Generic properties of random subgroups of free groups, and random group presentations

Frédérique Bassino, Cyril Nicaud, Pascal Weil

Paris-Nord, Paris-Est, Bordeaux

AMS-EMS-SPM, Porto, June 2015

Generic properties

- Given a tuple $\vec{h} = (h_1, \ldots, h_k)$ of elements of a free group $F(A)$, what can I say about the generic properties

Generic properties

- Given a tuple $\vec{h} = (h_1, \ldots, h_k)$ of elements of a free group $F(A)$, what can I say about the generic properties
  - of the subgroup $\langle h_1, \ldots, h_k \rangle \leq F(A)$?

Generic properties

- Given a tuple $\vec{h} = (h_1, \ldots, h_k)$ of elements of a free group $F(A)$, what can I say about the generic properties
  - of the subgroup $\langle h_1, \ldots, h_k \rangle \leq F(A)$?
  - of the group presentation $\langle A \mid h_1, \ldots, h_k \rangle$ (if the h_i are cyclically reduced)?

Generic properties

- Given a tuple $\vec{h} = (h_1, \ldots, h_k)$ of elements of a free group $F(A)$, what can I say about the generic properties
  - of the subgroup $\langle h_1, \ldots, h_k \rangle \leq F(A)$?
  - of the group presentation $\langle A \mid h_1, \ldots, h_k \rangle$ (if the h_i are cyclically reduced)?
- Depends on how we choose $\vec{h}$: length of the tuple, length of the words, distribution of tuples...

Generic properties

- Given a tuple $\vec{h} = (h_1, \ldots, h_k)$ of elements of a free group $F(A)$, what can I say about the generic properties
  - of the subgroup $\langle h_1, \ldots, h_k \rangle \leq F(A)$?
  - of the group presentation $\langle A \mid h_1, \ldots, h_k \rangle$ (if the h_i are cyclically reduced)?
- Depends on how we choose $\vec{h}$: length of the tuple, length of the words, distribution of tuples...
- On group presentations – Gromov (1987): "almost all groups are hyperbolic", proofs by Ol'shanskii, Champetier, Ollivier. Arzhantseva and Ol'shanskii (1996, 1997). On subgroups – Jutsikawa (2002). Etc. All relative to the uniform distribution.

Generic properties

- Given a tuple $\vec{h} = (h_1, \ldots, h_k)$ of elements of a free group $F(A)$, what can I say about the generic properties
  - of the subgroup $\langle h_1, \ldots, h_k \rangle \leq F(A)$?
  - of the group presentation $\langle A \mid h_1, \ldots, h_k \rangle$ (if the h_i are cyclically reduced)?
- Depends on how we choose $\vec{h}$: length of the tuple, length of the words, distribution of tuples...
- On group presentations – Gromov (1987): "almost all groups are hyperbolic", proofs by Ol'shanskii, Champetier, Ollivier. Arzhantseva and Ol'shanskii (1996, 1997). On subgroups – Jutsikawa (2002). Etc. All relative to the uniform distribution.
- Today, we concentrate on free groups, but there are also results on subgroups of hyperbolic groups, on 1-relator groups, etc. (Gilman, Kapovich, Miasnikov, Ollivier, Osin, Sapir, Schupp, Shpilrain, Špakulová, ...)

The density model (Gromov, Champetier, Ollivier)

- Let C_n be the set of cyclically reduced words of length n. Let d be a real number, $0 < d < 1$

The density model (Gromov, Champetier, Ollivier)

- Let C_n be the set of cyclically reduced words of length n. Let d be a real number, $0 < d < 1$
- Draw uniformly at random a $|C_n|^d$-tuple $\vec{h}$ of cyclically reduced words in C_n and consider $G = \langle A \mid \vec{h} \rangle$;

The density model (Gromov, Champetier, Ollivier)

- Let C_n be the set of cyclically reduced words of length n. Let d be a real number, $0 < d < 1$
- Draw uniformly at random a $|C_n|^d$-tuple $\vec{h}$ of cyclically reduced words in C_n and consider $G = \langle A \mid \vec{h} \rangle$;
- Generically... [that is: with probability tending to 1 when n tends to ∞]

The density model (Gromov, Champetier, Ollivier)

- Let C_n be the set of cyclically reduced words of length n. Let d be a real number, $0 < d < 1$
- Draw uniformly at random a $|C_n|^d$-tuple $\vec{h}$ of cyclically reduced words in C_n and consider $G = \langle A \mid \vec{h} \rangle$;
- Generically... [that is: with probability tending to 1 when n tends to ∞]
- Generically, small cancellation $C'(\lambda)$ holds if $d < \frac{\lambda}{2}$, and does not hold if $d > \frac{\lambda}{2}$; in particular, if $d < \frac{1}{12}$, $C'(\frac{1}{6})$ holds generically and G is hyperbolic

The density model (Gromov, Champetier, Ollivier)

- Let C_n be the set of cyclically reduced words of length n. Let d be a real number, $0 < d < 1$
- Draw uniformly at random a $|C_n|^d$-tuple $\vec{h}$ of cyclically reduced words in C_n and consider $G = \langle A \mid \vec{h} \rangle$;
- Generically... [that is: with probability tending to 1 when n tends to ∞]
- Generically, small cancellation $C'(\lambda)$ holds if $d < \frac{\lambda}{2}$, and does not hold if $d > \frac{\lambda}{2}$; in particular, if $d < \frac{1}{12}$, $C'(\frac{1}{6})$ holds generically and G is hyperbolic
- if $d < \frac{1}{2}$, G is generically infinite and hyperbolic, and if $d > \frac{1}{2}$, G is generically degenerate (trivial or $\mathbb{Z}/2\mathbb{Z}$)

- To discuss the properties of a subgroup of $F(A)$, we rely on a graphical representation:

Stallings graph of a subgroup of a free group

- To discuss the properties of a subgroup of $F(A)$, we rely on a graphical representation:

- $\Gamma(H)$, the Stallings graph of a finitely generated subgroup $H = \langle h_1, \ldots, h_k \rangle$, obtained by drawing loops labeled h_i around a base vertex, and then “folding”

Stallings graph of a subgroup of a free group

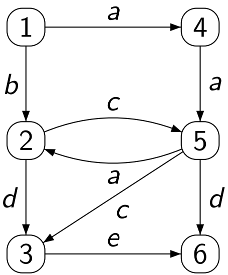
- To discuss the properties of a subgroup of $F(A)$, we rely on a graphical representation:
- $\Gamma(H)$, the Stallings graph of a finitely generated subgroup $H = \langle h_1, \ldots, h_k \rangle$, obtained by drawing loops labeled h_i around a base vertex, and then "folding"

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

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

Stallings graph of a subgroup of a free group

- To discuss the properties of a subgroup of $F(A)$, we rely on a graphical representation:
- $\Gamma(H)$, the Stallings graph of a finitely generated subgroup $H = \langle h_1, \ldots, h_k \rangle$, obtained by drawing loops labeled h_i around a base vertex, and then "folding"



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

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

Properties of subgroups 1/3

- Many properties of $H = \langle \vec{h} \rangle$ are encoded in $\Gamma(H)$: e.g. to find a basis for H, find a spanning tree in $\Gamma(H)$, and use the other edges as basis elements

Properties of subgroups 1/3

- Many properties of $H = \langle \vec{h} \rangle$ are encoded in $\Gamma(H)$: e.g. to find a basis for H, find a spanning tree in $\Gamma(H)$, and use the other edges as basis elements
- The central tree property = small initial cancellation = $\Gamma(\vec{h})$ consists of a central tree, and of one loop for each h_i connecting leaves of the tree

Properties of subgroups 1/3

- Many properties of $H = \langle\vec{h}\rangle$ are encoded in $\Gamma(H)$: e.g. to find a basis for H, find a spanning tree in $\Gamma(H)$, and use the other edges as basis elements
- The central tree property = small initial cancellation = $\Gamma(\vec{h})$ consists of a central tree, and of one loop for each h_i connecting leaves of the tree
- guaranteed if $\mathrm{lcp}(\vec{h}) < \frac{1}{2}\min\vec{h}$, where $\mathrm{lcp}(\vec{h})$ is the length of the least common prefix of the elements of $\vec{h}$ and $\vec{h}^{-1}$ and $\min\vec{h} = \min|h_i|$

Properties of subgroups 1/3

- Many properties of $H = \langle \vec{h} \rangle$ are encoded in $\Gamma(H)$: e.g. to find a basis for H, find a spanning tree in $\Gamma(H)$, and use the other edges as basis elements
- The central tree property = small initial cancellation = $\Gamma(\vec{h})$ consists of a central tree, and of one loop for each h_i connecting leaves of the tree
- guaranteed if $\mathrm{lcp}(\vec{h}) < \frac{1}{2} \min \vec{h}$, where $\mathrm{lcp}(\vec{h})$ is the length of the least common prefix of the elements of $\vec{h}$ and $\vec{h}^{-1}$ and $\min \vec{h} = \min |h_i|$
- If the central tree property holds, then $\vec{h}$ freely generates H

Properties of subgroups 2/3

- H is malnormal if $H^x \cap H = 1$ for every $x \notin H$: equivalently, no word labels a loop at two different vertices of $\Gamma(H)$

Properties of subgroups 2/3

- H is malnormal if $H^x \cap H = 1$ for every $x \notin H$: equivalently, no word labels a loop at two different vertices of $\Gamma(H)$
- Assume that the central tree property holds. Then a sufficient condition for malnormality can be expressed in terms of common factors occurring in the h_i:

Properties of subgroups 2/3

- H is malnormal if $H^x \cap H = 1$ for every $x \notin H$: equivalently, no word labels a loop at two different vertices of $\Gamma(H)$
- Assume that the central tree property holds. Then a sufficient condition for malnormality can be expressed in terms of common factors occurring in the h_i:
- if no word of length $\frac{1}{2}(\min \vec{h} - 2\text{lcp}(\vec{h}))$ occurs twice as a factor of the elements of $\vec{h}$ and $\vec{h}^{-1}$, then $\langle \vec{h} \rangle$ is malnormal

Properties of subgroups 3/3

- [Arzhantseva, Ol'shanskii, Jutsikawa] If we fix k and pick a k-tuple $\vec{h}$ of reduced words of length at most n uniformly at random, then exponentially generically, $\langle \vec{h} \rangle$ satisfies the central tree property and is malnormal

Properties of subgroups 3/3

- [Arzhantseva, Ol'shanskii, Jutsikawa] If we fix k and pick a k-tuple $\vec{h}$ of reduced words of length at most n uniformly at random, then exponentially generically, $\langle \vec{h} \rangle$ satisfies the central tree property and is malnormal
- [Bassino, Nicaud, W] The result actually holds for larger tuples. If we pick a $|B_n|^d$-tuple $\vec{h}$ of reduced words of length at most n uniformly at random, then exponentially generically $\langle \vec{h} \rangle$ has the central tree property if $d < \frac{1}{4}$, and is malnormal if $d < \frac{1}{8}$

Introducing non-uniform distributions 1/3

- On the set of tuples of reduced (or cyclically reduced) words, consider a sequence $(\mathbb{P}_n)_n$ of probability measures

Introducing non-uniform distributions 1/3

- On the set of tuples of reduced (or cyclically reduced) words, consider a sequence $(\mathbb{P}_n)_n$ of probability measures
- A subset $X \subseteq S$ is negligible (resp. generic) if $\lim_n \mathbb{P}_n(X) = 0$ (resp. $= 1$)

Introducing non-uniform distributions 1/3

- On the set of tuples of reduced (or cyclically reduced) words, consider a sequence $(\mathbb{P}_n)_n$ of probability measures
- A subset $X \subseteq S$ is negligible (resp. generic) if $\lim_n \mathbb{P}_n(X) = 0$ (resp. $= 1$)
- $\mathbb{P}_n$ can be the uniform distribution on all k-tuples of reduced words of length at most n, for some fixed integer k, or the uniform distribution on all $|C_n|^d$-tuples of words of length n, for some fixed $0 < d < 1$, as in the above models

Introducing non-uniform distributions 1/3

- On the set of tuples of reduced (or cyclically reduced) words, consider a sequence $(\mathbb{P}_n)_n$ of probability measures
- A subset $X \subseteq S$ is negligible (resp. generic) if $\lim_n \mathbb{P}_n(X) = 0$ (resp. $= 1$)
- $\mathbb{P}_n$ can be the uniform distribution on all k-tuples of reduced words of length at most n, for some fixed integer k, or the uniform distribution on all $|C_n|^d$-tuples of words of length n, for some fixed $0 < d < 1$, as in the above models
- We consider the following generalization. . .

Introducing non-uniform distributions 2/3

- Let $(\mathbb{R}_n)_n$ be a sequence of probability measures, with support the set of reduced words of length n

Introducing non-uniform distributions 2/3

- Let $(\mathbb{R}_n)_n$ be a sequence of probability measures, with support the set of reduced words of length n
- Let $(\mathbb{T}_n)_n$ be a sequence of probability measures on the set of tuples of integers. Then $(\mathbb{R}_n)$ and $(\mathbb{T}_n)$ define a sequence $(\mathbb{P}_n)$ of probability measures on the tuples of words as follows: first draw a tuple of lengths according to $\mathbb{T}_n$, and then, independently for each i in the tuple, draw a word of length i according to $\mathbb{R}_i$

Introducing non-uniform distributions 2/3

- Let $(\mathbb{R}_n)_n$ be a sequence of probability measures, with support the set of reduced words of length n
- Let $(\mathbb{T}_n)_n$ be a sequence of probability measures on the set of tuples of integers. Then $(\mathbb{R}_n)$ and $(\mathbb{T}_n)$ define a sequence $(\mathbb{P}_n)$ of probability measures on the tuples of words as follows: first draw a tuple of lengths according to $\mathbb{T}_n$, and then, independently for each i in the tuple, draw a word of length i according to $\mathbb{R}_i$
- We assume a rigidity property: let $\mathcal{P}(u)$ be the set of words starting with u. $(\mathbb{R}_n)$ is *prefix-heavy* if there exists $C > 0$, $\alpha < 1$ such that, for each n and each reduced word uv of length less than n, $\mathbb{R}_n[\mathcal{P}(uv) \mid \mathcal{P}(u)] \leq C\alpha^{|v|}$

Introducing non-uniform distributions 2/3

- Let $(\mathbb{R}_n)_n$ be a sequence of probability measures, with support the set of reduced words of length n
- Let $(\mathbb{T}_n)_n$ be a sequence of probability measures on the set of tuples of integers. Then $(\mathbb{R}_n)$ and $(\mathbb{T}_n)$ define a sequence $(\mathbb{P}_n)$ of probability measures on the tuples of words as follows: first draw a tuple of lengths according to $\mathbb{T}_n$, and then, independently for each i in the tuple, draw a word of length i according to $\mathbb{R}_i$
- We assume a rigidity property: let $\mathcal{P}(u)$ be the set of words starting with u. $(\mathbb{R}_n)$ is *prefix-heavy* if there exists $C > 0$, $\alpha < 1$ such that, for each n and each reduced word uv of length less than n, $\mathbb{R}_n[\mathcal{P}(uv) \mid \mathcal{P}(u)] \leq C\alpha^{|v|}$
- Examples: the uniform distribution on words of length n, the density model and the few-generator or few-relator models fall under this model

Introducing non-uniform distributions 3/3

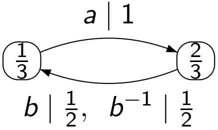
- Let $(\mathbb{R}_n)_n$ be a prefix-heavy sequence of probability measures on reduced words, with parameters (C, α)

Introducing non-uniform distributions 3/3

- Let $(\mathbb{R}_n)_n$ be a prefix-heavy sequence of probability measures on reduced words, with parameters (C, α)
- Let $0 < d < 1$. Draw an α^{-dn}-tuple $\vec{h}$ of reduced words of length at most n according to $(\mathbb{R}_n)_n$. Then, exponentially generically, $\langle \vec{h} \rangle$ has the central tree property if $d < \frac{1}{4}$, and is malnormal if $d < \frac{1}{8}$

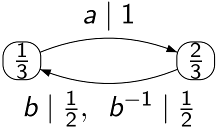
Introducing non-uniform distributions 3/3

- Let $(\mathbb{R}_n)_n$ be a prefix-heavy sequence of probability measures on reduced words, with parameters (C, α)
- Let $0 < d < 1$. Draw an α^{-dn}-tuple $\vec{h}$ of reduced words of length at most n according to $(\mathbb{R}_n)_n$. Then, exponentially generically, $\langle \vec{h} \rangle$ has the central tree property if $d < \frac{1}{4}$, and is malnormal if $d < \frac{1}{8}$
- Suppose in addition that $\liminf \mathbb{R}_n(C_n) > 0$, and let $\lambda < \frac{1}{2}$ and $d < \frac{\lambda}{2}$. Draw an α^{-dn}-tuple $\vec{h}$ of cyclically reduced words of length at most n according to $(\mathbb{R}_n)_n$. Then, exponentially generically, $\vec{h}$ has small cancellation property $C'(\lambda)$

- 

uniform distribution on the set of unique geodesics of $\mathrm{PSL}(2, \mathbb{Z})$: all reduced words on $\{a, b, b^{-1}\}$ without aa, bb or $b^{-1}b^{-1}$

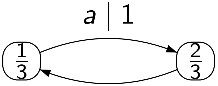
Markovian automata: a natural class of prefix-heavy sequences

- 

  uniform distribution on the set of unique geodesics of $\mathrm{PSL}(2,\mathbb{Z})$: all reduced words on $\{a, b, b^{-1}\}$ without aa, bb or $b^{-1}b^{-1}$

- A finite state automaton, with initial probability distribution on the states, and probabilities on the edges so that the sum of the weight of transitions going out of each vertex is 1

Markovian automata: a natural class of prefix-heavy sequences

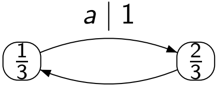


- $b \mid \frac{1}{2}, \quad b^{-1} \mid \frac{1}{2}$

uniform distribution on the set of unique geodesics of $\mathrm{PSL}(2, \mathbb{Z})$: all reduced words on $\{a, b, b^{-1}\}$ without aa, bb or $b^{-1}b^{-1}$

- A finite state automaton, with initial probability distribution on the states, and probabilities on the edges so that the sum of the weight of transitions going out of each vertex is 1
- Yields a probability measure $\mathbb{R}_n$ on the set of words of length n

Markovian automata: a natural class of prefix-heavy sequences

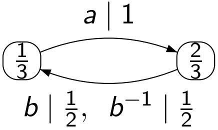


- $b \mid \frac{1}{2}, \;\; b^{-1} \mid \frac{1}{2}$

uniform distribution on the set of unique geodesics of $\mathrm{PSL}(2,\mathbb{Z})$: all reduced words on $\{a, b, b^{-1}\}$ without aa, bb or $b^{-1}b^{-1}$

- A finite state automaton, with initial probability distribution on the states, and probabilities on the edges so that the sum of the weight of transitions going out of each vertex is 1
- Yields a probability measure $\mathbb{R}_n$ on the set of words of length n
- Which is prefix-heavy if there is no loop of weight 1

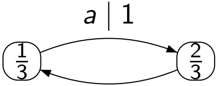
Markovian automata: a natural class of prefix-heavy sequences

- 

uniform distribution on the set of unique geodesics of $\mathrm{PSL}(2, \mathbb{Z})$: all reduced words on $\{a, b, b^{-1}\}$ without aa, bb or $b^{-1}b^{-1}$

- A finite state automaton, with initial probability distribution on the states, and probabilities on the edges so that the sum of the weight of transitions going out of each vertex is 1
- Yields a probability measure $\mathbb{R}_n$ on the set of words of length n
- Which is prefix-heavy if there is no loop of weight 1
- A parameter α ($\mathbb{R}_n[\mathcal{P}(uv) \mid \mathcal{P}(u)] \leq C\alpha^{|v|}$) can be computed from the automaton

Markovian automata: a natural class of prefix-heavy sequences



- $b \mid \frac{1}{2}, \;\; b^{-1} \mid \frac{1}{2}$

uniform distribution on the set of unique geodesics of $\mathrm{PSL}(2, \mathbb{Z})$: all reduced words on $\{a, b, b^{-1}\}$ without aa, bb or $b^{-1}b^{-1}$

- A finite state automaton, with initial probability distribution on the states, and probabilities on the edges so that the sum of the weight of transitions going out of each vertex is 1
- Yields a probability measure $\mathbb{R}_n$ on the set of words of length n
- Which is prefix-heavy if there is no loop of weight 1
- A parameter α ($\mathbb{R}_n[\mathcal{P}(uv) \mid \mathcal{P}(u)] \leq C\alpha^{|v|}$) can be computed from the automaton
- Up to α-density $\frac{1}{4}$ (resp. $\frac{1}{8}$), a tuple of reduced words of length at most n generates a subgroup with the central tree property (resp. a malnormal subgroup)

Irreducible Markovian automata, coincidence probability

- Assume that the Markovian automaton is *irreducible* (i.e. strongly connected)

Irreducible Markovian automata, coincidence probability

- Assume that the Markovian automaton is *irreducible* (i.e. strongly connected)
- Let $\pi_{[2]}$ be the *coincidence probability* of the hidden Markov chain = the modulus of the largest eigenvalue of the matrix of squares of the Markov chain

Irreducible Markovian automata, coincidence probability

- Assume that the Markovian automaton is *irreducible* (i.e. strongly connected)
- Let $\pi_{[2]}$ be the *coincidence probability* of the hidden Markov chain = the modulus of the largest eigenvalue of the matrix of squares of the Markov chain
- If the automaton does not consist of a single cycle, then $(\mathbb{R}_n)_n$ is prefix-heavy with parameter $\alpha = \sqrt{\pi_{[2]}}$

Ergodic Markovian automata, degeneracy at high density

- Now assume that the Markovian automaton is *ergodic*: irreducible + the gcd of the lengths of the loops is 1 + a mild but technical non-triviality hypothesis

Ergodic Markovian automata, degeneracy at high density

- Now assume that the Markovian automaton is *ergodic*: irreducible + the gcd of the lengths of the loops is 1 + a mild but technical non-triviality hypothesis
- Then $\liminf \mathbb{R}_n(C_n) > 0$. Let $\lambda < \frac{1}{2}$. General results show that a $\pi_{[2]}^{-dn}$-tuple of cyclically reduced words exponentially generically satisfies $C'(\lambda)$ if $d < \frac{\lambda}{4}$. We show that this holds up to $d < \frac{\lambda}{2}$

Ergodic Markovian automata, degeneracy at high density

- Now assume that the Markovian automaton is *ergodic*: irreducible + the gcd of the lengths of the loops is 1 + a mild but technical non-triviality hypothesis
- Then $\liminf \mathbb{R}_n(C_n) > 0$. Let $\lambda < \frac{1}{2}$. General results show that a $\pi_{[2]}^{-dn}$-tuple of cyclically reduced words exponentially generically satisfies $C'(\lambda)$ if $d < \frac{\lambda}{4}$. We show that this holds up to $d < \frac{\lambda}{2}$
- if $d > \frac{1}{2}$, then G is generically degenerated: more precisely = the free product of a trivial or 2-element group or $\mathbb{Z}$ (depending on the automaton), with $F(B)$, where B is the set of $a \in A$ such that neither a nor a^{-1} occurs in the automaton

Ergodic Markovian automata, degeneracy at high density

- Now assume that the Markovian automaton is *ergodic*: irreducible + the gcd of the lengths of the loops is 1 + a mild but technical non-triviality hypothesis
- Then $\liminf \mathbb{R}_n(C_n) > 0$. Let $\lambda < \frac{1}{2}$. General results show that a $\pi_{[2]}^{-dn}$-tuple of cyclically reduced words exponentially generically satisfies $C'(\lambda)$ if $d < \frac{\lambda}{4}$. We show that this holds up to $d < \frac{\lambda}{2}$
- if $d > \frac{1}{2}$, then G is generically degenerated: more precisely = the free product of a trivial or 2-element group or $\mathbb{Z}$ (depending on the automaton), with $F(B)$, where B is the set of $a \in A$ such that neither a nor a^{-1} occurs in the automaton
- In the Markovian automaton for the uniform distribution, $\pi_{[2]} = \frac{1}{2|A|-1}$, and we retrieve the classical results

Thank you for your attention!