COMMUTATOR LENGTH OF $[x, y]^n$: ALGEBRAIC AND GEOMETRIC APPROACHES

BRIAN RESZUTEK

Abstract. In F, the free group of rank 2 generated by x and y, we discuss the minimum number of commutators needed to express $[x, y]^n$, which has been shown to be $\lfloor \frac{n}{2} \rfloor + 1$ through both geometric and algebraic approaches. We outline these approaches to commutator length as well as the specific results they give for expressions of $[x, y]^n$ with $\lfloor \frac{n}{2} \rfloor + 1$.

Contents

1. Introduction

Let $[a, b] = aba^{-1}b^{-1}$ for any $a, b \in F$. For $w \in [F, F]$, the commutator subgroup of the free group, let $cl(w)$ be defined as the commutator length of w, the smallest integer n such that $w = [a_1, b_1][a_2, b_2]...[a_n, b_n]$ for some elements $a_1, b_1, a_2, ...b_n$ in F. Clearly this is only true if w is balanced with respect to the generators of F, x and y, that is, for each instance of the letter x in w, there is one of x^{-1} and vice versa. If the number of letters does not match the number of inverses, it is impossible to express the word as a product of commutators.

We first describe the geometric approach, which shows that the existence of a "tight maps" from a surface with boundary of genus n into the free group is equivalent with the image of the boundary having commutator length n. This is built upon with the concept of star graphs and/or fat graphs, alternate representations of surfaces whose boundaries represent words from the free group, which allow us to find $cl([x, y]^n)$.

We then examine the same results through the algebraic lens, which surprisingly uses only elementary algebra and the properties of the commutator to find an expression for $[x, y]^n$ with $\lfloor \frac{n}{2} \rfloor + 1$ commutators and prove that this is the best possible result.

2. Topological approach: Fatgraphs and Surfaces

It is known in geometry that a word $w \in F$ has $cl(w) = n$ if and only if there exists a "tight map" (defined by Culler in [1]) from a genus n surface S with boundary to the free group such that the image of ∂S is w. For this reason, $cl(w)$ is sometimes called the genus of w. The study of these maps leads to several representations of surfaces and maps into the free group based around a word w and a pairing p of the letters of w (a permutation such that $p = p^{-1}$ on $\{1, 2, ..., n\}$ where w has n letters in its reduced form). We will assume that pairings send instances of x to x^{-1}, y to y^{-1}, and vice-versa. This makes the pairing orientable. When a surface S with one boundary component is tapped into the free group by a tight map f, a pairing p is induced on letters of w, the word to which the boundary component is mapped. Such a pairing is determined by the tight map f when we consider the preimage of f, which produces arcs on S whose endpoints correspond to elements of the free group. The pairing p is given by these pairs of endpoints. Given a pairing p and word w, there exists at most one surface/map pair such that f maps ∂S to w and induces the pairing p (up to homeomorphism), so it is reasonable to represent tight maps forms that depend only on the word and pairing.

The first such representation is $\Delta(w, p)$, defined in [1], which can be described as a directed graph. The set of verticies consists of ordered pairs $(i, p(i))$ for each of the i letters in w. The graph has an edge from $(i, p(i))$ to $(p(i+1), i+1)$ for each i. This constructions works since p is an involution, so $(p(i), i)$ is a vertex for every i. In terms of S and f, this $\Delta(w, p)$ represents the one-complex obtained by cutting S along the arcs of f^{-1} and identifying each boundary components created by these arcs with a point. These are the same arcs that gave us the pairing p, which is why we label them using the ordered pairs $(i, p(i))$. This cuts the surface into several pieces, which are are the components of $\Delta(w, p)$ and their edges correspond to the pairs $(i, p(i))$ and $(p(i), i)$. This gives us the expression $\chi(S) = d - \frac{|w|}{2}$, where d is the number of components in the graph. An example is shown in figure 1 for $xxy^{-1}x^{-1}yx^{-2}y^{-1}xyx^{-2}yxy^{-1}x$.

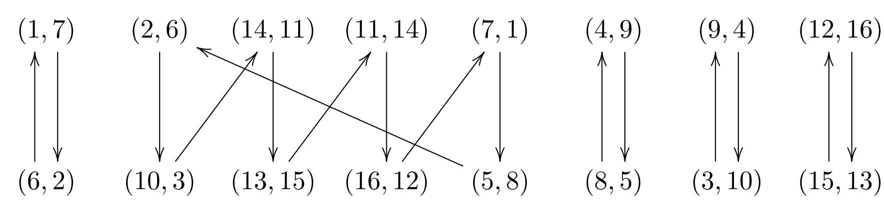

Figure 1. $\Delta(xxy^{-1}x^{-1}yx^{-2}y^{-1}xyx^{-2}yxy^{-1}x, p)$. Each vertex is an $i, p(i)$ pair, and p can be determined from these points (i.e., $p(1) = 7$, $p(2) = 6$, etc.) The graph has 5 components, so we conclude $\chi(S) = -3$ for this pairing, and thus $cl(xxy^{-1}x^{-1}yx^{-2}y^{-1}xyx^{-2}yxy^{-1}x) \leq 2$

.

The star graph $(\Sigma(w))$(introduced by Whitehead [2]) is a slightly different directed graph whose vertices are the letters in w with an edge from p to q for each instance of pq^{-1} in w. In this case, there is a correspondence between cycles of $\Sigma(w,p)$ and components of $\Delta(w,p)$, so we get a modified $\chi(S) = d - \frac{|w|}{2}$, where d is the number of cycles in the star graph. We can see this correspondence by identifying the vertices of $\Delta(w,p)$ whose first entry correspond to the same letter, as this produces the star graph for w. $\Sigma(xxy^{-1}x^{-1}yx^{-2}y^{-1}xyx^{-2}yxy^{-1}x)$ is shown in figure 2.

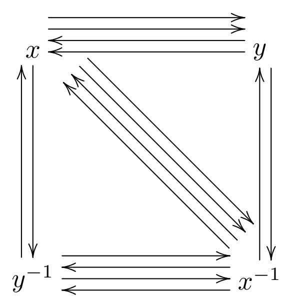

FIGURE 2. Star Graph of $xxy^{-1}x^{-1}yx^{-2}y^{-1}xyx^{-2}yxy^{-1}x$. The graph has 6 maximal cycles, so we conclude $\chi(S) = -2$, which tells us that the genus of S is greater than or equal to 1.5. Combined with the result from $\Delta(w,p)$ and the fact that commutator length must be an integer, we obtain $cl(xxy^{-1}x^{-1}yx^{-2}y^{-1}xyx^{-2}yxy^{-1}x) = 2$

The final representation is the fatgraph, a more direct visual representation of the boundary of a surface that has been partially deformed by a map into the free group, represented by the wedge product of 2 circles. The pairing is represented on rectangles that make up the fatgraph; there is one for each $(i, p(i))$ pair with i on one edge oriented in one direction and $p(i)$ on the other, running the opposite way. The fatgraph gives us another way to calculate the genus of S and thus $cl(w)$ via the Euler characteristic: $\chi(S) = J - R$ where J is the number of junctions in the fatgraph and R is the number of rectangles (this can be shown through triangulation of the fatgraph using two triangles for each rectangle and covering each junction with two fewer triangles than rectangles attached to it, which is done in [3]).

Using any one of these representations, we are able to see why $cl([x, y]^n) = \lfloor \frac{n}{2} \rfloor + 1$. The simplest to set up is the star graph. We label four vertices, x, y, x^{-1}, and y^{-1}, and run edges from x to y^{-1}, y to x, x^{-1} to y, and y^{-1} to x. This cycle repeats and touches x n times, once at the beginning and once for every subsequent commutator. Thus, we obtain n cycles, so $\chi(S) = n - 2n = -n$, since $|[x, y]^n| = 4n$. Since we assume S has one boundary component, we know that the genus is given by $\frac{-\chi(S)+1}{2} = \frac{n+1}{2}$. Genus and commutator length must be an integer, so this is equivalent to $\lfloor \frac{n}{2} \rfloor + 1$.

Alternately, we can generalize the construction of the fatgraph for $[x, y]^n$ by considering the minimum possible genus. Since $g = \frac{-\chi(S)+1}{2}$, and $\chi(S) = J - R$,

FIGURE 3. Fatgraph of $[x,y]^3$. $\chi(S) = 3-6 = -3$, so we obtain the result $cl([x,y]^3) \leq 3$, as the genus of the corresponding surface is 3.5. Equality will be a later result.

it is clear that we wish to minimize $\chi(S)$ and thus $cl([x,y]^n)$ by minimizing the number of rectangles that meet at each junction of the fatgraph. For a word of the form $[x,y]^n$, we see that the minimum number of rectangle ends that must be attached at each junction is 4. This can easily be demonstrated to be possible by drawing n of the junctions pictured and extending the rectangles to attach them together, being careful to make a continuous loop, and not multiple disconnected ones. Thus, we have $\chi(S) = \frac{4n}{4} - \frac{4n}{2} = -n$ (one rectangle for every two letters, and one junction for every two rectangles).
Again this implies $cl([x,y]^n) = \lfloor \frac{n}{2} \rfloor + 1$, since the genus must be an integer and we are able to explicitly construct the corresponding fatgraph in each case.

3. Culler's method for finding expressions of $cl(w)$ commutators

After proving this result, the next logical line of questioning is whether or not we can find formulas for expressing $[x,y]^n$, or any other $w \in F$, using the fewest possible commutators, a method for which is described in [1]. If we know the genus, k, of a word w, then we've already noted that there exists a surface S with genus k and one boundary element that is taken by a tight map f into the free group such that the boundary maps to w. As discussed earlier, this induces a pairing p on the letters of w, which, if we know the genus of w, we've likely already found. Using this pairing, we can construct another representation of S, this time as an $8k$ sided polygon. The sides alternate between letters of w, which represent arcs in the preimage of f, and arcs in ∂S. The polygon is initially constructed in several pieces that correspond to the components of $\Sigma(w,p)$ by beginning with a vertex i and drawing an edge from it to another vertex $p(i)$ (this edge represents w_i). This is continued by drawing another edge from $p(i)$ to $p(i)+1$ (this represents an

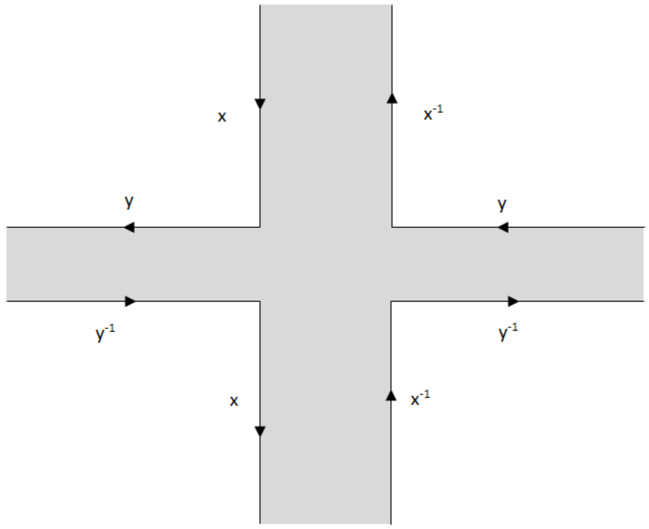

FIGURE 4. This junction is the smallest possible in terms of number of attached ends when constructing the fatgraph for $[x, y]^n$, since we must pair x with x^{-1} and y with y^{-1} and have the sequences of letters pictured.

arc in ∂S) and then continuing this pattern by drawing an edge from $p(i) + 1$ to $p(p(i) + 1)$ (which represents $w_{p(i)+1}$), alternating in this way between letters of w and unidentified pieces of the boundary. It is important to keep track of whether or not the construction is performed clockwise or counterclockwise. Also, rather than labeling inverses of letters, we identify the arcs of $f^{-1}(w)$ with the generator and an orientation: out of the polygon for the generator and into it for the inverse. This way, if we draw a path from the interior of the polygon across the arc, it will pass in "the right direction" for the generator and against the orientation for the inverse.

Gluing these components together by identifying common edges (being mindful to match their orientation) produces an $8k$ sided polygon. However, it might not be in the form we need to identify the k commutators whose product is w. This is obtained only when the identified sides of the polygon fit the pattern corresponding to the commutator: $a_1, b_1, a_1^{-1}, b_1^{-1}, a_2, ..., b_k^{-1}$. Note that these are just labels for the sides, based on whether the two vertices connected by a side increase in the direction of construction or against it. The actual words for the commutator are determined later. If the components cannot be glued together to fit this pattern, we "cut and paste" segments of the polygon, removing a section and gluing it back by identifying labeled edges, while keeping track of any arcs that run through the segment we moved. This produces two new unlabeled edges of the polygon and removes the two that are now stuck together, so we retain $8k$ sides.

Once the $8k$ sided polygon is arranged appropriately, we can identify the words by placing a basepoint on the surface of the polygon and tracing paths from that basepoint through each of the identified arcs, passing through each of them in order. That is, the path for the first word should pass through $(0, p(0))$ first, even if that is no longer an edge on the outside of the polygon, but an arc in the interior. Passing through edges on the outside "wraps around" and allows the path to return to the basepoint, producing a completed word. Each intersection the path makes with one of the labeled arcs represents a letter. If it crosses in the labeled direction, it is a generator, if it crosses against, it is an inverse.

Having thus identified the words in the commutators, we proceed around the edges in the direction of construction and obtain the k commutators.

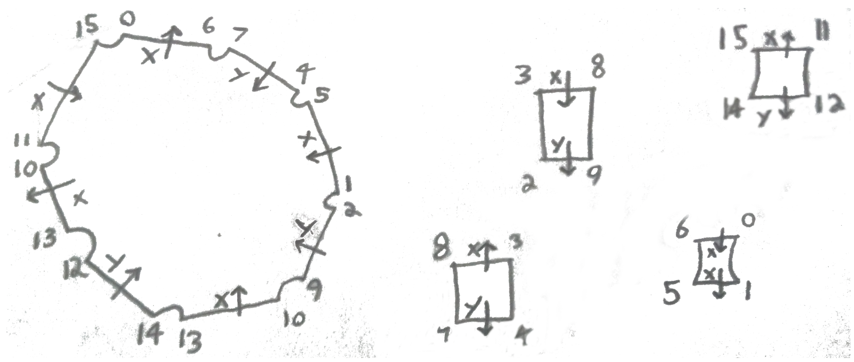

FIGURE 5. These are the polygons we get following the first steps of our procedure for $xxy^{-1}x^{-1}yx^{-2}y^{-1}xyx^{-2}yxy^{-1}x$. Note that each corresponds to a component from $\Delta(w,p)$.

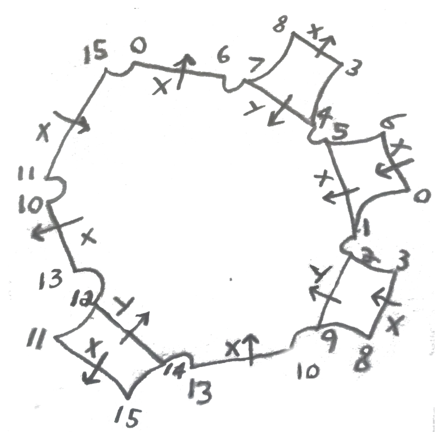

FIGURE 6. This shows the next step, where the polygons have been glued together by aligning similarly labeled edges.

4. Algebraic approach: Formulas for $[x,y]^n$

Surprisingly, all of the above findings can also be proven directly using only algebraic identities of the commutator. This does present a slight limitation of not

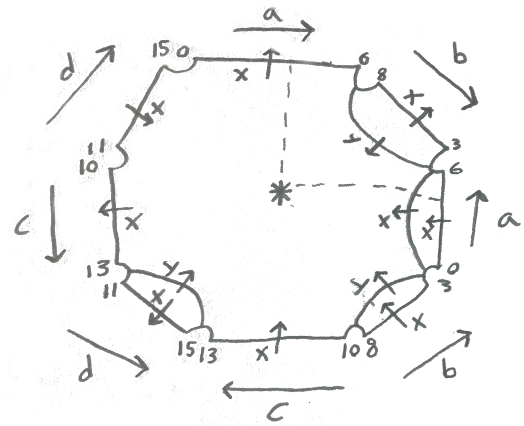

FIGURE 7. Finally we compress the glued polygon into one 8-sided figure, and trace loops from the basepoint to obtain $xxy^{-1}x^{-1}yx^{-2}y^{-1}xyx^{-2}yxy^{-1}x = [x^2, y^{-1}x^{-1}y][x^{-1}, x^{-1}y]$.

having a visual representation of the word or an exact idea of where the formulas come from. A greater problem is that these methods, outlined by Akhavan-Malayeri and Rhemtulla [4], do not give us a way to find an expression for any word in the free group in the minimum number of commutators, or even its commutator length, but only apply to $[x, y]^n$.

However, in this case, the algebraic results prove incredibly useful, giving us something that the geometric method could not: a general solution for the expression of $[x, y]^n$ in terms of $\lfloor \frac{n}{2} \rfloor + 1$ commutators. In fact, this result can even be extended to free groups with $2k$ generators and $([x_1, y_1][x_2, y_2]...[x_k, y_k])^n$, which can be expressed as $n(k-1) + \lfloor \frac{n}{2} \rfloor + 1$ commutators in that group.
Since the formula for new expressions of $[x, y]^n$ as a product of commutators gives an upper bound for $cl([x, y]^n)$, proving equality requires a lower bound:

Proposition 4.1. $cl([x, y]^n) \geq \lfloor \frac{n}{2} \rfloor + 1$

The proof requires the following definition, which is also used often in topological approaches to finding commutator length.

Definition 4.2. A *quasimorphism* is a function $f : G \to \mathbb{R}$ (for G a group) such that there exists $C \in \mathbb{R}$ with $|f(gh) - f(g) - f(h)| \leq C \; \forall g, h \in G$. We call the least upper bound of $|f(gh) - f(g) - f(h)|$ the *defect* of the quasimorphism.

Proof. For any word $m \in F$, let $h_m(w)$ be defined as the number of occurrences of m in w (without "wrapping around"). i.e. $h_{xy}(yxyyx) = 1$. Using this, let $f_m(w) = h_m(w) - h_{m^{-1}}(w)$.
It can easily be seen that f_m is a quasimorphism for all $m \in F$. Consider $m, g, h \in F$, and $|f_m(gh) - f_m(g)f_m(h)|$. If there is no cancellation when concatenating h and g, then the only difference can be copies of m or m^{-1} that straddle the gap between the two words, using letters from both h and g so that a subword appears in the concatenation that doesn't appear in h or g on its own. Such copies can only occur at most $l(m) - 1$ times, since each must have at least one letter in both h

and g to not have appeared in $f_m(g)$ or $f_m(h)$. If there is cancellation between h and g, then we can apply this same reasoning repeatedly by considering h and g to each be made of two words: h' or g', the component that doesn't cancel, and h_0 or g_0, which does. The difference between $f_m(hg)$ and $f_m(h)+f_m(g)$ thus comes from either copies of m or m^{-1} that use letters from both h' and g', which, as before, number less than or equal to $l(m)-1$, or copies of $m^{\pm 1}$ that straddle h' and h_0 (or g' and g_0). By the same reasoning, these two potential differences each contribute at most $l(m)-1$, so we obtain $|f_m(gh)-f_m(g)f_m(h)| \leq 3(l(m)-1)$. Note that we don't concern ourselves with $f_m(h_0)$ or $f_m(g_0)$, since we know $h_0 = g_0^{-1}$, and thus $f_m(h_0)+f_m(g_0) = 0$ and doesn't affect our calculations. Further, any linear combination of functions f_{m_i} is clearly a quasimorphism, since for the sum f, $|f(gh)-f(g)f(h)|$ is the absolute value of the sum of $f_{m_i}(gh)-f_{m_i}(g)f_{m_i}(h)$, which is bounded, as each component is.

This being established, let $f = f_{x^{-1}y^{-1}} + f_{y^{-1}x} + f_{xy} + f_{yx^{-1}}$ be a quasimorphism from the free group F on two generators to $\mathbb{R}$. We want to find an upper bound on the defect of f. As earlier, we can consider $h, g \in F$ and write $h = h'h_0$ and $g = g_0g' = h_0^{-1}g'$, where $hg = h'g'$ after cancellation. If we label the last letter of h and the first letters of h_0 and g, by a, b, and c we obtain the following identities:

$$f(h) = f(h') + f(h_0) + f(ab) \tag{4.3}$$

$$f(g) = f(g') - f(h_0) + f(b^{-1}c) \tag{4.4}$$

$$f(hg) = f(h') + f(g') + f(ac) \tag{4.5}$$

Thus, $f(hg)-f(h)-f(g) = f(ac)-f(ab)-f(b^{-1}c)$. Since each of the words (the four m_i) that determine the value of f have length two, each of $f(ac)$, $f(ab)$, and $f(b^{-1}c)$ have value either 1, 0, or -1. However, note that none of the m_i are squares. That is, each includes both a $x^{\pm 1}$ and a $y^{\pm 1}$. Therefore, it is not possible for all of ab, ac, and $b^{-1}c$ to be $m_i^{\pm 1}$ for some i, since if, for example, $a = x^{\pm 1}$ and $b = y^{\pm 1}$, then ab is some $m_i^{\pm 1}$, but one of ac or $b^{-1}c$ won't be, since c must be the "same letter" as either a or b, producing either a square or the identity, the value of f on which is 0. Therefore, $|f(hg)-f(h)-f(g)| \leq 2$ for all $h, g \in F$.

We now want to repeatedly apply this bound on the defect of f to find a bound for $|f(h_1h_2...h_n)-f(h_1)-f(h_2)-...-f(h_n)|$. Through induction, we see that it must be less than or equal to $n-1$ times the defect, so in the case of f, $2n$.
Also note that $f([x,y]^n) = n+n+n+n-1 = 4n-1$, since three of the m_i occur in $[x,y]$ and one straddles the product. Therefore, if we can express $[x,y]^n$ as a product of r commutators, then $f([a_1,b_1][a_2,b_2]...[a_r,b_r]) = 4n-1$. However, in combination with our results about the defect of f, we also know that $f(a_1b_1a_1^{-1}b_1^{-1}...a_rb_ra_r^{-1}b_r^{-1}) \leq 2(4r-1)$, since this is the upper bound we obtained earlier for the defect when multiplying $4r-1$ words, and, since $f(g) = -f(g^{-1})$, the sum $f(a_1)+f(a_1^{-1})+f(b_1)+...+f(b_1^{-1}) = 0$. Thus, since these two expressions are clearly equal, we obtain $4n-1 \leq 2(4r-1)$, and thus $r \geq \frac{n}{2}+\frac{1}{8}$. Since r must be an integer, this provides our result: $r \geq \lfloor \frac{m}{2} \rfloor + 1$. $\square$

With this lower bound in place, it is clear that a formula expressing $[x,y]^n$ using $\lfloor \frac{n}{2} \rfloor + 1$ commutators will prove equality. Doing that requires four commutator identities that be easily shown algebraically by expanding and simplifying the terms. (Note that for a word $w \in F$, w^x denotes xgx^{-1}). We express the identities in terms of several commutators in the free group.

$$
\begin{array}{ll}
A = [y, x^{-1}[y^{-1}, x]]^x & B = [x, y^2[x^{-1}, y^{-1}]]^{y^{-1}} \\
C = [y^{-1}, x^2[y, x^{-1}]]^{yx^{-1}} & D = [x^{-1}, y^{-2}[x, y]]^{yxy} \\
E = [y, x^{-2}[y^{-1}, x]]^{yxy^{-1}x} & F = [y^{-2}, yx] \\
G = [x^{-2}, xy^{-1}]^y & H = [y^2, y^{-1}x^{-1}]^{yx} \\
I = [x^2, x^{-1}y]^{yxy^{-1}} & J = B \cdot C \cdot D \cdot E \\
K = B \cdot G & L = B \cdot C \cdot H \\
M = B \cdot C \cdot DI &
\end{array}
$$

$$F[x,y]^2 = BG \tag{4.6}$$

$$G[x,y]^2 = CH \tag{4.7}$$

$$H[x,y]^2 = DI \tag{4.8}$$

$$I[x,y]^2 = EF^{[y,x]} \tag{4.9}$$

These identities are tedious to prove, but require only the basic properties of the commutator and conjugation. Having these, we only need to build on the formula $[x,y]^3 = AF$, which is actually equivalent to Culler's $[x,y]^3 = [xyx^{-1}, y^{-1}xyx^{-2}][y^{-2}, yx]$ (since $F = [y^{-2}, yx] = [y^{-1}xy, y^2]$). Note that we are looking to prove $cl([x,y]^n) = \lfloor \frac{n}{2} \rfloor + 1$, so we are really only concerned with finding formulas for odd values of n, seeing as we will still meet this minimum for even n by simply tacking an additional $[x,y]$ onto the beginning or end.

Upon examination of the four commutator identities, we notice that they each provide a way to reduce three commutators to two. Building on our formula for $[x,y]^3$ using these identities, a pattern slowly emerges:

$$[x,y]^3 = AF \tag{4.10}$$

$$[x,y]^5 = AF[x,y]^2 = ABG \tag{4.11}$$

$$[x,y]^7 = ABG[x,y]^2 = ABCH \tag{4.12}$$

$$[x,y]^9 = ABCH[x,y]^2 = ABCDI \tag{4.13}$$

$$[x,y]^{11} = ABCDI[x,y]^2 = ABCDEF^{[y,x]} = AJF^{[y,x]} \tag{4.14}$$

Since $[y,x]^{-1} = [x,y]$, this pattern continues with the last term conjugated by $[y,x]$.

$$[x,y]^{13} = AJF^{[y,x]}[x,y]^2 = AJ(F[x,y])^{[y,x]} = AJ(BG)^{[y,x]} = AJK^{[y,x]} \tag{4.15}$$

$$[x,y]^{15} = AJ(BG[x,y]^2)^{[y,x]} = AJ(BCH)^{[y,x]} = AJL^{[y,x]} \tag{4.16}$$

$$[x,y]^{17} = AJ(BCH[x,y]^2)^{[y,x]} = AJ(BCDI)^{[y,x]} = AJM^{[y,x]} \tag{4.17}$$

$$[x,y]^{19} = AJ(BCDI[x,y]^2)^{[y,x]} = AJ(BCDEF^{[y,x]})^{[y,x]} = AJJ^{[y,x]}F^{[y,x]^2}$$

$$(4.18)$$

Using a new piece of notation, we can describe this pattern as it continues for all values of n.
For $T = F, J, K, L$, or, M, let $T_i = T^{[y,x]^i}$. Then, continuing the pattern above, we see:

$$[x, y]^{8d+3} = AJ_0 \cdots J_{d-1}F_d \tag{4.19}$$
$$[x, y]^{8d+5} = AJ_0 \cdots J_{d-1}K_d \tag{4.20}$$
$$[x, y]^{8d+7} = AJ_0 \cdots J_{d-1}L_d \tag{4.21}$$
$$[x, y]^{8d+9} = AJ_0 \cdots J_{d-1}M_d \tag{4.22}$$

Expanding the multi commutator abbreviations, we see that each expression is indeed made up of $\lfloor \frac{n}{2} \rfloor + 1$ commutators, so these equations prove the result, using purely algebraic means.

Acknowledgments. I'd like to thank Henry Chan for providing guidance on my project and helping to revise this paper, as well as Alden Walker and Danny Calegari for helping me to understand the primary sources.

References

[1] Culler, Marc. "Using Surfaces to Solve Equations in Free Groups." *Topology.* (1981): 133-145.
[2] Whitehead, J.H.C. "On Certain Sets of Elements in a Free Group." *Pro. London Math. Soc.*, 41 (1936), pp. 48-56
[3] Walker, Alden. "Surfaces and Commutators." University of Chicago REU. Chicago. June-July 2013. Lecture series.
[4] Akhavan-Malayeri, Mehri, and Akbar Rhemtulla. "Products of Commutators in Free Groups." *International Journal of Algebra and Computation.* no. 2 (2003): 231-240.