1· Free groups and free presentations

> The words are all there ready; now we've got
> to get them in the right order. (Python)

A group G is generated by a subset X if each of its elements can be expressed as a product of members of $X^{\pm 1}$. Such a product is called a word, and a relation is an equation between two words. A set R of relations that hold in G defines the group if every relation that holds in G is a consequence of R. When this happens, we say that G is presented by X and R. This definition is made rigorous using the concept of a free group (essentially, a group having a set of generators between which there are no non-trivial relations), which is defined using a universal property. Having developed some elementary but important properties of free groups (such as their existence), we proceed to the fundamental theorem of §2, where Schreier's proof is given in detail and Nielsen's original method in outline. In §3, the definition of group presentation is made rigorous, and this is used to clarify the proof of the Nielsen-Schreier theorem by means of an anotated example. §4 explains how to pass from a group multiplication table to a presentation and from one presentation to another, as well as describing a presentation for a direct product of two groups.

§1. Elementary properties of free groups

The fundamental notion used in defining presentations of groups is that of a free group. As the definition suggests, the idea of freeness is applicable in algebraic situations other than group theory.

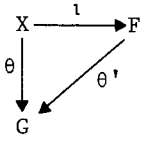
Definition 1. A group F is said to be *free* on a subset $X \subseteq F$ if, for any group G and any mapping $\theta : X \to G$, there is a

unique homomorphism $\theta' : F \to G$ such that

$$x\theta' = x\theta \qquad (1)$$

for all $x \in X$. The cardinality of X is called the *rank* of F.

Remark 1. There are various ways of expressing the property (1). We may say that θ' extends θ or that θ' agrees with θ on X or, letting $\iota : X \to F$ denote inclusion, that the following diagram is commutative:



In general, a diagram involving sets and mappings is called commutative if any two composite mappings, beginning at the same place and ending at the same place in the diagram, are equal. In this case, this boils down to the single assertion that $\iota\theta' = \theta$.

Remark 2. There is an analogy between this situation and a familiar one encountered in linear algebra; let V be a vector space over a field k and B a basis for V. Then for any vector space W over k and any mapping $\tau : B \to W$, there is a unique k-linear transformation $\tau' : V \to W$ extending τ. This property is known as 'extension by linearity' and can be used to *define* the notion of basis.

Remark 3. If we write 'abelian group' in place of 'group' in the two places where this word appears in Definition 1, we obtain the definition of a free abelian group. A free abelian group of rank ω is just the direct sum of ω infinite cyclic groups (proved for finite ω in Theorem 6.2).

Remark 4. By convention, we take E (the trivial group) to be free of rank 0, the subset X being empty. The infinite cyclic group $\{x^n \mid n \in Z\}$ is free of rank 1. We denote it by Z as it

is just the multiplicative version of the additive group of integers. Take $X = \{x\}$, and given

$$\left.\begin{array}{l} \theta : X \to G \\ \quad x \mapsto y \end{array}\right\},$$

simply define for all $n \in Z$,

$$x^n\theta' = y^n .$$

θ' is obviously a homomorphism extending θ, while if θ'' is another,

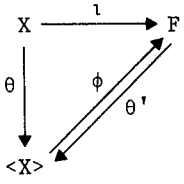
$$x^n\theta'' = (x\theta'')^n = y^n = (x\theta')^n = x^n\theta' ,$$

proving that θ' is unique.

Remark 5. There are one or two things to check before this definition can have any value. One can show for example that there does exist a free group of any given rank, and that the rank of a free group is well defined. These together with other elementary properties of free groups form the content of our first four theorems.

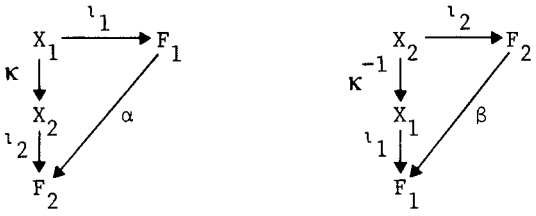
Theorem 1. (i) *If* F *is free on* X, *then* X *generates* F.
(ii) *Two free groups of the same rank are isomorphic.*
(iii) *Free groups of different ranks are not isomorphic.*

Proof. (i) Recall that if X is a subset of a group G, the intersection of all subgroups of G containing X is called the subgroup generated by X and written $\langle X \rangle$. We leave it as an exercise to show that this coincides with the set of all finite products of members of X and their inverses. Returning to the case in hand, we let $\langle X \rangle$ play the role of G in Definition 1, taking θ to be inclusion. Letting ϕ denote the inclusion of $\langle X \rangle$ in F, we have the following picture:



Since this diagram commutes, we have $\iota\theta'\phi = \theta\phi = \iota$, so that $\theta'\phi : F \to F$ extends ι. But so does 1_F, and so by the uniqueness part of Definition 1 (with ι, F in place of θ, G), we have $\theta'\phi = 1_F$, whence ϕ is onto and $\langle X \rangle = F$, as required.

(ii) Let F_j be free on X_j and let $\iota_j : X_j \to F_j$ denote inclusion, $j = 1,2$. Assume that $|X_1| = |X_2|$, so that there is a bijection $\kappa : X_1 \to X_2$. Let α, β be the homomorphisms extending $\kappa\iota_2, \kappa^{-1}\iota_1$ as in the following diagrams:



Now $\iota_1\alpha\beta = \kappa\iota_2\beta = \kappa\kappa^{-1}\iota_1 = \iota_1$, so that $\alpha\beta : F_1 \to F_1$ extends ι_1. But 1_{F_1} also extends ι_1, so uniqueness implies $\alpha\beta = 1_{F_1}$. Similarly, $\beta\alpha = 1_{F_2}$, and α is the required isomorphism.

(iii) Let F be free on a subset X with $|X| = \omega$, and let G be any group. Then it is the burden of Definition 1 that the mappings: $X \to G$ are in one-to-one correspondence with the homomorphisms: $F \to G$. Thus, there are exactly 2^ω homomorphisms from F to Z_2. Since this number is invariant under isomorphism, we see that 2^ω, and hence the rank ω, is determined by the isomorphism class of F.

Theorem 2. *There exists a free group of any given rank.*

Proof. We construct the 'group of words' $F = F(X)$ on a given set X, and prove that it is free of rank $|X|$. The free group F on a given set X is constructed as follows. Let $\hat{X} = \{\hat{x} | x \in X\}$ be any set in one-to-one correspondence with, and disjoint from, X and put $T = X \cup \hat{X}$. If T^n denotes the nth Cartesian power of T $(n = 0,1,2,\ldots)$, put $W = \bigcup_{n \geq 0} T^n$, the set of words in X. A word $w \in T^n$ is said to have length n, and the single element of T^0 is called the empty word and denoted by e. A word $w = (t_1,\ldots,t_n)$ in W is called reduced if there is no i between 1 and $n-1$ such that $\hat{t}_i = t_{i+1}$, interpreting $\hat{\hat{s}} = s$. Letting F be the set of reduced words, it is clear that $e \in F$ and $X \subseteq F$. The product of two reduced words of positive length

$$a = (x_1,\ldots,x_m),\ b = (y_1,\ldots,y_n)$$

is defined to be

$$(x_1,\ldots,x_{m-k},\ y_{k+1},\ldots,y_n)\ ,$$

where k is the largest integer such that none of the words

$$(x_m,y_1),\ldots,(x_{m-r+1},y_r)$$

are reduced, while $we = ew = w$ for any word w. It is clear that this defines a binary operation on F for which e is an identity and $(x_1,\ldots,x_m)^{-1} = (\hat{x}_m,\ldots,\hat{x}_1)$. The tricky bit, surprisingly enough, is the proof of the associative law. Now take three words in F:

$$a = (x_1,\ldots,x_\ell),\ b = (y_1,\ldots,y_m),\ c = (z_1,\ldots,z_n)\ .$$

If any of ℓ,m,n are zero, we clearly have $(ab)c = a(bc)$, so assume they are all positive. Supposing that the lengths of ab and bc are $\ell+m-2r$ and $m+n-2s$ respectively, we distinguish three cases. First, if $r+s < m$, both $(ab)c$ and $a(bc)$ are

equal to the reduced word

$$(x_1,\ldots,x_{\ell-r},y_{r+1},\ldots,y_{m-s},z_{s+1},\ldots,z_n) \quad ;$$

secondly, if $r+s = m$, both are equal to $\alpha\varepsilon$, where

$$\alpha = (x_1,\ldots,x_{\ell-r}),\ \varepsilon = (z_{s+1},\ldots,z_n) \quad .$$

Finally, in the case $r+s > m$, we define

$$\beta = (x_{\ell-r+1},\ldots,x_{\ell-m+s}) = (y_{m-s+1},\ldots,y_r)^{-1} = (z_{m-r+1},\ldots,z_s) \ ,$$

$$\gamma = (x_{\ell-m+s+1},\ldots,x_\ell) = (y_1,\ldots,y_{m-s})^{-1} \quad ,$$

$$\delta = (z_1,\ldots,z_{m-r}) = (y_{r+1},\ldots,y_m)^{-1} \quad .$$

Thus, $a = \alpha\beta\gamma$, $b = \gamma^{-1}\beta^{-1}\delta^{-1}$, $c = \delta\beta\varepsilon$, since the brackets can safely be ignored by the first case handled above. Now by the rule for forming products,

$$(ab)c = (\alpha\delta^{-1})(\delta\beta\varepsilon) = \alpha(\beta\varepsilon) \ , \quad \text{and}$$

$$a(bc) = (\alpha\beta\gamma)(\gamma^{-1}\varepsilon) = (\alpha\beta)\varepsilon \quad ,$$

and again by the first case, these both coincide with the reduced word $(x_1,\ldots,x_{\ell-m+s},z_{s+1},\ldots,z_n)$.

We now simplify the notation by dropping the commas and brackets and writing x^{-1} for $\hat{x}$ $(x \in X \cup \hat{X})$, so that if ι is the inclusion of X in F, all we have to do is check Definition 1 verbatim. If G is a group and $\theta : X \to G$ a mapping, define

$$e\theta' = e \ , \ x^{-1}\theta' = (x\theta)^{-1} \quad ,$$

$$(x_1\ldots x_n)\theta' = x_1\theta'\ldots x_n\theta' \quad ,$$

for any $x \in X$ and any reduced word $x_1\ldots x_n$. It is a routine

matter to check that θ' is a homomorphism extending θ. If θ'' is another, it must agree with θ' on X and since X plainly generates F, we must have $\theta' = \theta''$.

Theorem 3. *Let F be a group and X a subset of F; then F is free on X if and only if the following two conditions hold:*

(i) X generates F,

(ii) there is no non-trivial relation between the elements of X, that is, if for $n \in \mathbb{N}$, $x = x_1 \ldots x_n$ where for all i, either $x_i \in X$ or $x_i^{-1} \in X$, and for all i with $1 \leq i \leq n-1$, $x_i x_{i+1} \neq e$, then $x \neq e$.

Proof. First suppose that F is free on X, so that X generates F by Theorem 1(i). Now let $X' = \{x' | x \in X\}$ be an abstract copy of X and consider the group of words $F(X')$ as constructed in the proof of Theorem 2. By Definition 1, the priming map $: X \to F(X')$ extends to a homomorphism $: F \to F(X')$ under which any reduced word $x \in F$ is mapped to a reduced word in $F(X')$ of the same length. Thus no reduced word in F of length $n \leq 1$ can be e, since this certainly holds in $F(X')$.

For the converse, note that conditions (i) and (ii) imply that every member of F is uniquely expressible as a reduced word in $X \cup X^{-1}$. The freeness of F on X is now verified in just the same way as that of $F(X)$ on X in the final part of the proof of Theorem 2.

Theorem 4. *If X is a set of generators for a group G and $F(X)$ is the group of words in X, then there is an epimorphism $\theta: F(X) \to G$ fixing X elementwise. Every group is a homomorphic image of some free group.*

Proof. The required epimorphism is just the (unique) extension to the free group $F(X)$ of the inclusion $: X \to G$; it is onto because $X \subseteq \operatorname{Im} \theta \leq G$ and $\langle X \rangle = G$. The second assertion now follows from the simple observation that any group G *has* a set of generators, for example, $G = \langle G \rangle$.

EXERCISE 1. Let X be a subset of a group G. Prove that $\langle X \rangle$ is equal to the set of all finite products of members of X and their inverses. Deduce that if two homomorphisms from G to a group H agree on a set X of generators of G (i.e. $\langle X \rangle = G$), then they are equal.

EXERCISE 2. Given groups G and H, a subset X of G and a homomorphism $\theta: G \to H$, prove that $\langle X\theta \rangle = \langle X \rangle \theta$. Defining

$$d(G) = \min\{|X| \mid X \subseteq G, \langle X \rangle = G\},$$

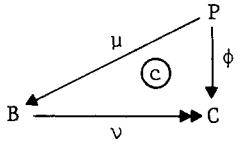
prove that for any homomorphic image H of G, $d(H) \leq d(G)$.

EXERCISE 3. Given a subset X of a group G, define the normal closure $\bar{X}$ of X to be the intersection of all *normal* subgroups of G containing X. Prove that $\bar{X}$ is just the set of all finite products of conjugates of members of X and their inverses. If H is a group and $\theta: G \to H$ an epimorphism, show that $\overline{X\theta} = \bar{X}\theta$.

EXERCISE 4. Let F be a free group of rank ω and G a group isomorphic to F. Prove that G is free of rank ω.

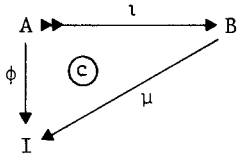
EXERCISE 5. A group G has a normal subgroup N such that G/N is free. Prove that G has a subgroup F such that $FN = G$ and $F \cap N = E$. (Such an F is called a *complement* for N in G.)

EXERCISE 6. Call a group P *projective* if given any epimorphism $\nu: B \to C$ of groups and any homomorphism $\phi: P \to C$, there is a homomorphism $\mu: P \to B$ such that $\phi = \mu\nu$:



Prove that P is projective if and only if P is free.

EXERCISE 7. Call a group I *injective* if given any monomorphism $\iota: A \to B$ of groups and any homomorphism $\phi: A \to I$, there is a homomorphism $\mu: B \to I$ such that $\phi = \iota\mu$:



Prove that I is injective if and only if I is trivial. (Hint (D.B.A. Epstein): Assume that the free group $A = F(x,y)$ is a subgroup of a group $B = \langle x,y,z \rangle$ in which $z^{-1}xz = y$, $z^{-1}yz = x$, $z^2 = e$.)

§2. The Nielsen-Schreier theorem

The first step in proving the Basis Theorem for finitely-generated abelian groups (see §6 below) is to show, at least in the case of finite rank, that subgroups of free abelian groups are free abelian, and that the rank of the subgroup does not exceed the rank of the group. This is a classical result of Dedekind, and our purpose here is to prove its non-abelian analogue, taking care to point out that the assertion about ranks does *not* hold in the non-abelian case. We shall consider the free group $F = F(X)$ on an arbitrary set X, invoking the Axiom of Choice to assert that X can be well-ordered. The intuitionistic reader is free to assume that X is finite, since this is the only case of practical interest to us.

The proof we give is essentially due to Schreier, and is divided into a number of steps, the most important for the sequel (§12) being embodied in Lemma 3.

1. The ordering of F. Given that X is well-ordered, so is X^{-1} and so also is $T = X \cup X^{-1} = X^{\pm 1}$; for example, if $x,y \in X$, define $x < y^{-1}$, and

$$x^{-1} < y^{-1} \iff x < y \ .$$

Now the elements of F are words of the form

$$w = x_1 \ldots x_n \ , \ x_i \in T \ , \ x_i x_{i+1} \neq e \ ;$$

we call n the *length* of the word w, $n = \ell(w)$, and take $\ell(e) = 0$. For $v,w \in F$, we define $v < w$ if $\ell(v) < \ell(w)$, and order words of equal length lexicographically, that is, if

$$v = x_1 \ldots x_n \neq w = y_1 \ldots y_n \ , \ x_i, y_i \in T \ ,$$

and m is least such that $x_m \neq y_m$, we define

$$v < w \iff x_m < y_m \ .$$

The result is easily seen to be a well-ordering of F. For example, if $X = \{x,y\}$, then with respect to the ordering $x < y < x^{-1} < y^{-1}$ of T, the first few elements of F are
$e < x < y < x^{-1} < y^{-1} < x^2 < xy < xy^{-1} < yx < y^2 < yx^{-1} < x^{-1}y$
$< x^{-2} < x^{-1}y^{-1} < y^{-1}x < y^{-1}x^{-1} < y^{-2} < x^3 < x^2y < x^2y^{-1} < xyx$.
Note that in any subgroup H of F, the least element is always e (since this is least in F) and if, for example, $H = \langle xyx \rangle$, the least element of the coset Hxy^2 is $x^{-1}y$. That F is well-ordered is particularly easy to see in the case of finite rank, for then there are only finitely many words of any given length.

Lemma 1. *Let*

$$w = x_1 \ldots x_n, \ x_i \in T, \ n \geq 1 \ ,$$

be a reduced word in F *; then, for* $v \in F$,

$$v < x_1 \ldots x_{n-1} \Rightarrow vx_n < w \ .$$