Turn Graphs and Extremal Surfaces in Free Groups

Noel Brady, Matt Clay, and Max Forester

Abstract. This note provides an alternate account of Calegari's rationality theorem for stable commutator length in free groups.

1. Introduction

The purpose of this note is to provide an alternate account of Calegari's main result from [**4**], establishing the existence of extremal surfaces for stable commutator length in free groups, via linear programming. The argument presented here is similar to that given in [**4**], except that we avoid using the theory of branched surfaces. Instead, the reduction to linear programming is achieved directly, using the combinatorics of words in the free group. We note that the specific linear programming problem resulting from the discussion here essentially agrees with that described in Example 4.34 of [**3**].

Acknowledgements. The authors would like to thank Dan Guralnik and Sang Rae Lee for helpful discussions during the course of this work.

2. Preliminaries

We start by giving a working definition of stable commutator length. Propositions 2.10 and 2.13 of [**3**] show that it is equivalent to the basic definition in terms of commutators or genus.

Definition 1. Let $G = \pi_1(X)$ and suppose $\gamma\colon S^1 \to X$ represents the conjugacy class of $a \in G$. The *stable commutator length of* a is given by

$$\mathrm{scl}(a) \;=\; \inf_S \frac{-\chi(S)}{2n(S)} \tag{1}$$

where S ranges over all singular surfaces $S \to X$ such that

- S is oriented and compact with $\partial S \neq \emptyset$
- S has no S^2 or D^2 components

2010 *Mathematics Subject Classification.* Primary 57M07, 20F65, 20J05.

Partially supported by NSF grants DMS-0906962 (Brady), DMS-1006898 (Clay), and DMS-0605137 (Forester).

- the restriction $\partial S \to X$ factors through γ; that is, there is a commutative diagram:

$$\begin{array}{ccc} \partial S & \longrightarrow & S \\ \downarrow & & \downarrow \\ S^1 & \xrightarrow{\gamma} & X \end{array}$$

- the restriction of the map $\partial S \to S^1$ to each connected component of ∂S is a map of *positive* degree

and where $n(S)$ is the total degree of the map $\partial S \to S^1$ (of oriented 1–manifolds).

A surface S satisfying the conditions above is called a *monotone admissible surface* in [**3**], abbreviated here as an *admissible surface*. Such a surface exists if and only if $a \in [G, G]$. If $a \notin [G, G]$ then by convention $\mathrm{scl}(a) = \infty$ (the infimum of the empty set).

A surface $S \to X$ is said to be *extremal* if it realizes the infimum in (1). Notice that if this occurs, then $\mathrm{scl}(a)$ is a rational number.

3. Singular surfaces in graphs

Let X be a graph with oriented 1–cells $\{e_\nu\}$. These edges may be formally considered as a generating set for the fundamental groupoid of X based at the vertices. These generators also generate the fundamental group F of X. Note that F is free, but the groupoid generators are not a basis unless X has only one vertex. (The reader may assume this latter property with no harm, in which case the fundamental groupoid is simply the fundamental group.)

Let $\gamma\colon S^1 \to X$ be a simplicial loop with no backtracking. There is a corresponding cyclically reduced word $w = x_1 \cdots x_\ell$ in the fundamental groupoid generators and their inverses. This word w represents a conjugacy class in $\pi_1(X)$, which we assume to be in $[F, F]$. Finally, let $S \to X$ be an admissible surface for w, as in Definition 1.

We are interested in computing $\chi(S)$ and $n(S)$, to estimate $\mathrm{scl}(w)$ from above. We are free to modify S if the resulting surface S' satisfies $\frac{-\chi(S')}{2n(S')} \leqslant \frac{-\chi(S)}{2n(S)}$, since this only strengthens the estimate.

Using transversality, the map $S \to X$ can be homotoped into a standard form, sometimes called a transverse map [**2**]. The surface is decomposed into pieces called 1–handles, which map to edges of X, and complementary regions, which map to vertices of X. Each 1–handle is a tubular neighborhood of a connected 1–dimensional submanifold, either an arc with endpoints on ∂S or a circle. The submanifold maps to the midpoint of an edge of X, and the fibers of the tubular neighborhood map over the edge, through its characteristic map. In particular, the boundary arcs or circles of the 1–handle (comprised of endpoints of fibers) map to vertices of X. A *transverse labeling* is a labeling of the fibers of 1–handles by fundamental groupoid generators, indicating which edge of X (and in which direction) the handle maps to. For more detail on putting maps into this form, see for instance [**6, 7, 5, 1**].

Let $S_0 \subset S$ be the codimension-zero submanifold obtained as the closure of the union of a collar neighborhood of ∂S and the 1–handles that meet ∂S. We will see that S_0 is the essential part of S, containing all of the relevant information. It is determined completely by ∂S, together with the additional data of which pairs

of edges in ∂S are joined by 1–handles. Note that ∂S_0 consists of ∂S together with additional components in the interior of S. These latter components will be called the *inner boundary* of S_0, denoted $\partial^- S_0$. Let S_1 be the closure of $S - S_0$. Note that $\partial S_1 = S_1 \cap S_0 = \partial^- S_0$. Figure 1 shows an example of S_0 for the word $w = aba^{-1}b^{-1}$. (The "turns" mentioned there are discussed in the next section.)

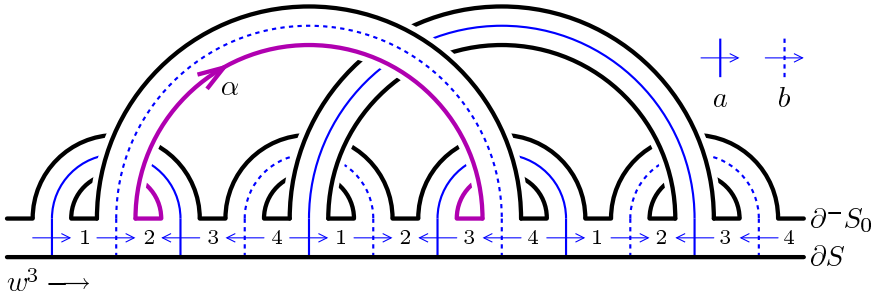

FIGURE 1. One possible S_0 for the word $w = aba^{-1}b^{-1}$, with outer boundary w^3. There are four turns, indicated by the numbers, each occurring three times. The boundary arc α leads from turn 2 to turn 3. This surface leads to an estimate $\mathrm{scl}(w) \leqslant 1/2$ (and in fact is extremal).

How large can $\chi(S)$ be? Note that $\chi(S) = \chi(S_0) + \chi(S_1)$ since S_0 and S_1 meet along circles. Also,

$$\chi(S_0) \;=\; \frac{-n(S)\,|w|}{2}$$

as can be seen by counting the 1–handles meeting ∂S: each 1–handle contributes -1 to $\chi(S_0)$ and occupies two edges in ∂S, of which there are $n(S)\,|w|$ in total. Finally, given S_0, the quantity $\chi(S_1)$ is largest when S_1 is a collection of disks. The number of disks is simply the number of components of $\partial^- S_0$. We can always replace S_1 by disks, since each component of ∂S_1 maps to a vertex of X and disks can be mapped to vertices also. Thus, after this modification, we have

$$\chi(S) \;=\; \frac{-n(S)\,|w|}{2} \;+\; \bigl|\pi_0(\partial^- S_0)\bigr|\,, \tag{2}$$

and therefore an upper bound for $\mathrm{scl}(w)$ is given by

$$\frac{-\chi(S)}{2n(S)} \;=\; \frac{|w|}{4} \;-\; \frac{|\pi_0(\partial^- S_0)|}{2n(S)}. \tag{3}$$

Indeed, $\mathrm{scl}(w)$ is precisely the infimum of the right hand side of (3) over all surfaces S_0 arising as above. (Note that $n(S)$ is determined by S_0.) Equation (3) essentially replaces the quantity $\chi(S)$ by the number of inner boundary components of S_0 in the computation of $\mathrm{scl}(w)$.

4. The turn graph

To help keep track of the inner boundary $\partial^- S_0$, we define the *turn graph*. Consider the word $w = x_1 \cdots x_\ell$. A *turn* in w is a position between two letters of w considered as a cyclic word. Thus there are $|w| = \ell$ turns. Each turn is labeled

by the length two subword $x_i x_{i+1}$ (or $x_\ell x_1$) of w which straddles the turn. Note that turns are not necessarily determined by their labels.

The *turn graph* $\Gamma(w)$ is a directed graph with vertices equal to the turns of w, and with a directed edge from turn i to turn j if $x_i^{-1} = x_{j+1}$. That is, if the label of a turn begins with the letter $a^{\pm 1}$, then there is a directed edge from this turn to every other turn whose label ends with $a^{\mp 1}$. Note that because w is cylically reduced, $\Gamma(w)$ has no loops.

The turn graph has a two-fold symmetry, or duality: if $e \in \Gamma(w)$ is an edge from turn i to turn j, then one verifies easily that there is also an edge $\bar{e}$ from turn $j+1$ to turn $i-1$, and moreover $\bar{\bar{e}} = e$. Figure 2 shows a turn graph and a dual edge pair.

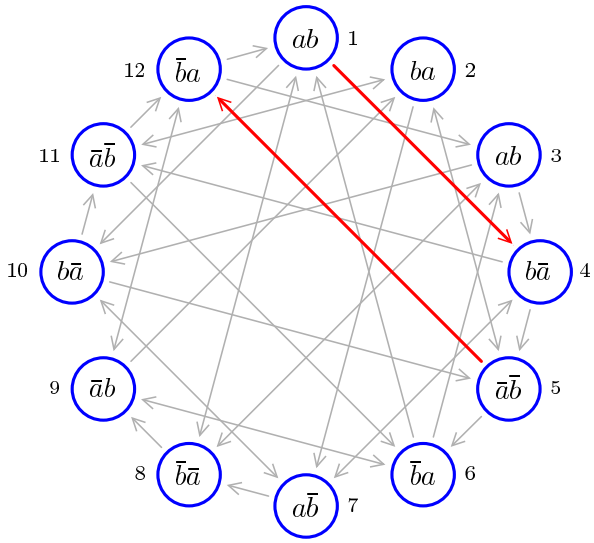

FIGURE 2. The turn graph for the word $w = abab\bar{a}\bar{b}a\bar{b}\bar{a}b\bar{a}\bar{b}$ (bar denotes inverse). The highlighted edges form a dual pair.

Turn circuits. Given the surface S_0, each inner boundary component can be described as follows. Traversing the curve in the positively oriented direction, one alternately follows 1–handles and visits turns of w positioned along ∂S; see again Figure 1 (this situation is the reason for the word "turn"). If a 1–handle leads from turn i to turn j, then the 1–handle bears the transverse label $x_i^{-1} = x_{j+1}$, and so there is an edge in $\Gamma(w)$ from turn i to turn j. The sequence of 1–handles traversed by the boundary component therefore yields a directed circuit in $\Gamma(w)$. In this way the inner boundary $\partial^- S_0$ gives rise to a finite collection (possibly with repetitions) of directed circuits in $\Gamma(w)$, called the *turn circuits* for S_0.

Recall that ∂S is labeled by $w^{n(S)}$ (possibly spread over several components), so there are $n(S)$ occurrences of each turn on ∂S. The turn circuits do not contain the information of which particular instances of turns are joined by 1–handles.

Turn surgery. There is a move one can perform on S_0 which is useful. Given two occurrences of turn i in S_0, cut the collar neighborhood of ∂S open along arcs positioned at the two turns, between the adjacent 1–handles; see Figure 3. These arcs both map to the same vertex of X. Now re-glue the four sides of the arcs,

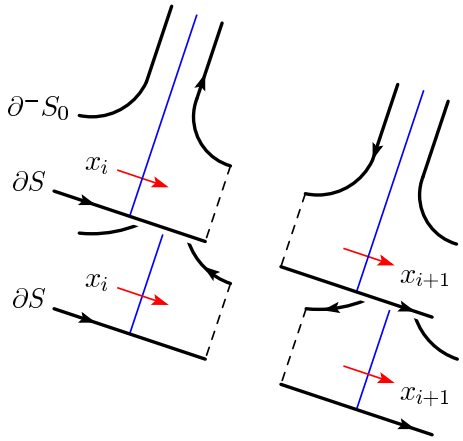

FIGURE 3. Cutting along two instances of turn i.

switching two of them. There is one way to do this which preserves orientations of S and of ∂S_0. The new surface is still admissible (that is, after capping off $\partial^- S_0$) and $n(S)$ is preserved.

The move changes both ∂S and $\partial^- S_0$, in each case either increasing or decreasing the number of connected components by one. If both instances of the turn occupy the same component, then the move splits this component into two, with each occupied by one of the turns. Otherwise, the move joins the two components occupied by the turns into one.

DEFINITION 2. An admissible surface S is *taut* if every component of $\partial^- S_0$ visits each turn at most once. In terms of the turn graph, this means that every turn circuit for S_0 is embedded in $\Gamma(w)$. Let $\mathscr{T}(w)$ be the set of taut admissible surfaces for w.

Any admissible surface S can be made taut by performing a finite number of turn surgeries, each increasing the number of inner boundary components of ∂S_0. Since $n(S)$ remains constant, the quantity (3) will only decrease. Hence we have the following result:

LEMMA 3. *There is an equality*

$$\mathrm{scl}(w) \quad = \quad \inf_{S \in \mathscr{T}(w)} \frac{|w|}{4} - \frac{|\pi_0(\partial^- S_0)|}{2n(S)}.$$

5. Weight vectors and linear optimization

Let $\{\alpha_1, \ldots, \alpha_k\}$ be the set of embedded directed circuits in $\Gamma(w)$. For each taut admissible surface S let $u_i(S)$ be the number of occurrences of α_i among the turn circuits of S_0, and let $u(S) \in \mathbb{R}^k$ be the non-negative integer vector $(u_1(S), \ldots, u_k(S))$. We call $u(S)$ the *weight vector* for S.

For each vertex v and edge e of $\Gamma(w)$, there are linear functions

$$F_v \colon \mathbb{R}^k \to \mathbb{R}, \quad F_e \colon \mathbb{R}^k \to \mathbb{R}$$

whose values on the i^{th} standard basis vector $(0, \ldots, 0, 1, 0, \ldots 0)$ are given by the number of times α_i passes through the vertex v (respectively, over the edge e).

Since α_i is embedded, these numbers will be 0 or 1, although this is not important. For the taut surface S, if $e \in \Gamma(w)$ is an edge from turn i to turn j, then $F_e(u(S))$ counts the number of times $\partial^- S_0$ follows a 1–handle from turn i to turn j. Similarly, if $v \in \Gamma(w)$ is turn i, then $F_v(u(S))$ counts the number of occurrences of turn i on $\partial^- S_0$ (which is $n(S)$, as observed earlier).

REMARK 4. For taut admissible surfaces, the functions $|\pi_0(\partial^- S_0)|$ and $n(S)$ both factor as

$$\mathscr{T}(w) \xrightarrow{u} \mathbb{R}^k \longrightarrow \mathbb{R}$$

where the second map is linear, with integer coefficients. In the case of $|\pi_0(\partial^- S_0)|$ the second map is given by $(u_1, \ldots, u_k) \mapsto \sum_i u_i$, and in the case of $n(S)$, the second map is simply F_v (for any vertex $v \in \Gamma(w)$).

By (2) it follows that the function $-\chi(S)$ also factors as above, through an integer coefficient linear function $\mathbb{R}^k \to \mathbb{R}$.

LEMMA 5. *Every weight vector $u(S)$ satisfies the linear equation*

$$F_e(u(S)) = F_{\bar{e}}(u(S))$$

for every dual pair $e, \bar{e}$ of edges in $\Gamma(w)$.

PROOF. Suppose e leads from turn i to turn j (so $\bar{e}$ leads from turn $j+1$ to turn $i-1$). If a 1–handle has a boundary arc representing e then the other side of the 1–handle represents $\bar{e}$. Hence both sides of the equation count the number of 1–handles of S_0 joining occurrences of x_i and x_{j+1} in ∂S. □

This lemma has a converse:

PROPOSITION 6. *If $u \in \mathbb{R}^k - \{0\}$ has non-negative integer entries and satisfies the linear equations*

$$F_e(u) = F_{\bar{e}}(u) \quad \textit{for all dual pairs } e, \bar{e} \tag{4}$$

then u is the weight vector of a taut admissible surface.

PROOF. Suppose $u = (u_1, \ldots, u_k)$. For each i let D_i be a polygonal disk with $|\alpha_i|$ sides. Label the oriented boundary of D_i by the edges and vertices of α_i. That is, sides are labeled by edges of $\Gamma(w)$, and corners are labeled by turns. Note that there are no monogons, since $\Gamma(w)$ has no loops. To form the taut admissible surface S, take u_i copies of D_i for each i. For each dual edge pair e, $\bar{e}$ the total number of edges labeled e among the D_i's will equal the number of edges labeled $\bar{e}$, by (4). Hence the sides of the disks can be joined in dual pairs to form a closed oriented surface.

However, this is not how S is formed. Instead, whenever two disks were to be joined along sides labeled e and $\bar{e}$, insert an oriented rectangle, with sides labeled by e, x_{j+1}, $\bar{e}$, x_i (here, e leads from turn i to turn j, and $\bar{e}$ from turn $j+1$ to turn $i-1$). See Figure 4. The opposite sides labeled by e and $\bar{e}$ are joined to the appropriate sides of the disks, and the remaining two sides become part of the boundary of S. Each rectangle can be transversely labeled by a fundamental groupoid generator (equal to $x_i^{-1} = x_{j+1}$), and then the rectangles become 1–handles in the resulting surface S.

Note that the side of a rectangle labeled x_i has neighboring polygonal disk corners labeled $i-1$ and i. Following this edge along ∂S, the next edge must be labeled x_{i+1} (adjacent to i and $i+1$); see again Figure 4. Hence each component

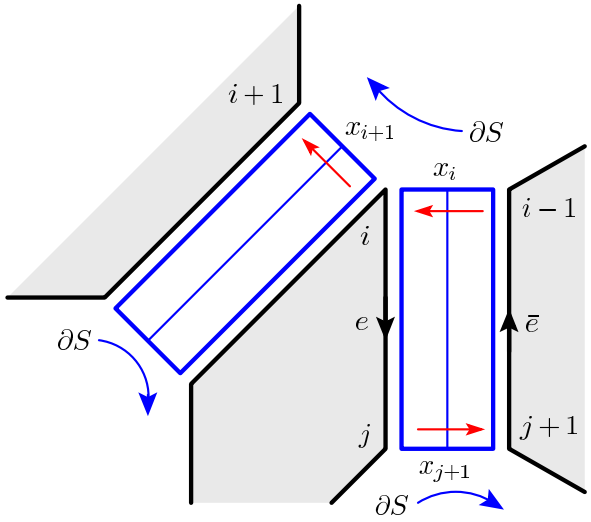

FIGURE 4. Building S from $(u_1, \ldots, u_k)$.

of ∂S is labeled by a positive power of w. There are no S^2 components since no component of S is closed, and no D^2 components, since an outermost 1–handle on such a disk would have to bound a monogon. The map $S \to X$ is defined on the rectangles according to the transverse labels (each maps to an edge of X) and the disks map to vertices. Now S is admissible, and by construction, the turn circuits will all be instances of the circuits α_i, so S is taut. □

THEOREM 7 (Calegari). *If X is a graph and $a \in [\pi_1(X), \pi_1(X)]$ then there exists an extremal surface $S \to X$ for a. Moreover, there is an algorithm to construct S. In particular,* $\mathrm{scl}(a)$ *is rational and computable.*

PROOF. Let w be the cyclically reduced word representing the conjugacy class of a, as defined in Section 3. By Remark 4 the function $\frac{-\chi(S)}{2n(S)}$ factors as

$$\mathscr{T}(w) \xrightarrow{u} \mathbb{R}^k \longrightarrow \mathbb{R}$$

where the second map is a ratio of linear functions $A(u)/B(u)$ with integer coefficients. Let $P \subset \mathbb{R}^k$ be the polyhedron defined by the (integer coefficient) linear equations (4) and the inequalities $B(u) > 0$ and $u_i \geqslant 0$, $i = 1, \ldots, k$. Lemma 5 and Proposition 6 together imply that the image of $u\colon\ \mathscr{T}(w) \to \mathbb{R}^k$ is precisely $P \cap \mathbb{Z}^k$. Hence

$$\mathrm{scl}(w) \quad = \quad \inf_{u \in P \cap \mathbb{Z}^k} A(u)/B(u).$$

Note that P and $A(u)/B(u)$ are projectively invariant. Normalizing $B(u)$ to be 1, we have

$$\mathrm{scl}(w) \quad = \quad \inf_{u \in P' \cap \mathbb{Q}^k} A(u) \tag{5}$$

where P' is the rational polyhedron $P \cap B^{-1}(1)$. Note that P' is a closed set.

From Remark 4 and equation (2) the function A is given by

$$A(u) \;=\; \frac{|w|}{2}F_v(u) + \sum_i u_i\,,$$

which has strictly positive values on the standard basis vectors. Hence A achieves a minimum on P', along a non-empty rational sub-polyhedron. The vertices of this sub-polyhedron are rational points realizing the infimum in (5). Hence there exist extremal surfaces for w, and $\mathrm{scl}(w)$ is rational. An extremal surface can be constructed explicitly from a rational solution $u \in P' \cap \mathbb{Q}^k$, by first multiplying by an integer to obtain a minimizer for $A(u)/B(u)$ in $P \cap \mathbb{Z}^k$, and then applying the procedure given in the proof of Proposition 6. Lastly, we note that from the word w it is straightforward to algorithmically construct the turn graph $\Gamma(w)$, the equations (4), and the polyhedron P'. □

References

1. Noel Brady and Max Forester, *Density of isoperimetric spectra*, Geom. Topol. **14** (2010), no. 1, 435–472. MR 2578308
2. S. Buoncristiano, C. P. Rourke, and B. J. Sanderson, *A geometric approach to homology theory*, Cambridge University Press, Cambridge, 1976, London Mathematical Society Lecture Note Series, No. 18. MR 0413113 (54 #1234)
3. Danny Calegari, *scl*, MSJ Memoirs, vol. 20, Mathematical Society of Japan, Tokyo, 2009. MR 2527432
4. ———, *Stable commutator length is rational in free groups*, J. Amer. Math. Soc. **22** (2009), no. 4, 941–961. MR 2525776
5. Marc Culler, *Using surfaces to solve equations in free groups*, Topology **20** (1981), no. 2, 133–145. MR 605653 (82c:20052)
6. C. P. Rourke, *Presentations and the trivial group*, Topology of low-dimensional manifolds (Proc. Second Sussex Conf., Chelwood Gate, 1977), Lecture Notes in Math., vol. 722, Springer, Berlin, 1979, pp. 134–143. MR 547460 (81a:57001)
7. John R. Stallings, *A graph-theoretic lemma and group-embeddings*, Combinatorial group theory and topology (Alta, Utah, 1984), Ann. of Math. Stud., vol. 111, Princeton Univ. Press, Princeton, NJ, 1987, pp. 145–155. MR 895613 (88k:20056)

Mathematics Department, University of Oklahoma, Norman, OK 73019, USA
E-mail address: nbrady@math.ou.edu

Mathematics Department, Allegheny College, Meadville, PA 16335, USA
E-mail address: mclay@allegheny.edu

Mathematics Department, University of Oklahoma, Norman, OK 73019, USA
E-mail address: forester@math.ou.edu