Limits of Rauzy graphs and horocyclic products of trees

Preliminary version

P.-H. Leemann

May 19, 2016

The main result of this paper is the computation of the Benjamini-Schramm limit of the Rauzy graphs associated to a finite subshift. We show that it is a measure supported on horocyclic product of two trees.

1 Introduction

Let $\mathcal{A}$ be a finite alphabet on $k \geq 2$ letters. For convenience, say $\mathcal{A} = \{0, 1, \dots, k-1\}$. We note $\mathcal{A}^n$ the set of sequences of n letters and $\mathcal{A}^*$ the set of finite sequences (it is the free monoid on $\mathcal{A}$). We also define $\mathcal{A}^{\mathbf{N}}$, respectively $\mathcal{A}^{-\mathbf{N}}$, to be the set of right, respectively left, infinite sequences on $\mathcal{A}$

Def of oriented graphs, positive and negative edges, derangement

Let $F \subseteq \mathcal{A}^*$ be a finite set of forbidden words. We are interested in $\mathcal{A}^*_F$ (and in $\mathcal{A}^n_F$), the set of all words (of length n) without subwords belonging to F. Up to changing $\mathcal{A}$, we can suppose that $F \subseteq \mathcal{A}^2$. In the following , we will be interested only in F such that for all $a \in \mathcal{A}$, there exists b and c such that ab and ca do not belong to F. That is, any element of $\mathcal{A}^*_F$ is always left and right extendable.

Really necessary ?

Given an alphabet $\mathcal{A}$ and a set of forbidden words F, we can construct an infinite family of oriented labeled *Rauzy graphs* $R_{F,n}$ in the following way. The vertex set of $R_{F,n}$ is $\mathcal{A}^n_F$ and the edge set is $\mathcal{A}^{n+1}_F$, where the edge $x_1 \dots x_{n+1}$ is labeled by x_{n+1} and has initial vertex $x_1 \dots x_n$ and final vertex $x_2 \dots x_{n+1}$.

For example, $R_{\emptyset,n} = B_n$ the oriented de Bruijn graph on $\mathcal{A}$.

For $F \subseteq \mathcal{A}^2$, it is also natural to define an infinite oriented labeled (non connected) graph $R_{F,\infty}$ with vertex set $\mathcal{A}^{\mathbf{N}}_F \times \mathcal{A}^{-\mathbf{N}}_F$ and with an edge labeled by i from $(x_1x_2\dots; \dots x_{-2}x_{-1})$ to $(x_2x_3\dots; \dots x_{-1}i)$. We also define the subgraph $R'_{F,\infty} \subseteq R_{F,\infty}$ as the complete subgraph on the vertex set consisting of paires (ξ, η) such that at least one of the sequence is not eventually periodic. We have that $R'_{F,\infty}$ is a disjoint union of connected components of $R_{F,\infty}$.

As an immediate consequence of the definitions we have the following.

Lemma 1. *For every $F \subseteq \mathcal{A}^2$,*

1. *If $n \geq 2$, then $R_{F,n}$ is a complete subgraph of B_n;*

2. *If $n \geq 1$, then $R_{F,n+1} \cong \mathcal{L}(R_{F,n})$, the line graph of $R_{F,n}$.*

Proof. If $x_1 \dots x_n$ is a vertex in $R_{F,n}$, then $x_2 \dots x_n i$ is in $R_{F,n}$ if and only if $x_n i \notin F$, if and only if $x_1 x_2 \dots x_n i$ is an edge in $R_{F,n}$.

The isomorphism is given by $\phi\colon \mathcal{L}(R_{F,n}) \to R_{F,n+1}$ described in Figure 1. $\square$

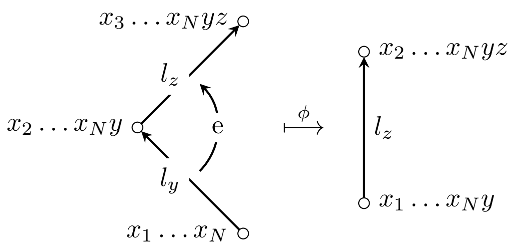

Figure 1: The edge e in the line graph of $R_{F,n}$ and its image by ϕ.

Since we have, for each F, a infinite family of finite graphs $R_{F,n}$, it is natural to ask for the existence of some kind of limit for this family. In the context of finite non-rooted graphs, the more natural notion of limit is the Benjamini-Schramm limit, also called local limit or weak limit. A rooted graph, is a pair (G, v) where G is a graph and v, the root, is a vertex of G. There are natural analogs of this definition for oriented, labeled or oriented labeled graphs. Let $\mathcal{G}_\bullet$ denotes the space of oriented, labeled, rooted graphs, up to labeled root preserving isomorphisms. This is a metric space where the distance between (G, v) and (H, w) is 2^{-r} with r is the biggest integer such that $\mathrm{Ball}_G(v, r)$ and $\mathrm{Ball}_H(w, r)$ are isomorphic. For any integer d, the space $\mathcal{G}_{\bullet,d}$ of marked graph with degree bounded by d is compact. To a sequence of finite graphs G_n, it is possible to attach a sequence of probability measure μ_n on $\mathcal{G}_\bullet$, where μ_n consist of choosing uniformly the root of G_n. The sequence (G_n) is said to have the Benjamini-Schramm limit μ if the sequence μ_n weakly converges to μ. Every sequence of finite graphs of uniformly bounded degree admits a converging subsequence.

Rewrite this part.

Be careful of the orientation. Add references + convergence by sampling + compactness.

Def of Ball as the ball in the underlying non-oriented graph

Question 1. Does the Benjamini-Schramm limit of $R_{F,n}$ exist and it is possible to compute it?

We will show that, under some mild assumptions, any connected component of $R'_{F,\infty}$ is a limit of the sequence $(R_{F,n})_{n\geq 1}$. Before doing that, we show that connected components of $R_{F,\infty}$ are horocyclic products of tree.

Definition 1. Let $F \subseteq \mathcal{A}^2$ such that $\mathcal{A}_F^{\mathbf{N}}$ is non empty and let $\xi = (x_1 x_2 \dots)$ be an element of $\mathcal{A}_F^{\mathbf{N}}$. We will construct a directed infinite tree $T_{F,\xi}$. Start with an infinite ray, directed upstair, where vertices are labeled by $x_1, x_2, \dots$. Under each vertex v already constructed and labeled by x, for each $i \in \mathcal{A}$ such that $ix \notin F$, if there is not already a vertex with label i under v, attach a new vertex with label i. Repeat this construction

This is the tree of sequences that are co-finals with ξ. Maybe use this as a definition.

infinitely many time to obtain a tree. Finally, we put an orientation on $T_{F,\xi}$ such that each edge is going up and we label each edge by its initial vertex.

Maybe not label this one?

If $\eta = \dots x_{-2}x_{-1}$ is an element of $\mathcal{A}_F^{-\mathbf{N}}$, we can similarly construct a directed infinite tree $T_{F,\eta}$. Start with an infinite ray, directed upstair , where vertices are labeled by $x_{-1}, x_{-2}, \dots$. Under each vertex v already constructed and labeled by x, for each $i \in \mathcal{A}$ such that $xi \notin F$, if there is not already a vertex with label i under v, attach a new vertex with label i. Repeat this construction infinitely many time to obtain a tree. Finally, we put an orientation on $T_{F,\eta}$ such that each edge is going down and we label each edge by its terminal vertex.

not very clear: directed upstair, but the edges are going down

See Figure 2 for an example. An important but rather obvious property of $T_{F,\xi}$ is that all paths (and infinite and biinfinite rays) consisting of positive edges are labeled by elements in $\mathcal{A}_F^*$ (or in $\mathcal{A}_F^{\mathbf{N}} \cup \mathcal{A}_F^{\mathbf{Z}}$ for rays). Moreover, infinite rays consisting of positive edges in $T_{F,\xi}$ (respectively in $T_{F,\xi}$) are exactly the right infinite words in $\mathcal{A}_F$ that are cofinal with ξ (respectively coinitial with η).

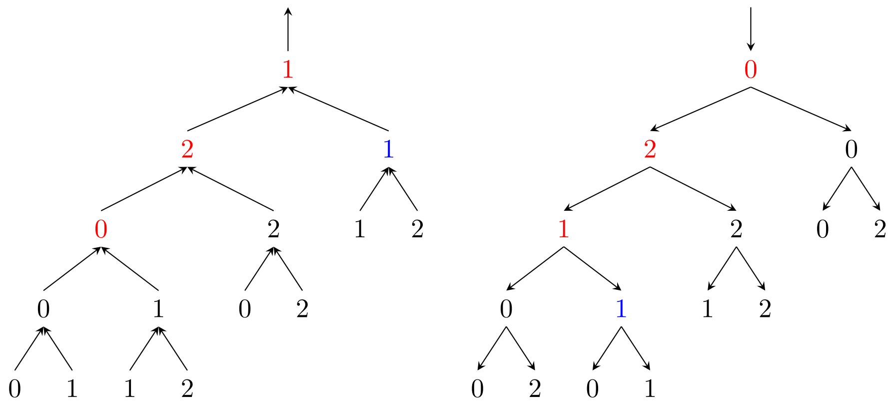

Figure 2: Part of the tree $T_{\{01,12,20\},021\dots}$ and $T_{\{01,12,20\},\dots 021}$ on the alphabet $\{0,1,2\}$. Labeling of the edges are not shown. In blue, the image of $w = (11\dots;\dots 0211)$.

For an infinite tree T, and a right infinite ray ξ, it is possible to define the *Busemann rank function* $\mathfrak{h}\colon V(X) \to \mathbf{Z}$ by $\mathfrak{h}(v) := d(p_\xi(v), \xi_1) - d(p_\xi(v), v)$, where $p_\xi(v)$ is the projection of v onto ξ. This naturally endows T with an orientation: the initial vertex of an edge between v and w is the vertex with the lower Busemann rank. If η is a left infinite ray, we define $\mathfrak{h}$ by $\mathfrak{h}(v) := -d(p_\eta(v), \eta_1) + d(p_\eta(v), v)$

we already defined an orientation on T...

Definition 2. Let (T_1, ξ) be a pair consisting of a infinite tree with a right infinite ray and (T_2, η) be a pair consisting of a infinite tree with a left infinite ray. The horocyclic product $\mathrm{DL}((T_1,\xi),(T_2,\eta))$ of (T_1,ξ) and (T_2,η) is graph with vertex set $\{(v,w) \in V(T_1) \times V(T_2) \mid \mathfrak{h}(v) = \mathfrak{h}(w)\}$. There is an edge from (v,w) to (x,z) if and only if there is an edge from v to x and an edge from w to z. This graph is naturally rooted at (ξ_1, η_1)

Explain the notation.

It is a connected component of the tensor product.

If T_2 is a labeled graphs, then the horocyclic product is a labeled graph where the label of (e,f) is the label of f.

This is one possibility to define the label. Justify!

Lemma 2. *Let* $(\xi, \eta) = (x_1 x_2 \dots ; \dots x_{-2} x_{-1})$ *be a vertex in* $R'_{F,\infty}$. *Let* p *and* q *be two paths with initial vertex* v *and with same final vertex. Then,* $\operatorname{der}(p) = \operatorname{der}(q)$.

Proof. Observe that positive edges shift both sequences $x_1 x_2 \dots$ and $\dots x_{-2} x_{-1}$ to the left (and change a finite number of digits), while negative edges shift the sequence to the right.

Therefore, if $\tau(p) = \tau(q)$ but $\operatorname{der}(p) \neq \operatorname{der}(q)$, both sequences ξ and η are eventually periodic, which is absurd. $\square$

Definition 3. Let v be a vertex in $R'_{F,\infty}$. We define the associated *rank function*: $\operatorname{rk}_v \colon (R'_{F,\infty}, v)^0 \to \mathbf{Z}$ by $\operatorname{rk}_v(w) = \operatorname{der}(p)$ where p is any path from v to w.

By the last lemma, this is well defined. The following formula directly follows from definitions and will be useful later.

Lemma 3. *Let* w *be a vertex in* $(R'_{F,\infty}, v)^0$. *If* $\operatorname{rk}_v(w) \geq 0$, *we have*

$$w = (y_{1+\operatorname{rk}_v(w)} \dots y_n v_{n+1} \dots ; \dots v_{-m} y_{-m+1} \dots y_{-1} z_1 \dots z_{\operatorname{rk}_v(w)})$$

for some y_j*'s and* z_j*'s in* $\mathcal{A}$, *while if* $\operatorname{rk}_v(w) \leq 0$ *we have*

$$w = (z_1 \dots z_{-\operatorname{rk}_v(w)} y_1 \dots y_n v_{n+1} \dots ; \dots v_{-m} y_{-m+1} \dots y_{-1+\operatorname{rk}_v(w)}).$$

Theorem 1. *For each vertex* $v = (\xi, \eta)$ *in* $R'_{F,\infty}$, *there is a labeled and rooted isomorphism between* $(R'_{F,\infty}, v)^0$ *and the horocyclic product* $\operatorname{DL}(T_{F,\xi}, T_{F,\eta})$.

give a better name

Proof. For a vertex x in $T_{F,\xi}$, the projection of x on ξ is noted by $p_\xi(x)$.

already defined

Let w be a vertex in $(R'_{F,\infty}, v)^0$. By the last lemma, if $\operatorname{rk}_v(w) \geq 0$, then w determines a vertex in $T_{F,\xi}$: the unique vertex x such that $p_\xi(x) = v_{n+1}$ and the label of vertices between x and v_{n+1} is given by $y_{1+\operatorname{rk}_v(w)} \dots y_n$. This also determines a vertex in $T_{F,\eta}$: the unique vertex y such that $p_\eta(y) = v_{-m}$ and the label of vertices between x and v_{n+1} is given by $y_{-m+1} \dots y_{-1} z_1 \dots z_{\operatorname{rk}_v(w)}$. Under this correspondance, we have

$$\begin{aligned} \mathfrak{h}(x) &= n - (n - \operatorname{rk}_v(w)) = \operatorname{rk}_v(w) \\ \mathfrak{h}(y) &= -(m-1) + (\operatorname{rk}_v(w) + m - 1) = \operatorname{rk}_v(w) \end{aligned}$$

A similar argument takes care of the case $\operatorname{rk}_v(w) \leq 0$.

Hence, we have a function ϕ from vertices of $(R'_{F,\infty}, v)^0$ to vertices of $\operatorname{DL}(T_{F,\xi}, T_{F,\eta})$. By definition, ϕ preserves root and is injective.

In $R'_{F,\infty}$, there is an outgoing edge e from w with label i if and only if $w_{-1} i$ does not belongs to F, if and only if in $\operatorname{DL}(T_{F,\xi}, T_{F,\eta})$ there is an outgoing edge f from $\phi(w)$ with label i. Moreover, in this case, the terminal vertex of e is $(w_2 \dots ; \dots w_{-1} i)$ which is sent by ϕ to the terminal vertex of f. Therefore, the function ϕ naturally extends to a labeled graph homomorphism, which is thus locally bijective on outgoing edges. The injectivity of ϕ implies that it is locally injective and thus locally injective on ingoing edges.

It remains to show that ϕ is locally surjective on ingoing edges. Indeed, in this case ϕ is locally bijective and hence surjective. If w is a vertex in $R'_{F,\infty}$, then the number of ingoing edges is $|\{i \mid i w_1 \notin F\}|$, which is exactly the number of ingoing edges for $\phi(w)$. Since ϕ is locally injective on ingoing edges, it is also locally bijective on ingoing edges. $\square$

For an illustration of the proof, see Figure 2. It shows the two tree corresponding to $v = (021\ldots,\ldots 021)$ and in blue the vertices corresponding to $w = (1v_3\ldots;\ldots v_{-1}1)$. This w is connected to v by at least two paths, given by

$$v \xrightarrow{1} (v_2\ldots;\ldots v_{-1}1) \xrightarrow{j} (v_3\ldots;\ldots v_{-1}1j) \xleftarrow{j} w$$

for $j \in 0, 1$.

Now that we have a description of connected components of $R'_{F,\infty}$ as horocyclic products, we are going to investigate the relations between the $R_{F,n}$'s and $R'_{F,\infty}$. In order to do that, we are going to construct a labeled tree where (some of the) vertices of level n will correspond to vertices of $R_{F,n}$.

Definition 4. Let $F \subseteq \mathcal{A}^2$ be a set of forbidden words. The associated infinite rooted tree T_F is constructed inductively on level as follows. The root is the empty set. Every vertex $x_1\ldots x_{2n}$ of level $2n$ has as children all the $x_1\ldots x_n i x_{n+1}\ldots x_{2n}$ such that $x_n i \notin F$. On the other hand, every vertex vertex $x_1\ldots x_{2n+1}$ of level $2n+1$ has as children all the $x_1\ldots x_{n+1} i x_{n+2}\ldots x_{2n+1}$ such that $i x_{n+1} \notin F$. The vertices in $\mathcal{A}_F^*$ are called *good vertices* and the ones not in $\mathcal{A}_F^*$ the *bad vertices*. Put later in the text ?

Observe that in a bad vertex $x_1\ldots x_n$, the only forbidden word that appears is in the middle, i.e. is $x_{\lceil \frac{n}{2}\rceil} x_{\lceil \frac{n}{2}\rceil+1}$. See Figure 3 for an example.

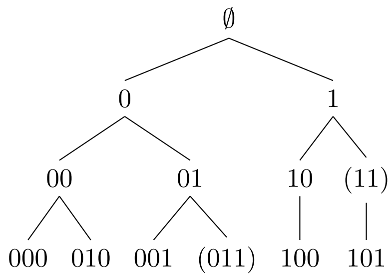

Figure 3: The top of the tree $T_{\{11\}}$ on the alphabet $\{0,1\}$. Bad vertices are parenthesized.

By construction, good vertices of level n are in one-to-one correspondance with vertices of $R_{F,n}$. On the other hand, infinite rays in T_F are parametrized by a pair of infinite sequences $(x_1x_2\ldots;\ldots x_{-2}x_{-1})$ and are in one-to-one correspondance with vertices (ξ,η) in $R_{F,\infty}$.

Definition 5. Let v be a vertex in $R'_{F,\infty}$ and $r \in \mathbf{N}$ an integer. For all $n \geq 2r+2$, define a function on vertices

$$\begin{aligned}\pi_n = \pi_{n,r} \colon \ &\mathrm{Ball}_{R'_{F,\infty}}(v,r) \to R_{\emptyset,n}\\ &w = (w_1\ldots;\ldots w_{-1}) \mapsto (w_1\ldots w_{\lceil \frac{n}{2}\rceil - \mathrm{rk}_v(w)} w_{-\lfloor \frac{n}{2}\rfloor - \mathrm{rk}_v(w)}\ldots w_{-1})\end{aligned}$$

Lemma 4. *Let $v \in R'_{F,\infty}$. For all $n \geq 2r+2$, the function π_n*

1. *Is well-defined and has values in* $\mathrm{Ball}_{R_{\emptyset,n}}(\pi_n(v), r)$*;*
2. *Naturally extends to a labeled graph homomorphism;*
3. *Is locally injective;*
4. *Is locally bijective if* $v_{\lceil \frac{n}{2} \rceil} v_{-\lfloor \frac{n}{2} \rfloor} \notin F$*;*
5. *We have* $\mathrm{Im}(\pi_n) = \mathrm{Ball}_{R_{F,n}}(\pi_n(v), r)$ *if and only if* $v_{\lceil \frac{n}{2} \rceil} v_{-\lfloor \frac{n}{2} \rfloor} \notin F$*.*

Proof. We will prove the assertions for $2n$, in this case $n = \lceil \frac{2n}{2} \rceil = \lfloor \frac{2n}{2} \rfloor \geq r + 1$. The proofs for $2n + 1$ are similar. Write at least once the prof to check

Since w is at distance at most r from v, it has rank between $-r$ and r. Hence, π_{2n} is well-defined.

If there is an edge labeled by i with initial vertex $w = (w_1 \dots ; \dots w_{-1})$, then it has final vertex $u = (w_2 \dots ; \dots w_{-1} i)$ and $\mathrm{der}(u) = \mathrm{der}(w) + 1$. Therefore, these vertices are send respectively onto $(w_1 \dots w_{n-\mathrm{rk}_v(w)} w_{-n-\mathrm{rk}_v(w)} \dots w_{-1})$ and $(w_2 \dots w_{n-\mathrm{rk}_v(w)} w_{-n-\mathrm{rk}_v(w)} \dots w_{-1} i)$. Hence, there is a natural extension of π_{2n} to a labeled graph homomorphism. This directly implies that it has values in $\mathrm{Ball}_{R_{\emptyset,2n}}(\pi_{2n}(v), r)$.

The local injectivity for outgoing edges follow from the fact that π_{2n} preserves the labeling and that for each vertex there is at most one outgoing edge with label i. On the other hand, for any vertex $w \in R'_{F,\infty}$, the set of incoming edges (they all have label w_{-1}) is $\{(jw_1 \dots ; \dots w_{-2} \,|\, jw_1 \notin F\}$. Since $2n \geq 2r + 2$, these edges are sent onto $j \dots w_{-2}$ which are pairwise distinct.

If $v_n v_{-n}$ is in F, then $\pi_{2n}(v)$ is not in $\mathcal{A}_F^*$. On the other hand, if $v_n v_{-n}$ is not in F, we have $\mathrm{Im}(\pi_{2n}) \subseteq R_{F,2n}$. Indeed, in this case, a vertex w in $\mathrm{Ball}_{R_{F,\infty}}(v, r)$ is of the form $(* \cdots * \, v_{r+1} \dots v_n \dots ; \dots v_{-n} \dots v_{-r-1} * \cdots *)$ and is therefore send onto $* \cdots * \, v_{r+1} \dots v_n v_{-n} \dots v_{-r-1} * \cdots *$. Since w belongs to $R_{F,\infty}$, there is no forbidden words in $* \cdots * \, v_{r+1} \dots v_n$ and in $v_{-n} \dots v_{-r-1} * \cdots *$ and $v_n v_{-n}$ is not forbidden by assumption.

It remains to show that if $v_n v_{-n}$ is not in F, then π_{2n} is locally surjective. Indeed, in this case π_{2n} is locally bijective and hence surjective. If w is a vertex in $R_{F,\infty}$, then the number of outgoing edges is $|\{i \,|\, w_{-1} i \notin F\}|$, which is exactly the number of outgoing edges for $\pi_{2n}(w)$. The same kind of relation holds for ingoing edges. Since π_{2n} is locally injective, it is therefore locally bijective. $\square$

We want to prove that the sequence of $\left(R_{F,n}, \pi_n(v)\right)^0$'s (or at least a subsequence of it) locally converges to $\left(R_{F,\infty}, v\right)^0$ for n going to infinity. Therefore, we want to prove that for all r, there exist n big enough such that π_n is an isomorphism onto $\mathrm{Ball}_{R_{F,n}}(\pi_n(v), r)$. That is, π_n injective and, by the last lemma, $v_{\lceil \frac{n}{2} \rceil} v_{-\lfloor \frac{n}{2} \rfloor}$ is not in F.

Lemma 5. *For all vertex* v *of* $R'_{F,\infty}$ *and all* r*, there exists* n_0 *such that for all* $n \geq n_0$*, the function* π_n *is injective.*

Proof. Let w and u be two distinct vertices in the ball or radius r around v. Suppose that $\mathrm{rk}_v(w) = \mathrm{rk}_v(u) = i \geq 0$. In this case we have

$$w = (y_{1+i} \dots y_r v_{r+1} \dots ; \dots v_{-r-1} y_{-r} \dots y_{-1} z_1 \dots z_i)$$

and a similar formula for u. This implies that $w_j = u_j$ for all $j \geq r - i$ and all $j \leq -r - i$, and therefore that there exists $-r - i \leq j \leq r - i$ such that $w_j \neq u_j$. By definition of π_n, we have $\pi_n(w) \neq \pi_n(u)$ as soon as $\lfloor \frac{n}{2} \rfloor \geq r$, i.e. as soon as $n \geq 2r$. If $\mathrm{rk}_v(w) = \mathrm{rk}_v(u) = i \leq 0$, a similar proof gives the same result.

Suppose now that $\mathrm{rk}_v(u) - \mathrm{rk}_v(w) = i \neq 0$. By Lemma 3, if $\pi_n(u) = \pi_n(w)$, then for all $r < j \leq \lfloor \frac{n}{2} \rfloor - i$ we have $v_j = v_{j+i}$ and $v_{-j} = v_{-j-i}$. That is, between the $r + 1^{\text{th}}$ and the $\lfloor \frac{n}{2} \rfloor^{\text{th}}$ digit, both sequences $(v_1 v_2 \dots)$ and $(v_{-1} v_{-2} \dots)$ are periodic of period i. Since at least one of this sequence is not eventually periodic, there exists $n_i(v)$ such that for all $n \geq n_i(v)$, we have $\pi_n(u) \neq \pi_n(w)$.

Finally, for $n_0 = 2 \cdot \max_{-r \leq i \leq r}\{r + 1, n_i\}$ we have that π_{n_0} is injective and so is π_n for $n \geq n_0$. □

Observe that in the proof we can chose $n_0 = 2j$, where j is the smallest integer $j \geq r+1$ such that at least one of the sequences $(v_{r+1} v_{r+2} \dots v_j)$ or $(v_{-r-1} v_{-r-2} \dots v_{-j})$ is not $r!$ periodic. In order to obtain a better, but less useful in practice, estimate, one can replace the $r!$ periodicity by $\mathrm{lcm}_{l \leq r}\{l\}$ periodicity.

Proposition 1. *Let $v = (w_1, w_2, w_3, \dots)$ be a ray in T_F. Suppose that infinitely many* Careful with the notation *w_i are good and let w_{m_j} be the corresponding subsequence. Suppose that either $v_1 \dots$ or $\dots v_{-1}$ is not eventually periodic. Then the sequence $(R_{F,m_j}, w_{m_j})^0$ locally converges to $(R_{F,\infty}, v)^0$ as j goes to infinity.*

Proof. The ray $v = (v_1 \dots ; \dots v_{-1})$ correspond to the vertex $(v_1 \dots ; \dots v_{-1}) \in R_{F,\infty})$. If at least one of the sequence is not eventually periodic, then $(v_1 \dots ; \dots v_{-1}$ is in $R_{F,\infty}$. Under this correspondance, $w_m = \pi_m(v) = (v_1 \dots v_{\lceil \frac{n}{2} \rceil} v_{-\lfloor \frac{n}{2} \rfloor} \dots v_{-1})$. It is parenthesized if and only if $v_{\lceil \frac{n}{2} \rceil} v_{-\lfloor \frac{n}{2} \rfloor}$ belongs to F. An application of Lemmas 4 and 5 finishes the proof. □

By the discussion after Definition 4, points of ∂T_F are in bijection with vertices of $R_{F,\infty}$. Therefore, we have a map $f \colon \partial T_F \to \mathcal{G}_\bullet$, where $f(v)$ is the equivalence class of $(R_{F,\infty}, v)^0$. Since the function f is continuous and thus measurable, for any measure μ on ∂T_F, the pushforward $f_*(\mu)$ of μ gives a measure on $\mathcal{G}_\bullet$.

Definition 6. For every n, it is possible to define a function μ_n on the the vertices of T_F in the following way. Let $\{v_1 \dots v_m, u_1 \dots u_l\}$ be the vertices of level n of T_F where the v_i are the good ones. Then

$$\mu_n(w) := \begin{cases} \frac{|\{i \mid v_i \leq w\}|}{m} & \text{if } w \text{ is of level at most } n \\ \frac{1}{m} \frac{1}{|\{v \leq v_i \mid v \text{ of same level as } w\}|} & \text{if } w \leq v_i \text{ for some } i \\ 0 & \text{otherwise (i.e. } w \leq u_i \text{ for some } i) \end{cases}$$

That is, if w is of level at most n, then μ_n simply count the percentage of good vertices of level n which are under w. On the other hand, for $w \leq v$ with v of level n, then μ_n is $\mu_n(v)$ times the uniform measure on $(T_F)_v$. Define $(T_F)_v$

It is immediate that the restriction of μ_n on any level defines a measure. Moreover, these measure are compatible in the sense that for all v and all j bigger than the level of v, we have $\mu_n(v) = \sum_{w \leq v, w \in \mathcal{L}(j)} \mu_n(w)$. Therefore, μ_n induces a measure on ∂T_F, also noted μ_n, by $\mu_n(C_v) := \mu_n(v)$ where $C_v = \{w \mid w \leq v\}$ is a cylinder set.

Since ∂T_F is a compact metric space, the set of measure on it is compact for the weak* topology. In particular, since the cylinder sets are clopen, for any converging sequence $\nu_n \to \nu$, the Portmanteau theorem implies that for any $v \in T_F$, the $\lim \nu_m(C_v)$ exists and is equal to $\nu(C_v)$.

We will show that under a weak assumption on $F \subseteq \mathcal{A}^2$, the limit points of the sequence $(\mu_n)_n$ of Definition 6 give raise to Benjamini-Schramm limit points for the sequence of finite graphs $(R_{F,n})_n$.

Let $T = T_F$. Since for any n, the measure μ_n on ∂T is cylindrical, it can be represented by putting some weight on the edges of T such that for each vertex v, the measure $\mu_n(C_v)$ is the product of the weight of edges on the unique path from v to the root. It is immediate that a subsequence $(\mu_{n_i})_i$ is weakly convergent if and only for each v and each child w of v, the sequence of weight of the edge from v to w is convergent.

Definition 7. For any vertex $v = (v_1 \dots v_n)$ of $T = T_F$, define its *type* to be $v_{\lceil \frac{n}{2} \rceil} \overline{v}_{\lceil \frac{n}{2} \rceil + 1}$ if v is of even level and $\overline{v}_{\lceil \frac{n}{2} \rceil} v_{\lceil \frac{n}{2} \rceil + 1}$ if v is of odd level, with the convention that the type of the root is $\emptyset$ and the type of a vertex a of the first level is $\overline{a}$.

By construction of T, the type of the children of v depends only on the type of v. Indeed, if v has type $\overline{a}b$, then it has one descendant of each type $a\overline{c}$ such that cb does not belongs to F. On the other hand, if v has type $a\overline{b}$, then it has one descendant of each type $\overline{c}b$ such that ac does not belongs to F. For any two type t and t', write $t' \lessdot t$ if a vertex with type t has a vertex with type t' has a child.

Use this to encode into a finite graph and use Perron-Frobenius to show convergence

Lemma 6. *If the subsequence $(\mu_{n_j})_j$ weakly converges to μ^j, then the weight of the edge between v and w depends only on their types.*

Proof. For a vertex v of level m and type t, let $G_t(n)$ be the number of good vertices of level $n + m$ that are under v. This is depends only on t on not on v or m. For example, $G_t(0) = 0$ if t corresponds to a bad vertex and 1 otherwise. For any type t, we have

$$G_t(n+1) = \sum_{t' \lessdot t} G_{t'}(n).$$

Now, if v is a vertex of type t and of level m, w a child of type t' and $n \geq m+1$, then the measure μ_n gives the weight $G_{t'}(n-m-1)/G_t(n-m)$ to the edge from v to w. This implies that the for all m and all $t' \lessdot t$ (such that there is a vertex of level m and type t), the sequence $G_{t'}(n_j - m - 1)/G_t(n_j - m)$ is convergent. $\square$

Definition 8. A set of forbidden words F is *too restrictif* if there exists i and j in $\mathcal{A}$ such that there is at most one pair of infinite sequence $(ix_2 \dots ; \dots x_{-2}j)$ in $\mathcal{A}^{\mathbf{N}} \times \mathcal{A}^{-\mathbf{N}}$.

Proposition 2. *Suppose that F is not too restrictif and that $(\mu_{n_j})_j$ is a subsequence weakly converging to a measure μ on ∂T_F. Then,* *1.* $\mu(R'_{F,\infty}) = 1$;

2. *the Benjamini-Schramm limit of the* R_{F,n_j} *is* $f_*(\mu)$.

Proof. Firstly, observe that since any ray is a countable intersection of cylinder, which are clopen, if $(\mu_{n_j})_j$ weakly converges to μ, then it also punctually converges to μ.

The set F is not too restrictif if and only if for every vertex v in $T = T_F$ there is at least two infinite rays passing trough v. This is in turn equivalent to the fact that for every vertex v in T_F there is uncountably many rays passing trough v. Therefore, for all n, and all ray ω in ∂T, we have $\mu_n(\omega) = 0$ and thus $\mu(\omega) = 0$. This implies that $\mu(R'_{F,\infty}) = 1$ since $R_{F,\infty} \setminus R'_{F,\infty}$ is a countable union of rays.

Now, we look at

$$A := \{\omega \in \partial T \mid \text{only finitely many vertices on } \omega \text{ are good}\}.$$

It is possible to write $A = \bigcup_{m \geq 0} A_m$ where

$$A_m := \{\omega_1\omega_2 \ldots \mid \omega_m \text{ is good}, \forall i > m, \omega_i \text{ is bad}\}$$

Since the alphabet is finite, there exist a constant $c < 1$ such that for all vertex v in T, the proportion of children of v that are bad is at most c. Hence, for all $n_j > m$ we have $\mu_{n_j}(A_m) \leq c^{n_j - m}$ tends to 0 when n_j tends to infinity. This implies that $\mu(A_m) = 0$ and $\mu(A) = 0$. Need a condition on F for that

Now, let $E := R'_{F,\infty} \setminus A \subseteq \partial T$. This is a measurable subset, with $\mu(E) = 1$. For any finite rooted subgraph α and any positive radius r, let $B := \{\omega \in \partial T = R_{F,\infty} \mid \mathrm{Ball}(\omega, r) \cong \alpha\}$. We have $B = (B \cap E) \cup (B \setminus E)$. Since $\mu(E) = 1$, the subset $B \setminus E$ is measurable and of measure 0. On the other hand, if $\omega = \omega_1\omega_2 \ldots$ is in $B \cap E$, then by Proposition 1 there exists m such that

$$\mathrm{Ball}_{R_{F,m}}(\omega_m, r) \cong \mathrm{Ball}_{R_{F,\infty}}(\omega, r) \cong \alpha$$

By Lemmas 4 and 5, if δ is in $C_{\omega_m} \cap E$, then $\mathrm{Ball}_{R_{F,\infty}}(\omega, r)$ is also isomorphic to $\mathrm{Ball}_{R_{F,m}}(\omega_m, r)$. This implies that $B \cap E$ can be written as the intersection of E with an union of cylinder and is therefore measurable. We conclude that B is measurable and $\mu(B \cap E) = \mu(B) = \mathbf{P}(f_*(\mu), \alpha, r)$.

By Proposition 1, for any ω in E, we have rooted convergence of the graphs along ω.This implies that Rewrite the proof

$$\mathbf{P}(\mu_{n_j}, \alpha, r) \to \mathbf{P}(f_*(\mu), \alpha, r)$$

□

Corollary 1. *Let* $F \subseteq \mathcal{A}^2$. *For* $a \in \mathcal{A}$, *define* $i_a := |\{b \mid ab \in F\}|$ *and* $o_a := |\{a \mid bs \in F\}|$. *Suppose that for all* a *and* b *in* $\mathcal{A}$ *we have* $i_a = i_b$ *and* $o_a = o_b$. *Suppose moreover that all subsequence of* μ_n *admits a weakly converging subsequence. Then*

1. $i_a = o_a$;

2. *The Benjamini-Schramm limit of the Rauzy graphs $R_{F,n}$ is a Dirac measure concentrated on* $\mathrm{DL}(d,d)$*, where* $d = |\mathcal{A}| - i_a$*.*

Proof. Let $k = |\mathcal{A}|$. Then $|F| = k \cdot i_a = o_a \cdot k$ and the first assertion is proven.

It is clear that for any ray (ξ, ω) in $R_{F,\infty}$, the trees $T_{F,\xi}$ and $T_{F,\eta}$ are both $d+1$ regular. Therefore, $\mathrm{DL}(T_{F,\xi}, T_{F,\eta}) \cong \mathrm{DL}(d,d)$.

By assumption, there exists a partition of $\mathbf{N}$ in infinite set $\mathbf{N} = \sqcup X_j$ such that for all j, the sequence $(\mu_n)_{n \in X_j}$ weakly converges to a measure μ^j. But each of this measure, the measure $f_*(\mu^j)$ is the Dirac measure on $\mathrm{DL}(d,d)$. Therefore, the whole sequence $R_{F,n}$ converge to the Dirac measure on $\mathrm{DL}(d,d)$. □

Example 1. Corollary 1 directly applies to De Bruijn graphs, with $F = \emptyset$ and $i_a = o_a = 0$.

On the other hand, there are examples of F satisfying the hypothesis such that for all n, the graph $R_{F,n}$ is not isomorphic to De Bruijn graphs. For example, for $\mathcal{A} = \{0,1,2\}$ and $F = \{01, 12, 20\}$. In this case, $i_a = 1 = o_a$, the limit is $\mathrm{DL}(2,2)$ and $R_{F,n}$ are 4 regular graphs with exactly 3 loops (on $0\dots0$, $1\dots1$ and $2\dots2$). But the only De Bruijn graphs with 3 loops are the $B_{3,n}$ which are 6 regular.