The Nielsen-Schreier Theorem in Homotopy Type Theory

Andrew W Swan

Carnegie Mellon University

July 17, 2021

Theorem (Nielsen-Schreier)

Every subgroup of a free group is itself a free group.

Theorem (Nielsen-Schreier)

Every subgroup of a free group is itself a free group.

- Original direct proofs were long and unintuitive.

Theorem (Nielsen-Schreier)

Every subgroup of a free group is itself a free group.

- Original direct proofs were long and unintuitive.
- Later proofs e.g. by Baer-Levi and Chevalley-Herbrand use ideas from algebraic topology to provide easier to understand proofs.

Theorem (Nielsen-Schreier)

Every subgroup of a free group is itself a free group.

- Original direct proofs were long and unintuitive.
- Later proofs e.g. by Baer-Levi and Chevalley-Herbrand use ideas from algebraic topology to provide easier to understand proofs.
- In HoTT we can use ideas from algebraic topology without needing to develop the theory of topological spaces and fundamental groups, resulting in a proof that is both intuitive and easy to formalise.

Definition

A *group* is a pointed type $(BG, \mathtt{base})$ such that BG is 1-truncated and connected.
A *homomorphism* $(BG, \mathtt{base}_G) \to (BH, \mathtt{base}_H)$ is a pointed map.

Definition

A *group* is a pointed type $(BG, \mathtt{base})$ such that BG is 1-truncated and connected.
A *homomorphism* $(BG, \mathtt{base}_G) \to (BH, \mathtt{base}_H)$ is a pointed map.

Note that the identity type $\mathtt{base} =_{BG} \mathtt{base}$ is a set, and has an binary operation given by path concatenation.

Theorem (Buchholtz-Van Doorn-Rijke)

The category of groups is equal to the category of sets with associative binary operation with inverses and identity (groups in the more traditional sense).

Definition (Favonia-Harper)

Let $(BG, \mathtt{base})$ be a group. A *covering space* of $(BG, \mathtt{base})$ is a map $X : BG \to \mathbf{hSet}$.

We can use covering spaces to understand subgroups:

Definition (Favonia-Harper)

Let $(BG, \mathtt{base})$ be a group. A *covering space* of $(BG, \mathtt{base})$ is a map $X : BG \to \mathbf{hSet}$.

We can use covering spaces to understand subgroups:

A *pointed covering space* is a covering space $X : BG \to \mathbf{hSet}$ together with a point $x_0 : X(\mathtt{base})$.

Definition (Favonia-Harper)

Let $(BG, \mathtt{base})$ be a group. A *covering space* of $(BG, \mathtt{base})$ is a map $X : BG \to \mathbf{hSet}$.

We can use covering spaces to understand subgroups:

A *pointed covering space* is a covering space $X : BG \to \mathbf{hSet}$ together with a point $x_0 : X(\mathtt{base})$.

A covering space $X : BG \to \mathbf{hSet}$ is *connected* if the total space $\sum_{z:BG} X(z)$ is a connected type.

Definition (Favonia-Harper)

Let $(BG, \mathtt{base})$ be a group. A *covering space* of $(BG, \mathtt{base})$ is a map $X : BG \to \mathbf{hSet}$.

We can use covering spaces to understand subgroups:

A *pointed covering space* is a covering space $X : BG \to \mathbf{hSet}$ together with a point $x_0 : X(\mathtt{base})$.

A covering space $X : BG \to \mathbf{hSet}$ is *connected* if the total space $\sum_{z:BG} X(z)$ is a connected type.

A *subgroup* of $(BG, \mathtt{base})$ is a pointed connected covering space.

Definition (Favonia-Harper)

Let $(BG, \mathtt{base})$ be a group. A *covering space* of $(BG, \mathtt{base})$ is a map $X : BG \to \mathbf{hSet}$.

We can use covering spaces to understand subgroups:

A *pointed covering space* is a covering space $X : BG \to \mathbf{hSet}$ together with a point $x_0 : X(\mathtt{base})$.

A covering space $X : BG \to \mathbf{hSet}$ is *connected* if the total space $\sum_{z:BG} X(z)$ is a connected type.

A *subgroup* of $(BG, \mathtt{base})$ is a pointed connected covering space.

The *underlying group* of a subgroup is the total space $\sum_{z:BG} X(z)$ together with the point $(\mathtt{base}, x_0)$.

Definition (Favonia-Harper)

Let $(BG, \mathtt{base})$ be a group. A *covering space* of $(BG, \mathtt{base})$ is a map $X : BG \to \mathbf{hSet}$.

We can use covering spaces to understand subgroups:

A *pointed covering space* is a covering space $X : BG \to \mathbf{hSet}$ together with a point $x_0 : X(\mathtt{base})$.

A covering space $X : BG \to \mathbf{hSet}$ is *connected* if the total space $\sum_{z:BG} X(z)$ is a connected type.

A *subgroup* of $(BG, \mathtt{base})$ is a pointed connected covering space.

The *underlying group* of a subgroup is the total space $\sum_{z:BG} X(z)$ together with the point $(\mathtt{base}, x_0)$.

We refer to the set $X(\mathtt{base})$ as the *index* of the subgroup.

Definition (Kraus-Altenkirch)

For any set A the free group on A is the higher inductive type BF_A defined as follows:

1. BF_A contains a point $\mathtt{base}$
2. For every $a : A$, there is a path $\mathtt{loop}(a) : \mathtt{base} =_{BF_A} \mathtt{base}$
3. 1-truncation

Definition (Kraus-Altenkirch)

For any set A the free group on A is the higher inductive type BF_A defined as follows:

1. BF_A contains a point $\mathtt{base}$
2. For every $a : A$, there is a path $\mathtt{loop}(a) : \mathtt{base} =_{BF_A} \mathtt{base}$
3. 1-truncation

They showed this satisfies the usual universal property for free groups:

$$\begin{tikzcd} A \arrow[r, "i"] \arrow[dr, "\forall f"'] & \mathtt{base} =_{BF_A} \mathtt{base} \arrow[d, dotted, "\mathtt{ap}_h(\mathtt{base})"] \\ & \mathtt{base} =_{BG} \mathtt{base} \end{tikzcd} \qquad \begin{tikzcd} (BF_A, \mathtt{base}) \arrow[d, dotted, "\exists! h"] \\ (BG, \mathtt{base}) \end{tikzcd}$$

We can now see the HoTT formulation of the Nielsen-Schreier theorem:

Theorem

Let A be a set and let $X : BF_A \to \mathbf{hSet}$ be a subgroup of the free group $(BF_A, \mathtt{base})$ (with point x_0).

Then the underlying group of the subgroup, $\sum_{z:BF_A} X(z)$ is merely equivalent to the free group BF_B for some set B.

We will see a constructive proof when the index $X(\mathtt{base})$ of the subgroup is finite, which has also been formalised in Agda. The full version requires the axiom of choice.

Definition

A *graph* is a pair of sets E, V, together with a pair of maps $\pi_0, \pi_1 : E \to V$. We refer to the elements of V as *vertices*, the elements of E as *edges* and for each edge $e : E$ we call $\pi_0(e)$ and $\pi_1(e)$ the *endpoints* of e.

Definition

A *graph* is a pair of sets E, V, together with a pair of maps $\pi_0, \pi_1 : E \to V$. We refer to the elements of V as *vertices*, the elements of E as *edges* and for each edge $e : E$ we call $\pi_0(e)$ and $\pi_1(e)$ the *endpoints* of e.

The *coequalizer* of a graph $E \rightrightarrows V$ is the higher inductive type V/E generated as follows.

1. For each vertex $v : V$, V/E contains a point $[v] : V/E$.
2. For each edge $e : E$, V/E contains a path $\mathrm{edge}(e) : [\pi_0(e)] = [\pi_1(e)]$.

Definition

A *graph* is a pair of sets E, V, together with a pair of maps $\pi_0, \pi_1 : E \to V$. We refer to the elements of V as *vertices*, the elements of E as *edges* and for each edge $e : E$ we call $\pi_0(e)$ and $\pi_1(e)$ the *endpoints* of e.

The *coequalizer* of a graph $E \rightrightarrows V$ is the higher inductive type V/E generated as follows.

1. For each vertex $v : V$, V/E contains a point $[v] : V/E$.
2. For each edge $e : E$, V/E contains a path $\mathrm{edge}(e) : [\pi_0(e)] = [\pi_1(e)]$.

We will refer to the 1-truncation of the coequalizer, $\|V/E\|_1$ as the *geometric realization* of the graph.

Definition

A *graph* is a pair of sets E, V, together with a pair of maps $\pi_0, \pi_1 : E \to V$. We refer to the elements of V as *vertices*, the elements of E as *edges* and for each edge $e : E$ we call $\pi_0(e)$ and $\pi_1(e)$ the *endpoints* of e.

The *coequalizer* of a graph $E \rightrightarrows V$ is the higher inductive type V/E generated as follows.

1. For each vertex $v : V$, V/E contains a point $[v] : V/E$.
2. For each edge $e : E$, V/E contains a path $\mathrm{edge}(e) : [\pi_0(e)] = [\pi_1(e)]$.

We will refer to the 1-truncation of the coequalizer, $\|V/E\|_1$ as the *geometric realization* of the graph.

In particular for any A, BF_A is the geometric realization of a graph with one vertex and an edge for each element of A.

The proof of the Nielsen-Schreier theorem proceeds in two steps:

1. For any subgroup of a free group, the underlying group is the geometric realization of a graph.
2. Under certain assumptions the geometric realization of a graph is a free group.

As a special case of flattening for coequalizers, we have the following lemma:

Lemma

Let $E \rightrightarrows V$ be a graph and $X : V/E \to \mathbf{Type}$ a family of types on its coequalizer. We define a graph $E_X \rightrightarrows V_X$ as follows:

$$
\begin{aligned}
V_X &:= \sum_{v:V} X([v]) \\
E_X &:= \sum_{e:E} X([\pi_0(e)]) \\
\pi_0(e,x) &:= (\pi_0(e), x) \\
\pi_1(e,x) &:= \mathtt{edge}(e)_*(x)
\end{aligned}
$$

Then $\sum_{z:V/E} X(z) \simeq V_X/E_X$.

Applying to the graph $A \rightrightarrows 1$ and "1-truncating" we get the first part of the Nielsen-Schreier theorem:

Theorem

Let A be a set, $(BF_A, \mathtt{base})$ the free group on A and $X : BF_A \to \mathbf{hSet}$ a covering space on $(BF_A, \mathtt{base})$. Then we have the following equivalence:

$$\sum_{z:BF_A} X(z) \simeq \|X(\mathtt{base})/(A \times X(\mathtt{base}))\|_1$$

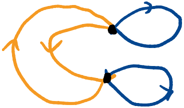
Applying to the graph $A \rightrightarrows 1$ and "1-truncating" we get the first part of the Nielsen-Schreier theorem:

Theorem

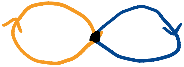
Let A be a set, $(BF_A, \mathtt{base})$ the free group on A and $X : BF_A \to \mathbf{hSet}$ a covering space on $(BF_A, \mathtt{base})$. Then we have the following equivalence:

$$\sum_{z:BF_A} X(z) \simeq \|X(\mathtt{base})/(A \times X(\mathtt{base}))\|_1$$

E.g. Here is an index 2 subgroup for $A = \{a, b\}$



$\sum_{z:BF_A} X(z)$



BF_A

We now need to show that the geometric realization of a graph is a free group. For this we need a bit more graph theory. We note that we can naturally formulate some important concepts in graph theory using the geometric realization.

Let $E \rightrightarrows V$ be a graph.

Definition

Given $v, v' : V$, a *path* from v to v' is an element of $[v] = [v']$ in the geometric realization.

Let $E \rightrightarrows V$ be a graph.

Definition

Given $v, v' : V$, a *path* from v to v' is an element of $[v] = [v']$ in the geometric realization.

$E \rightrightarrows V$ is *connected* if its geometric realization is a connected type. I.e. it is merely inhabited and there merely exists a path from any vertex to any other vertex.

Let $E \rightrightarrows V$ be a graph.

Definition

Given $v, v' : V$, a *path* from v to v' is an element of $[v] = [v']$ in the geometric realization.

$E \rightrightarrows V$ is *connected* if its geometric realization is a connected type. I.e. it is merely inhabited and there merely exists a path from any vertex to any other vertex.

$E \rightrightarrows V$ is a *tree* if its geometric realization is contractible. Equivalently the geometric realization is connected and 0-truncated. I.e. the graph is connected and any *cycle* (path from a vertex to itself) is trivial.

Let $E \rightrightarrows V$ be a graph.

Definition

Given $v, v' : V$, a *path* from v to v' is an element of $[v] = [v']$ in the geometric realization.

$E \rightrightarrows V$ is *connected* if its geometric realization is a connected type. I.e. it is merely inhabited and there merely exists a path from any vertex to any other vertex.

$E \rightrightarrows V$ is a *tree* if its geometric realization is contractible. Equivalently the geometric realization is connected and 0-truncated. I.e. the graph is connected and any *cycle* (path from a vertex to itself) is trivial.

A *spanning tree* is an embedding $E' \hookrightarrow E$ with decidable image such that the graph $E' \rightrightarrows V$ is a tree.

Lemma

If a graph has a spanning tree then its geometric realization is equivalent to a free group.

Intuitively we contract the spanning tree down to a point, leaving the remaining edges as loops from the point to itself. Formally, since E' is decidable, it has a complement $\neg E'$, and we can compute as follows.

$$\begin{aligned} V/E &\simeq V/(E' + \neg E') \\ &\simeq (V/E')/\neg E' \\ &\simeq 1/\neg E' \end{aligned}$$

Finally we need to construct the spanning tree. This uses the following key lemma.

Lemma

Let $E \rightrightarrows V$ be a connected graph, where V decomposes as a coproduct of inhabited types $V \simeq V_0 + V_1$. Then there merely exists an edge $e : E$ such that $\pi_0(e)$ and $\pi_1(e)$ lie in different components of V.

To illustrate the proof we assume the law of excluded middle (the constructive proof is no longer but slightly less intuitive).

Proof.

The partition $V \simeq V_0 + V_1$ determines a "colouring" $c : V \to 2$. Assume for a contradiction that there is no edge e with $\pi_0(e)$ and $\pi_1(e)$ lying in different components of V.

Proof.

The partition $V \simeq V_0 + V_1$ determines a "colouring" $c : V \to 2$. Assume for a contradiction that there is no edge e with $\pi_0(e)$ and $\pi_1(e)$ lying in different components of V. Then for all e we have $c(\pi_0(e)) = c(\pi_1(e))$. Hence c extends to a function c' on V/E:

$$\begin{tikzcd} V \arrow[r, "c"] \arrow[d, "{[-]}"'] & 2 \\ V/E \arrow[ur, dotted, "c'"'] & \end{tikzcd}$$

Proof.

The partition $V \simeq V_0 + V_1$ determines a "colouring" $c : V \to 2$. Assume for a contradiction that there is no edge e with $\pi_0(e)$ and $\pi_1(e)$ lying in different components of V. Then for all e we have $c(\pi_0(e)) = c(\pi_1(e))$. Hence c extends to a function c' on V/E:

$$\begin{tikzcd} V \arrow[r, "c"] \arrow[d, "{[-]}"'] & 2 \\ V/E \arrow[ur, dotted, "c'"'] & \end{tikzcd}$$

We assumed that both components of V are inhabited. Let $v_0, v_1 : V$ be such that $c(v_0) = 0$ and $c(v_1) = 1$. By connectedness, there merely exists a path $[v_0] = [v_1]$. But then we have $c(v_0) = c'([v_0]) = c'([v_1]) = c(v_1)$, giving a contradiction. $\square$

Lemma

Let $E \rightrightarrows V$ be a connected graph and suppose that either of the following conditions.

1. *V is finite and E has decidable equality.*
2. *The axiom of choice holds.*

Then $E \rightrightarrows V$ has a spanning tree.

In both cases we build up the spanning tree in stages by "iterating" the key lemma.

Finally combining the lemma with the first part of the theorem we get the full theorem:

Theorem

Suppose that A is a set and $X : BF_A \to \mathbf{hSet}$ a subgroup, and that either of the following conditions holds.

1. *A has decidable equality and the index $X(\mathtt{base})$ is finite*
2. *the axiom of choice*

Then the underlying group $\sum_{z:BF_A} X(z)$ is equivalent to a free group.

1. Basic ideas in group theory and graph theory can be naturally formulated in homotopy type theory, making essential use of higher inductive types and univalence.
2. The finite index version of the Nielsen-Schreier theorem has a completely constructive proof in HoTT and the full version can be proved using AC.
3. AC is strictly necessary: there is a boolean ∞-topos where it is false, the "∞-Schanuel topos".

For more details see the paper:
Swan, *On the Nielsen-Schreier theorem in homotopy type theory*, arXiv:2010.01187

and the Agda formalisation:
`https://github.com/awswan/nielsenschreier-hott`.

Thank you for your attention!