FREE GROUPS, LENGTHS AND COMPUTER PROOFS

SIDDHARTHA GADGIL

Indian Institute of Science

JOINT WITH

the rest of (spontaneous) polymath 14

THE POLYMATH 14 PARTICIPANTS

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

- On Saturday, December 16, 2017, Terrence Tao posted on his blog a question, which Apoorva Khare had asked him.

> *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.

LENGTH FUNCTIONS

- Fix a group G, i.e. a set with an associative product with inverses.
- A *pseudo-length function* $l : G \to [0, \infty)$ is a function such that:
  - $l(e) = 0$.
  - $l(g^{-1}) = l(g)$, for all $g \in G$.
  - (Triangle inequality) $l(gh) \leq l(g) + l(h)$, for all $g, h \in G$.
- A *length function* is a pseudo-length function such that $l(g) > 0$ whenever $g \neq e$ (positivity condition).

CONJUGACY-INVARIANCE AND HOMOGENEITY

- A pseudo-length function l is said to be conjugacy invariant if $l(ghg^{-1}) = l(h)$ for all $g, h \in G$.
- Conjugacy often corresponds to change of coordinates.
- l is said to be *homogeneous* if $l(g^n) = nl(g)$ for all $g \in G$.
- Homogeneity was motivated by norms on vector spaces.

LENGTH FUNCTIONS ON $\mathbb{Z}^2$

- The group $\mathbb{Z}^2 = \{(n, m) : n, m \in \mathbb{Z}\}$ with $(n_1, m_1) + (n_2, m_2) = (n_1 + n_2, m_1, m_2)$.
- Two length functions on $\mathbb{Z}^2$ are:
  - $l_1((a, b)) = |a| + |b|$,
  - $l_2((a, b)) = \sqrt{a^2 + b^2}$
- These are homogeneous (and conjugacy invariant).
- The function $l'((a, b)) = \sqrt{|a|} + \sqrt{|b|}$ is a length function, but not homogeneous.
- The function $l''((a, b)) = |a - b|$ is a homogeneous pseudo-length function which is not positive.

GROUPS WITHOUT HOMOGENEOUS LENGTH FUNCTIONS

- If l is a homogeneous length function on G and $g \in G$ is such that $g^n = e$ then $l(g) = 0$ as $nl(g) = l(g^n) = 0$.
- Thus, if G has *torsion*, i.e., there exists $g \neq e \in G$, such that $g^n = e$ for some $n > 0$ and $g \neq e$, then there is no homogeneous length function on G.
- In particular finite groups have no homogeneous length functions.

FREE GROUP $\mathbb{F}_2$ ON α, β

- Consider words in four letters α, β, $\bar{\alpha}$ and $\bar{\beta}$.
- We multiply words by concatenation, e.g. $\alpha\beta \cdot \bar{\alpha}\beta = \alpha\beta\bar{\alpha}\beta$.
- Two words are regarded as equal if they can be related by cancelling adjacent letters that are inverses, e.g. $\alpha\beta\bar{\beta}\alpha\beta = \alpha\alpha\beta$.
- This gives a group;
  - the identity is the empty word
  - every word has an inverse, e.g., $(\alpha\bar{\beta}\alpha\beta\beta)^{-1} = \bar{\beta}\bar{\beta}\bar{\alpha}\beta\bar{\alpha}$.

WORD LENGTH ON $\mathbb{F}_2$

- The *word length* l_0 on $\mathbb{F}_2$ is the length of the shortest word representing an element.
- Indeed any word is equivalent to a unique *reduced* word, whose length is the word length.
- The word length is not conjugacy-invariant as $l_0(\alpha\beta\alpha^{-1}) = 3 \neq 1 = l_0(\beta)$.
- The word length is also not homogeneous, e.g., $l_0((\alpha\beta\alpha^{-1})^2) = l_0(\alpha\beta\beta\alpha^{-1}) = 4 \neq 2l_0(\alpha\beta\alpha^{-1})$.

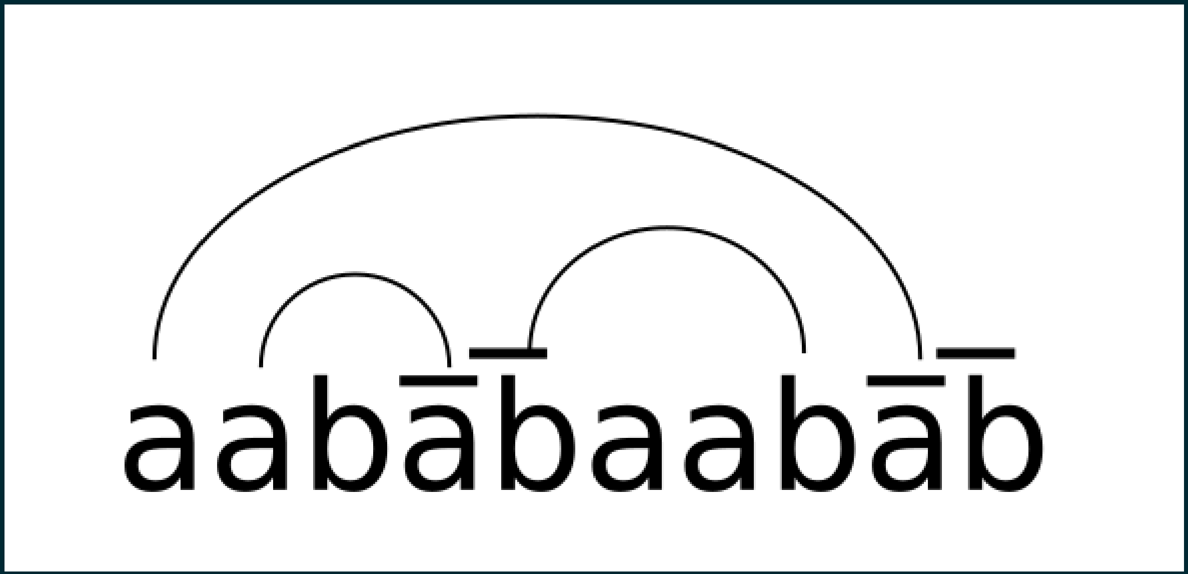
PULLBACK PSEUDO-LENGTH ON $\mathbb{F}_2$

- We define a homomorphism $ab : \mathbb{F}_2 \to \mathbb{Z}^2$.
  - For a word g, let $n_\alpha(g)$ and $n_\beta(g)$ be the number of α's and β's counted with sign.
  - Define $ab(g) = (n_\alpha(g), n_\beta(g))$, e.g. $ab(\alpha\alpha\beta\bar{\alpha}\bar{\beta}\alpha\bar{\beta}) = (2, -1)$.
- We get a pullback pseudo-length l on $\mathbb{F}_2$ by $l_{ab}(g) = l_1(ab(g)) = |n_\alpha(g)| + |n_\beta(g)|$.
- As l_1 is homogeneous, so is l_{ab}.
- However, this is *not* a length function as $l_{ab}([\alpha, \beta]) = 0$, where $[g, h] = ghg^{-1}h^{-1}$ (violating positivity).

THE MAIN RESULTS

- **Question:** Is there a homogeneous length function on the free group on two generators?
- **Answer:** No; we in fact describe all homogeneous pseudo-lengths.
- **Theorem:** Any homogeneous pseudo-length function on a group G is the pullback of a pseudo-length on its abelianization $G/[G, G]$.
- **Corollary:** If G is not abelian (e.g. $G = \mathbb{F}_2$) there is no homogeneous length function on G.
- In fact, every homogeneous pseudo-length is the pullback of a *norm* on a vector space.

LENGTHS FROM NON-CROSSING MATCHINGS



- Given a word in $\mathbb{F}_2$, we consider *matchings* such that
  - letters are paired with their inverses,
  - there are no *crossings*
- The *energy* is the number of unmatched letters.

WATSON-CRICK LENGTH ON $\mathbb{F}_2$

- For $g \in \mathbb{F}_2$, define the Watson-Crick length $l_{WC}(g)$ to be the minimum number of unmatched letters over all non-crossing matchings.
- l_{WC} is unchanged under cancellation (hence well-defined on $\mathbb{F}_2$) and conjugacy invariant.
- In fact it is the *maximal* normalized conjugacy-invariant length function on $\mathbb{F}$.
- However, if $g = \alpha[\alpha, \beta]$, then $l_{WC}(g^2) = 4$, but $l_{WC}(g) = 3$, violating homogeneity.

THE GREAT BOUND CHASE

- By Tuesday morning, most people were convinced that there are no homogeneous length functions on the free group, and probably $l([\alpha, \beta]) = 0$ for homogeneous pseudo-lengths.
- There was a steady improvements in the combinatorial/analytic bounds on $l([\alpha, \beta])$.
- These seemed to be stuck above 1 (as observed by Khare) - but eventually broke this barrier (work of Fritz, Khare, Nielsen, Silberman, Tao).
- At this stage, my approach diverged.

COMPUTER ASSISTED PROOFS: BEYOND (SYMBOLIC) COMPUTATION, ENUMERATION?

- We can recursively compute the Watson-Crick length $l_{WC}(g)$ for a word g.
- This gives an upper bound on all conjugacy-invariant normalized lengths.
- This can be combined with using homogenity.
- Using this, I obtained an upper bound of about 0.85 on $l(\alpha, \beta)$.
- This was upgraded to a (computer) checkable proof.
- On Thursday morning, I posted an in principle human readable proof of a bound.

Back to presentation (HomogeneousLengths.html#gross)

Gross proof time

Here is a computer generated proof of a bound on the length of the commutator abāb̄ for a *linear* norm on the free group with the lengths of the generators bounded above by 1.

1. |ā| ≤ 1.0
2. |b̄āb| ≤ 1.0 using |ā| ≤ 1.0
3. |b̄| ≤ 1.0
4. |ab̄ā| ≤ 1.0 using |b̄| ≤ 1.0
5. |āb̄abāb̄| ≤ 2.0 using |āb̄a| ≤ 1.0 and |bāb̄| ≤ 1.0
6. |a| ≤ 1.0
7. |bab̄| ≤ 1.0 using |a| ≤ 1.0
8. |b| ≤ 1.0
9. |āba| ≤ 1.0 using |b| ≤ 1.0
10. |abāb̄ab| ≤ 2.0 using |abā| ≤ 1.0 and |b̄ab| ≤ 1.0
11. |āabāb̄aba| ≤ 2.0 using |abāb̄ab| ≤ 2.0
12. |bāb̄aabāb̄abā| ≤ 3.0 using |bāb̄| ≤ 1.0 and |aabāb̄abā| ≤ 2.0
13. |āb̄abāb̄aabāb̄abā| ≤ 4.0 using |āb̄a| ≤ 1.0 and |bāb̄aabāb̄abā| ≤ 3.0
14. |b̄āb̄abāb̄aabāb̄abāb| ≤ 4.0 using |āb̄abāb̄aabāb̄abā| ≤ 4.0
15. |bāb̄abāb̄aabāb̄abāb̄aabāb̄abā| ≤ 6.0 using |bāb̄abāb̄aabāb̄abāb̄| ≤ 4.0 and |aabāb̄abā| ≤ 2.0
16. |b̄abāb̄a| ≤ 2.0 using |b̄ab| ≤ 1.0 and |āb̄a| ≤ 1.0
17. |ab̄abāb̄aā| ≤ 2.0 using |b̄abāb̄a| ≤ 2.0
18. |āb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abā| ≤ 8.0 using |āb̄abāb̄aa| ≤ 2.0 and |bāb̄abāb̄aabāb̄abāb̄aabāb̄abā| ≤ 6.0
19. |b̄āb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abāb| ≤ 8.0 using |āb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abā| ≤ 8.0
20. |bāb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abā| ≤ 10.0 using |bāb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄| ≤ 8.0 and |aabāb̄abā| ≤ 2.0
21. |āb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abā| ≤ 12.0 using |āb̄abāb̄aa| ≤ 2.0 and |bāb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abā| ≤ 10.0
22. |b̄āb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abāb̄aabāb̄abāb| ≤ 12.0 using

- The computer-generated proof was studied by Pace Nielsen, who extracted the *internal repetition* trick.
- This was extended by Nielsen and Fritz and generalized by Tao; from this Fritz obtained the key lemma:
- Let $f(m, k) = l(x^m[x, y]^k)$. Then $f(m, k) \leq \frac{f(m-1,k)+f(m+1,k-1)}{2}$.
- A probabilistic argument of Tao finished the proof.

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

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

- **Lemma:** $f(m,k) \leq \frac{f(m-1,k)+f(m+1,k-1)}{2}$, where $f(m,k) = l(x^m[x,y]^k)$.
- In other words, for Y a Rademacher random variable, i.e., Y is 1 or -1 each with probability $1/2$, $f(m,k) \leq E(f((m,k-1/2)+Y(1,-1/2)))$.
- Hence for Y_i i.i.d. Rademacher random variables, $f(0,n) \leq E(f((Y_1+Y_2+\cdots+Y_{2n})(1,-1/2)))$
- By triangle inequality, $f(a,b) \leq A\|(a,b)\|$.
- $Y_1+Y_2+\ldots Y_{2n}$ has mean $(0,0)$ and variance $2n$, so $E(\|Y_1+Y_2+\ldots Y_{2n}\|) \leq B\sqrt{n}$.
- We deduce that $l([x,y]^n) = f(0,n) \leq C\sqrt{n}$, hence $l([x,y]) = 0$.

COMPUTER BOUNDS AND PROOFS

- For $g = ah, a \in \{\alpha, \beta, \bar{\alpha}, \bar{\beta}\}$, the length $l_{WC}(g)$ is the minimum of:
  - $1 + l_{WC}(h)$: corresponding to a unmatched.
  - $\min\{l_{WC}(x) + l_{WC}(y) : h = x\bar{a}y\}$.
- We can describe a minimal non-crossing matching by a similar recursive definition.
- A similar recursion gives a *proof* of a bound on $l(g)$ for g a homogeneous, conjugacy-invariant length with $l(\alpha) = l(\beta) = 1$.
- We can also use homogeneity for selected elements and powers to bound the length function l.

CONCLUSIONS

- If we view applications of the axioms as moves, the computer proof helped in identifying composite moves that could be applied recursively.
- The principal difficulty in finding computer proofs often lies in choosing the useful complexity increasing moves (here homogeneity).
- In this case, these were chosen mainly on mathematical considerations.
- However, a general heuristic we see is to seek useful *families* of complexity-increasing moves.

CONCLUSIONS

- The unusual feature of the use of computers here was that 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.
- The use of computers was based on the idea of proofs as mathematical objects which can be computed, following foundations based on Homotopy Type Theory.