NIELSEN-SCHREIER THEOREM

AABID SEEYAL ABDUL KHARIM

Abstract. This paper gives a proof of the Nielsen-Schreier theorem, that every subgroup of a free group is a free group itself. We do so by showing that any subgroup generated by an N-reduced subset of a free group is a free group itself, and that every subgroup of a free group can be generated by an N-reduced subset of the free group.

1. Introduction

In 1921, Nielsen proved, using a method of cancellation arguments, that every finitely generated subgroup of a free group is itself a free group. Schreier later proved that *every* subgroup of a free group is itself a free group, but with a somewhat different method. The latter is known as the Nielsen-Schreier theorem. In this paper, we give a proof of the Nielsen-Schreier Theorem, using a method similar to that used by Nielsen. To do so, we use the following theorem.

Theorem 1.1 ([LS], Proposition 1.9)**.** Let X be a subset of a group G. Then X is a basis for a free subgroup $\langle X\rangle$ of G if and only if no product $w = x_1...x_n$ is trivial, where $n \neq 1$, $x_i \in X^{\pm}$, and all $x_i x_{i+1} \neq 1$.

Remark 1.2. Note that [LS] does not mention in the statement of the theorem that the subgroup is $\langle X\rangle$, which we have included, but this is mentioned in the proof.

We use standard notation, as used in [LS], unless specifically mentioned or defined otherwise (like in Notation 1.3). We also use *words* to refer to both the elements of a free group and the words that represent those elements.

Notation 1.3. We denote $w = w_1...w_n$ *reduced* by $w = \overline{w_1...w_n}$ (note that this also means $|w| = |w_1|+...+|w_n|$). (Similarly, we also use this notation in the following manner: $w = w_1...w_x\overline{w_{x+1}w_{x+2}...w_n}$ when w_x does not cancel completely in $w_1...w_x$, and w_x and w_{x+1} don't cancel in $w_x w_{x+1}$, which means $|w| = |w_1...w_{x+1}| + |w_{x+2}| + ... + |w_n|$.)

In Section 2, we define N-reduced subsets of a free group, and prove that any subgroup generated by an N-reduced subset is a free group. In Section 3, using a well-order that we define, we prove that every subgroup of a free group can be generated by an N-reduced subset.

2. N-reduced subsets of Free Groups

Definition 2.1. Let x be a word. $L(x) := l$, where $x = \overline{lr}$ and $|l| = \frac{|x|+1}{2}$.

Remark 2.2. $L(x)$ is just the left half of x. Note that $L(x^{-1})$ is just the inverse of the right half of x.

Definition 2.3. Let $X \subseteq$ free group F. X is said to be *N-reduced*, if and only if, for every $x, y, z \in X^{\pm}$,:

(N0) $x \neq 1$
(N1) $xy \neq 1$ implies $|xy| \geq |x|, |y|$
(N2) $xy \neq 1$ and $yz \neq 1$ implies $|xyz| > |x| - |y| + |z|$

Remark 2.4. If the initial condition is met, (N1) means, given $x = \overline{ap^{-1}}$, $y = \overline{ab}$, $xy = \overline{pb}$, that $|p| \leq |L(x)|, |L(y)|$, i.e., not more than half of each word is cancelled.
Additionally, it means, given $x = \overline{ap^{-1}}$, $y = \overline{pbq^{-1}}$, $z = \overline{q^{-1}c}$ (where q and p are the largest cancellations), if (N2) is also satisfied, then $b \neq 1$ and $xyz = \overline{abc}$, i.e., some part of y remains uncancelled, otherwise $b = 1$ (, since, $|xyz| = |x| - |y| + |z| + |b| > |x| - |y| + |z|$ does not hold), $xy = \overline{aq^{-1}}$ and $yz = \overline{pc}$, where $p \neq q$, and $|p| = |q| \leq |L(x)|, |L(z)|$ (by (N1)).

Lemma 2.5. *Let $X \subseteq$ free group F be N-reduced, if $w = x_1...x_n$, $n \geq 3$, $x_i \in X^{\pm}$, and all $x_i x_{i+1} \neq 1$ then $w = x_1...a\overline{bc}$, $|b|, |c| > 0$ (equivalently, $|w| = |x_1...a| + |bc|$, $|bc| > 0$) where $x_{n-2} = \overline{ap^{-1}}$, $x_{n-1} = \overline{pbq^{-1}}$, $x_n = \overline{qc}$, and $x_{n-2}x_{n-1}x_n = \overline{abc}$.*

Proof. Let $w = w = x_1...x_n$, $n \geq 3$.
Since all $x_i x_{i+1} \neq 1$, by (N1) and (N2), $x_{n-2} = \overline{ap^{-1}}$, $x_{n-1} = \overline{pbq^{-1}}$, $x_n = \overline{q^{-1}c}$, such that $x_{n-2}x_{n-1}x_n = \overline{abc}$ and $|b| > 0$.
By (N0), $|q^{-1}c| > 0$, and, by (N1), $|q^{-1}| \leq |L(q^{-1}c)|$, which implies $|c| > 0$.
(Similarly, $|a| \geq |L(q^{-1}c)| > 0$, and a does not cancel completely in $x_1...a$ when $n > 3$ by (N2).)
Therefore, since $x_{n-2}x_{n-1} = \overline{ab}$ and $x_{n-1}x_n = \overline{bc}$, $w = x_1...a\overline{bc}$, where $|b|, |c| > 0$. ∎

Lemma 2.6. *Let $X \subseteq$ free group F be N-reduced, then no product $w = x_1...x_n$ is trivial, where $n \geq 1$, $x_i \in X^{\pm}$, and all $x_i x_{i+1} \neq 1$.*

Proof. Case: $n = 1$
By (N0), $w = x_1 \neq 1$.
Case: $n = 2$
Since all $x_i x_{i+1} \neq 1$, this means $w = x_1 x_2 \neq 1$.
Case: $n \geq 3$
Since X is N-reduced, $n \geq 3$, $x_i \in X^{\pm}$, and all $x_i x_{i+1} \neq 1$, by Lemma 2.5, $|w| = |x_1...x_{n-2}a| + |bc|$ where $|bc| > 0$, which implies, $|w| > 0$, which means $w \neq 1$. ∎

Lemma 2.7. *Let $X \subseteq$ free group F. If X is N-reduced, then $\langle X \rangle$ is a free group.*

Proof. Assume X is N-reduced.
Therefore, by Lemma 2.6, no product $w = x_1...x_n$ is trivial, where $n \geq 1$, $x_i \in X^{\pm}$, and all $x_i x_{i+1} \neq 1$, which means, by Theorem 1.1, $\langle X \rangle$ is a free group. ∎

3. Nielsen-Schreier theorem

Notation 3.1. We denote a lexicographic order, sometimes, more specifically, referred to as a shortlex order, on (the reduced form of) words, using $<_{\text{lex}}$.

Remark 3.2. Note that if $p <_{\text{lex}} q$ and $|pa| = |qb|$ then $pa <_{\text{lex}} qb$.

Definition 3.3. We define a well-ordering on sets of words $\{w_1, w_1^{-1}\} <_{L^\pm} \{w_2, w_2^{-1}\}$ if and only if:

$$\min(L(w_1), L(w_1^{-1})) <_{\text{lex}} \min(L(w_2), L(w_2^{-1})), \text{ or}$$
$$\text{(if they are equal) } \max(L(w_1), L(w_1^{-1})) <_{\text{lex}} \max(L(w_2), L(w_2^{-1}))$$

Notation 3.4. Since xRw if and only if $x = w^{-1}$ is an equivalence relation, we can (and will) denote $\{w_1, w_1^{-1}\} <_{L^\pm} \{w_2, w_2^{-1}\}$ using $w_1 <_{L^\pm} w_2$ without ambiguity.

Remark 3.5. Note that since $<_{L^\pm}$ is defined on $<_{\text{lex}}$, if $|p| < |q|$ then $p <_{L^\pm} q$.

(It may be more useful to refer to the following two lemmas while reading the proof of Lemma 3.11).

Lemma 3.6. Let words $x_1 = \overline{pc}$, $x_2 = \overline{qc}$, where $|p| = |q| \leq |L(x_1)|, |L(x_2)|$, if $p <_{\text{lex}} q$ then $x_1 <_{L^\pm} x_2$.

Proof. Assume $p <_{\text{lex}} q$.
Since, $|p| = |q| \leq |L(x_1)|, |L(x_2)|$, let $L(x_1^{-1}) = L(x_2^{-1}) = \overline{c_2^{-1}}$ and $L(x_1) = \overline{pc_1}$, $L(x_2) = \overline{qc_1}$.
<u>Case: $c_2^{-1} <_{\text{lex}} pc_1$</u>
Then $c_2^{-1} = \min(pc_1, c_2^{-1}) = \min(qc_1, c_2^{-1})$.
Therefore, $pc_1 = \max(pc_1, c_2^{-1}) <_{\text{lex}} \max(qc_1, c_2^{-1}) = qc_1$, since $p <_{\text{lex}} q$, means $x_1 <_{L^\pm} x_2$.
<u>Case: $pc_1 <_{\text{lex}} c_2^{-1}$</u>
If $c_2^{-1} <_{\text{lex}} qc_1$, then $pc_1 = \min(pc_1, c_2^{-1}) <_{\text{lex}} \min(qc_1, c_2^{-1}) = c_2^{-1}$, which means $x_1 <_{L^\pm} x_2$.
Otherwise, $qc_1 <_{\text{lex}} c_2^{-1}$, which implies,
$pc_1 = \min(pc_1, c_2^{-1}) <_{\text{lex}} \min(qc_1, c_2^{-1}) = qc_1$, since $p <_{\text{lex}} q$, which means $x_2 <_{L^\pm} x_1$.
Thus, $x_2 <_{L^\pm} x_1$. □

Lemma 3.7. Let words $x_1 = \overline{ap^{-1}}$, $x_2 = \overline{aq^{-1}}$, where $|p^{-1}| = |q^{-1}| \leq |L(x_1)|, |L(x_2)|$, if $q <_{\text{lex}} p$ then $x_2 <_{L^\pm} x_1$.

Proof. Assume $q <_{\text{lex}} p$.
Since, $|p^{-1}| = |q^{-1}| \leq |L(x_1)|, |L(x_2)|$, let $L(x_1) = L(x_2) = \overline{a_1}$ and $L(x_1^{-1}) = \overline{pa_2{}^{-1}}$, $L(x_2^{-1}) = \overline{qa_2{}^{-1}}$.
<u>Case: $a_1 <_{\text{lex}} qa_2{}^{-1}$</u>
Then $a_1 = \min(pa_2{}^{-1}, a_1) = \min(a_1, qa_2{}^{-1})$.
Therefore, $qa_2{}^{-1} = \max(a_1, qa_2{}^{-1}) <_{\text{lex}} \max(a_1, pa_2{}^{-1}) = pa_2{}^{-1}$, since $q <_{\text{lex}} p$, means $x_2 <_{L^\pm} x_1$.
<u>Case: $qa_2{}^{-1} <_{\text{lex}} a_1$</u>

If $a_1 <_{\text{lex}} p{a_2}^{-1}$, then $q{a_2}^{-1} = \min(a_1, q{a_2}^{-1}) <_{\text{lex}} \min(a_1, p{a_2}^{-1}) = a_1$, which means $x_2 <_{L^{\pm}} x_1$.
Otherwise, $p{a_2}^{-1} <_{\text{lex}} a_1$, which implies,
$q{a_2}^{-1} = \min(a_1, q{a_2}^{-1}) <_{\text{lex}} \min(a_1, p{a_2}^{-1}) = p{a_2}^{-1}$, since $q <_{\text{lex}} p$, which means $x_2 <_{L^{\pm}} x_1$. □

Definition 3.8. Let G be a subgroup of the free group F.

(1) For each $g \in G$, $G_g := \langle \{h \in G \mid h <_{L^{\pm}} g\} \rangle$.
(2) $X_G := \{g \in G \mid g \notin G_g\}$

Lemma 3.9. Let G be a subgroup of the free group F. $\langle X_G \rangle = G$

Proof. Assume $\langle X_G \rangle \neq G$, for the sake of contradiction, meaning $\exists x \in G$ such that $x \notin \langle X_G \rangle$. Let g be the least such element (i.e. $g \in G/\langle X_G \rangle$ and $\forall x \in G/\langle X_G \rangle \wedge x \neq g.\ g <_{L^{\pm}} x$).

Therefore, for all $h <_{L^{\pm}} g$ in G, $h \in \langle X_G \rangle$, and since $g \notin \langle X_G \rangle$ (by definition of g), $g \notin \langle \{h \in G \mid h <_{L^{\pm}} g\} \rangle$, which means $g \notin G_g$. This implies, by definition of X_G, $g \in X_G$, implying $g \in \langle X_G \rangle$, which is a contradiction.

Thus, $\langle X_G \rangle = G$. □

Lemma 3.10. Let G be a subgroup of the free group F, and $x, y \in G, x \neq y$. If $x <_{L^{\pm}} y$ and $xy <_{L^{\pm}} y$, then $y \notin X_G$.

Proof. Assume $x <_{L^{\pm}} y$.
Let $x, y \in G$, then $xy \in G$. Additionally, $(x^{-1})(xy) = y$. This implies $y \in \langle \{x, xy\} \rangle$, which implies $y \in \langle \{h \in G \mid h <_{L^{\pm}} g\} \rangle$, since $x <_{L^{\pm}} y$ and $xy <_{L^{\pm}} y$, which means $y \in G_y$.

Thus, by definition of X_G, $y \notin X_G$. □

Lemma 3.11. Let G be a subgroup of the free group F. X_G is N-reduced

Proof. $1 \in G$, and, $1 \in G_g$, for all $g \in G$, which implies $1 \in G_1$, which means $1 \notin X_G$.
Therefore, X_G satisfies (N0).

Assume X_G does not satisfy (N1), for the sake of contradiction, which means $\exists x, y \in X_G, xy \neq 1$ such that $|xy| < |x|$ or $|xy| < |y|$. $x \neq y$ (, since $|xx| < x$ is not possible for elements in a free group, and $x \in F$), and we show that in every case, there is a contradiction.

Case: $x <_{L^{\pm}} y$

This implies, $|xy| < |x| \leq |y|$, which means $xy <_{L^{\pm}} y$. Since $x <_{L^{\pm}} y$ and $xy <_{L^{\pm}} y$, by Lemma 3.10, $y \notin X_G$, which is a contradiction.

Case: $y <_{L^{\pm}} x$

This implies, $|xy| < |y| \leq |x|$, which means $xy <_{L^{\pm}} x$. Since $y <_{L^{\pm}} x$ and $xy <_{L^{\pm}} x$, by Lemma 3.10, $x \notin X_G$, which is a contradiction.
Therefore, X_G satisfies (N1).

Assume X_G does not satisfy (N2), for the sake of contradiction, which means $\exists x, y, z \in X_G, xy \neq 1 \wedge yz \neq 1$ such that $|xyz| \leq |x| - |y| + |z|$. This means, since X_G satisfies (N1), which means $|xy| \geq |y|$ and $|yz| \geq |z|$, we have $x = \overline{ap^{-1}}$, $y = \overline{pq^{-1}}$, $z = \overline{qc}$, such that $xy = \overline{aq^{-1}}$ and $yz = \overline{pc}$. $p \neq q$, (, and note $|p| = |q| \leq |L(x)|, |L(z)|$, which also implies $x \neq y \neq z$), and we show that in every case, there is a contradiction.

Case: $p <_{\text{lex}} q$

$y = \overline{pq^{-1}}$, $z = \overline{qc}$, which means $yz = \overline{pc}$, which implies, since $p <_{\text{lex}} q$, by Lemma 3.6, $yz <_{L^{\pm}} z$.
$y <_{L^{\pm}} z \vee z <_{L^{\pm}} y$, therefore, since $yz <_{L^{\pm}} z$, by Lemma 3.10, this implies $z \notin X_G \vee y \notin X_G$, which is a contradiction.
Case: $q <_{\text{lex}} p$
$x = \overline{ap^{-1}}$, $y = \overline{pq^{-1}}$, which means $xy = \overline{aq^{-1}}$, which implies, since $q <_{\text{lex}} p$, by Lemma 3.7, $xy <_{L^{\pm}} x$.
$y <_{L^{\pm}} x \vee x <_{L^{\pm}} y$, therefore, since $xy <_{L^{\pm}} x$, by Lemma 3.10, this implies $x \notin X_G \vee y \notin X_G$, which is a contradiction.
Therefore, X_G satisfies (N2).

Thus, X_G is N-reduced. □

Lemma 3.12 (Nielsen-Schreier theorem)**.** Every subgroup of a free group is a free group itself.

Proof. Let G be a subgroup of a free group F.
Therefore, $G = \langle X_G \rangle$, by Lemma 3.9, and X_G is N-reduced, by Lemma 3.11.
Thus, by Lemma 2.7, G is a free group. □

References

[LS] Lyndon, R. C., & Schupp, P. E. (2001). *Combinatorial Group theory.* Berlin, Heidelberg, New York: Springer. 1