Random subgroups of free groups

Pascal Weil

LaBRI, CNRS et Université de Bordeaux

Joint work with
Frédérique Bassino, Cyril Nicaud – LIPN, LIGM

CALIN, LIPN, Décembre 2017

General context 1/2

- Free group on (finite) alphabet A, $F(A) = F_r$ (if $r = |A|$): the group generated by A with no relations

General context 1/2

- Free group on (finite) alphabet A, $F(A) = F_r$ (if $r = |A|$): the group generated by A with no relations
- Identified with the language of reduced words: a word on alphabet $A \cup A^{-1}$ is reduced if it contains no factor of the form aa^{-1} or $a^{-1}a$

General context 1/2

- Free group on (finite) alphabet A, $F(A) = F_r$ (if $r = |A|$): the group generated by A with no relations
- Identified with the language of reduced words: a word on alphabet $A \cup A^{-1}$ is reduced if it contains no factor of the form aa^{-1} or $a^{-1}a$
- The group operation is $u \cdot v = \mathrm{red}(uv)$, $(uv)^{-1} = v^{-1}u^{-1}$

General context 1/2

- Free group on (finite) alphabet A, $F(A) = F_r$ (if $r = |A|$): the group generated by A with no relations
- Identified with the language of reduced words: a word on alphabet $A \cup A^{-1}$ is reduced if it contains no factor of the form aa^{-1} or $a^{-1}a$
- The group operation is $u \cdot v = \mathrm{red}(uv)$, $(uv)^{-1} = v^{-1}u^{-1}$
- Let $K \leq F(A)$: then K is rational in $F(A)$ if and only if the set of reduced words representing K is rational in $(A \cup A^{-1})^*$ (Benois, 1969)

General context 1/2

- Free group on (finite) alphabet A, $F(A) = F_r$ (if $r = |A|$): the group generated by A with no relations
- Identified with the language of reduced words: a word on alphabet $A \cup A^{-1}$ is reduced if it contains no factor of the form aa^{-1} or $a^{-1}a$
- The group operation is $u \cdot v = \text{red}(uv)$, $(uv)^{-1} = v^{-1}u^{-1}$
- Let $K \leq F(A)$: then K is rational in $F(A)$ if and only if the set of reduced words representing K is rational in $(A \cup A^{-1})^*$ (Benois, 1969)
- A subgroup $H \leq F(A)$ is finitely generated if and only H is rational (Anisimov and Seifert, 1975)

General context 2/2

Study the lattice of finitely generated (*fg*) subgroups of $F(A) = F_r$ (if $r = |A|$), algorithmically and asymptotically

- random generation – if algorithmically efficient: test of conjectures, exploration

General context 2/2

Study the lattice of finitely generated (fg) subgroups of $F(A) = F_r$ (if $r = |A|$), algorithmically and asymptotically

- random generation – if algorithmically efficient: test of conjectures, exploration
- statistical (or asymptotic) properties: evaluation of the frequency of certain properties: genericity, negligibility

General context 2/2

Study the lattice of finitely generated (*fg*) subgroups of $F(A) = F_r$ (if $r = |A|$), algorithmically and asymptotically

- random generation – if algorithmically efficient: test of conjectures, exploration
- statistical (or asymptotic) properties: evaluation of the frequency of certain properties: genericity, negligibility
- Motivations: algorithmic complexity and cryptography + curiosity

General context 2/2

Study the lattice of finitely generated (*fg*) subgroups of $F(A) = F_r$ (if $r = |A|$), algorithmically and asymptotically

- random generation – if algorithmically efficient: test of conjectures, exploration
- statistical (or asymptotic) properties: evaluation of the frequency of certain properties: genericity, negligibility
- Motivations: algorithmic complexity and cryptography + curiosity
- Gromov, Arjantseva, Ol'shanskii, Kapovich, Miasnikov, Schupp, Shpilrain, Ollivier, Jitsukawa, ...

Which distribution on fg subgroups?

- Classical approach: a subgroup is generated by a random tuple of reduced words. A k-tuple (few-generators), or a s_n^d-tuple, where $s_n =$ cardinality of the sphere of radius n and $0 < d < 1$ (Gromov's density model)

Which distribution on fg subgroups?

- Classical approach: a subgroup is generated by a random tuple of reduced words. A k-tuple (few-generators), or a s_n^d-tuple, where $s_n =$ cardinality of the sphere of radius n and $0 < d < 1$ (Gromov's density model)
- Today: a different approach. Every fg subgroup H of $F(A)$ is characterized by a finite A-labeled graph, called the *Stallings graph* of H.

Which distribution on fg subgroups?

- Classical approach: a subgroup is generated by a random tuple of reduced words. A k-tuple (few-generators), or a s_n^d-tuple, where $s_n =$ cardinality of the sphere of radius n and $0 < d < 1$ (Gromov's density model)
- Today: a different approach. Every fg subgroup H of $F(A)$ is characterized by a finite A-labeled graph, called the *Stallings graph* of H.
- This graph is efficiently computable (Touikan), opens the way to countless efficient (and elegant) decision or computation algorithms on fg subgroups. A natural finite discrete structure attached to a subgroup.

Which distribution on fg subgroups?

- Classical approach: a subgroup is generated by a random tuple of reduced words. A k-tuple (few-generators), or a s_n^d-tuple, where $s_n =$ cardinality of the sphere of radius n and $0 < d < 1$ (Gromov's density model)
- Today: a different approach. Every fg subgroup H of $F(A)$ is characterized by a finite A-labeled graph, called the *Stallings graph* of H.
- This graph is efficiently computable (Touikan), opens the way to countless efficient (and elegant) decision or computation algorithms on fg subgroups. A natural finite discrete structure attached to a subgroup.
- The idea: use these graphs to define what a random subgroup is. There are finitely many possible Stallings graphs with n vertices: draw one uniformly at random.

Stallings graph of a finitely generated subgroup

$\Gamma(H)$, the Stallings graph of a finitely generated subgroup H: the *interesting part* of the Schreier graph $\Gamma(G;H)$ – a picture of H and a *unique* representation

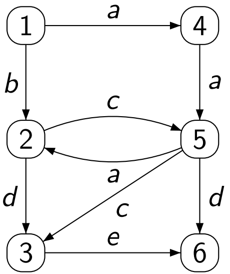
Stallings graph of a finitely generated subgroup

$\Gamma(H)$, the Stallings graph of a finitely generated subgroup H: the *interesting part* of the Schreier graph $\Gamma(G; H)$ – a picture of H and a *unique* representation

$$H = \langle h_1, h_2, h_3, h_4 \rangle$$

$$\begin{aligned} h_1 &= a^3 b^{-1} \\ h_2 &= a^3 c a^{-2} \\ h_3 &= a^2 c d^{-1} b^{-1} \\ h_4 &= a^2 d e^{-1} d^{-1} b^{-1} \end{aligned}$$

$\Gamma(H)$, the Stallings graph of a finitely generated subgroup H: the *interesting part* of the Schreier graph $\Gamma(G; H)$ – a picture of H and a *unique* representation

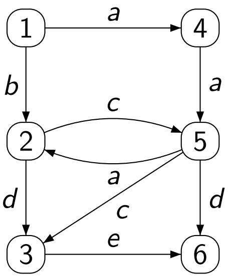


$$H = \langle h_1, h_2, h_3, h_4 \rangle$$

$$h_1 = a^3 b^{-1}$$
$$h_2 = a^3 c a^{-2}$$
$$h_3 = a^2 c d^{-1} b^{-1}$$
$$h_4 = a^2 d e^{-1} d^{-1} b^{-1}$$

Stallings graph of a finitely generated subgroup

$\Gamma(H)$, the Stallings graph of a finitely generated subgroup H: the *interesting part* of the Schreier graph $\Gamma(G;H)$ – a picture of H and a *unique* representation



$H = \langle h_1, h_2, h_3, h_4 \rangle$

rank $= E - V + 1$

conjugation, finite index

intersection of subgroups, malnormality

effective separability

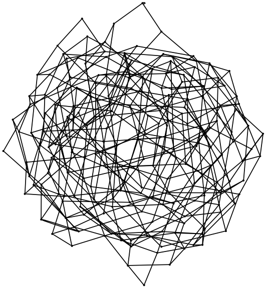
What does a random subgroup look like?

- So: there is a finite number of Stallings graph with n vertices. Draw one uniformly at random to get a random subgroup of size n

What does a random subgroup look like?

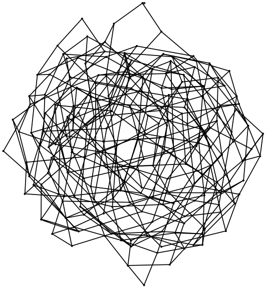
- So: there is a finite number of Stallings graph with n vertices. Draw one uniformly at random to get a random subgroup of size n

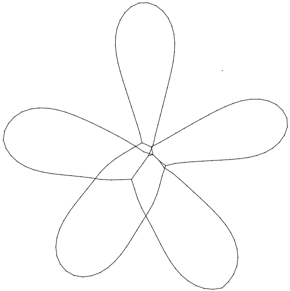
A picture with $n = 200$



What does a random subgroup look like?

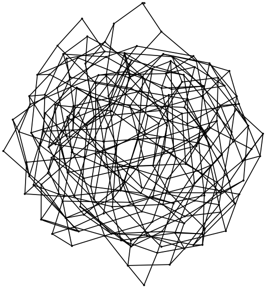
- So: there is a finite number of Stallings graph with n vertices. Draw one uniformly at random to get a random subgroup of size n

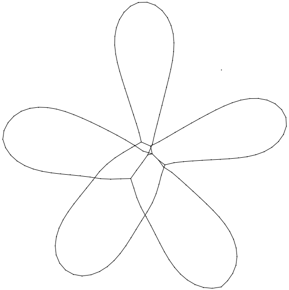




What does a random subgroup look like?

- So: there is a finite number of Stallings graph with n vertices. Draw one uniformly at random to get a random subgroup of size n





- Many more edges, many more cycles in the graph based distribution. Higher rank, lesser probability of malnormality, etc.

How do we investigate random Stallings graphs?

- Characterize Stallings graphs as discrete objects:

How do we investigate random Stallings graphs?

- Characterize Stallings graphs as discrete objects:
- finite graphs with a base vertex

How do we investigate random Stallings graphs?

- Characterize Stallings graphs as discrete objects:
- finite graphs with a base vertex
- connected

How do we investigate random Stallings graphs?

- Characterize Stallings graphs as discrete objects:
- finite graphs with a base vertex
- connected
- with a locally injective A-labeling

How do we investigate random Stallings graphs?

- Characterize Stallings graphs as discrete objects:
- finite graphs with a base vertex
- connected
- with a locally injective A-labeling
- every vertex has valency at least 2, except maybe the base vertex.

How do we investigate random Stallings graphs?

- Characterize Stallings graphs as discrete objects:
- finite graphs with a base vertex
- connected
- with a locally injective A-labeling
- every vertex has valency at least 2, except maybe the base vertex.
- There are many! although estimating that number is non-trivial

General strategy to draw a random Stallings graph

- View Stallings graphs as purely combinatorial objects: a collection $(f_a)_{a \in A}$ of partial injections $[n] \to [n]$, subject to the connectedness and no vertex of valency 1 (global) constraints

General strategy to draw a random Stallings graph

- View Stallings graphs as purely combinatorial objects: a collection $(f_a)_{a \in A}$ of partial injections $[n] \to [n]$, subject to the connectedness and no vertex of valency 1 (global) constraints
- Draw a random partial injection f_a of $[n]$, independently for each letter $a \in A$

General strategy to draw a random Stallings graph

- View Stallings graphs as purely combinatorial objects: a collection $(f_a)_{a \in A}$ of partial injections $[n] \to [n]$, subject to the connectedness and no vertex of valency 1 (global) constraints
- Draw a random partial injection f_a of $[n]$, independently for each letter $a \in A$
- If the $(f_a)_{a \in A}$ do not induce an admissible graph (with base vertex 1), reject and repeat

General strategy to draw a random Stallings graph

- View Stallings graphs as purely combinatorial objects: a collection $(f_a)_{a \in A}$ of partial injections $[n] \to [n]$, subject to the connectedness and no vertex of valency 1 (global) constraints
- Draw a random partial injection f_a of $[n]$, independently for each letter $a \in A$
- If the $(f_a)_{a \in A}$ do not induce an admissible graph (with base vertex 1), reject and repeat
- What needs to be done is explain how one draws random partial injections, and

General strategy to draw a random Stallings graph

- View Stallings graphs as purely combinatorial objects: a collection $(f_a)_{a \in A}$ of partial injections $[n] \to [n]$, subject to the connectedness and no vertex of valency 1 (global) constraints
- Draw a random partial injection f_a of $[n]$, independently for each letter $a \in A$
- If the $(f_a)_{a \in A}$ do not induce an admissible graph (with base vertex 1), reject and repeat
- What needs to be done is explain how one draws random partial injections, and
- to estimate the probability of non-admissibility – we show that it tends to 0 as n tends to infinity

Strategy to draw a random injection

- A size n partial injection (*i.e.*, a partial injection $[n] \to [n]$) is a disjoint union of orbits that are either cycles, or sequences (non-empty)

Strategy to draw a random injection

- A size n partial injection (*i.e.*, a partial injection $[n] \to [n]$) is a disjoint union of orbits that are either cycles, or sequences (non-empty)
- Compute the distribution of sizes of orbits (cycles and sequences), and the distribution of cycles vs. sequences for each size of orbits

Strategy to draw a random injection

- A size n partial injection (*i.e.*, a partial injection $[n] \to [n]$) is a disjoint union of orbits that are either cycles, or sequences (non-empty)
- Compute the distribution of sizes of orbits (cycles and sequences), and the distribution of cycles vs. sequences for each size of orbits
- Draw a size m of an orbit, decide whether it is a cycle or a sequence; and draw another random partial injection of size $n - m$

A versatile tool: exponential generating series (EGS)

- EGS of structures $\mathcal{A}$: $\sum_{n\geq 0} \frac{a_n}{n!} z^n$ if there are a_n structures $\mathcal{A}$ of size n

A versatile tool: exponential generating series (EGS)

- EGS of structures $\mathcal{A}$: $\sum_{n\geq 0} \frac{a_n}{n!} z^n$ if there are a_n structures $\mathcal{A}$ of size n
- Example: for sequences. There are $n!$ sequences of size n. EGS is $\sum_n z^n = \frac{1}{1-z}$

A versatile tool: exponential generating series (EGS)

- EGS of structures $\mathcal{A}$: $\sum_{n\geq 0} \frac{a_n}{n!} z^n$ if there are a_n structures $\mathcal{A}$ of size n
- Example: for sequences. There are $n!$ sequences of size n. EGS is $\sum_n z^n = \frac{1}{1-z}$
- A calculus of EGSs (Flajolet, Sedgewick, etc): if $A(z)$ and $B(z)$ are the EGS for structures $\mathcal{A}$ and $\mathcal{B}$, then

A versatile tool: exponential generating series (EGS)

- EGS of structures $\mathcal{A}$: $\sum_{n\geq 0} \frac{a_n}{n!} z^n$ if there are a_n structures $\mathcal{A}$ of size n
- Example: for sequences. There are $n!$ sequences of size n. EGS is $\sum_n z^n = \frac{1}{1-z}$
- A calculus of EGSs (Flajolet, Sedgewick, etc): if $A(z)$ and $B(z)$ are the EGS for structures $\mathcal{A}$ and $\mathcal{B}$, then
- structures $\mathcal{A}$ or $\mathcal{B}$: $A(z) + B(z)$

A versatile tool: exponential generating series (EGS)

- EGS of structures $\mathcal{A}$: $\sum_{n\geq 0} \frac{a_n}{n!} z^n$ if there are a_n structures $\mathcal{A}$ of size n
- Example: for sequences. There are $n!$ sequences of size n. EGS is $\sum_n z^n = \frac{1}{1-z}$
- A calculus of EGSs (Flajolet, Sedgewick, etc): if $A(z)$ and $B(z)$ are the EGS for structures $\mathcal{A}$ and $\mathcal{B}$, then
- structures $\mathcal{A}$ or $\mathcal{B}$: $A(z) + B(z)$
- sequences of structures $\mathcal{A}$: $1 + A(z) + A^2(z) + \cdots = \frac{1}{1-A(z)}$

A versatile tool: exponential generating series (EGS)

- EGS of structures $\mathcal{A}$: $\sum_{n\geq 0} \frac{a_n}{n!} z^n$ if there are a_n structures $\mathcal{A}$ of size n
- Example: for sequences. There are $n!$ sequences of size n. EGS is $\sum_n z^n = \frac{1}{1-z}$
- A calculus of EGSs (Flajolet, Sedgewick, etc): if $A(z)$ and $B(z)$ are the EGS for structures $\mathcal{A}$ and $\mathcal{B}$, then
- structures $\mathcal{A}$ or $\mathcal{B}$: $A(z) + B(z)$
- sequences of structures $\mathcal{A}$: $1 + A(z) + A^2(z) + \cdots = \frac{1}{1-A(z)}$
- cycles of structures $\mathcal{A}$: $\log\left(\frac{1}{1-A(z)}\right)$

A versatile tool: exponential generating series (EGS)

- EGS of structures $\mathcal{A}$: $\sum_{n\geq 0} \frac{a_n}{n!} z^n$ if there are a_n structures $\mathcal{A}$ of size n
- Example: for sequences. There are $n!$ sequences of size n. EGS is $\sum_n z^n = \frac{1}{1-z}$
- A calculus of EGSs (Flajolet, Sedgewick, etc): if $A(z)$ and $B(z)$ are the EGS for structures $\mathcal{A}$ and $\mathcal{B}$, then
- structures $\mathcal{A}$ or $\mathcal{B}$: $A(z) + B(z)$
- sequences of structures $\mathcal{A}$: $1 + A(z) + A^2(z) + \cdots = \frac{1}{1-A(z)}$
- cycles of structures $\mathcal{A}$: $\log\left(\frac{1}{1-A(z)}\right)$
- sets of structures $\mathcal{A}$: $\exp(A(z))$

Exponential generating series of partial injections

- The EGS for a single point is z. The EGS for sequences is $\frac{1}{1-z}$, and for non-empty sequences $\frac{1}{1-z} - 1 = \frac{z}{1-z}$

Exponential generating series of partial injections

- The EGS for a single point is z. The EGS for sequences is $\frac{1}{1-z}$, and for non-empty sequences $\frac{1}{1-z} - 1 = \frac{z}{1-z}$
- The EGS for cycles is $\log\left(\frac{1}{1-z}\right)$

Exponential generating series of partial injections

- The EGS for a single point is z. The EGS for sequences is $\frac{1}{1-z}$, and for non-empty sequences $\frac{1}{1-z} - 1 = \frac{z}{1-z}$
- The EGS for cycles is $\log\left(\frac{1}{1-z}\right)$
- The EGS for partial injections is $I(z) = \exp\left(\frac{z}{1-z} + \log(\frac{1}{1-z})\right) = \frac{1}{1-z}\exp(\frac{z}{1-z})$

Exponential generating series of partial injections

- The EGS for a single point is z. The EGS for sequences is $\frac{1}{1-z}$, and for non-empty sequences $\frac{1}{1-z} - 1 = \frac{z}{1-z}$
- The EGS for cycles is $\log\left(\frac{1}{1-z}\right)$
- The EGS for partial injections is $I(z) = \exp\left(\frac{z}{1-z} + \log(\frac{1}{1-z})\right) = \frac{1}{1-z}\exp(\frac{z}{1-z})$
- Let $I(z) = \sum_n \frac{I_n}{n!} z^n$. We will be interested in an asymptotic equivalent of the coefficients of $I(z)$

Connectedness is generic

- Partial injections $I(z) = \sum \frac{I_n}{n!} z^n = \frac{1}{1-z} \exp(\frac{z}{1-z})$

Connectedness is generic

- Partial injections $I(z) = \sum \frac{I_n}{n!} z^n = \frac{1}{1-z} \exp(\frac{z}{1-z})$
- r-tuples of partial injections: $1 + J(z)$, with $J(z) = \sum_{n \geq 1} \frac{I_n^r}{n!} z^n$

Connectedness is generic

- Partial injections $I(z) = \sum \frac{I_n}{n!} z^n = \frac{1}{1-z} \exp(\frac{z}{1-z})$
- r-tuples of partial injections: $1 + J(z)$, with $J(z) = \sum_{n \geq 1} \frac{I_n^r}{n!} z^n$
- Let $C(z)$ be the EGS of connected r-tuples: then $1 + J(z) = \exp C(z)$, so $C(z) = \log(1 + J(z)) = \sum_n \frac{C_n}{n!} z^n$

Connectedness is generic

- Partial injections $I(z) = \sum \frac{I_n}{n!} z^n = \frac{1}{1-z} \exp(\frac{z}{1-z})$
- r-tuples of partial injections: $1 + J(z)$, with $J(z) = \sum_{n \geq 1} \frac{I_n^r}{n!} z^n$
- Let $C(z)$ be the EGS of connected r-tuples: then $1 + J(z) = \exp C(z)$, so $C(z) = \log(1 + J(z)) = \sum_n \frac{C_n}{n!} z^n$
- Then $\mathbb{P}(\text{connected}_n) = \frac{C_n}{I_n^r}$.

Connectedness is generic

- Partial injections $I(z) = \sum \frac{I_n}{n!} z^n = \frac{1}{1-z} \exp(\frac{z}{1-z})$
- r-tuples of partial injections: $1 + J(z)$, with $J(z) = \sum_{n \geq 1} \frac{I_n^r}{n!} z^n$
- Let $C(z)$ be the EGS of connected r-tuples: then $1 + J(z) = \exp C(z)$, so $C(z) = \log(1 + J(z)) = \sum_n \frac{C_n}{n!} z^n$
- Then $\mathbb{P}(\mathrm{connected}_n) = \frac{C_n}{I_n^r}$.
- Then… dive into complex analysis!

Connectedness is generic

- Partial injections $I(z) = \sum \frac{I_n}{n!} z^n = \frac{1}{1-z} \exp(\frac{z}{1-z})$
- r-tuples of partial injections: $1 + J(z)$, with $J(z) = \sum_{n\geq 1} \frac{I_n^r}{n!} z^n$
- Let $C(z)$ be the EGS of connected r-tuples: then $1 + J(z) = \exp C(z)$, so $C(z) = \log(1 + J(z)) = \sum_n \frac{C_n}{n!} z^n$
- Then $\mathbb{P}(\text{connected}_n) = \frac{C_n}{I_n^r}$.
- Then... dive into complex analysis!

Use a theorem of Bender (with $F(z, y) = \log(1 + y)$)

Let $F(z, y)$ is a real function, analytic at $(0, 0)$. Let $J(z) = \sum_{n>0} j_n z^n$, $C(z) = \sum_{n>0} c_n z^n$ and $D(z) = \sum_{n>0} d_n z^n$ with $C(z) = F(z, J(z))$ and $D(z) = \frac{\partial F}{\partial y}(z, J(z))$. If $j_{n-1} = o(j_n)$ and there exists such that $\sum_{k=s}^{n-s} |j_k j_{n-k}| = \mathcal{O}(j_{n-s})$, then $c_n = \sum_{k=0}^{s-1} d_k j_{n-k} + \mathcal{O}(j_{n-s})$.

Connectedness is generic

- Partial injections $I(z) = \sum \frac{I_n}{n!} z^n = \frac{1}{1-z} \exp(\frac{z}{1-z})$
- r-tuples of partial injections: $1 + J(z)$, with $J(z) = \sum_{n \geq 1} \frac{I_n^r}{n!} z^n$
- Let $C(z)$ be the EGS of connected r-tuples: then $1 + J(z) = \exp C(z)$, so $C(z) = \log(1 + J(z)) = \sum_n \frac{C_n}{n!} z^n$
- Then $\mathbb{P}(\text{connected}_n) = \frac{C_n}{I_n^r}$.
- Then... dive into complex analysis!
- $\mathbb{P}(\text{connected}_n) = 1 - \frac{2^r}{n^{r-1}} + o(\frac{1}{n^{r-1}})$

Connectedness is generic

- Partial injections $I(z) = \sum \frac{I_n}{n!} z^n = \frac{1}{1-z} \exp(\frac{z}{1-z})$
- r-tuples of partial injections: $1 + J(z)$, with $J(z) = \sum_{n\geq 1} \frac{I_n^r}{n!} z^n$
- Let $C(z)$ be the EGS of connected r-tuples: then $1 + J(z) = \exp C(z)$, so $C(z) = \log(1 + J(z)) = \sum_n \frac{C_n}{n!} z^n$
- Then $\mathbb{P}(\mathrm{connected}_n) = \frac{C_n}{I_n^r}$.
- Then… dive into complex analysis!
- $\mathbb{P}(\mathrm{connected}_n) = 1 - \frac{2^r}{n^{r-1}} + o(\frac{1}{n^{r-1}})$
- Generically, every r-tuple of partial injections is connected

Number of sequences, admissibility 1/2

- For a given partial injection f_a, a point in $[n]$ is either isolated (a sequence of length 1), or an extremity of a sequence, or has arity 2 in the graph of f_a

Number of sequences, admissibility 1/2

- For a given partial injection f_a, a point in $[n]$ is either isolated (a sequence of length 1), or an extremity of a sequence, or has arity 2 in the graph of f_a
- A vertex has arity 1 if it is an extremity for one letter and isolated for all the other letters.

Number of sequences, admissibility 1/2

- For a given partial injection f_a, a point in $[n]$ is either isolated (a sequence of length 1), or an extremity of a sequence, or has arity 2 in the graph of f_a
- A vertex has arity 1 if it is an extremity for one letter and isolated for all the other letters.
- The number of extremities, and of isolated points can be bounded above and under in terms of the number of sequences in the partial injection

Number of sequences, admissibility 1/2

- For a given partial injection f_a, a point in $[n]$ is either isolated (a sequence of length 1), or an extremity of a sequence, or has arity 2 in the graph of f_a
- A vertex has arity 1 if it is an extremity for one letter and isolated for all the other letters.
- The number of extremities, and of isolated points can be bounded above and under in terms of the number of sequences in the partial injection
- So: study the random variable sequence$_n$, which counts the number of sequences in a partial injection: use an analogous calculus for bivariate EGSs, to study $I(z,u) = \sum_{n,k} \frac{I_{n,k}}{n!} z^n u^k$, where $I_{n,k}$ is the number of partial injections of size n with k sequences

Number of sequences, admissibility 2/2

- $$I(z, u) = \frac{1}{1-z} \exp\left(\frac{zu}{1-z}\right)$$

Number of sequences, admissibility 2/2

- $$I(z,u) = \frac{1}{1-z}\exp(\frac{zu}{1-z})$$
- More complex analysis (and more complicated!) shows that $\mathbb{E}(\text{sequence}_n) = \sqrt{n} + o(\sqrt{n})$, with standard deviation $o(\sqrt{n})$

Number of sequences, admissibility 2/2

- $$I(z,u) = \frac{1}{1-z}\exp\left(\frac{zu}{1-z}\right)$$
- More complex analysis (and more complicated!) shows that $\mathbb{E}(\text{sequence}_n) = \sqrt{n} + o(\sqrt{n})$, with standard deviation $o(\sqrt{n})$
- This gives bounds to the expected number of isolated points and extremities, and we use Chebyshev to show that the probability that a vertex has valency 1 is $o(1)$

Number of sequences, admissibility 2/2

- $$I(z,u) = \frac{1}{1-z}\exp(\frac{zu}{1-z})$$
- More complex analysis (and more complicated!) shows that $\mathbb{E}(\text{sequence}_n) = \sqrt{n} + o(\sqrt{n})$, with standard deviation $o(\sqrt{n})$
- This gives bounds to the expected number of isolated points and extremities, and we use Chebyshev to show that the probability that a vertex has valency 1 is $o(1)$
- Generically, every r-tuple of partial injections is admissible

Number of sequences, admissibility 2/2

- $$I(z,u) = \frac{1}{1-z}\exp(\frac{zu}{1-z})$$
- More complex analysis (and more complicated!) shows that $\mathbb{E}(\text{sequence}_n) = \sqrt{n} + o(\sqrt{n})$, with standard deviation $o(\sqrt{n})$
- This gives bounds to the expected number of isolated points and extremities, and we use Chebyshev to show that the probability that a vertex has valency 1 is $o(1)$
- Generically, every r-tuple of partial injections is admissible
- and this justifies the rejection algorithm

Consequences

- Since the number of sequences of f_a has expected value $\sqrt{n}$, the number of a-labeled edge has expected value $n - \sqrt{n}$

Consequences

- Since the number of sequences of f_a has expected value $\sqrt{n}$, the number of a-labeled edge has expected value $n - \sqrt{n}$
- The expected rank of a random subgroup of size n is $E - V + 1$, that is,

$$(|A| - 1)n - |A|\sqrt{n} + 1$$

Consequences

- Since the number of sequences of f_a has expected value $\sqrt{n}$, the number of a-labeled edge has expected value $n - \sqrt{n}$
- The expected rank of a random subgroup of size n is $E - V + 1$, that is,

$$(|A| - 1)n - |A|\sqrt{n} + 1$$

- Also: $\frac{I_n}{n!} \sim \frac{1}{\sqrt{2e\pi}} n^{-\frac{1}{4}} e^{2\sqrt{n}}$ [saddlepoint asymptotics]

Consequences

- Since the number of sequences of f_a has expected value $\sqrt{n}$, the number of a-labeled edge has expected value $n - \sqrt{n}$
- The expected rank of a random subgroup of size n is $E - V + 1$, that is,

$$(|A| - 1)n - |A|\sqrt{n} + 1$$

- Also: $\frac{I_n}{n!} \sim \frac{1}{\sqrt{2e\pi}} n^{-\frac{1}{4}} e^{2\sqrt{n}}$ [saddlepoint asymptotics]
- The number of size n subgroups in F_r is equivalent to

$$n!^{r-1} \frac{n^{1-r/4} e^{2r\sqrt{n}}}{(2\sqrt{e\pi})^r}$$

How to randomly draw a size n partial injection 1/2

- A size n partial injection is a disjoint union of orbits that are either cycles, or sequences: compute the distribution of sizes of orbits (cycles and sequences), and the distribution of cycles vs. sequences for each size of orbits

How to randomly draw a size n partial injection 1/2

- A size n partial injection is a disjoint union of orbits that are either cycles, or sequences: compute the distribution of sizes of orbits (cycles and sequences), and the distribution of cycles vs. sequences for each size of orbits
- Draw at random the size k of an orbit, decide whether it is a cycle or a sequence; and draw another random partial injection of size $n - m$

How to randomly draw a size n partial injection 1/2

- A size n partial injection is a disjoint union of orbits that are either cycles, or sequences: compute the distribution of sizes of orbits (cycles and sequences), and the distribution of cycles vs. sequences for each size of orbits
- Draw at random the size k of an orbit, decide whether it is a cycle or a sequence; and draw another random partial injection of size $n - m$
- More calculus on EGSs: pick at random a component of a random partial injection. Then the probability that this component has size k is $\frac{I_{n-k}}{I_n}(k+1)\frac{(n-1)!}{(n-k)!}$,

How to randomly draw a size n partial injection 1/2

- A size n partial injection is a disjoint union of orbits that are either cycles, or sequences: compute the distribution of sizes of orbits (cycles and sequences), and the distribution of cycles vs. sequences for each size of orbits
- Draw at random the size k of an orbit, decide whether it is a cycle or a sequence; and draw another random partial injection of size $n - m$
- More calculus on EGSs: pick at random a component of a random partial injection. Then the probability that this component has size k is $\frac{I_{n-k}}{I_n}(k+1)\frac{(n-1)!}{(n-k)!}$,
- and the probability that a size k component is a sequence is $\frac{k}{k+1}$

How to randomly draw a size n partial injection 2/2

- How to pick at random a size $k \in [n]$, according to the distribution where $p_k = \frac{I_{n-k}}{I_n}(k+1)\frac{(n-1)!}{(n-k)!}$?

How to randomly draw a size n partial injection 2/2

- How to pick at random a size $k \in [n]$, according to the distribution where $p_k = \frac{I_{n-k}}{I_n}(k+1)\frac{(n-1)!}{(n-k)!}$?
- Requires a pre-computation phase, to compute the I_k $(k \leq n)$.

How to randomly draw a size n partial injection 2/2

- How to pick at random a size $k \in [n]$, according to the distribution where $p_k = \frac{I_{n-k}}{I_n}(k+1)\frac{(n-1)!}{(n-k)!}$?
- Requires a pre-computation phase, to compute the I_k $(k \leq n)$.
- We have $I(z) = \sum_n \frac{I_n}{n!} z^n = \frac{1}{1-z} \exp(\frac{z}{1-z})$ and $I'(z) = \sum_n \frac{I_{n+1}}{n!} z^n$, we find that

$$(1-z)^2 I'(z) = (2-z)I(z) \text{ and}$$

$$I_n = 2nI_{n-1} - (n-1)^2 I_{n-2} \text{ with } I_0 = 1 \text{ and } I_1 = 2.$$

Complexity

- It looks complicated!. . .

Complexity

- It looks complicated!...
- But it is fast

Complexity

- It looks complicated!...
- But it is fast
- In the RAM model, the pre-computation is $\mathcal{O}(n)$ and each random draw is $\mathcal{O}(n)$

Complexity

- It looks complicated!...
- But it is fast
- In the RAM model, the pre-computation is $\mathcal{O}(n)$ and each random draw is $\mathcal{O}(n)$
- In the bit (or logarithmic cost) complexity, I_n requires space and time $\mathcal{O}(n \log n)$. The pre-computation in $\mathcal{O}(n^2 \log n)$ and each random draw is in $\mathcal{O}(n^2 \log^2 n)$

Generic and negligible properties

- H is malnormal if, for each $x \notin H$, $x^{-1}Hx \cap H = 1$. This property is negligible

Generic and negligible properties

- H is malnormal if, for each $x \notin H$, $x^{-1}Hx \cap H = 1$. This property is negligible
- Why? H is not malnormal if there exists $u \neq 1$ and two vertices $x \neq y$ such that u labels a loop at x and at y. This will be the case, for instance, if for some letter, the partial injection f_a has a cycle of length ≥ 2

Generic and negligible properties

- H is malnormal if, for each $x \notin H$, $x^{-1}Hx \cap H = 1$. This property is negligible
- Why? H is not malnormal if there exists $u \neq 1$ and two vertices $x \neq y$ such that u labels a loop at x and at y. This will be the case, for instance, if for some letter, the partial injection f_a has a cycle of length ≥ 2
- With probablility tending to e^{-r}, H contains a conjugate of a letter.

Generic and negligible properties

- H is malnormal if, for each $x \notin H$, $x^{-1}Hx \cap H = 1$. This property is negligible
- Why? H is not malnormal if there exists $u \neq 1$ and two vertices $x \neq y$ such that u labels a loop at x and at y. This will be the case, for instance, if for some letter, the partial injection f_a has a cycle of length ≥ 2
- With probablility tending to e^{-r}, H contains a conjugate of a letter.
- H is *minimal* if for every automorphism φ of $F(A)$, $\varphi(H)$ is not smaller than H (in terms of the number of vertices of its Stallings graph). This is a generic property

Thank you for your attention!