by a quantity called *cogrowth*: the smaller the cogrowth, the larger the spectral gap (see also [C]). So this is an important quantity from the combinatorial, probabilistic and operator-algebraic point of view (see [GdlH] or [W] and the references therein for an overview).

In [Gri] (Theorem 7.1) and [Ch93], Grigorchuk and Champetier show that groups defined by a presentation satisfying the small cancellation condition, or a weaker assumption in the case of Champetier, with long enough relators (depending on the number of relators in in the presentation), has a cogrowth exponent arbitrarily close to $1/2$ (the smallest possible value), hence a spectral gap almost as large as that of the free group with same number of generators.

We get the same conclusion for *generic* groups in a precise probabilistic meaning: that of the density model of random groups introduced in [Gro93], which we briefly recall in section 1.2. (Note that in the density model of random groups, if $d>0$ the number of relators is exponentially large and so Grigorchuk's and Champetier's results do not apply). Recall from [Gro93] that above density $d_{\text{crit}}=1/2$, random groups are very probably trivial.

Theorem 1 – *Let $0\leqslant d<1/2$ be a density parameter and let G be a random group on $m\geqslant 2$ generators at density d and length ℓ.*

Then, for any $\varepsilon > 0$, the probability that the cogrowth exponent of G lies in the interval $[1/2; 1/2+\varepsilon]$ tends to 1 as $\ell \to \infty$.

In particular, this provides a new large class of groups having a large spectral gap.

This theorem cannot be interpreted by saying that as the relators are very long, the geometry of the group is trivial up to scale ℓ. Indeed, cogrowth is an asymptotic invariant and thus takes into account the very non-trivial geometry of random groups at scale ℓ (see paragraph "locality of cogrowth" below). This is crudely examplified by the collapse of the group when density is too large.

Our primary motivation is the study of generic properties of groups. The study of random groups emerged from an affirmation of Gromov in [Gro87] that "almost every group is hyperbolic". Since the pioneer work of Champetier ([Ch95]) and Ol'shanskiĭ ([Ols]) it has been flourishing, now having connections with lots of topics in group theory such as property T, the Baum-Connes conjecture, small cancellation, the isomorphism problem, the Haagerup property, planarity of Cayley graphs...

The density model of random groups (which we recall in section 1.2), introduced in [Gro93], is very rich in allowing a precise control of the number of relators to be put in the group (and it actually allows this number to be very large). It has proven to be very fruitful, as random groups at different densities can have different properties (e.g. property T). See [Gh] and [Oll] for a general discussion of random groups and the density model, and [Gro93] for an enlightening presentation of the initial intuition behind this model.

Cogrowth of random quotients. A generic group is simply a random quotient of a free group[1]. More generally, we show that, when taking a random quotient of a torsion-free hyperbolic group, the cogrowth of the resulting group is very close to that of the initial group. Recall from [Oll] that a random quotient of a torsion-free hyperbolic group is very probably trivial above some critical density d_{crit}, which precisely depends on the cogrowth of the group (see Theorem 7 in section 1.2 below).

Theorem 2 – *Let G_0 be a non-elementary, torsion-free hyperbolic group generated by the elements $a_1^{\pm 1}, \dots, a_m^{\pm 1}$. Let η be the cogrowth exponent of G_0 with respect to this generating set.*

Let $0 \leqslant d < d_{\text{crit}}$ be a density parameter and let G be a random quotient (either by plain or reduced random words) of G_0 at density d and length ℓ.

Then, for any $\varepsilon > 0$, the probability that the cogrowth exponent of G lies in the interval $[\eta; \eta+\varepsilon]$ tends to 1 when $\ell \to \infty$.

[1] There is a very interesting and intriguing parallel approach to generic groups, developed by Champetier in [Ch00], which consists in considering the topological space of all group presentations with a given number of generators. See [P] for a description of connections of this approach with other problems in group theory.

Of course Theorem 1 is a particular case of Theorem 2. Also, since the cogrowth and gross cogrowth exponent can be computed from each other by the Grigorchuk formula (see section 1.1), this implies that the gross cogrowth exponent does not change either.

This answers a very natural question arising from [Oll]: indeed, it is known that for each torsion-free hyperbolic group, the critical density d_{crit}, below which random quotients are infinite and above which they are trivial, is equal to 1 minus the cogrowth exponent (resp. 1 minus the gross cogrowth exponent) for a quotient by random reduced words (resp. random plain words). So wondering what happens to the cogrowth exponent after a random quotient is very natural.

Knowing that cogrowth does not change much allows in particular to iterate the operation of taking a random quotient. These iterated quotients are the main ingredient in the construction by Gromov ([Gro03]) of a counter-example to the Baum-Connes conjecture with coefficients (see also [HLS]). Without the stability of cogrowth, in order to get the crucial cogrowth control necessary to build these iterated quotients Gromov had to use a very indirect and non-trivial way involving property T (which allows uniform control of cogrowth over all infinite quotients of a group); this could be avoided with our argument. So besides their interest as generic properties of groups, the results presented here could be helpful in the field.

Remark 3 – Theorem 2 only uses the two following facts: that the random quotient axioms of [Oll] are satisfied, and that there is a local-to-global principle for cogrowth in the random quotient. So in particular the result holds under slightly weaker conditions than torsion-freeness of G_0, as described in [Oll] ("harmless torsion").

Locality of cogrowth in hyperbolic groups. As one of our tools we use a result about locality of cogrowth in hyperbolic groups. Cogrowth is an asymptotic invariant, and large relations in a group can change it noticeably. But in hyperbolic groups, if the hyperbolicity constant is known, it is only necessary to evaluate cogrowth in some ball in the group to get a bound for cogrowth of the group (see Proposition 8). So in this case cogrowth is accessible to computation.

In the case of random quotients by relators of length ℓ, this principle shows that it is necessary to check cogrowth up to words of length at most $A\ell$ for some constant A (which depends on density and actually tends to infinity when d is close to the critical density), so that geometry of the quotient matters up to scale ℓ but not at higher scales.

This result may have independent interest.

About the proofs. The proofs make heavy use of the techniques developed in [Ch93] and [Oll]. We hope to have included precise enough reminders.

As often in hyperbolic group theory, the general case is very involved but lots of ideas are already present in the case of the free group. So in order to help understand the structure of the argument, we first present a proof in the case of the free group (Theorem 1), and then the proof of Theorem 2 for any torsion-free hyperbolic group.

Also, the proofs for random quotients by reduced and plain random words are very similar. They can be treated at once using the general but heavy terminology of [Oll]. We rather chose to present the proof of Theorem 1 in the case of reduced words (for which it seems to be more natural) and of Theorem 2 in the case of plain words.

Acknowledgments. I would like to thank Étienne Ghys and Pierre Pansu for very helpful discussions and many comments on the text. Pierre Pansu especially insisted that I should go on with this question at a time when I had no ideas about it. Lots of the ideas presented here emerged during my stay at the École normale supérieure de Lyon in Spring 2003, at the invitation of Damien Gaboriau and Étienne Ghys. I am very grateful to all the team of the math department there for their warmth at receiving me.

1 Definitions and notations

1.1 Cogrowth, gross cogrowth, spectral gap

These are variants around the same ideas. The spectral radius of the random walk operator on a group was studied by Kesten in [K], and cogrowth was defined later, simultaneously by Grigorchuk ([Gri]) and Cohen ([C]). See [GdlH] for an overview of results and open problems about these quantities and other, related ones.

So let G be an infinite group generated by the elements $a_1^{\pm}, \ldots, a_m^{\pm 1}$. Let W_ℓ be the set of words w of length ℓ in the letters $a_1^{\pm}, \ldots, a_m^{\pm 1}$ such that w is equal to e in the group G. Let $W'_\ell \subset W_\ell$ be the set of *reduced* words in W_ℓ. (Note that W'_ℓ is empty if G is freely generated by $a_1, \ldots, a_m$.) Denote the cardinal of a set by $|.|$.

Definition 4 (Cogrowth exponent) – *The* cogrowth exponent *of G with respect to $a_1, \ldots, a_m$ is defined as*

$$\eta = \lim_{\substack{\ell\to\infty \\ \ell \text{ even}}} \frac{1}{\ell} \log_{2m-1} |W'_\ell|$$

or $\eta = 1/2$ if G is freely generated by $a_1, \ldots, a_m$.

The gross cogrowth exponent *of G with respect to $a_1, \ldots, a_m$ is defined as*

$$\theta = \lim_{\substack{\ell\to\infty \\ \ell \text{ even}}} \frac{1}{\ell} \log_{2m} |W_\ell|$$

So the cogrowth exponent is the logarithm in base $2m - 1$ of the cogrowth as defined by Grigochuk and Cohen. The exponents η and θ always lie in the interval $[1/2;1]$, with equality only in case of η of a free group. Amenability of G is equivalent to $\eta = 1$ and to $\theta = 1$.

It is shown in the references mentioned above that the limit exists. We have to take ℓ even in case there are no relations of odd length in the group (in which case W_ℓ is empty).

The convention for the free group is justified by the following Grigorchuk formula ([Gri], Theorem 4.1):

$$(2m)^\theta = (2m-1)^\eta + (2m-1)^{1-\eta}$$

which allows to compute one exponent knowing the other (also using that these are at least $1/2$), and shows that η and θ vary the same way. Given that θ is well-defined for a free group, the formula yields $\eta(F_m) = 1/2$. As this is also the convention which makes all our statements valid without isolating the case of a free group, we strongly plead for this being the right convention.

The cogrowth exponent is also the exponent of growth of the kernel of the natural map from the free group F_m to G sending a_i to a_i.

The probability of return to e in time t of the simple random walk on G (with respect to the generators $a_1^{\pm1}, \ldots, a_m^{\pm1}$) is of course equal to $|W_t|/(2m)^t$. So $(2m)^{\theta-1}$ is also the spectral radius of the random walk operator on $L^2(G)$ defined by $Mf(x) = \frac{1}{2m}\sum f(xa_i^{\pm1})$. This is the form studied by Kesten ([K]), who denotes by λ this spectral radius.

Since the discrete Laplacian on G is equal to the operator $\mathrm{Id} - M$, $1-(2m)^{\theta-1}$ is also equal to $\min(\lambda_1, 2-\lambda_\infty)$ where λ_1 is the smallest and λ_∞ the largest eigenvalue of the Laplacian acting on $L^2(G)$. (The problems of λ_∞ and of parity of ℓ in the definition can be avoided by considering lazy random walks.) In particular, if θ (or η) is small then the spectral gap λ_1 is large.

The cardinals of the sets W_ℓ of course satisfy the superadditivity property $|W_{\ell+\ell'}| \geqslant |W_\ell|\,|W_{\ell'}|$. This implies that for any ℓ we have an exact (instead of asymptotic) bound $|W_\ell| \leqslant (2m)^{\theta\ell}$. For cogrowth this is not exactly but almost true, due to reduction problems, and we have $\left|W'_{\ell+\ell'+2}\right| \geqslant |W'_\ell|\,|W'_{\ell'}|$ and the exact inequality $|W'_\ell| \leqslant (2m-1)^{\eta\ell+2}$. We will often implicitly use these inequalities in the sequel.

1.2 The density model of random groups

A random group is a quotient of a free group $F_m = \langle a_1, \ldots, a_m \rangle$ by (the normal closure of) a randomly chosen set $R \subset F_m$. Typically R is viewed as a set of words in the letters $a_i^{\pm1}$. So defining a random group is giving a law for R.

More generally, given a group G_0 generated by the elements $a_1^{\pm1}, \ldots, a_m^{\pm1}$, and given a set R of random words in these generators we define a random quotient of G_0 by $G = G_0/\langle R \rangle$.

The density model which we now define allows a precise control on the size of R: the bigger the size of R, the smaller the random group. For comparison, remember the number of words of length ℓ in $a_1^{\pm1}, \ldots, a_m^{\pm1}$ is $(2m)^\ell$, and the number of reduced words is $(2m)(2m-1)^{\ell-1} \approx (2m-1)^\ell$.

In the whole text we suppose $m \geqslant 2$.

Definition 5 (Density model of random groups or quotients) – *Let G_0 be a group generated by the elements $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$. Let $0 \leqslant d \leqslant 1$ be a density parameter.*

Let R be a set of $(2m)^{d\ell}$ randomly chosen words of length ℓ (resp. a set of $(2m-1)^{d\ell}$ randomly chosen reduced words of length ℓ), uniformly and independently picked among all those words.

We call the group $G = G_0/\langle R\rangle$ a random quotient *of G_0 by plain random words (resp. by reduced random words), at density d, at length ℓ.*

In case G_0 is the free group F_m and reduced words are taken, we simply call G a random group.

In this definition, we can also replace "words of length ℓ " by "words of length between ℓ and $\ell+C$ " for any constant C; the theorems presented thereafter remain valid. In [Oll], section 4, we describe generalizations of these models.

The interest of the density model was established by the following theorem of Gromov, which shows a sharp phase transition between infinity and triviality of random groups.

Theorem 6 (M. Gromov, [Gro93]) – *Let $d < 1/2$. Then with probability tending to 1 as ℓ tends to infinity, random groups at density d are infinite hyperbolic.*

Let $d > 1/2$. Then with probability tending to 1 as ℓ tends to infinity, random groups at density d are either $\{e\}$ or $\mathbb{Z}/2\mathbb{Z}$.

(The occurrence of $\mathbb{Z}/2\mathbb{Z}$ is of course due to the case when ℓ is even; this disappears if one takes words of length between ℓ and $\ell + C$ with $C \geqslant 1$.)

Basically, $d\ell$ is to be interpreted as the "dimension" of the random set R (see the discussion in [Gro93]). As an illustration, if $L < 2d\ell$ then very probably there will be two relators in R sharing a common subword of length L. Indeed, the dimension of the couples of relators in R is $2d\ell$, whereas sharing a common subword of length L amounts to L "equations", so the dimension of those couples sharing a subword is $2d\ell - L$, which is positive if $L < 2d\ell$. This "shows" in particular that at density d, the small cancellation condition $C'(2d)$ is satisfied.

Since a random quotient of a free group is hyperbolic, one can wonder if a random quotient of a hyperbolic group is still hyperbolic. The answer is basically yes, and the critical density in this case is linked to the cogrowth exponent of the initial group.

Theorem 7 (Y. Ollivier, [Oll]) – *Let G_0 be a non-elementary, torsion-free hyperbolic group, generated by the elements $a_1^{\pm 1}, \ldots, a_m^{\pm 1}$, with cogrowth exponent η and gross cogrowth exponent θ.*

Let $0 \leqslant d \leqslant 1$ be a density parameter, and set $d_{crit} = 1 - \theta$ (resp. $d_{crit} = 1 - \eta$).

If $d < d_{crit}$, then a random quotient of G_0 by plain (resp. reduced) random words is infinite hyperbolic, with probability tending to 1 as ℓ tends to infinity.

If $d > d_{crit}$, then a random quotient of G_0 by plain (resp. reduced) random words is either $\{e\}$ or $\mathbb{Z}/2\mathbb{Z}$, with probability tending to 1 as ℓ tends to infinity.

This is the context in which Theorem 2 is to be understood.

1.3 Hyperbolic groups and isoperimetry of van Kampen diagrams

Let G be a group given by the finite presentation $\langle a_1, \dots, a_m \mid R \rangle$. Let w be a word in the $a_i^{\pm 1}$'s. We denote by $|w|$ the number of letters of w, and by $\|w\|$ the distance from e to w in the Cayley graph of the presentation, that is, the minimal length of a word representing the same element of G as w.

Let λ be the maximal length of a relation in R.

We refer to [LS] for the definition and basic properties of van Kampen diagrams. Remember that a word represents the neutral element of G if and only if it is the boundary word of some van Kampen diagram. If D is a van Kampen diagram, we denote its number of faces by $|D|$ and its boundary length by $|\partial D|$.

It is known ([Sh]) that G is hyperbolic if and only if there exists a constant $C_1 > 0$ such that for any (reduced) word w representing the neutral element of G, there exists a van Kampen diagram with boundary word w, and with at most $|w|/C_1$ faces. This can be reformulated as: for any word w representing the neutral element of G, there exists a van Kampen diagram with boundary word w satisfying the isoperimetric inequality

$$|\partial D| \geqslant C_1 |D|$$

We are going to use a *homogeneous* way to write this inequality. The above form compares the boundary length of a van Kampen diagram to its number of faces. This amounts to comparing a length with a number, which is not very well-suited for geometric arguments, especially when dealing with groups having relations of very different lengths.

So let D be a van Kampen diagram w.r.t. the presentation and define the *area* of D to be

$$\mathcal{A}(D) = \sum_{f \text{ face of } D} |\partial f|$$

which is also the number of external edges (not couting "filaments") plus twice the number of internal ones. This has, heuristically speaking, the homogeneity of a length.

It is immediate to see that if D satisfies $|\partial D| \geqslant C_1 |D|$, then we have $|\partial D| \geqslant C_1 \mathcal{A}(D)/\lambda$ (recall λ is the maximal length of a relation in the presentation). Conversely, if $|\partial D| \geqslant C_2 \mathcal{A}(D)$, then $|\partial D| \geqslant C_2 |D|$. So we can express the isoperimetric inequality using $\mathcal{A}(D)$ instead of $|D|$.

Say a diagram is *minimal* if it has minimal area for a given boundary word. So G is hyperbolic if and only if there exists a constant $C > 0$ such that every minimal van Kampen diagram satisfies the isoperimetric inequality

$$|\partial D| \geqslant C\,\mathcal{A}(D)$$

This formulation is homogeneous, that is, it compares a length to a length. This inequality is the one that naturally arises in $C'(\alpha)$ small cancellation theory (with $C = 1 - 6\alpha$) as well as in random groups at density d (with $C = \frac{1}{2} - d$). So in these contexts the value of C is naturally linked with some parameters of the presentation.

This kind of isoperimetric inequality is also the one appearing in the assumptions of Champetier in [Ch93], in random quotients of hyperbolic groups (cf. [Oll]) and in the (infinitely presented) limit groups constructed by Gromov in [Gro03]. So we think this is the right way to write the isoperimetric inequality when the lengths of the relators are very different.

2 Locality of cogrowth in hyperbolic groups

The goal of this section is to show that in a hyperbolic group, in order to estimate cogrowth (which is an asymptotic invariant), it is enough to check only words of bounded length, where the bound depends on the quality of the isoperimetric inequality in the group.

Everything here is valid, *mutatis mutandis*, for cogrowth and gross cogrowth.

Here $G = \langle\, a_1, \dots, a_m \mid R \,\rangle$ ($m \geqslant 2$) is a hyperbolic group and W_ℓ is the set of reduced words of length ℓ in the $a_i^{\pm 1}$ equal to e in G. Let also λ be the maximal length of a relation in R.

As explained above, hyperbolicity of G amounts to the existence of some constant $C > 0$ such that any minimal van Kampen diagram D over this presentation satisfies the isoperimetric inequality

$$|\partial D| \geqslant C\mathcal{A}(D)$$

We will prove the following.

Proposition 8 – *Suppose that, for some $A > 1$, for any $A\lambda/4 \leqslant \ell \leqslant A\lambda$ one has*

$$|W_\ell| \leqslant (2m-1)^{\eta\ell}$$

for some $\eta \geqslant 1/2$.

Then for any $\ell \geqslant A\lambda/4$,

$$|W_\ell| \leqslant (2m-1)^{\eta\ell(1+o(1)_{A\to\infty})}$$

where the constant implied in $o(1)$ depends only on C.

It follows from the proof that actually $1 + o(1) \leqslant \exp \frac{200}{C\sqrt{A}}$, so that is it enough to take $A \approx 40000/C^2$ for a good result.

Proof –

First we need some simple lemmas.

The *distance to boundary* of a face of a van Kampen diagram is the minimal length of a sequence of faces adjacent by an edge, beginning with the given face and ending with a face adjacent to the boundary (so that a boundary face is at distance 1 from the boundary).

Set $\alpha = 1/\log(1/(1-C)) \leqslant 1/C$, where we can suppose $C \leqslant 1$.

Lemma 9 – *Let D be a minimal van Kampen diagram. Then D can be written as a disjoint union $D = D_1 \cup D_2$ (with maybe D_2 not connected) such that each face of D_1 is at distance at most $\alpha \log(\mathcal{A}(D)/\lambda)$ from the boundary of D, and D_2 has area at most λ.*

Proof – Since D is minimal it satisfies the isoperimetric inequality $|\partial D| \geqslant C\mathcal{A}(D)$. Thus, the cumulated area of the faces of D which are adjacent to the boundary is at least $C\mathcal{A}(D)$, and so the cumulated area of the faces at distance at least 2 from the boundary is at most $(1-C)\mathcal{A}(D)$.

Applying the same reasoning to the (maybe not connected) diagram obtained from D by removing the boundary faces, we get by induction that the cumulated area of the faces of D lying at distance at least k from the boundary is at most $(1-C)^{k-1}\mathcal{A}(D)$. Taking $k = 1+\alpha\log(\mathcal{A}(D)/\lambda)$ (rounded up to the nearest integer) provides the desired decomposition. □

In the sequel we will neglect divisibility problems (such as the length of a diagram being a multiple of 4).

Lemma 10 – *Let D be a minimal van Kampen diagram. D can be partitioned into two diagrams D', D'' by cutting it along a path of length at most $\lambda + 2\alpha\lambda\log(\mathcal{A}(D)/\lambda)$ such that each of D' and D'' contains at least one quarter of the boundary of D.*

(Here a *path* in a diagram is meant to be a path in its 1-skeleton.)

Proof – Consider the decomposition $D = D_1 \cup D_2$ of the previous lemma, and first suppose that D_2 is empty, so that any face of D_1 lies at distance at most $\alpha\lambda\log(\mathcal{A}(D)/\lambda)$ from the boundary.

Let L be the boundary length of D and mark four points A, B, C, D on ∂D at distance $L/4$ of each other. As D is $\alpha\log(\mathcal{A}(D)/\lambda)$-narrow, there exists a path of length at most $2\alpha\lambda\log(\mathcal{A}(D)/\lambda)$ joining either a point of AB to a point of CD or a point of AD to a point of BC, which provides the desired cutting.

Now if D_2 was not empty, first retract each connected component of D_2 to a point: the reasoning above shows that there exists a path of length at most $2\alpha\lambda\log(\mathcal{A}(D)/\lambda)$ joining either a point of AB to a point of CD or a point of AD to a point of BC, *not counting the length in D_2*. But since the sum of the lengths of the faces of D_2 is at most λ, the cumulated length of the travel in D_2 is at most λ, hence the lemma. □

The cardinal of the W_ℓ's (almost in the case of cogrowth, see above) satisfy the supermultiplicativity property $|W_\ell| \geqslant |W_{\ell-L}|\,|W_L|$. Using narrowness of diagrams we are able to show a converse inequality, which will enable us to control cogrowth.

Corollary 11 – *We have, up to parity problems,*

$$\begin{aligned} |W_\ell| &\leqslant \sum_{\ell/4\leqslant \ell'\leqslant 3\ell/4} \left|W_{\ell'+2\alpha\lambda\log(\ell/C\lambda)+\lambda}\right| \left|W_{\ell-\ell'+2\alpha\lambda\log(\ell/C\lambda)+\lambda}\right| \\ &\leqslant \frac{\ell}{\lambda}\max_{\ell/4\leqslant \ell'\leqslant 3\ell/4} \left|W_{\ell'+2\alpha\lambda\log(\ell/C\lambda)+3\lambda}\right| \left|W_{\ell-\ell'+2\alpha\lambda\log(\ell/C\lambda)+3\lambda}\right| \end{aligned}$$

Proof – Any word in W_ℓ is the boundary word of some (minimal) van Kampen diagram D with boundary length ℓ, and so the first inequality follows from the previous lemma, together with the inequality $\mathcal{A}(D)\leqslant |\partial D|/C$.

The last inequality uses the fact that, up to moving the cutting points by at most λ, we can assume that the lengths involved are multiples of λ, hence the factor ℓ/λ in front of the max and the increase of the lengths by 2λ. □

Now for the proof of Proposition 8 proper.

First, choose ℓ between $A\lambda$ and $4A\lambda/3$. By Corollary 11 and the assumptions, we have

$$|W_\ell|\leqslant (2m-1)^{\eta(\ell+4\alpha\lambda\log(\ell/C\lambda)+6\lambda)+\log_{2m-1}(\ell/\lambda)}$$

Let B be a number (depending on C) such that

$$4\alpha\log(B/C)+6+\frac{1}{\eta}\log_{2m-1}B\leqslant B$$

(noting that $m\geqslant 2$, $\eta\geqslant 1/2$ and $\alpha\leqslant 1/C$ one can check that $B\geqslant 144/C^2$ is enough). It is then easy to check that for $B'\geqslant B$ one has

$$4\alpha\log(B'/C)+6+\frac{1}{\eta}\log_{2m-1}B'\leqslant 2\sqrt{B'B}$$

Thus, if $\ell\geqslant A\lambda$ and $A\geqslant B$ we have

$$|W_\ell|\leqslant (2m-1)^{\eta\left(\ell+2\lambda\sqrt{AB}\right)}\leqslant (2m-1)^{\eta\ell\left(1+2\sqrt{B/A}\right)}$$

We have just shown that if $|W_\ell|\leqslant (2m-1)^{\eta\ell}$ for $\ell\leqslant A\lambda$, then $|W_\ell|\leqslant (2m-1)^{\eta\ell\left(1+2\sqrt{B/A}\right)}$ for $\ell\leqslant (4A/3)\lambda$. Thus, iterating the process shows that for $\ell\leqslant (4/3)^k A\lambda$ we have

$$|W_\ell|\leqslant (2m-1)^{\eta\ell\prod_{0\leqslant i<k}\left(1+2\sqrt{\frac{B}{A}}\left(\frac{3}{4}\right)^{i/2}\right)}$$

and we are done as the product $\prod_i\left(1+2\sqrt{\frac{B}{A}}\left(\frac{3}{4}\right)^{i/2}\right)$ converges to some value tending to 1 when $A\to\infty$; if one cares, its value is less than $\exp\frac{200}{C\sqrt{A}}$. □

3 Application to random groups: the free case

Here we first treat the case when the initial group G_0 is the free group F_m on m generators. This will serve as a template for the more complex general case.

So let $G = \langle a_1, \ldots, a_m \mid R \rangle$ be a random group at density d, with R a set of $(2m-1)^{d\ell}$ random reduced words.

We have to evaluate the number of reduced words of a given length L which represent the trivial element in G. Any such word is the boundary word of some van Kampen diagram D with respect to the set of relators R. We will proceed as follows: for any diagram D involving n relators, we will evaluate the expected number of n-tuples of random relators from R that make it a van Kampen diagram. We will show that this expected number is controlled by the boundary length L of the diagram, and this will finally allow to control the number of van Kampen diagrams of boundary length L.

We call a van Kampen diagram *non-filamenteous* if each of its edges lies on the boundary on some face. Each diagram can be decomposed into non-filamenteous components linked by filaments. For the filamenteous part we will use the estimation from [Ch93], one step of which counts the number of ways in which the different non-filamenteous parts can be glued together to form a van Kampen diagram.

So we first focus on non-filamenteous diagrams, for which a genuinely new argument has to be given compared to [Ch93] (since the number of relators here is unbounded).

We first suppose that we care only about diagrams with at most K faces, for some K to be chosen later. (We will of course use the locality of cogrowth principle to remove this assumption.)

3.1 Fulfilling of diagrams

So let D be a non-filamenteous van Kampen diagram. Let $|D|$ be its number of faces and let $n \leqslant |D|$ be the number of different relators it involves. Let m_i, $1 \leqslant i \leqslant n$ be the number of times the i-th relator appears in D, where we choose to enumerate the relators in decreasing order of multiplicity, that is, $m_1 \geqslant m_2 \geqslant \ldots \geqslant m_n$. Let also D_i be the subdiagram of D made of relators $1, 2, \ldots, i$ only, so that $D = D_n$.

It is shown in [Oll] (section 2.2) that to this diagram we can associate numbers $d_1, \ldots, d_n$ such that

- The probability that i given random relators fulfill D_i is less than $(2m-1)^{d_i - id\ell}$; consequently, the probability that there exists an i-tuple of relators in R fulfilling D_i is less than $(2m-1)^{d_i}$.
- The following isoperimetric inequality holds :

$$|\partial D| \geqslant (1-2d)\ell\, |D| + 2 \sum d_i (m_i - m_{i+1})$$

So for a given n-tuple of random relators, the probability that this n-tuple fulfills D is at most $(2m-1)^{\inf(d_i - id\ell)}$. So, as there are $(2m-1)^{nd\ell}$ n-tuples of relators in R, the expected number S of n-tuples fulfilling D in R is at most

$$S \leqslant (2m-1)^{nd\ell + \inf(d_i - id\ell)}$$

which so turns out to be not only an upper bound for the probability of D to be fulfillable but rather an estimate of the number of ways in which it is. (The probabilities that two n-tuples fulfill the diagram are independent only when the n-tuples are disjoint, but expectation is linear anyway.)

Set $d'_i = d_i - id\ell$. Then, rewriting the isoperimetric inequality above and using that $m_i - m_{i+1} \geqslant 0$ yields

$$\begin{aligned}
|\partial D| &\geqslant (1-2d)\ell\,|D| + 2\sum (d'_i + id\ell)(m_i - m_{i+1}) \\
&= (1-2d)\ell\,|D| + 2d\ell \sum m_i + 2\sum d'_i(m_i - m_{i+1}) \\
&= \ell\,|D| + 2(\inf d'_i)\sum (m_i - m_{i+1}) + 2\sum (d'_i - \inf d'_i)(m_i - m_{i+1}) \\
&\geqslant \ell\,|D| + 2m_1 \inf d'_i \\
&\geqslant \ell\,|D| + 2\inf d'_i
\end{aligned}$$

and consequently

$$\mathbb{E}S \leqslant (2m-1)^{nd\ell + \inf d'_i} \leqslant (2m-1)^{|D|d\ell + \frac{1}{2}(|\partial D| - |D|\ell)} = (2m-1)^{\frac{1}{2}(|\partial D| - (1-2d)\ell|D|)}$$

Of course this also holds for filamenteous diagrams because the faces are the same but $|\partial D|$ is even greater. So the conclusion is:

Proposition 12 – *For any reduced van Kampen diagram D, the expected number of ways it can be fulfilled by random relators at density d is at most $(2m-1)^{\frac{1}{2}(|\partial D| - (1-2d)\ell|D|)}$.*

By Markov's inequality, the probability to pick a random presentation R for which $S \geqslant (2m-1)^{\varepsilon\ell}\mathbb{E}S$ is less than $(2m-1)^{-\varepsilon\ell}$. Since the number of diagrams with less than K faces grows subexponentially in ℓ, we have shown:

Proposition 13 – *For any fixed integer K and any $\varepsilon > 0$, with probability exponentially close to 1 as $\ell \to \infty$, for each (non-filamenteous) van Kampen diagram with at most K faces, the number of ways to fulfill it with relators of R is at most $(2m-1)^{\frac{1}{2}(|\partial D| - (1-2d-\varepsilon)\ell|D|)}$.*

In particular, taking $\varepsilon < (\frac{1}{2} - d)/2$, this is less than $(2m-1)^{|\partial D|/2}$.

3.2 Evaluation of the cogrowth

We now conclude using the general scheme of [Ch93], together with Proposition 8 which allows to check only a finite number of diagrams.

Consider a reduced word w in the generators $a_i^{\pm 1}$, representing e in the random group. This word is the boundary word of some van Kampen diagram D which may have filaments.

Choose $\varepsilon > 0$. We are going to show that with probability exponentially close to 1 when $\ell \to \infty$, the number of such words w is at most $(2m-1)^{(1/2+\varepsilon)|w|}$.

We know from [Oll] (Section 2.2) that up to exponentially small probability in ℓ, we can suppose that any diagram satisfies the inequality

$$|\partial D| \geqslant C\ell\,|D|$$

where C depends only on the density d (basically $C = 1/2 - d$ divided by the constants appearing in the Cartan-Hadamard-Gromov theorem, see [Oll]) and not on ℓ.

Now we use Proposition 8. We are facing a group G in which all relations are of length ℓ. Consider a constant A given by Proposition 8 such that if we know that $|W_L| \leqslant (2m-1)^{L(1/2+\varepsilon/2)}$ for $L \leqslant A\ell$, then we know that $|W_L| \leqslant (2m-1)^{L(1/2+\varepsilon)}$ for any L. Such an A depends only on the isoperimetry constant C.

So we suppose that our word w has length at most $A\ell$. We have $|w| = |\partial D| \geqslant C\ell\,|D|$ and in particular, $|D| \leqslant A/C$, which is to say, we have to consider only diagrams with a number of faces bounded independently of ℓ.

So set $K = A/C$, which most importantly does not depend on ℓ. After Proposition 13, we can assume (up to exponentially small probability) that for any non-filamenteous diagram D' with at most K faces, the number of ways to fulfill it with relators of the random presentation is at most $(2m-1)^{|\partial D'|/2}$.

Back to our word w read on the boundary of some diagram D. Decompose D into filaments and connected non-filamenteous parts D_i. The word w is determined by the following data: a set of relators from the random presentation R fulfilling the D_i's, a set of reduced words to put on the filaments, the combinatorial choice of the diagrams D_i, and the combinatorial choice of how to connect the D_i's using the filaments.

The combinatorial part is precisely the one analyzed in [Ch93]. It is shown there (section "Premier pas") that if each D_i satisfies $|\partial D_i| \geqslant L$, the combinatorial factor controlling the connecting of the D_i's by the filaments and the sharing of the length $|\partial D|$ between the filaments and the D_i's is less than

$$\frac{|w|}{L}\,|w|\,(eL)^{2|w|/L}(2eL)^{|w|/L}(3eL)^{2|w|/L}$$

Observe that for L large enough this behaves like $(2m-1)^{|w|O(\log L/L)}$.

Here each diagram D_i satisfies $|\partial D_i| \geqslant C\ell\,|D_i| \geqslant C\ell$, so setting $L = C\ell$, each D_i has boundary length at least L. In particular, $O(\log L/L) = O(\log \ell/\ell)$.

The number of components D_i is obviously at most $|w|/L$. Each component has at most K faces since D itself has. So the number of choices for the combinatorial choices of the diagrams D_i's is at most $N(K)^{|w|/L}$ where $N(K)$ is the (finite!) number of planar graphs with at most K faces. This behaves like $(2m-1)^{|w|O(1/L)}$.

Now the number of ways to fill the D_i's with relators from the random presentation is, after Proposition 13, at most $\prod (2m-1)^{|\partial D_i|/2} = (2m-1)^{\sum|\partial D_i|/2}$.

The last choice to take into account is the choice of reduced words to put on the filaments. The total length of the filaments is $\frac{1}{2}(|w| - \sum |\partial D_i|)$ (each edge of a filament counts twice in the boundary), thus the number of ways to fill in the filaments is at most $(2m-1)^{\frac{1}{2}(|w|-\sum|\partial D_i|)}$.

So the total number of possibilities for w is

$$(2m-1)^{|w| O(\log \ell/\ell) + \frac{1}{2}(|w|-\sum|\partial D_i|)+\sum|\partial D_i|/2}$$

and if we take ℓ large enough, this will be at most $(2m-1)^{|w|(1/2+\varepsilon/2)}$, after what we conclude by Proposition 8.

This proves Theorem 1.

4 The non-free case

Now we deal with random quotients of a non-elementary torsion-free hyperbolic group G_0. We are going to give the proof in the case of a random quotient by plain random words, the case of a quotient by random reduced words being similar.

So let G_0 be a non-elementary torsion-free hyperbolic group given by the presentation $\langle a_1, \ldots, a_m \mid Q \rangle$ ($m \geqslant 2$), with the relations in Q having length at most λ. Let θ be the gross cogrowth of G_0 w.r.t. this generating set. Let $G = G_0/\langle R \rangle$ be a random quotient of G_0 by a set R of $(2m)^{d\ell}$ randomly chosen words of length ℓ. Also set $\beta = 1 - \theta$, so that the random quotient axioms of [Oll] (section 4) are satisfied.

We have to show that the number of boundary words of van Kampen diagrams of a given boundary length L grows slower than $(2m)^{L(\theta+\varepsilon)}$. This time, since we are going to give a proof in the case of gross cogrowth rather than cogrowth, we will not have many problems with filaments: the counting of filaments is already included in the knowledge of gross cogrowth of G_0.

For a van Kampen diagram D, let D'' be the subdiagram made of faces bearing "new" relators in R, and D' be the part made of faces bearing "old" relators in Q. By Proposition 32 of [Oll], we know that very probably G is hyperbolic and that its isoperimetric inequality takes the form

$$|\partial D| \geqslant \kappa\ell\,|D''| + \kappa'\,|D'|$$

whenever D is reduced and D' is minimal, with $\kappa, \kappa' > 0$ and where, most importantly, κ and κ' do not depend on ℓ. By definition of $\mathcal{A}(D)$, this can be rewritten as $|\partial D| \geqslant C\mathcal{A}(D)$ with $C = \min(\kappa, \kappa'/\lambda)$.

Fix some $\varepsilon > 0$ and let A be the constant provided by Proposition 8 applied to G, having the property that if we know that gross cogrowth is at most $\theta + \varepsilon/2$ up to words of length $A\ell$, then we know that gross cogrowth is at most $\theta + \varepsilon$. This A depends on ε, C and G_0 but not on ℓ. Thanks to this and the isoperimetric inequality, we only have to consider diagrams of boundary length at most $A\ell$ hence area at most $A\ell/C$. In particular the number of new relators $|D''|$ is at most A/C. So for all the sequel set

$$K = A/C$$

which, most importantly, does not depend on ℓ. This is the maximal size of diagrams we have to consider, thanks to the local-global principle.

4.1 Reminder from [Oll]

In this context, it is proven in [Oll] that the van Kampen diagram D can be seen as a "van Kampen diagram at scale ℓ with respect to the new relators, with equalities modulo G_0". More precisely, this can be stated as follows: (we refer to [Oll] for the definition of "strongly reduced" diagrams; the only thing to know here is that for any word equal to e in G, there exists a strongly reduced van Kampen diagram with this word as its boundary word).

Proposition 14 ([Oll], section 6.6) – *Let $G_0 = \langle\, S \mid Q \,\rangle$ be a non-elementary hyperbolic group, let R be a set of words of length ℓ, and consider the group $G = G_0/\langle R\rangle = \langle\, S \mid Q \cup R \,\rangle$.*

Let $K \geqslant 1$ be an arbitrarily large integer and let $\varepsilon_1, \varepsilon_2 > 0$ be arbitrarily small numbers. Take ℓ large enough depending on $G_0, K, \varepsilon_1, \varepsilon_2$.

Let D be a van Kampen diagram with respect to the presentation $\langle\, S \mid Q \cup R \,\rangle$, which is strongly reduced, of area at most $K\ell$. Let also D' be the subdiagram of D which is the union of the 1-skeleton of D and of those faces of D bearing relators in Q (so D' is a possibly non-simply connected van Kampen diagram with respect to G_0), and suppose that D' is minimal.

We will call worth-considering *such a van Kampen diagram.*

Let $w_1, \dots, w_p$ be the boundary (cyclic) words of D', so that each w_i is either the boundary word of D or a relator in R.

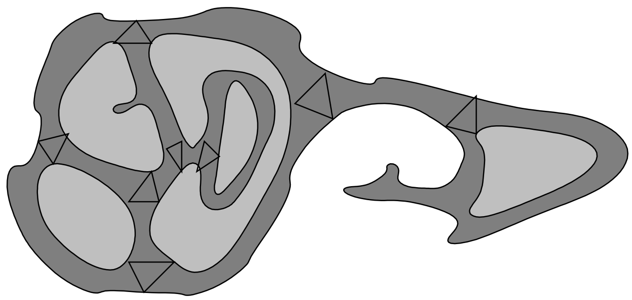
Then there exists an integer $k \leqslant 3K/\varepsilon_2$ and words $x_2, \dots, x_{2k+1}$ such that:

- *Each x_i is a subword of some cyclic word w_j;*
- *As subwords of the w_j's, the x_i's are disjoint and their union exhausts a proportion at least $1 - \varepsilon_1$ of the total length of the w_j's.*
- *For each $i \leqslant k$, there exists words δ_1, δ_2 of length at most $\varepsilon_2(|x_{2i}| + |x_{2i+1}|)$ such that $x_{2i}\delta_1 x_{2i+1}\delta_2 = e$ in G_0.*
- *If two words x_{2i}, x_{2i+1} are subwords of the boundary words of two faces of D bearing the same relator $r^{\pm 1} \in R$, then, as subwords of r, x_{2i} and x_{2i+1} are either disjoint or equal with opposite orientations (so that the above equality reads $x\delta_1 x^{-1}\delta_2 = e$).*

The couples (x_{2i}, x_{2i+1}) are called translators. *Translators are called* internal, internal-boundary *or* boundary-boundary *according to whether x_{2i} and x_{2i+1} is a subword of some w_j which is a relator in R or the boundary word of D.*

(There are slight differences between the presentation here and that in [Oll]. Therein, boundary-boundary translators did not have to be considered: they were eliminated earlier in the process, before section 6.6, because they have a positive contribution to boundary length, hence always improve isoperimetry and do not deserve consideration in order to prove hyperbolicity. Moreover, in [Oll] we further distinguished "commutation translators" for the kind of internal translator with $x_{2i} = x_{2i+1}^{-1}$, which we need not do here.)

Translators appear as dark strips on the following figure:



Remark 15 – The number of ways to partition the words w_i into translators is at most $(2K\ell)^{12K/\varepsilon_2}$, because each w_i can be determined by its starting- and endpoint, which can be given as numbers between 1 and $2K\ell$ which is an upper bound for the cumulated length of the w_i's (since the area of D is at most $K\ell$). For fixed K and ε_2 this grows subexponentially in ℓ.

Remark 16 – Knowing the words x_i, the number of possibilities for the boundary word of the diagram is at most $(6K/\varepsilon_2)!$ (choose which subwords x_i make the boundary word of the diagram, in which order), which does not depend on ℓ for fixed K and ε_2.

We need another notion from [Oll], namely, that of *apparent length* of an element in G_0. This basically answers the question: If this element were obtained through a random walk at time t, what would be a reasonable value of t? This accounts for the fact that, unlike in the free group, the hitting probability of an element in the group does not depend only on the norm of this element.

Apparent length is defined in [Oll] in a more general setting, with respect to a measure on the group, which is here the measure obtained after a simple random walk with respect to the given set of generators $a_1, \ldots, a_m$. We only give here what the definition amounts to in our context.

Definition 17 (Definition 36 of [Oll]) – *Let x be a word. Let $\varepsilon_2 > 0$. Let L be an integer. Let $p_L(xuyv = e)$ be the probability that, for a random word y of length L, there exists elements $u, v \in G_0$ of norm at most $\varepsilon_2(|x| + L)$ such that $xuyv = e$ in G_0.*

The apparent length *of x at test-length L is*

$$\mathbb{L}_L(x) = -\frac{1}{1-\theta} \log_{2m} p_L(xuyv = e) - L$$

The apparent length *of x is*

$$\mathbb{L}(x) = \min\left(\|x\| \frac{\theta}{1-\theta}, \min_{0\leqslant L\leqslant K\ell} \mathbb{L}_L(x) \right)$$

where we recall ℓ is the length of the relators in a random presentation.

(The first term $\|x\|\,\theta/(1-\theta)$ is an easy upper bound for $\mathbb{L}_{\|x\|}(x)$, and so if $\|x\| \leqslant K\ell$ then the first term in the min is useless.)

It is shown in [Oll], section 6.7, that in a randomly chosen presentation at density d and length ℓ, all subwords of the relators have apparent length at most 4ℓ, with probability exponentially close to 1 as $\ell \to \infty$. So from now on we suppose that this is indeed the case.

We further need the notion of a *decorated abstract van Kampen diagram* (which was implicitly present in the free case when we mentioned the probability that some diagram "is fulfilled by random relators"), which is inspired by Proposition 14: it carries the combinatorial information about how the relators and boundary word of a diagram were cut into subwords in order to make the translators.

Definition 18 (Decorated abstract van Kampen diagram) – *Let $K \geqslant 1$ be an arbitrarily large integer and let $\varepsilon_1, \varepsilon_2 > 0$ be arbitrarily small numbers. Let I_ℓ be the cyclically ordered set of ℓ elements.*

A decorated abstract van Kampen diagram *$\mathcal{D}$ is the following data:*

- *An integer $|\mathcal{D}| \leqslant K$ called its* number of faces.
- *An integer $|\partial\mathcal{D}| \leqslant K\ell$ called its* boundary length.
- *An integer $n \leqslant |\mathcal{D}|$ called its* number of distinct relators.
- *An application $r^{\mathcal{D}}$ from $\{1, \ldots, |\mathcal{D}|\}$ to $\{1, \ldots, n\}$; if $r^{\mathcal{D}}(i) = r^{\mathcal{D}}(j)$ we will say that* faces i and j bear the same relator.
- *An integer $k \leqslant 3K/\varepsilon_2$ called the* number of translators *of $\mathcal{D}$.*
- *For each integer $2 \leqslant i \leqslant 2k+1$, a set of the form $\{j_i\} \times I'_i$ where either j_i is an integer between 1 and $|\mathcal{D}|$ and I'_i is an oriented cyclic subinterval of I_ℓ, or $j_i = |\mathcal{D}| + 1$ and I'_i is a subinterval of $I_{|\partial\mathcal{D}|}$; this is called an* (internal) subword of the j_i-th face *in the first case, or a* boundary subword *in the second case.*
- *For each integer $1 \leqslant i \leqslant k$ such that $j_{2i} \leqslant |\mathcal{D}|$, an integer between 0 and 4ℓ called the* apparent length of the $2i$-th subword.

such that

- *The sets $\{j_i\} \times I'_i$ are all disjoint and the cardinal of their union is at least $(1-\varepsilon_1)\left(|\mathcal{D}|\,\ell + |\partial\mathcal{D}|\right)$.*

- *For all $1 \leqslant i \leqslant k$ we have $j_{2i} \leqslant j_{2i+1}$ (this can be ensured by maybe swapping them).*

- *If two faces j_{2i} and j_{2i+1} bear the same relator, then either I'_{2i} and I'_{2i+1} are disjoint or are equal with opposite orientations.*

This way, Proposition 14 ensures that any worth-considering van Kampen diagram D with respect to $G_0/\langle R\rangle$ defines a decorated abstract van Kampen diagram $\mathcal{D}$ in the way suggested by terminology (up to rounding the apparent lengths to the nearest integer; we neglect this problem). We will say that $\mathcal{D}$ is *associated to* D. Remark 15 tells that the number of decorated abstract van Kampen diagrams grows subexponentially with ℓ (for fixed K).

Given a decorated abstract van Kampen diagram $\mathcal{D}$ and n given relators $r_1, \ldots, r_n$, we say that these relators *fulfill* $\mathcal{D}$ if there exists a worth-considering van Kampen diagram D with respect to $G_0/\langle r_1, \ldots, r_n\rangle$, such that the associated decorated abstract van Kampen diagram is $\mathcal{D}$. Intuitively speaking, the relators $r_1, \ldots, r_n$ can be "glued modulo G_0 in the way described by $\mathcal{D}$".

So we want to study which diagrams can probably be fulfilled by random relators in R. The main conclusion from [Oll] is that these are those with large boundary length, hence hyperbolicity of the quotient $G_0/\langle R\rangle$. Here for cogrowth we are rather interested in the number of ways to fulfill an abstract diagram with given boundary length.

4.2 Cogrowth of random quotients

So now let R again be a set of $(2m)^{d\ell}$ random relators. Let $\mathcal{D}$ be a given decorated abstract van Kampen diagram. Recall we set $K = A/C$. The free parameters ε_1 and ε_2 will be chosen later.

We will show (Proposition 21) that, up to exponentially small probability in ℓ, the number of different boundary words of worth-considering van Kampen diagrams D such that $\mathcal{D}$ is associated to D, is at most $(2m)^{\theta|\partial\mathcal{D}|(1+\varepsilon/2)}$.

Further notations. Let n be the number of distinct relators in $\mathcal{D}$. For $1 \leqslant a \leqslant n$, let m_a be the number of times the a-th relator appears in $\mathcal{D}$. Up to reordering, we can suppose that the m_a 's are non-increasing. Also to avoid trivialities take n minimal so that $m_n \geqslant 1$.

Let also P_a be the probability that, if a words $r_1, \ldots, r_a$ of length ℓ are picked at random, there exist $n - a$ words $r_{a+1}, \ldots, r_n$ of length ℓ such that the relators $r_1, \ldots, r_n$ fulfill $\mathcal{D}$. The P_a 's are of course a non-increasing sequence of probabilities. In particular, P_n is the probability that a random n-tuple of relators fulfills $\mathcal{D}$.

Back to our set R of $(2m)^{d\ell}$ randomly chosen relators. Let P^a be the probability that there exist a relators $r_1, \ldots, r_a$ in R, such that there exist words $r_{a+1}, \ldots, r_n$

of length ℓ such that the relators $r_1, \ldots, r_n$ fulfill $\mathcal{D}$. Again the P^a 's are a non-increasing sequence of probabilities and of course we have

$$P^a \leqslant (2m)^{ad\ell} P_a$$

since the $(2m)^{ad\ell}$ factor accounts for the choice of the a-tuple of relators in R.

The probability that there exists a worth-considering van Kampen diagram D with respect to the random presentation R, such that $\mathcal{D}$ is associated to D, is by definition less than P^a for any a. In particular, if for some $\mathcal{D}$ we have $P^a \leqslant (2m)^{-\varepsilon'\ell}$, then with probability exponentially close to 1 when $\ell \to \infty$, $\mathcal{D}$ is not associated to any worth-considering van Kampen diagram of the random presentation. Since, by Remark 15, the number of possibilities for $\mathcal{D}$ grows subexponentially with ℓ, we can sum this over $\mathcal{D}$ and conclude that for any $\varepsilon' > 0$, with probability exponentially close to 1 when $\ell \to \infty$ (depending on ε'), all decorated abstract van Kampen diagrams $\mathcal{D}$ associated to some worth-considering van Kampen diagram of the random presentation satisfy $P^a \geqslant (2m)^{-\varepsilon'\ell}$ and in particular

$$P_a \geqslant (2m)^{-ad\ell-\varepsilon'\ell}$$

which we assume from now on.

We need to define one further quantity. Keep the notations of Definition 18. Let $1 \leqslant a \leqslant n$ and let $1 \leqslant i \leqslant k$ where k is the number of translators of $\mathcal{D}$. Say that the i-th translator is half finished at time a if $r^{\mathcal{D}}(j_{2i}) \leqslant a$ and $r^{\mathcal{D}}(j_{2i+1}) > a$, that is, if one side of the translator is a subword of a relator $r_{a'}$ with $a' \leqslant a$ and the other of $r_{a''}$ with $a'' > a$. Now let A_a be the sum of the apparent lengths of all translators which are half finished at time a. In particular, A_n is the sum of the apparent lengths of all subwords $2i$ such that $2i$ is an internal subword and $2i+1$ is a boundary subword of $\mathcal{D}$.

The proof. In this context, equation $(\star)$ (section 6.8) of [Oll] reads

$$A_a - A_{a-1} \geqslant m_a \left(\ell(1-\varepsilon'') + \frac{\log_{2m} P_a - \log_{2m} P_{a-1}}{\beta} \right)$$

where ε'' tends to 0 when our free parameters $\varepsilon_1, \varepsilon_2$ tend to 0 (and ε'' also absorbs the $o(\ell)$ term in [Oll]). Also recall that in the model of random quotient by plain random words, we have

$$\beta = 1 - \theta$$

by Proposition 15 of [Oll].

Setting $d'_a = \log_{2m} P_a$ and summing over a we get, using $\sum m_a = |\mathcal{D}|$, that

$$\begin{aligned} A_n &\geqslant \left(\sum m_a\right) \ell\,(1-\varepsilon'') + \frac{1}{\beta} \sum m_a (d'_a - d'_{a-1}) \\ &= |\mathcal{D}|\,\ell(1-\varepsilon'') + \frac{1}{\beta} \sum d'_a (m_a - m_{a+1}) \end{aligned}$$

Now recall we saw above that for any $\varepsilon' > 0$, taking ℓ large enough we can suppose that $P_a \geqslant (2m)^{-ad\ell-\varepsilon'\ell}$, that is, $d'_a + ad\ell + \varepsilon'\ell \geqslant 0$. Hence

$$\begin{aligned} A_n &\geqslant |\mathcal{D}|\,\ell(1-\varepsilon'') + \frac{1}{\beta}\sum(d'_a + ad\ell + \varepsilon'\ell)(m_a - m_{a+1}) \\ &\quad -\frac{1}{\beta}\sum(ad\ell + \varepsilon'\ell)(m_a - m_{a+1}) \\ &= |\mathcal{D}|\,\ell(1-\varepsilon'') + \frac{1}{\beta}\sum(d'_a + ad\ell + \varepsilon'\ell)(m_a - m_{a+1}) - \frac{d\ell}{\beta}\sum m_a - \frac{\varepsilon'\ell}{\beta}m_1 \\ &\geqslant |\mathcal{D}|\,\ell(1-\varepsilon'') + \frac{d'_n + nd\ell + \varepsilon'\ell}{\beta}m_n - \frac{d\ell + \varepsilon'\ell}{\beta}\sum m_a \end{aligned}$$

where the last inequality follows from the fact that we chose the order of the relators so that $m_a - m_{a+1} \geqslant 0$.

So using $m_n \geqslant 1$ we finally get

$$A_n \geqslant |\mathcal{D}|\,\ell\left(1 - \varepsilon'' - \frac{d+\varepsilon'}{\beta}\right) + \frac{d'_n + nd\ell}{\beta}$$

Suppose the free parameters ε_1, ε_2 and ε' are chosen small enough so that $1 - \varepsilon'' - (d+\varepsilon')/\beta \geqslant 0$ (remember that ε'' is a function of $\varepsilon_1, \varepsilon_2$ and K; we will further decrease ε_1 and ε_2 later). This is possible since by assumption we take the density d to be less than the critical density β. This is the only, but crucial, place where density plays a role. Thus the first term in the inequality above is non-negative and we obtain the simple inequality $A_n \geqslant (d'_n + nd\ell)/\beta$.

Proposition 19 – *Up to exponentially small probability in ℓ, we can suppose that any worth-considering decorated abstract van Kampen diagram $\mathcal{D}$ satisfies*

$$A_n(\mathcal{D}) \geqslant \frac{d'_n(\mathcal{D}) + nd\ell}{\beta}$$

This we now use to evaluate the number of possible boundary words for van Kampen diagrams associated with $|\mathcal{D}|$.

Remember that, by definition, d'_n is the log-probability that n random relators $r_1, \ldots, r_n$ fulfill $\mathcal{D}$. As there are $(2m)^{nd\ell}$ n-tuples of random relators in R (by definition of the density model), by linearity of expectation the expected number of n-tuples of relators in R fulfilling $\mathcal{D}$ is $(2m)^{nd\ell+d'_n}$, hence the interest of an upper bound for $d'_n + nd\ell$.

By the Markov inequality, for given $\mathcal{D}$ the probability to pick a random set R such that the number of n-tuples of relators of R fulfilling $\mathcal{D}$ is greater than $(2m)^{nd\ell+d'_n+C\varepsilon\ell/4}$, is less than $(2m)^{-C\varepsilon\ell/4}$. By Remark 15 the number of possibilities for $\mathcal{D}$ is subexponential in ℓ, and so, using Proposition 19 we get

Proposition 20 – *Up to exponentially small probability in ℓ, we can suppose that for any worth-considering decorated abstract van Kampen diagram $\mathcal{D}$, the* *number of n-tuples of relators in R fulfilling $\mathcal{D}$ is at most*

$$(2m)^{\beta A_n(\mathcal{D})+C\varepsilon\ell/4}$$

Now let D be a van Kampen diagram associated to $\mathcal{D}$. Given $\mathcal{D}$ we want to evaluate the number of different boundary words for D. Recall Proposition 14: the boundary word of D is determined by giving two words for each boundary-boundary translator, and one word for each internal-boundary translator, this latter one being subject to the apparent length condition imposed in the definition of $\mathcal{D}$. By Remark 16, the number of ways to combine these subwords into a boundary word for D is controlled by K and ε_2 (independently of ℓ).

So let (x_{2i}, x_{2i+1}) be a boundary-boundary translator in D. By Proposition 14 (definition of translators) there exist words δ_1, δ_2 of length at most $\varepsilon_2(|x_{2i}|+|x_{2i+1}|)$ such that $x_{2i}\delta_1 x_{2i+1}\delta_2 = e$ in G_0. So $x_{2i}\delta_1 x_{2i+1}\delta_2$ is a word representing the trivial element in G_0, and by definition of θ the number of possibilities for (x_{2i}, x_{2i+1}) is at most $(2m)^{\theta(|x_{2i}|+|x_{2i+1}|)(1+2\varepsilon_2)}$.

Now let (x_{2i}, x_{2i+1}) be an internal-boundary translator. The apparent length of x_{2i} is imposed in the definition of $\mathcal{D}$. The subword x_{2i} is an internal subword of D, and so by definition is a subword of some relator $r_i \in R$. So if the relators in D are given, x_{2i} is determined. But knowing x_{2i} still leaves open lots of possibilities for x_{2i+1}. This is where apparent length comes into play.

Since $y = x_{2i+1}$ is a boundary word of D one has $|y| \leqslant A\ell \leqslant K\ell$. So by definition we have $\mathbb{L}(x_{2i}) \leqslant \mathbb{L}_{|y|}(x_{2i})$. By definition of translators there exist words u and v of length at most $\varepsilon_2\ell$ such that $x_{2i}uyv = e$ in G_0. By definition of $\mathbb{L}_{|y|}(x_{2i})$, if y' is a random word of length $|y|$, then the probability that $x_{2i}uy'v = e$ in G_0 is $(2m)^{-(1-\theta)\left(|y|+\mathbb{L}_{|y|}(x_{2i})\right)} \leqslant (2m)^{-(1-\theta)(|y|+\mathbb{L}(x_{2i}))}$. This means that the total number of words y' of length $|y|$ such that there exists u, v with $x_{2i}uyv = e$ is at most $(2m)^{|y|}(2m)^{-(1-\theta)(|y|+\mathbb{L}(x_{2i}))} = (2m)^{\theta|y|-(1-\theta)\mathbb{L}(x_{2i})}$. So, given x_{2i}, the number of possibilities for $y = x_{2i+1}$ is at most this number.

So if the relators in R fulfilling $\mathcal{D}$ are fixed, the number of possible boundary words for D is the product of $(2m)^{\theta(|x_{2i}|+|x_{2i+1}|)(1+2\varepsilon_2)}$ for all boundary-boundary translators (x_{2i}, x_{2i+1}), times the product of $(2m)^{\theta|x_{2i+1}|-(1-\theta)\mathbb{L}(x_{2i})}$ for all internal-boundary translators (x_{2i}, x_{2i+1}), times the number of ways to order these subwords (which is subexponential in ℓ by Remark 16), times the number of possibilities for the parts of the boundary of D not belonging to any translator, which by Proposition 14 have total length not exceeding $\varepsilon_1 K\ell$.

Now the sum of $|x_{2i}|+|x_{2i+1}|$ for all boundary-boundary translators (x_{2i}, x_{2i+1}), plus the sum of $|x_{2i+1}|$ for all internal-boundary translators, is $|\partial\mathcal{D}|$ (maybe up to $\varepsilon_1 K\ell$). And the sum of $\mathbb{L}(x_{2i})$ for all internal-boundary translators is A_n by definition.

So given $\mathcal{D}$ and given a n-tuple of relators fulfilling $\mathcal{D}$, the number of possibilities for the boundary word of D is at most

$$(2m)^{\theta|\partial\mathcal{D}|(1+2\varepsilon_2)-(1-\theta)A_n+\varepsilon_1 K\ell}$$

up to a subexponential term in ℓ. By Proposition 20 (remember $\beta = 1 - \theta$), if we include the choices of the relators fulfilling $\mathcal{D}$ the number of possibilities is at most

$$(2m)^{\theta|\partial\mathcal{D}|(1+2\varepsilon_2)+\varepsilon_1 K\ell + C\varepsilon\ell/4}$$

If we choose $\varepsilon_2 \leqslant \varepsilon/16$ and $\varepsilon_1 \leqslant \varepsilon C/8K$ so that (using $|\partial D| \geqslant C\ell\,|\mathcal{D}| \geqslant C\ell$ for any fulfillable abstract diagram) the sum of the corresponding terms is less than $\varepsilon\,|\partial D|\,/4$ (note that this choice does not depend on ℓ) and if we remember that, after Remark 15, the number of choices for $\mathcal{D}$ is subexponential in ℓ, we finally get:

Proposition 21 – *Up to exponentially small probability in ℓ, the number of different boundary words of worth-considering van Kampen diagrams of a random presentation with given boundary length L, is at most*

$$(2m)^{\theta L(1+\varepsilon/2)}$$

But remember the discussion at the beginning of section 4 (where we invoked Proposition 8): it is enough to show that gross cogrowth is at most $\theta + \varepsilon/2$ for words of length L between $A\ell/4$ and $A\ell$. Any such word is the boundary word of a van Kampen diagram of area at most $K\ell$, hence is the boundary word of some worth-considering van Kampen diagram. This ends the proof of Theorem 2.

References

[C] J.M. Cohen, *Cogrowth and Amenability of Discrete Groups*, J. Funct. Anal. **48** (1982), 301–309.

[Ch93] C. Champetier, *Cocroissance des groupes à petite simplification*, Bull. London Math. Soc. **25** (1993), No 5, 438–444.

[Ch95] C. Champetier, *Propriétés statistiques des groupes de présentation finie*, J. Adv. Math. **116** (1995), No 2, 197–262.

[Ch00] C. Champetier, *L'espace des groupes de type fini*, Topology **39** (2000), No 4, 657–680.

[GdlH] R.I. Grigorchuk, P. de la Harpe, *On problems related to growth, entropy, and spectrum in group theory*, Dynam. Control Systems **3** (1997), No 1, 51–89.

[Gh] É. Ghys, *Groupes aléatoires*, séminaire Bourbaki **916** (2003).

[Gri] R.I. Grigorchuk, *Symmetrical Random Walks on Discrete Groups*, in *Multicomponent Random Systems*, ed. R.L. Dobrushin, Ya.G. Sinai, Adv. Prob. Related Topics **6**, Dekker (1980), 285–325.

[Gro87] M. Gromov, *Hyperbolic Groups*, in *Essays in group theory*, ed. S.M. Gersten, Springer (1987), 75–265.

[Gro93] M. Gromov, *Asymptotic Invariants of Infinite Groups*, in *Geometric group theory*, ed. G. Niblo, M. Roller, Cambridge University Press, Cambridge (1993).

[Gro03] M. Gromov, *Random Walk in Random Groups*, Geom. Funct. Anal. **13** (2003), No 1, 73–146.

[HLS] N. Higson, V. Lafforgue, G. Skandalis, *Counterexamples to the Baum-Connes conjecture*, Geom. Funct. Anal. **12** (2002), No 2, 330–354.

[K] H. Kesten, *Symmetric Random Walks on Groups*, Trans. Amer. Math. Soc. **92** (1959), 336–354.

[LS] R.C. Lyndon, P.E. Schupp, *Combinatorial Group Theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete **89**, Springer (1977).

[Oll] Y. Ollivier, *Sharp phase transition theorems for hyperbolicity of random groups*, GAFA, Geom. Funct. Anal. **14** (2004), No 3, 595–679.

[Ols] A.Yu. Ol'shanskiĭ, *Almost Every Group is Hyperbolic*, Int. J. Algebra Comput. **2** (1992), No 1, 1–17.

[P] F. Paulin, *Sur la théorie élémentaire des groupes libres*, Séminaire Bourbaki **922** (2003).

[Sh] H. Short et al., in *Group Theory from a Geometrical Viewpoint*, ed. É. Ghys, A. Haefliger, A. Verjovsky, World Scientific (1991).

[W] W. Woess, *Random Walks on Infinite Graphs and Groups*, Cambridge Tracts in Mathematics **138**, Cambridge University Press (2000).