Homogeneous length functions on Groups

A PolyMath adventure

Siddhartha Gadgil

Department of Mathematics,
Indian Institute of Science.

January 31, 2019

- 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. The Quest

Outline

1. The Question
2. The Quest
3. Computer Bounds and Proofs

Outline

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

Outline

The Question

Groups

- A Group G is a set together with

Groups

- A Group G is a set together with
  - an associative binary operation $G \times G \to G$,

Groups

- A Group G is a set together with
  - an associative binary operation $G \times G \to G$,
  - an identity e such that $g \cdot e = e \cdot g = g$ for all $g \in G$,

Groups

- A Group G is a set together with
  - an associative binary operation $G \times G \to G$,
  - an identity e such that $g \cdot e = e \cdot g = g$ for all $g \in G$,
  - an inverse function $g \mapsto g^{-1}$ such that $g \cdot g^{-1} = g^{-1} \cdot g = e$ for all $g \in G$.

Groups

- A Group G is a set together with
  - an associative binary operation $G \times G \to G$,
  - an identity e such that $g \cdot e = e \cdot g = g$ for all $g \in G$,
  - an inverse function $g \mapsto g^{-1}$ such that $g \cdot g^{-1} = g^{-1} \cdot g = e$ for all $g \in G$.
- Integers $\mathbb{Z}$ with the addition operation form a group.

Groups

- A Group G is a set together with
  - an associative binary operation $G \times G \to G$,
  - an identity e such that $g \cdot e = e \cdot g = g$ for all $g \in G$,
  - an inverse function $g \mapsto g^{-1}$ such that $g \cdot g^{-1} = g^{-1} \cdot g = e$ for all $g \in G$.
- Integers $\mathbb{Z}$ with the addition operation form a group.
- Pairs of real numbers with componentwise addition form the group $\mathbb{R}^2$.

Groups

- A Group G is a set together with
  - an associative binary operation $G \times G \to G$,
  - an identity e such that $g \cdot e = e \cdot g = g$ for all $g \in G$,
  - an inverse function $g \mapsto g^{-1}$ such that $g \cdot g^{-1} = g^{-1} \cdot g = e$ for all $g \in G$.
- Integers $\mathbb{Z}$ with the addition operation form a group.
- Pairs of real numbers with componentwise addition form the group $\mathbb{R}^2$.
- For $n \geq 1$, $n \times n$ real matrices with determinant 1 form a group (called $Sl(n, \mathbb{R})$).

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 $g, h \in G$.
- If G is abelian ($gh = hg$ for all $g, h \in G$) this holds.

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 Free Group $\langle \alpha, \beta \rangle$

- Consider words in $S = \{\alpha, \beta, \alpha^{-1}, \beta^{-1}\}$, where we think of α^{-1} and β^{-1} as simply formal symbols.

The Free Group $\langle \alpha, \beta \rangle$

- Consider words in $S = \{\alpha, \beta, \alpha^{-1}, \beta^{-1}\}$, where we think of α^{-1} and β^{-1} as simply formal symbols.
- We regard two words as equal if they are related by a sequence of moves given by cancellation of pairs of adjacent letters that are inverses of each other.

The Free Group $\langle \alpha, \beta \rangle$

- Consider words in $S = \{\alpha, \beta, \alpha^{-1}, \beta^{-1}\}$, where we think of α^{-1} and β^{-1} as simply formal symbols.
- We regard two words as equal if they are related by a sequence of moves given by cancellation of pairs of adjacent letters that are inverses of each other.
- For example, $\alpha\beta\beta^{-1}\alpha\beta\alpha^{-1} = \alpha\alpha\beta\alpha^{-1}$.

The Free Group $\langle \alpha, \beta \rangle$

- Consider words in $S = \{\alpha, \beta, \alpha^{-1}, \beta^{-1}\}$, where we think of α^{-1} and β^{-1} as simply formal symbols.
- We regard two words as equal if they are related by a sequence of moves given by cancellation of pairs of adjacent letters that are inverses of each other.
- For example, $\alpha\beta\beta^{-1}\alpha\beta\alpha^{-1} = \alpha\alpha\beta\alpha^{-1}$.
- Formally, we define an equivalence relation and consider the corresponding quotient.

The Free group $\langle \alpha, \beta \rangle$

- The group $\langle \alpha, \beta \rangle$ as a set consists of words in S up to the equivalence given above.

The Free group $\langle \alpha, \beta \rangle$

- The group $\langle \alpha, \beta \rangle$ as a set consists of words in S up to the equivalence given above.
- Multiplication in $\langle \alpha, \beta \rangle$ is given by concatenation, i.e.

$$(\xi_1 \xi_2 \dots \xi_n) \cdot (l'_1 l'_2 \dots l'_m) = \xi_1 \xi_2 \dots \xi_n l'_1 l'_2 \dots l'_m$$

The Free group $\langle \alpha, \beta \rangle$

- The group $\langle \alpha, \beta \rangle$ as a set consists of words in S up to the equivalence given above.
- Multiplication in $\langle \alpha, \beta \rangle$ is given by concatenation, i.e.

$$(\xi_1 \xi_2 \dots \xi_n) \cdot (l'_1 l'_2 \dots l'_m) = \xi_1 \xi_2 \dots \xi_n l'_1 l'_2 \dots l'_m$$

- The identity e is the empty word.

The Free group $\langle \alpha, \beta \rangle$

- The group $\langle \alpha, \beta \rangle$ as a set consists of words in S up to the equivalence given above.
- Multiplication in $\langle \alpha, \beta \rangle$ is given by concatenation, i.e.

$$(\xi_1 \xi_2 \dots \xi_n) \cdot (l'_1 l'_2 \dots l'_m) = \xi_1 \xi_2 \dots \xi_n l'_1 l'_2 \dots l'_m$$

- The identity e is the empty word.
- The inverse of an element is obtained by inverting letters and reversing the order, i.e., $(\xi_1 \xi_2 \dots \xi_n)^{-1} = \xi_n^{-1} \dots \xi_2^{-1} \xi_1^{-1}$.

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* $g \in \langle \alpha, \beta \rangle$, $n \in \mathbb{Z}$.

The Quest

Some observations

- By counting the number of occurences of α and β with sign, we get a *homomorphism* $\varphi : \langle \alpha, \beta \rangle \to \mathbb{Z}^2$.

Some observations

- By counting the number of occurences of α and β with sign, we get a *homomorphism* $\varphi : \langle\alpha, \beta\rangle \to \mathbb{Z}^2$.
- The length $l_{\mathbb{Z}^2}(x, y) = |x| + |y|$ on $\mathbb{Z}^2$ induces a homogeneous, conjugacy-invariant pseudo-length $\bar{l}(g) = l_{\mathbb{Z}^2}(\varphi(g))$ on $\langle\alpha, \beta\rangle$;

Some observations

- By counting the number of occurences of α and β with sign, we get a *homomorphism* $\varphi : \langle \alpha, \beta \rangle \to \mathbb{Z}^2$.
- The length $l_{\mathbb{Z}^2}(x, y) = |x| + |y|$ on $\mathbb{Z}^2$ induces a homogeneous, conjugacy-invariant pseudo-length $\bar{l}(g) = l_{\mathbb{Z}^2}(\varphi(g))$ on $\langle \alpha, \beta \rangle$; however, as $\varphi(\alpha\beta\alpha^{-1}\beta^{-1}) = (0, 0)$, $\bar{l}(\alpha\beta\alpha^{-1}\beta^{-1}) = 0$.

Some observations

- By counting the number of occurences of α and β with sign, we get a *homomorphism* $\varphi : \langle\alpha, \beta\rangle \to \mathbb{Z}^2$.
- The length $l_{\mathbb{Z}^2}(x, y) = |x| + |y|$ on $\mathbb{Z}^2$ induces a homogeneous, conjugacy-invariant pseudo-length $\bar{l}(g) = l_{\mathbb{Z}^2}(\varphi(g))$ on $\langle\alpha, \beta\rangle$; however, as $\varphi(\alpha\beta\alpha^{-1}\beta^{-1}) = (0, 0)$, $\bar{l}(\alpha\beta\alpha^{-1}\beta^{-1}) = 0$.
- (Fritz) Homogeneity implies conjugacy invariant.

Some observations

- By counting the number of occurences of α and β with sign, we get a *homomorphism* $\varphi : \langle \alpha, \beta \rangle \to \mathbb{Z}^2$.
- The length $l_{\mathbb{Z}^2}(x, y) = |x| + |y|$ on $\mathbb{Z}^2$ induces a homogeneous, conjugacy-invariant pseudo-length $\bar{l}(g) = l_{\mathbb{Z}^2}(\varphi(g))$ on $\langle \alpha, \beta \rangle$; however, as $\varphi(\alpha\beta\alpha^{-1}\beta^{-1}) = (0, 0)$, $\bar{l}(\alpha\beta\alpha^{-1}\beta^{-1}) = 0$.
- (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 a construction using non-crossing matchings,

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 a construction using non-crossing matchings, but this was not homogeneous;

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 a construction using non-crossing matchings, but this was not homogeneous;
  - 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 a construction using non-crossing matchings, but this was not homogeneous;
  - 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 a construction using non-crossing matchings, but this was not homogeneous;
  - 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.

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 $|\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}ab\bar{a}\bar{b}| \leq 12.859649122807017$
- $|ab\bar{a}\bar{b}| \leq 0.8152734778121775$ 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 $f(m, k) = l(x^m(xyx^{-1}y^{-1})^k)$. *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 $f(m,k) = l(x^m(xyx^{-1}y^{-1})^k)$*. Then*

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

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

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 $L(g) := \min(\{\lambda_0\} \cup \Lambda)$.

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 (in place of $L(g)$) 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 (in place of $L(g)$) 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 (in place of $L(g)$) recursively in the above algorithm.
- We computed such bounds in interactive sessions.
- The words used were $\alpha(\alpha\beta\alpha^{-1}\beta^{-1})^k$, chosen based on non-homogeneity of the conjugacy-invariant length function l_{WC} based on non-crossing matchings.

From bounds to Proofs

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, which could be viewed as embedded 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, which could be viewed as embedded in Homotopy Type Theory; which is a system of foundations of mathematics related to topology.

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, which could be viewed as embedded in Homotopy Type Theory; which is a system of foundations of mathematics related to topology.
- In this case, we can instead view our algorithm as just keeping track of relevant inequalities.

Domain specific foundations in scala

- Proofs were represented as objects of a specific type.
- The correctness was independent of discovery.

Domain specific foundations in scala

- Proofs were represented as objects of a specific type.
- The correctness was independent of discovery.

```scala
sealed abstract class LinNormBound(val word: Word, val bound: Double)

final case class Gen(n: Int) extends LinNormBound(Word(Vector(n)), 1)

final case class ConjGen(n: Int, pf: LinNormBound) extends
      LinNormBound(n +: pf.word :+ (-n), pf.bound)

final case class Triang(
  pf1: LinNormBound, pf2: LinNormBound) extends
      LinNormBound( pf1.word ++ pf2.word, pf1.bound + pf2.bound)

final case class PowerBound(
  baseword: Word, n: Int, pf: LinNormBound) extends
    LinNormBound(baseword, pf.bound/n){require(pf.word == baseword.pow(n))}

final case object Empty extends LinNormBound(Word(Vector()), 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

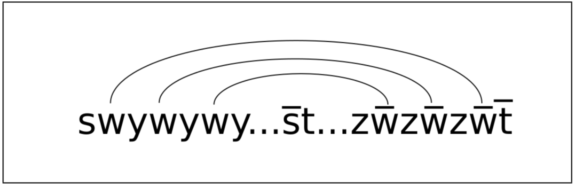
If $x = s(wy)s^{-1} = t(zw^{-1})t^{-1}$, *we have* $l(x) \leq \frac{l(y)+l(z)}{2}$.

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* $l(x) \leq \frac{l(y)+l(z)}{2}$.

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

$$swywywy\ldots\bar{s}t\ldots z\bar{w}z\bar{w}z\bar{w}\bar{t}$$

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

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.

Epilogue

On the computer proof

- A limitation was that the elements for which we applied homogeneity were selected by hand.

On the computer proof

- A limitation was that the elements for which we applied homogeneity were selected by hand.
- More importantly, in our representations of proofs, the bounds were only for concrete group elements.

On the computer proof

- A limitation was that the elements for which we applied homogeneity were selected by hand.
- More importantly, in our representations of proofs, the bounds were only for concrete group elements.
- In particular, we could not

On the computer proof

- A limitation was that the elements for which we applied homogeneity were selected by hand.
- More importantly, in our representations of proofs, the bounds were only for concrete group elements.
- In particular, we could not
    - represent inequalities for expressions,

On the computer proof

- A limitation was that the elements for which we applied homogeneity were selected by hand.
- More importantly, in our representations of proofs, the bounds were only for concrete group elements.
- In particular, we could not
  - represent inequalities for expressions,
  - use induction.

On the computer proof

- A limitation was that the elements for which we applied homogeneity were selected by hand.
- More importantly, in our representations of proofs, the bounds were only for concrete group elements.
- In particular, we could not
  - represent inequalities for expressions,
  - use induction.
- Would want proof in complete foundations;

On the computer proof

- A limitation was that the elements for which we applied homogeneity were selected by hand.
- More importantly, in our representations of proofs, the bounds were only for concrete group elements.
- In particular, we could not
  - represent inequalities for expressions,
  - use induction.
- Would want proof in complete foundations; which I completed a few days after the PolyMath proof (in my own implementation of HoTT).

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, $l(xyx^{-1}y^{-1}) \leq 4c$ 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, $l(xyx^{-1}y^{-1}) \leq 4c$ 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 equivalent to a pullback from $G/[G, G]$.

Afterword

- This work became PolyMath 14, and has been published in *Algebra & Number Theory*.

Afterword

- This work became PolyMath 14, and has been published in *Algebra & Number Theory*.
- The work was a spontaneous collaboration across (at least) three continents, and a range of skills.

Afterword

- This work became PolyMath 14, and has been published in *Algebra & Number Theory*.
- The work was a spontaneous collaboration across (at least) three continents, and a range of skills.
- 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*.
- The work was a spontaneous collaboration across (at least) three continents, and a range of skills.
- 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.