**Free Groups**

## 1. Existence of Free Groups

Let $\mathscr{A}$ be a set, called an **alphabet**. For each $a \in \mathscr{A}$, let the symbol $a^{-1}$ be called the **inverse** of $a$. For simplicity, we assume that for all $a \in \mathscr{A}$, $a^{-1} \notin \mathscr{A}$. We let $a^{+1} = (a^{-1})^{-1} = a$, for each $a \in \mathscr{A}$. Let $\mathscr{A}^{-1} = \{a^{-1} : a \in \mathscr{A}\}$. A **word** in $\mathscr{A} \cup \mathscr{A}^{-1}$ is a finite sequence of elements of $\mathscr{A} \cup \mathscr{A}^{-1}$. The **empty word** is the sequence with no terms. Let $\mathscr{W}$ be the set of all words in $\mathscr{A} \cup \mathscr{A}^{-1}$. For each non-empty word $w \in \mathscr{W}$, there exist elements $s_1, \ldots, s_n \in \mathscr{A}$ and $\epsilon_1, \ldots, \epsilon_n \in \{-1, +1\}$ such that $w$ is the sequence $(s_1^{\epsilon_1}, s_2^{\epsilon_2}, \ldots, s_n^{\epsilon_n})$. We write

$$w = s_1^{\epsilon_1} s_2^{\epsilon_2} \cdots s_n^{\epsilon_n}.$$

The **length** of $w$ is equal to $n$.

**Example 1.1.** Let $\mathscr{A} = \{a, b\}$. Here are some examples of words. They are all different.

- $a$
- $b$
- $ab$
- $ba$
- $aab$
- $aaba^{-1}$
- $abb^{-1}bbbaaaa^{-1}$

Given words $w = s_1^{\epsilon_1} s_2^{\epsilon_2} \cdots s_n^{\epsilon_n}$ and $u = t_1^{\delta_1} t_2^{\delta_2} \cdots t_n^{\delta_n}$, we define the **concatenation** to be the word

$$wu = s_1^{\epsilon_1} s_2^{\epsilon_2} \cdots s_n^{\epsilon_n} t_1^{\delta_1} t_2^{\delta_2} \cdots t_n^{\delta_n}$$

Observe that concatenation is an associative binary operation on $\mathscr{W}$. We desire to turn $\mathscr{W}$ into a group $F(\mathscr{A})$. It will be called the **free group** on $\mathscr{A}$.

Suppose that $w = s_1^{\epsilon_1} s_2^{\epsilon_2} \cdots s_n^{\epsilon_n}$ is a word such that there is a $k \in \{1, \ldots, n-1\}$ with $s_k = s_{k+1}$ and $\epsilon_k = -\epsilon_{k+1}$. That is, the $k$th letter of $w$ is the inverse of the $(k+1)$st letter. Define the word:

$$w' = s_1^{\epsilon_1} s_2^{\epsilon_2} \cdots s_{k-1}^{\epsilon_{k-1}} s_{k+2}^{\epsilon_{k+2}} \cdots s_n^{\epsilon_n}$$

to be a **reduction** of $w$. If $n = 2$, then $w'$ is the empty word. The inverse of reduction is **insertion**.

Define a relation $\sim$ on $\mathscr{W}$ by declaring $w \sim w'$ if $w'$ is obtained from $w$ by a finite sequence of deletions and insertions. Observe that $\sim$ is an equivalence relation on $\mathscr{W}$. Let $F(\mathscr{A})$ be the quotient set. Define a binary operation, called **concatenation**, on $F(\mathscr{A})$ by

$$[u][w] = [uw].$$

**Lemma 1.2.** Concatenation is well-defined on $F(\mathscr{A})$.

*Proof.* Suppose that $u \sim u'$ and $w \sim w'$. We must show that $uw \sim u'w'$. Let $\alpha, \beta$ be the sequence of reductions and insertions producing $u'$ from $u$ and $w'$ from $w$, respectively. Observe that we can apply $\alpha$ to the word $uw$ to obtain $u'w$. Similarly, we may apply $\beta$ (or rather the sequence obtained by shifting the indices in $\beta$ by the length of $u'$) to $u'w$ to obtain $u'w'$. Thus, there is a sequence of reductions and insertions producing $u'w'$ from $uw$. Hence, $uw \sim u'w'$. □

**Theorem 1.3.** *The set $F(\mathscr{A})$ is a group with concatenation as the operation and the empty word as the identity.*

*Proof.* Since concatenation is well-defined, $F(\mathscr{A})$ is closed under concatenation. Now we consider the other group axioms. Let $[\mathbb{1}] \in F(G)$ be the equivalence class of the empty word $\mathbb{1} \in \mathscr{W}$. In $\mathscr{W}$, we have $u\mathbb{1} = \mathbb{1}u = u$ for every $u \in \mathscr{W}$. Thus, passing to $F(G)$, the element $[\mathbb{1}]$ is an identity.

Let $f \in F(\mathscr{G})$. Choose $w \in \mathscr{W}$ such that $f = [w]$ and write

$$w = s_1^{\epsilon_1} s_2^{\epsilon_2} \cdots s_n^{\epsilon_n}.$$

Define

$$w' = s_n^{-\epsilon_n} s_{n-1}^{-\epsilon_{n-1}} \cdots s_1^{-\epsilon_1}.$$

It's easy to verify that there is a sequence of reductions of both $ww'$ and $w'w$ to the empty word. Thus, $[w]^{-1} = [w']$ in $F(\mathscr{A})$.

Finally, we consider associativity. Let $f, g, h \in F(\mathscr{A})$ and let $u, v, w \in W$ be words representing them. Recall that $(uv)w = u(vw) = uvw$ in $\mathscr{W}$ since concatenation in $\mathscr{W}$ is obviously associative. Thus,

$$(fg)h = [uv][w] = [(uv)w] = [uvw].$$

Likewise,

$$f(gh) = [u][vw] = [u(vw)] = [uvw].$$

Thus, concatenation in $F(\mathscr{A})$ is associative and so $F(\mathscr{A})$ is a group. □

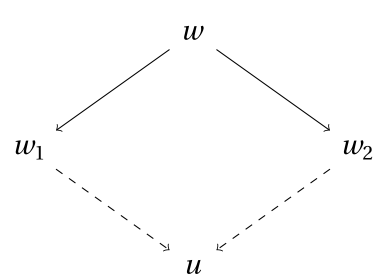
## 2. Normal Forms

Notice that if $w'$ is a reduction of $w$, then the length of $w'$ is strictly less than the length of $w$. Since the length of a word is a non-negative integer, we cannot perform infinitely reductions on a word. If $w$ is a word for which no reductions are possible, we say that $w$ is a **reduced** word.

**Question:** Is it possible for there to be a word $w$ and two sequences of reductions applied to $w$ which arrive at *different* reduced words?

For the purposes of this document, write $w \to w'$ if $w'$ is a reduction of $w$. Observe that the relation $\to$ makes $\mathscr{W}$ into a directed graph.

**Lemma 2.1.** Suppose that $w$ is a word and that $w_1$ and $w_2$ are different reductions of $w$. Then there exists a word $u$ which is a reduction of both $w_1$ and $w_2$.




*Proof.* Let

$$w = s_1^{\epsilon_1} s_2^{\epsilon_2} \cdots s_n^{\epsilon_n}.$$

Since $w$ has two different reductions, $n \geq 3$. Suppose that $w_1$ is obtained by cancelling $s_k^{\epsilon_k}$ and $s_{k+1}^{\epsilon_{k+1}} = s_k^{-\epsilon_k}$ and $w_2$ is obtained by cancelling $s_\ell^{\epsilon_\ell}$ and $s_{\ell+1}^{\epsilon_{\ell+1}} = s_\ell^{-\epsilon_\ell}$, with $\ell \neq k$.

We claim that $|k-\ell| \geq 2$. For simplicity, suppose that $s_k = a$ and $\epsilon_k = +1$ and that $k < \ell$. If $\ell = k+1$, then we have the subword

$$s_k^{\epsilon_k} s_{k+1}^{\epsilon_{k+1}} s_{k+2}^{\epsilon_{k+2}} = a\,a^{-1}\,a.$$

To form $w_1$ we cancel the first two letters of the subword; to form $w_2$, we cancel the second two. In either case, we are left with just $a$. Consequently, $w_1 = w_2$. This contradicts our assumption that $w_1$ and $w_2$ are different. □

We can now prove an important "normal form" theorem for free groups.

**Theorem 2.2** (Free Group Normal Form)**.** *Suppose that $w \in \mathscr{W}$. Then there exists a unique reduced word $w'$ such that $w'$ is obtained by a sequence of reductions on $w$.*

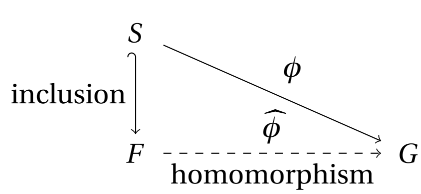
*Proof.* Existence follows from the fact that each reduction decreases the length of the word and that the length of a word is a non-negative integer. We concentrate on showing uniqueness. We prove uniqueness by induction on the length of $w \in \mathscr{W}$. If the length of $w$ is 0, then $w$ is the empty word. The empty word is reduced and so the result follows.

Suppose that the theorem is true for all words of length at most $k \in \mathbb{N} \cup \{0\}$. We prove it for words of length $k+1$. Let $w$ be a word of length $k+1$. If $w$ is reduced, then it is the unique reduced word created by a sequence of reductions applied to $w$; we may assume, therefore, that $w$ is not reduced. If there is a unique word $u$ obtained by reducing $w$, then the inductive hypothesis applied to $u$, guarantees that there is a unique reduced word $w'$ obtained by a sequence of reductions applied to $u$. Since every sequence of reductions applied to $w$ passes through $u$, the word $w'$ is the unique reduced word obtained by a sequence of reductions on $w$.

Suppose, therefore, that $w_1$ and $w_2$ are distinct words, both obtained by a reduction of $w$. By the inductive hypothesis applied to $w_1$ and $w_2$, there are unique reduced words $w'_1$ and $w'_2$ obtained by sequences of reductions applied to $w_1$ and $w_2$ respectively. We desire to show $w'_1 = w'_2$. By the lemma, there is a word $u$ obtained by reducing both $w_1$ and $w_2$. Choose a sequence $\alpha$ of reductions converting $u$ into a reduced word $u'$. The reduction of $w_1$ to $u$ followed by $\alpha$ is a sequence of reductions applied to $w_1$ resulting in the reduced word $u'$. Thus, $u' = w'_1$. Similarly, we may conclude that $u' = w'_2$. Hence, $w'_1 = w'_2$. Since this applies to all distinct words obtained by reducing $w$, there is a unique reduced word obtained by a sequence of reductions applied to $w$. By induction, the theorem holds. □

## 3. The Universal Property and Group Presentations

Let $F$ be a group with $S \subset F$. We say that $(F, S)$ has the **homomorphism extension universal property** if for every group $G$ and every function $\phi\colon S \to G$, there is a *unique* homomorphism $\widehat{\phi}\colon F \to G$ such that $\widehat{\phi}(s) = \phi(s)$ for every $s \in S$. In terms of commutative diagrams:




The idea behind this universal property is analogous to the process in linear algebra of definining linear maps by specifying what they do on a basis. It turns out that free groups are exactly the groups with the homomorophism extension universal property.

**Theorem 3.1.** *Suppose that $(F, S)$ has the homomorphism extension universal property. Then there exists an alphabet $\mathscr{A}$ and an isomorphism $F \to F(\mathscr{A})$ taking $S$ to $\mathscr{A}$. Conversely, if $F(\mathscr{A})$ is the free group on the alphabet $\mathscr{A}$, then $(F(\mathscr{A}), \mathscr{A})$ has the homomorphism extension universal property.*

*Proof.* We start by showing that each free group has the universal property. Let $\mathscr{A}$ be an alphabet and $F(\mathscr{A})$ the free group on $\mathscr{A}$. Let $S(\mathscr{A}) = \{[a] : a \in \mathscr{A}\}$ and suppose that $\phi\colon S(\mathscr{A}) \to G$ is a function onto some group $G$. Each element $g \in F(\mathscr{A})$ is represented by a unique reduced word

$$w = s_1^{\epsilon_1} s_2^{\epsilon_2} \cdots s_n^{\epsilon_n}$$

with $s_i \in \mathscr{A}$ and $\epsilon_i \in \{\pm 1\}$ for all $i$. Define

$$\widehat{\phi}(g) = \phi(s_1)^{\epsilon_1} \phi(s_2)^{\epsilon_2} \cdots \phi(s_n)^{\epsilon_n}.$$

Notice that $\widehat{\phi}$ extends $\phi$ and is well-defined and is a homomorphism. Since every homomorphism must respect products and inverses, it is the unique homomorphism extending $\phi$. Thus, $(F(\mathscr{A}), \mathscr{A})$ has desired universal property.

Now suppose that $(F, S)$ has the homomorphism extension universal property. Let $\mathscr{A} = S$. We begin by verifying that for all $a \in \mathscr{A} = S$, $a^{-1} \notin \mathscr{A}$. Suppose, to the contrary, that $a, a^{-1} \in S$. Let $G$ be the integers with addition and define $\phi\colon S \to \mathbb{Z}$ by letting $\phi(a) = +1$ and $\phi(s) = 0$ for all $s \in S \setminus \{a\}$. If $a \neq a^{-1}$, we have $\phi(a^{-1}) = 0$ but $\phi(a) = 1$. Since, 0 is not the inverse of 1 in $\mathbb{Z}$, the function $\phi$ cannot be extended to a group homomorphism $\widehat{\phi}\colon F \to \mathbb{Z}$. Similarly, if $a = a^{-1}$, we have $\phi(a) + \phi(a^{-1}) = \phi(a) + \phi(a) = 2 \neq 0$. Thus, again, $\phi$ cannot be extended to a group homomorphism $\widehat{\phi}\colon F \to \mathbb{Z}$. Thus, $\mathscr{A}$ is a permissible alphabet.

By hypothesis, $(F, S)$ has the homomorphism extension universal property. Let $\phi\colon S \to F(\mathscr{A})$ be the map defined by $\phi(s) = [s]$ for every $s \in S$. By definition, there exists a group homomorphism $\widehat{\phi}\colon F \to F(\mathscr{A})$ extending $\phi$. We claim that $\widehat{\phi}$ is an isomorphism. We use the fact that $(F(\mathscr{A}), \mathscr{A})$ has the universal property.

For each $a \in \mathscr{A}$, let $\psi([a]) = a \in S$. Since each $a \in \mathscr{A}$ is reduced, this is well-defined. Since $(F(\mathscr{A}), S(\mathscr{A}))$ satisfies the universal property, we can uniquely extend $\psi$ to a homomorphism $\widehat{\psi}\colon F(\mathscr{A}) \to F$. Observe that $\widehat{\psi} \circ \widehat{\phi}\colon F \to F$ is a homomorphism and that for each $s \in S$, $\widehat{\psi} \circ \widehat{\phi}(s) = s$. The identity map $\mathrm{id}\colon F \to F$ is also a group homomorphism taking each $s \in S$ to itself. By the uniqueness of the extension of maps $S \to S$ to homomorphisms $F \to F$, we must have

$$\widehat{\psi} \circ \widehat{\phi} = \mathrm{id}.$$

Similarly, $\widehat{\phi} \circ \widehat{\psi}\colon F(\mathscr{A}) \to F(\mathscr{A})$ is a group homomorphism taking each $[a] \in S(\mathscr{A})$ to itself. The identity is another such homomorphism. By uniqueness of extensions, $\widehat{\phi} \circ \widehat{\psi} = \mathrm{id}$. Thus, $\widehat{\phi}$ and $\widehat{\psi}$ are inverses and so $\widehat{\phi}$ is an isomorphism. □

**Corollary 3.2.** Suppose that $G$ is a group generated by $S \subset G$. Let $\mathscr{W}$ be the set of words in $\mathscr{A} \cup \mathscr{A}^{-1}$ where $\mathscr{A} = S$. Then there exists a subset $R \subset \mathscr{W}$ such that

$$G \cong \mathscr{F}(\mathscr{A}) / \langle R \rangle$$

where $\langle R \rangle$ is the smallest normal subgroup of $G$ containing $R$.

In the context of the corollary we write $G = \langle S | R \rangle$ and the set $R$ is called a set of **relations** for $G$ with respect to the generating set $S$. The generating set $S$ together with the relations $R$ is called a **presentation** of the group $G$. It is a **finite presentation** if both $S$ and $R$ are finite sets.

*Proof.* Let $S(\mathscr{A})$ be the set of elements $[a] \in F(\mathscr{A})$ such that $a \in \mathscr{A}$. Notice each $a \in F(\mathscr{A})$ is reduced. Thus, the function $\phi\colon S(\mathscr{A}) \to S$ defined by $\phi(a) = a \in S \subset G$ for each $a \in \mathscr{A}$ is well-defined. Since $(F(\mathscr{A}), S(\mathscr{A}))$ has the universal property, the function $\phi$ extends to a homomorphism $\widehat{\phi}\colon F(\mathscr{A}) \to G$. Let $N$ be it's kernel. Standard algebra shows that $G \cong F(\mathscr{A})/N$. Let $R$ be the set of reduced words representing elements in $N$. Observe that $N = \langle R \rangle$. □

In the previous proof, we took $R$ to be the set of all reduced words representing elements of the kernel of $\widehat{\phi}$. Observe that if $u, w \in R$, then the reduced word representing $[uw]$ is also in $R$. But this is overkill, since simply knowing that $N$ is a subgroup is enough to guarantee that if $u, w \in R$ then $[uw] \in N$. Generally, we want to take $R$ to be as small as possible.

**Definition.** Suppose that $G$ is a group and that $N \lhd G$ is a normal subgroup. A subset $R \subset N$ **normally generates** $N$ if whenever $U \vartriangle G$ is a normal subgroup such that $R \subset U$ then $N < U$.

The next lemma shows that normally generated subgroups exist.

**Lemma 3.3.** Suppose that $G$ is a group and that $R \subset G$. Then there exists a normal subgroup $N \lhd G$ such that $N = \langle R \rangle$.

We say that $N$ is the **normal closure** of $R$ and that $R$ **normally generates** $N$.

*Proof.* Let $\mathscr{H}$ be the set of all normal subgroups of $G$ containing $R$ as a subset. Since $G \in \mathscr{H}$, the set $\mathscr{H}$ is non-empty. Define $N = \bigcap_{H \in \mathscr{H}} H$. Since the intersection of subgroups is always a subgroup, $N$ is a subgroup of $G$. We show that it is normal.

Let $n \in N$ and $g \in G$. We must show $gng^{-1} \in N$. Since $n \in \bigcap_{H \in \mathscr{H}} H$, the element $n \in H$ for every $H \in \mathscr{H}$. Since each $H \in \mathscr{H}$ is normal, $gng^{-1} \in H$ for every $H \in \mathscr{H}$. Consequently, $gng^{-1} \in N$.

Now we show that $N$ is as small as possible. Suppose that $U$ is a normal subgroup of $G$ such that $R \subset U$. Then $U \in \mathscr{H}$. Since $U$ is one of the subgroups in the intersection forming $N$, $N < U$. □

**Corollary 3.4.** Let $S$ be a set and let $\mathscr{W}$ be the set of words in $S$ and $S^{-1}$. Then for every subset $R \subset \mathscr{W}$, there is a group

$$\langle S | R \rangle$$

*Proof.* Let $\mathscr{A} = S$ and let $N$ be the normal closure of $R$. Then the group we are after is $\mathscr{F}(\mathscr{A})/N$. □

**Example 3.5.** The group $D_\infty$ has presentation $\langle a, b | a^2, b^2 \rangle$

**Example 3.6.** The group $\mathbb{Z}^2$ has presentation $\langle a, b | aba^{-1}b^{-1} \rangle$.

Consider the following fundamental questions:

(1) (Triviality Problem) Given a group $G$ with presentation $\langle S|R \rangle$, is there an algorithm to determine if $G$ is the trivial group?
(2) (Isomorphism Problem - Tietz 1908) Given groups $G = \langle S|R \rangle$ and $G' = \langle S'|R' \rangle$, is there an algorithm to determine if $G$ is isomorphic to $G'$?
(3) (Word Problem - Dehn 1910) Given a group $G = \langle S|R \rangle$ and a word $w \in (S \cup S^{-1})^*$, is there an algorithm to determine if $[w] = \mathbb{1}$ in $G$?
(4) (Conjugacy Problem - Dehn 1911) Given a group $G = \langle S|R \rangle$ and words $w, w' \in (S \cup S^{-1})^*$, is there an algorithm to determine if $[w]$ is conjugate to $[w']$ in $G$?

Observe that these problems are solvable for free groups with their usual presentation. However, Boone and Novikov proved the word problem unsolvable in 1955/56.Tthe isomorphism problem was proved unsolvable by Adian and Rabin (1958).

For more on the history and details on the constructions of Boone and Novikov (and others) see:

*The word problem and the isomorphism problem for groups*, John Stillwell. *Bull. AMS* (6) No. 1, 1982.