Lengths on Free groups

Siddhartha Gadgil

Department of Mathematics,
Indian Institute of Science.

January 16, 2020

- On Saturday, December 16, 2017, Terrence Tao posted on his blog a question, from Apoorva Khare.

- On Saturday, December 16, 2017, Terrence Tao posted on his blog a question, from Apoorva Khare.

Question

Is there a homogeneous, (conjugacy invariant) length function on the free group on two generators?

- On Saturday, December 16, 2017, Terrence Tao posted on his blog a question, from Apoorva Khare.

Question

Is there a homogeneous, (conjugacy invariant) length function on the free group on two generators?

- Six days later, this was answered in a collaboration involving several mathematicians (and a computer).

- On Saturday, December 16, 2017, Terrence Tao posted on his blog a question, from Apoorva Khare.

Question

Is there a homogeneous, (conjugacy invariant) length function on the free group on two generators?

- Six days later, this was answered in a collaboration involving several mathematicians (and a computer).
- This the story of the answer and its discovery.

PolyMath 14 Participants

- Tobias Fritz, MPI MIS
- Siddhartha Gadgil, IISc, Bangalore
- Apoorva Khare, IISc, Bangalore
- Pace Nielsen, BYU
- Lior Silberman, UBC
- Terence Tao, UCLA

Outline

1. The Question

Outline

1. The Question
2. Some lengths

Outline

Outline

1. The Question
2. Some lengths
3. The Quest
4. The Theorem and Proof

Outline

1. The Question
2. Some lengths
3. The Quest
4. The Theorem and Proof
5. Computer Bounds and Proofs

Outline

1. The Question
2. Some lengths
3. The Quest
4. The Theorem and Proof
5. Computer Bounds and Proofs
6. Epilogue

The Question

Length functions

- A pseudo-length function on a group G is a function $l : G \to [0, \infty)$ such that

Length functions

- A pseudo-length function on a group G is a function $l : G \to [0, \infty)$ such that
  - $l(e) = 0$, where $e \in G$ is the identity,

Length functions

- A pseudo-length function on a group G is a function $l : G \to [0, \infty)$ such that
  - $l(e) = 0$, where $e \in G$ is the identity,
  - $l(g^{-1}) = l(g)$ for all $g \in G$ (**symmetry**),

Length functions

- A pseudo-length function on a group G is a function $l : G \to [0, \infty)$ such that
  - $l(e) = 0$, where $e \in G$ is the identity,
  - $l(g^{-1}) = l(g)$ for all $g \in G$ (**symmetry**),
  - $l(gh) \leq l(g) + l(h)$ for all $g, h \in G$ (the **triangle inequality**).

Length functions

- A pseudo-length function on a group G is a function $l : G \to [0, \infty)$ such that
  - $l(e) = 0$, where $e \in G$ is the identity,
  - $l(g^{-1}) = l(g)$ for all $g \in G$ (**symmetry**),
  - $l(gh) \leq l(g) + l(h)$ for all $g, h \in G$ (the **triangle inequality**).
- A pseudo-length function l on a group G is said to be a length function if $l(g) > 0$ for all $g \in G \setminus \{e\}$.

Length functions

- A pseudo-length function on a group G is a function $l : G \to [0, \infty)$ such that
  - $l(e) = 0$, where $e \in G$ is the identity,
  - $l(g^{-1}) = l(g)$ for all $g \in G$ (**symmetry**),
  - $l(gh) \leq l(g) + l(h)$ for all $g, h \in G$ (the **triangle inequality**).
- A pseudo-length function l on a group G is said to be a length function if $l(g) > 0$ for all $g \in G \setminus \{e\}$.
- Norms on vector spaces, such as $l(x, y) = \sqrt{x^2 + y^2}$ on $\mathbb{R}^2$, are length functions.

Homogeneity and Conjugacy invariance

- A pseudo-length function l on a group G is said to be homogeneous if $l(g^n) = nl(g)$ for all $g \in G$, $n \in \mathbb{Z}$.

Homogeneity and Conjugacy invariance

- A pseudo-length function l on a group G is said to be homogeneous if $l(g^n) = nl(g)$ for all $g \in G$, $n \in \mathbb{Z}$.
- Norms are homogeneous – indeed Apoorva's question was motivated by generalizing *stochastic inequalities* from Vector spaces with norms.

Homogeneity and Conjugacy invariance

- A pseudo-length function l on a group G is said to be homogeneous if $l(g^n) = nl(g)$ for all $g \in G$, $n \in \mathbb{Z}$.
- Norms are homogeneous – indeed Apoorva's question was motivated by generalizing *stochastic inequalities* from Vector spaces with norms.
- A pseudo-length function l on a group G is said to be conjugacy invariant if $l(ghg^{-1}) = l(h)$ for all $g, h \in G$

Homogeneity and Conjugacy invariance

- A pseudo-length function l on a group G is said to be homogeneous if $l(g^n) = nl(g)$ for all $g \in G$, $n \in \mathbb{Z}$.
- Norms are homogeneous – indeed Apoorva's question was motivated by generalizing *stochastic inequalities* from Vector spaces with norms.
- A pseudo-length function l on a group G is said to be conjugacy invariant if $l(ghg^{-1}) = l(h)$ for all $g, h \in G$

Homogeneity and Conjugacy invariance

- A pseudo-length function l on a group G is said to be homogeneous if $l(g^n) = nl(g)$ for all $g \in G$, $n \in \mathbb{Z}$.
- Norms are homogeneous – indeed Apoorva's question was motivated by generalizing *stochastic inequalities* from Vector spaces with norms.
- A pseudo-length function l on a group G is said to be conjugacy invariant if $l(ghg^{-1}) = l(h)$ for all $g, h \in G$ – if G is abelian every pseudo-length function is conjugacy-invariant.

Lengths and Metrics

- Given a length $l : G \to \mathbb{R}$ on a group G, we can define a metric on G by $d(x, y) = l(x^{-1}y)$.

Lengths and Metrics

- Given a length $l : G \to \mathbb{R}$ on a group G, we can define a metric on G by $d(x, y) = l(x^{-1}y)$.
- This is left-invariant, i.e., $d(gx, gy) = d(x, y)$ for all $g, x, y \in G$.

Lengths and Metrics

- Given a length $l : G \to \mathbb{R}$ on a group G, we can define a metric on G by $d(x, y) = l(x^{-1}y)$.
- This is left-invariant, i.e., $d(gx, gy) = d(x, y)$ for all $g, x, y \in G$.
- Conversely any left invariant metric gives a length $l(g) := d(e, g)$, with $d(x, y) = l(x^{-1}y)$.

Lengths and Metrics

- Given a length $l : G \to \mathbb{R}$ on a group G, we can define a metric on G by $d(x, y) = l(x^{-1}y)$.
- This is left-invariant, i.e., $d(gx, gy) = d(x, y)$ for all $g, x, y \in G$.
- Conversely any left invariant metric gives a length $l(g) := d(e, g)$, with $d(x, y) = l(x^{-1}y)$.
- The metric d associated to l is right-invariant, (i.e., $d(xg, yg) = d(x, y)$ for all $g, x, y \in G$) if and only if l is conjugacy invariant.

The Question

The Question

Question (Apoorva Khare via Terence Tao)

Is there a function $l : \langle \alpha, \beta \rangle \to [0, \infty)$ *on the free group on two generators such that*

The Question

Question (Apoorva Khare via Terence Tao)

Is there a function $l : \langle \alpha, \beta \rangle \to [0, \infty)$ *on the free group on two generators such that*

- $l(g) = 0$ *if and only if* $g = e$ *(positivity).*

The Question

Question (Apoorva Khare via Terence Tao)

Is there a function $l : \langle \alpha, \beta \rangle \to [0, \infty)$ on the free group on two generators such that

- $l(g) = 0$ *if and only if* $g = e$ (*positivity*).
- $l(g^{-1}) = l(g)$ *for all* $g \in \langle \alpha, \beta \rangle$.

The Question

Question (Apoorva Khare via Terence Tao)

Is there a function $l : \langle \alpha, \beta \rangle \to [0, \infty)$ on the free group on two generators such that

- $l(g) = 0$ *if and only if* $g = e$ (*positivity*).
- $l(g^{-1}) = l(g)$ *for all* $g \in \langle \alpha, \beta \rangle$.
- $l(gh) \leq l(g) + l(h)$ *for all* $g, h \in \langle \alpha, \beta \rangle$.

The Question

Question (Apoorva Khare via Terence Tao)

Is there a function $l : \langle \alpha, \beta \rangle \to [0, \infty)$ on the free group on two generators such that

- $l(g) = 0$ *if and only if* $g = e$ *(positivity).*
- $l(g^{-1}) = l(g)$ *for all* $g \in \langle \alpha, \beta \rangle$*.*
- $l(gh) \leq l(g) + l(h)$ *for all* $g, h \in \langle \alpha, \beta \rangle$*.*
- $l(ghg^{-1}) = l(h)$ *for all* $g, h \in \langle \alpha, \beta \rangle$*.*

The Question

Question (Apoorva Khare via Terence Tao)

Is there a function $l : \langle \alpha, \beta \rangle \to [0, \infty)$ on the free group on two generators such that

- $l(g) = 0$ *if and* ***only if*** $g = e$ (***positivity***).
- $l(g^{-1}) = l(g)$ *for all* $g \in \langle \alpha, \beta \rangle$.
- $l(gh) \leq l(g) + l(h)$ *for all* $g, h \in \langle \alpha, \beta \rangle$.
- $l(ghg^{-1}) = l(h)$ *for all* $g, h \in \langle \alpha, \beta \rangle$.
- $l(g^n) = nl(g)$ *for all* $g \in \langle \alpha, \beta \rangle$, $n \in \mathbb{Z}$.

Some lengths

Word length

- The *word length* $l_w(g)$ *of an element* $g \in \langle \alpha, \beta \rangle$ *is the number of letters in the unique reduced word representing* g.

Word length

- The *word length* $l_w(g)$ *of an element* $g \in \langle \alpha, \beta \rangle$ *is the number of letters in the unique reduced word representing* g*.*
- The word length is not conjugacy invariant as $l_w(\alpha\beta\alpha^{-1}) = 3 \neq 1 = l(\beta)$.

Word length

- The *word length* $l_w(g)$ *of an element* $g \in \langle \alpha, \beta \rangle$ *is the number of letters in the unique reduced word representing* g.
- The word length is not conjugacy invariant as $l_w(\alpha\beta\alpha^{-1}) = 3 \neq 1 = l(\beta)$.
- It is also not homogeneous as $l_w((\alpha\beta\alpha^{-1})^2) = l_w(\alpha\beta^2\alpha^{-1}) = 4 \neq 2l_w(\alpha\beta\alpha^{-1})$.

A pullback length

- Consider the abelianization homomorphism $ab : \langle \alpha, \beta \rangle \to \mathbb{Z}^2$.

A pullback length

- Consider the abelianization homomorphism $ab : \langle \alpha, \beta \rangle \to \mathbb{Z}^2$.
- The length $l_{\mathbb{Z}^2}(x, y) = |x| + |y|$ on $\mathbb{Z}^2$ induces a homogeneous, conjugacy-invariant pseudo-length $\bar{l}(g) = l_{\mathbb{Z}^2}(ab(g))$ on $\langle \alpha, \beta \rangle$.

A pullback length

- Consider the abelianization homomorphism $ab : \langle \alpha, \beta \rangle \to \mathbb{Z}^2$.
- The length $l_{\mathbb{Z}^2}(x, y) = |x| + |y|$ on $\mathbb{Z}^2$ induces a homogeneous, conjugacy-invariant pseudo-length $\bar{l}(g) = l_{\mathbb{Z}^2}(ab(g))$ on $\langle \alpha, \beta \rangle$.
- However this is not a length as $ab(\alpha\beta\alpha^{-1}\beta^{-1}) = (0, 0)$, $\bar{l}(\alpha\beta\alpha^{-1}\beta^{-1}) = 0$.

Pullback lengths

- In general, let $\varphi : G \to H$ be a homomorphism and $l_H : H \to [0, \infty)$ is a pseudo-length on H.

Pullback lengths

- In general, let $\varphi : G \to H$ be a homomorphism and $l_H : H \to [0, \infty)$ is a pseudo-length on H.
- We get a pseudo-length l_G on G given by $l_G(g) = l_H(\varphi(g))$.

Pullback lengths

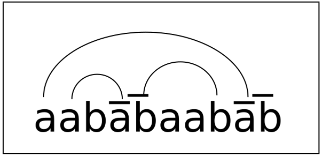
- In general, let $\varphi : G \to H$ be a homomorphism and $l_H : H \to [0, \infty)$ is a pseudo-length on H.
- We get a pseudo-length l_G on G given by $l_G(g) = l_H(\varphi(g))$.
- Homogeneity and conjugacy-invariance are inherited by l_G from l_H.

Pullback lengths

- In general, let $\varphi : G \to H$ be a homomorphism and $l_H : H \to [0, \infty)$ is a pseudo-length on H.
- We get a pseudo-length l_G on G given by $l_G(g) = l_H(\varphi(g))$.
- Homogeneity and conjugacy-invariance are inherited by l_G from l_H.
- But l_G satisfies positivity if and only if $l_H|_{\phi(G)}$ satisfies positivity and φ is injective.

Non-crossing matchings

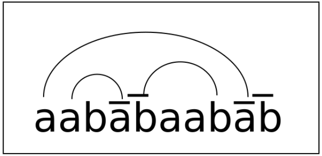
- Consider *non-crossing matchings* for a word in the letters α, β, α^{-1}, and β^{-1};



  - letters can only be matched with their inverses,
  - there are no crossings.

Non-crossing matchings

- Consider *non-crossing matchings* for a word in the letters α, β, α^{-1}, and β^{-1};



  - letters can only be matched with their inverses,
  - there are no crossings.
- The *energy* is the number of unmatched letters.

Watson-Crick length

- For a word w in $\{\alpha, \beta, \alpha^{-1}, \beta^{-1}\}$ consider the *minimum* number of unmatched letters over all non-crossing matchings.

Watson-Crick length

- For a word w in $\{\alpha, \beta, \alpha^{-1}, \beta^{-1}\}$ consider the *minimum* number of unmatched letters over all non-crossing matchings.
- **Proposition:** This depends only on the equivalence class $[w] \in \langle \alpha, \beta \rangle$.

Watson-Crick length

- For a word w in $\{\alpha, \beta, \alpha^{-1}, \beta^{-1}\}$ consider the *minimum* number of unmatched letters over all non-crossing matchings.
- **Proposition:** This depends only on the equivalence class $[w] \in \langle \alpha, \beta \rangle$.
- Hence we have an induced length $l_{WC} : \langle \alpha, \beta \rangle \to [0, \infty)$.

Watson-Crick length

- For a word w in $\{\alpha, \beta, \alpha^{-1}, \beta^{-1}\}$ consider the *minimum* number of unmatched letters over all non-crossing matchings.
- **Proposition:** This depends only on the equivalence class $[w] \in \langle \alpha, \beta \rangle$.
- Hence we have an induced length $l_{WC} : \langle \alpha, \beta \rangle \to [0, \infty)$.
- **Proposition:** The length l_{WC} is conjugacy-invariant.

Watson-Crick length

- **Proposition:** The Watson-Crick length is the *maximal* normalized conjugacy-invariant length, i.e.,

Watson-Crick length

- **Proposition:** The Watson-Crick length is the *maximal* normalized conjugacy-invariant length, i.e.,
  - let $l : \langle \alpha, \beta \rangle \to [0, \infty)$ be any conjugacy-invariant pseudo-length,

Watson-Crick length

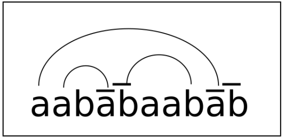
- **Proposition:** The Watson-Crick length is the *maximal* normalized conjugacy-invariant length, i.e.,
  - let $l : \langle \alpha, \beta \rangle \to [0, \infty)$ be any conjugacy-invariant pseudo-length,
  - assume $l(\alpha) \leq 1$ and $l(\beta) \leq 1$, then

Watson-Crick length

- **Proposition:** The Watson-Crick length is the *maximal* normalized conjugacy-invariant length, i.e.,
  - let $l : \langle \alpha, \beta \rangle \to [0, \infty)$ be any conjugacy-invariant pseudo-length,
  - assume $l(\alpha) \leq 1$ and $l(\beta) \leq 1$, then
  - for all $g \in \langle \alpha, \beta \rangle$, $l(g) \leq l_{WC}(g)$.

Watson-Crick length

- **Proposition:** The Watson-Crick length is the *maximal* normalized conjugacy-invariant length, i.e.,
  - let $l : \langle \alpha, \beta \rangle \to [0, \infty)$ be any conjugacy-invariant pseudo-length,
  - assume $l(\alpha) \leq 1$ and $l(\beta) \leq 1$, then
  - for all $g \in \langle \alpha, \beta \rangle$, $l(g) \leq l_{WC}(g)$.
- However l_{WC} is not homogeneous; if $g = \alpha[\alpha, \beta]$, then $l_{WC}(g) = 3$ but $l_{WC}(g^2) = 4$.



The Quest

Some observations

- Groups with torsion have no homogeneous length functions. Namely, if $g^n = e$,

$$l(g) = \frac{l(g^n)}{n} = \frac{l(e)}{n} = 0.$$

Some observations

- Groups with torsion have no homogeneous length functions. Namely, if $g^n = e$,

$$l(g) = \frac{l(g^n)}{n} = \frac{l(e)}{n} = 0.$$

- (Fritz) Homogeneity implies conjugacy invariant.

Some observations

- Groups with torsion have no homogeneous length functions. Namely, if $g^n = e$,

$$l(g) = \frac{l(g^n)}{n} = \frac{l(e)}{n} = 0.$$

- (Fritz) Homogeneity implies conjugacy invariant.
- (Tao, Khare) Homogeneity follows from $l(g^2) \geq 2l(g)$ for all $g \in \langle \alpha, \beta \rangle$.

The Quest

- Over the first 4-5 days after the question was posted,
  - there were many (failed, but instructive) attempts to construct such length functions;

The Quest

- Over the first 4-5 days after the question was posted,
  - there were many (failed, but instructive) attempts to construct such length functions;
  - in particular I focussed on l_{WC} (along with homogenization, Kobayashi construction);

The Quest

- Over the first 4-5 days after the question was posted,
  - there were many (failed, but instructive) attempts to construct such length functions;
  - in particular I focussed on l_{WC} (along with homogenization, Kobayashi construction);
  - the failures of various constructions led to the feeling that $l(\alpha\beta\alpha^{-1}\beta^{-1}) = 0$ for homogeneous pseudo-lengths;

The Quest

- Over the first 4-5 days after the question was posted,
  - there were many (failed, but instructive) attempts to construct such length functions;
  - in particular I focussed on l_{WC} (along with homogenization, Kobayashi construction);
  - the failures of various constructions led to the feeling that $l(\alpha\beta\alpha^{-1}\beta^{-1}) = 0$ for homogeneous pseudo-lengths;
  - increasingly sharp bounds and methods of combining bounds were found, but there was no visible path to proving $l(\alpha\beta\alpha^{-1}\beta^{-1}) = 0$.

The Quest

- Over the first 4-5 days after the question was posted,
  - there were many (failed, but instructive) attempts to construct such length functions;
  - in particular I focussed on l_{WC} (along with homogenization, Kobayashi construction);
  - the failures of various constructions led to the feeling that $l(\alpha\beta\alpha^{-1}\beta^{-1}) = 0$ for homogeneous pseudo-lengths;
  - increasingly sharp bounds and methods of combining bounds were found, but there was no visible path to proving $l(\alpha\beta\alpha^{-1}\beta^{-1}) = 0$.
- On Thursday morning I posted a proof of a computer-assisted bound on $l(\alpha\beta\alpha^{-1}\beta^{-1})$.

Proof which I posted online

Proof which I posted online

Proof of a bound on $l(\alpha\beta\alpha^{-1}\beta^{-1})$ for l a homogeneous, conjugacy invariant length function with $l(\alpha), l(\beta) \leq 1$.

Proof which I posted online

Proof of a bound on $l(\alpha\beta\alpha^{-1}\beta^{-1})$ for l a homogeneous, conjugacy invariant length function with $l(\alpha), l(\beta) \leq 1$.

- $|\bar{a}| \leq 1.0$
- $|\bar{b}\bar{a}b| \leq 1.0$ using $|\bar{a}| \leq 1.0$
- $|\bar{b}| \leq 1.0$
- $|a\bar{b}\bar{a}| \leq 1.0$ using $|\bar{b}| \leq 1.0$
- $|\bar{a}\bar{b}ab\bar{a}\bar{b}| \leq 2.0$ using $|\bar{a}\bar{b}a| \leq 1.0$ and $|b\bar{a}\bar{b}| \leq 1.0$
- ... (119 lines)
- $|ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}| \leq 13.859649122807017$ using $|ab\bar{a}| \leq 1.0$ and $|\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}| \leq 12.859649122807017$
- $|ab\bar{a}\bar{b}| \leq 0.8152734778121775$ using $|ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}| \leq 13.859649122807017$ by taking 17th power.

Proof which I posted online

Proof of a bound on $l(\alpha\beta\alpha^{-1}\beta^{-1})$ for l a homogeneous, conjugacy invariant length function with $l(\alpha), l(\beta) \leq 1$.

- $|\bar{a}| \leq 1.0$
- $|\bar{b}\bar{a}b| \leq 1.0$ using $|\bar{a}| \leq 1.0$
- $|\bar{b}| \leq 1.0$
- $|a\bar{b}\bar{a}| \leq 1.0$ using $|\bar{b}| \leq 1.0$
- $|\bar{a}\bar{b}ab\bar{a}\bar{b}| \leq 2.0$ using $|\bar{a}\bar{b}a| \leq 1.0$ and $|b\bar{a}\bar{b}| \leq 1.0$
- ... (119 lines)
- $|ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}| \leq 13.859649122807017$ using $|ab\bar{a}| \leq 1.0$ and $|\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}| \leq 12.859649122807017$
- $|ab\bar{a}\bar{b}| \leq 0.8152734778121775$ using $|ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}| \leq 13.859649122807017$ by taking 17th power.

i.e., $l(\alpha, \beta) \leq 0.8152734778121775$

- The computer-generated proof was studied by Pace Nielsen, who extracted the internal repetition trick.

- The computer-generated proof was studied by Pace Nielsen, who extracted the internal repetition trick.
- This was extended by Pace Nielsen and Tobias Fritz and generalized by Terence Tao.

- The computer-generated proof was studied by Pace Nielsen, who extracted the internal repetition trick.
- This was extended by Pace Nielsen and Tobias Fritz and generalized by Terence Tao.
- From this Fritz obtained the key lemma:

- The computer-generated proof was studied by Pace Nielsen, who extracted the internal repetition trick.
- This was extended by Pace Nielsen and Tobias Fritz and generalized by Terence Tao.
- From this Fritz obtained the key lemma:

Lemma

Let $f(m,k) = l(x^m[x,y]^k)$. *Then*

$$f(m,k) \leq \frac{f(m-1,k) + f(m+1,k-1)}{2}.$$

- The computer-generated proof was studied by Pace Nielsen, who extracted the internal repetition trick.
- This was extended by Pace Nielsen and Tobias Fritz and generalized by Terence Tao.
- From this Fritz obtained the key lemma:

Lemma

Let $f(m,k) = l(x^m[x,y]^k)$. Then

$$f(m,k) \leq \frac{f(m-1,k) + f(m+1,k-1)}{2}.$$

- Using Probability, Tao showed $l(\alpha\beta\alpha^{-1}\beta^{-1}) = 0$.

The Theorem and Proof

The main results

Theorem

For any group G, every homogeneous pseudo-length $l : G \to \mathbb{R}$ is the pullback of a homogeneous pseudo-length on the abelianization $G/[G, G]$.

The main results

Theorem

For any group G, every homogeneous pseudo-length $l : G \to \mathbb{R}$ is the pullback of a homogeneous pseudo-length on the abelianization $G/[G, G]$.

Corollary

If G is not abelian (e.g. $G = \mathbb{F}_2$) there is no homogeneous length function on G.

Internal Repetition trick

Lemma

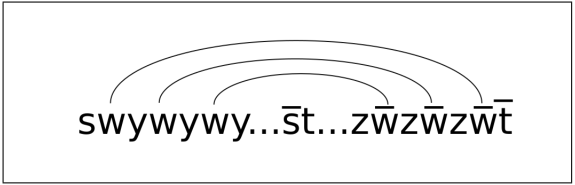
If $x = s(wy)s^{-1} = t(zw^{-1})t^{-1}$, we have $l(x) \leq \frac{l(y)+l(z)}{2}$.

Internal Repetition trick

Lemma

If $x = s(wy)s^{-1} = t(zw^{-1})t^{-1}$, we have $l(x) \leq \frac{l(y)+l(z)}{2}$.

- $$\begin{aligned} l(x^n x^n) &= l(s(wy)^n s^{-1} t(zw^{-1})^n t^{-1}) \\ &\leq n(l(y) + l(z)) + 2(l(s) + l(t)) \end{aligned}$$

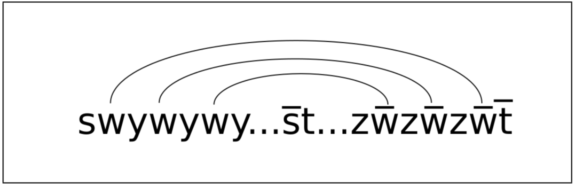


Internal Repetition trick

Lemma

If $x = s(wy)s^{-1} = t(zw^{-1})t^{-1}$*, we have* $l(x) \leq \frac{l(y)+l(z)}{2}$.

- $$\begin{aligned} l(x^n x^n) &= l(s(wy)^n s^{-1} t(zw^{-1})^n t^{-1}) \\ &\leq n(l(y)+l(z)) + 2(l(s)+l(t)) \end{aligned}$$



- Use $l(x) = \frac{l(x^n x^n)}{2n}$ and take limits.

The key inequality

- The above lemma says that if $x \sim wy$ and $x \sim zw^{-1}$, then $l(x) \leq \frac{l(y)+l(z)}{2}$.

The key inequality

- The above lemma says that if $x \sim wy$ and $x \sim zw^{-1}$, then $l(x) \leq \frac{l(y)+l(z)}{2}$.
- We can now deduce $f(m,k) \leq \frac{f(m-1,k)+f(m+1,k-1)}{2}$.

The key inequality

- The above lemma says that if $x \sim wy$ and $x \sim zw^{-1}$, then $l(x) \leq \frac{l(y)+l(z)}{2}$.
- We can now deduce $f(m,k) \leq \frac{f(m-1,k)+f(m+1,k-1)}{2}$.
- Namely, observe that $x^m[x,y]^k$ is conjugate to both $x(x^{m-1}[x,y]^k)$ and $(y^{-1}x^m[x,y]^{k-1}xy)x^{-1}$.

The key inequality

- The above lemma says that if $x \sim wy$ and $x \sim zw^{-1}$, then $l(x) \leq \frac{l(y)+l(z)}{2}$.
- We can now deduce $f(m,k) \leq \frac{f(m-1,k)+f(m+1,k-1)}{2}$.
- Namely, observe that $x^m[x,y]^k$ is conjugate to both $x(x^{m-1}[x,y]^k)$ and $(y^{-1}x^m[x,y]^{k-1}xy)x^{-1}$.
- Hence $l(x^m[x,y]^k) \leq \frac{l(x^{m-1}[x,y]^k)+l(y^{-1}x^m[x,y]^{k-1}xy)}{2}$.

The key inequality

- The above lemma says that if $x \sim wy$ and $x \sim zw^{-1}$, then $l(x) \leq \frac{l(y)+l(z)}{2}$.
- We can now deduce $f(m,k) \leq \frac{f(m-1,k)+f(m+1,k-1)}{2}$.
- Namely, observe that $x^m[x,y]^k$ is conjugate to both $x(x^{m-1}[x,y]^k)$ and $(y^{-1}x^m[x,y]^{k-1}xy)x^{-1}$.
- Hence $l(x^m[x,y]^k) \leq \frac{l(x^{m-1}[x,y]^k)+l(y^{-1}x^m[x,y]^{k-1}xy)}{2}$.
- Since $y^{-1}x^m[x,y]^{k-1}xy$ is conjugate to $x^{m+1}[x,y]^{k-1}$, the claim follows.

Tao's probability theory argument

- The inequality $f(m, k) \leq \frac{f(m-1,k)+f(m+1,k-1)}{2}$ can be interpreted as the average of f being non-decreasing along the random walk on $\mathbb{Z}^2$ where we move by $(-1, 0)$ or $(1, -1)$ with equal probability.

Tao's probability theory argument

- The inequality $f(m,k) \leq \frac{f(m-1,k)+f(m+1,k-1)}{2}$ can be interpreted as the average of f being non-decreasing along the random walk on $\mathbb{Z}^2$ where we move by $(-1,0)$ or $(1,-1)$ with equal probability.
- The average displacement of a step is $(0,-1/2)$.

Tao's probability theory argument

- The inequality $f(m,k) \leq \frac{f(m-1,k)+f(m+1,k-1)}{2}$ can be interpreted as the average of f being non-decreasing along the random walk on $\mathbb{Z}^2$ where we move by $(-1,0)$ or $(1,-1)$ with equal probability.
- The average displacement of a step is $(0,-1/2)$.
- Hence taking $2n$ steps starting at $(0,n)$ gives an upper bound for $f(0,2n) = l((\alpha\beta\alpha^{-1}\beta^{-1})^n)$ by the average length for a distribution centered at the origin.

Tao's probability theory argument

- The inequality $f(m,k) \leq \frac{f(m-1,k)+f(m+1,k-1)}{2}$ can be interpreted as the average of f being non-decreasing along the random walk on $\mathbb{Z}^2$ where we move by $(-1,0)$ or $(1,-1)$ with equal probability.
- The average displacement of a step is $(0,-1/2)$.
- Hence taking $2n$ steps starting at $(0,n)$ gives an upper bound for $f(0,2n) = l((\alpha\beta\alpha^{-1}\beta^{-1})^n)$ by the average length for a distribution centered at the origin.
- This was bounded using the Chebyshev inequality.

Computer Bounds and Proofs

Bounds from Conjugacy invariance

- Fix a conjugacy-invariant, normalized length function $l : \langle \alpha, \beta \rangle \to \mathbb{R}$, i.e. with $l(\alpha), l(\beta) \leq 1$.

Bounds from Conjugacy invariance

- Fix a conjugacy-invariant, normalized length function $l : \langle \alpha, \beta \rangle \to \mathbb{R}$, i.e. with $l(\alpha), l(\beta) \leq 1$.
- Let $g = \xi_1 \xi_2 \dots \xi_n$ with $n \geq 1$.

Bounds from Conjugacy invariance

- Fix a conjugacy-invariant, normalized length function $l : \langle \alpha, \beta \rangle \to \mathbb{R}$, i.e. with $l(\alpha), l(\beta) \leq 1$.
- Let $g = \xi_1 \xi_2 \dots \xi_n$ with $n \geq 1$.
  - By the triangle inequality

$$l(g) \leq 1 + l(\xi_2 \xi_3 \dots \xi_n).$$

Bounds from Conjugacy invariance

- Fix a conjugacy-invariant, normalized length function $l : \langle \alpha, \beta \rangle \to \mathbb{R}$, i.e. with $l(\alpha), l(\beta) \leq 1$.
- Let $g = \xi_1\xi_2 \dots \xi_n$ with $n \geq 1$.
  - By the triangle inequality
  $$l(g) \leq 1 + l(\xi_2\xi_3 \dots \xi_n).$$
  - If $\xi_k = \xi_1^{-1}$, by the triangle inequality and conjugacy invariance
  $$l(g) \leq l(\xi_2\xi_3 \dots \xi_{k-1}) + l(\xi_{k+1}\xi_{k+2} \dots \xi_n)$$
  as $l(\xi_1\xi_2 \dots \xi_k) = l(\xi_1\xi_2 \dots \xi_{k-1}\xi_1^{-1}) = l(\xi_2\xi_2 \dots \xi_{k-1})$.

The recursive algorithm

For $g \in F$, compute $L(g)$ such that $l(g) \leq L(g)$ by:

- If $g = e$ is the empty word, define $L(g) := 0$.

The recursive algorithm

For $g \in F$, compute $L(g)$ such that $l(g) \leq L(g)$ by:

- If $g = e$ is the empty word, define $L(g) := 0$.
- If $g = \xi_1$ has exactly one letter, define $L(g) := 1$.

The recursive algorithm

For $g \in F$, compute $L(g)$ such that $l(g) \leq L(g)$ by:

- If $g = e$ is the empty word, define $L(g) := 0$.
- If $g = \xi_1$ has exactly one letter, define $L(g) := 1$.
- If $g = \xi_1 \xi_2 \ldots \xi_n$ has at least two letters:

The recursive algorithm

For $g \in F$, compute $L(g)$ such that $l(g) \leq L(g)$ by:

- If $g = e$ is the empty word, define $L(g) := 0$.
- If $g = \xi_1$ has exactly one letter, define $L(g) := 1$.
- If $g = \xi_1\xi_2 \dots \xi_n$ has at least two letters:
  - let $\lambda_0 = 1 + L(\xi_2\xi_3 \dots \xi_n)$ (computed recursively).

The recursive algorithm

For $g \in F$, compute $L(g)$ such that $l(g) \leq L(g)$ by:

- If $g = e$ is the empty word, define $L(g) := 0$.
- If $g = \xi_1$ has exactly one letter, define $L(g) := 1$.
- If $g = \xi_1\xi_2 \dots \xi_n$ has at least two letters:
  - let $\lambda_0 = 1 + L(\xi_2\xi_3 \dots \xi_n)$ (computed recursively).
  - let Λ be the (possibly empty) set

$$\{L(\xi_2\xi_3 \dots \xi_{k-1}) + L(\xi_{k+1}\xi_{k+2} \dots \xi_n) : 2 \leq k \leq n, \xi_k = \xi_1^{-1}\}$$

The recursive algorithm

For $g \in F$, compute $L(g)$ such that $l(g) \leq L(g)$ by:

- If $g = e$ is the empty word, define $L(g) := 0$.
- If $g = \xi_1$ has exactly one letter, define $L(g) := 1$.
- If $g = \xi_1 \xi_2 \dots \xi_n$ has at least two letters:
  - let $\lambda_0 = 1 + L(\xi_2 \xi_3 \dots \xi_n)$ (computed recursively).
  - let Λ be the (possibly empty) set

    $$\{L(\xi_2 \xi_3 \dots \xi_{k-1}) + L(\xi_{k+1} \xi_{k+2} \dots \xi_n) : 2 \leq k \leq n, \xi_k = \xi_1^{-1}\}$$
  - define $L(g) := \min(\{\lambda_0\} \cup \Lambda)$.

The recursive algorithm

For $g \in F$, compute $L(g)$ such that $l(g) \leq L(g)$ by:

- If $g = e$ is the empty word, define $L(g) := 0$.
- If $g = \xi_1$ has exactly one letter, define $L(g) := 1$.
- If $g = \xi_1 \xi_2 \dots \xi_n$ has at least two letters:
  - let $\lambda_0 = 1 + L(\xi_2 \xi_3 \dots \xi_n)$ (computed recursively).
  - let Λ be the (possibly empty) set
    $$\{L(\xi_2 \xi_3 \dots \xi_{k-1}) + L(\xi_{k+1} \xi_{k+2} \dots \xi_n) : 2 \leq k \leq n, \xi_k = \xi_1^{-1}\}$$
  - define $L(g) := \min(\{\lambda_0\} \cup \Lambda)$.
- In general, we may also have elements g for which $L(g)$ is given (or previously computed).

Ad hoc bounds using Homogeneity

- For chosen $g \in \langle \alpha, \beta \rangle$, $n \geq 1$, homogeneity gives $l(g) \leq L(g^n)/n$ for l a normalized, homogeneous length function on $\langle \alpha, \beta \rangle$.

Ad hoc bounds using Homogeneity

- For chosen $g \in \langle \alpha, \beta \rangle$, $n \geq 1$, homogeneity gives $l(g) \leq L(g^n)/n$ for l a normalized, homogeneous length function on $\langle \alpha, \beta \rangle$.
- Further, we can use this (in general improved) bound (as $L(g)$) recursively in the above algorithm.

Ad hoc bounds using Homogeneity

- For chosen $g \in \langle \alpha, \beta \rangle$, $n \geq 1$, homogeneity gives $l(g) \leq L(g^n)/n$ for l a normalized, homogeneous length function on $\langle \alpha, \beta \rangle$.
- Further, we can use this (in general improved) bound (as $L(g)$) recursively in the above algorithm.
- We computed such bounds in interactive sessions.

Ad hoc bounds using Homogeneity

- For chosen $g \in \langle \alpha, \beta \rangle$, $n \geq 1$, homogeneity gives $l(g) \leq L(g^n)/n$ for l a normalized, homogeneous length function on $\langle \alpha, \beta \rangle$.
- Further, we can use this (in general improved) bound (as $L(g)$) recursively in the above algorithm.
- We computed such bounds in interactive sessions.
- The words used were $\alpha[\alpha, \beta]^k$, chosen based on non-homogeneity of the function l_{WC}.

From Bounds to Proofs

- Rather than (recursively) generating just bounds, we can recursively generate proofs of bounds.

From Bounds to Proofs

- Rather than (recursively) generating just bounds, we can recursively generate proofs of bounds.
- These were in terms of domain specific foundations, embeddable in Homotopy Type Theory.

From Bounds to Proofs

- Rather than (recursively) generating just bounds, we can recursively generate proofs of bounds.
- These were in terms of domain specific foundations, embeddable in Homotopy Type Theory.
- Objects of mathematics, meta-mathematics and algorithms/programs are all first-class, e.g., proofs could be arguments and values of functions.

From Bounds to Proofs

- Rather than (recursively) generating just bounds, we can recursively generate proofs of bounds.
- These were in terms of domain specific foundations, embeddable in Homotopy Type Theory.
- Objects of mathematics, meta-mathematics and algorithms/programs are all first-class, e.g., proofs could be arguments and values of functions.
- In this case, we can instead view our algorithm as just keeping track of relevant inequalities.

Epilogue

Quasification

- The function $l : G \to [0, \infty)$ is a quasi-pseudo-length function if there exists $c \in \mathbb{R}$ such that $l(gh) \leq l(g) + l(h) + c$, for all $g, h \in G$.

Quasification

- The function $l : G \to [0, \infty)$ is a quasi-pseudo-length function if there exists $c \in \mathbb{R}$ such that $l(gh) \leq l(g) + l(h) + c$, for all $g, h \in G$.
- We see that for a homogeneous quasi-pseudo-length function, $l([x, y]) \leq 4c$ for all $x, y \in G$.

Quasification

- The function $l : G \to [0, \infty)$ is a quasi-pseudo-length function if there exists $c \in \mathbb{R}$ such that $l(gh) \leq l(g) + l(h) + c$, for all $g, h \in G$.
- We see that for a homogeneous quasi-pseudo-length function, $l([x, y]) \leq 4c$ for all $x, y \in G$.
- For a group with vanishing stable commutator length, e.g. $G = Sl(3, \mathbb{Z})$, any homogeneous quasi-pseudo-length function is bounded distance from a pullback from $G/[G, G]$.

Questions about the computer proof

- How much did the proof depend on expert knowledge?

Questions about the computer proof

- How much did the proof depend on expert knowledge?
- Was finding the proof a fluke or was it likely to be found? How much trial and error was needed?

Questions about the computer proof

- How much did the proof depend on expert knowledge?
- Was finding the proof a fluke or was it likely to be found? How much trial and error was needed?
- What about rounding off errors?

Questions about the computer proof

- How much did the proof depend on expert knowledge?
- Was finding the proof a fluke or was it likely to be found? How much trial and error was needed?
- What about rounding off errors?
- Are there any general lessons for finding computer proofs, especially without expert knowledge?

Questions about the computer proof

- How much did the proof depend on expert knowledge?
- Was finding the proof a fluke or was it likely to be found? How much trial and error was needed?
- What about rounding off errors?
- Are there any general lessons for finding computer proofs, especially without expert knowledge?
- We can use the *families* $g_k = \alpha[\alpha, \beta]^k$, $k = 1, 2, 6$ and use $l(g_k) \leq \frac{l(g_k)^n}{n}$ with $n = 1, 2, \ldots, 20$.

Afterword

- This work became PolyMath 14, and has been published in *Algebra & Number Theory* (with an account of the computer proving published in the *Journal of Automated Reasoning*).

Afterword

- This work became PolyMath 14, and has been published in *Algebra & Number Theory* (with an account of the computer proving published in the *Journal of Automated Reasoning*).
- A computer generated but human readable proof was read, understood, generalized and abstracted by mathematicians to obtain the key lemma in an interesting mathematical result;

Afterword

- This work became PolyMath 14, and has been published in *Algebra & Number Theory* (with an account of the computer proving published in the *Journal of Automated Reasoning*).
- A computer generated but human readable proof was read, understood, generalized and abstracted by mathematicians to obtain the key lemma in an interesting mathematical result;

Afterword

- This work became PolyMath 14, and has been published in *Algebra & Number Theory* (with an account of the computer proving published in the *Journal of Automated Reasoning*).
- A computer generated but human readable proof was read, understood, generalized and abstracted by mathematicians to obtain the key lemma in an interesting mathematical result; this is perhaps the first time this has happened.