CHAPTER 1

FREE GROUPS

The fundamental notion underlying the theory of group presentations is that of a free group. Roughly speaking, a group F is called free if it has a subset X with the property that every element of F can be written uniquely is a product of elements of X and their inverses. The uniqueness here means that if two such products look different, then they are different, so that no non-trivial relations hold between elements of F. This is made precise in the definition given below, which suggests (correctly) that the idea of freeness is also applicable in algebraic situations other than group theory.

1. Definition and elementary properties

Definition 1. A group F is said to be *free on a subset* $X \subseteq F$ if, given any group G and any map $\theta : X \to G$, there is a unique homomorphism $\theta' : F \to G$ extending θ, that is, having the property that $x\theta' = x\theta$ for all $x \in X$, or that the diagram

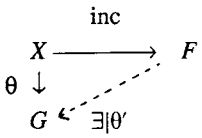

Fig. 1

is commutative. Then X is called a *basis* of F and $|X|$ the *rank* of F, written $r(F)$.

Remarks. 1. This accords with our intuitive idea of freeness in so far as:

(i) if there were a "relation" $w = w'$ among the members of $X^{\pm}: = \{x, x^{-1} \mid x \in X\}$, then we could find a group G with corresponding elements (under θ) for which that relation does not hold; then θ' would have to map $e = w'w^{-1} \in F$ to an element of G other than the identity, which is impossible;

(ii) if the elements of X did not generate F, the extension of θ would only be defined as far as the subgroup $<X>$ of F generated by X and thereafter would be arbitrary, violating the uniqueness.

In this sense then,

$$\text{existence of } \theta' \Rightarrow \text{ no relations in } X^{\pm},$$
$$\text{uniqueness of } \theta' \Rightarrow X \text{ generates } F.$$

2. There is a strong analogy here with the notion of "extension by linearity" in the theory of vector spaces: if B is a basis for a vector space V and $\theta : B \to W$ any map into a vector space W, there is a unique extension of θ to a linear transformation from V into W. Here, the existence and uniqueness of the extension are guaranteed by the two defining properties of a basis, namely, that the elements of B are linearly independent and span V, respectively.

3. Prefacing the word "group" by the adjective "abelian" in the two places where it occurs in the above definition yields another, that of *free abelian group*. This is even more like the situation in linear algebra, for if A is a free abelian group with basis $X = \{x_1, ..., x_n\}$ say, then every element of A is uniquely expressible in the form $\sum_{i=1}^{n} k_i x_i$, $k_i \in \mathbb{Z}$, $1 \le i \le n$. Such an expression is called a $\mathbb{Z}$-linear combination, and we have an isomorphism

$$\left.\begin{array}{ccc} A & \longrightarrow & \mathbb{Z} \oplus \mathbb{Z} \oplus ... \oplus \mathbb{Z} \\ \Sigma k_i x_i & \longmapsto & (k_1, k_2, ..., k_n) \end{array}\right\}.$$

In fact, the only conceptual difference between a free abelian group of rank n and a vector space of dimension n is that in the former, the coefficients $\mathbb{Z}$ do not form a field.

4. Now let F be an arbitrary group, $X \subseteq F$, and G any group. Let Hom (F, G) denote the set of homomorphisms from F into G, Map (X,G) the set of maps from X into G, and

$$\left.\begin{array}{l} \rho\text{: Home } (F,G) \longrightarrow \text{Map } (X,G) \\ (F \xrightarrow{\phi} G) \mapsto (X \xrightarrow{\text{inc}} F \xrightarrow{\phi} G) \end{array}\right\} \qquad (1)$$

the restriction map. Then

$$\rho \text{ is surjective} \Leftrightarrow \forall\, \theta, \exists\, \theta' \text{ as in Definition 1},$$
$$\rho \text{ is injective} \Leftrightarrow \theta', \text{ if it exists, is unique.}$$

Thus, F is free on X if and only if the map ρ of (1) is a bijection for any group G.

5. While a free group may have many different bases, they all turn out to have the same number of elements, so that the rank is well-defined. This will be proved directly, along with the converse that a free group is determined up to isomorphism by its rank. The

existence of free groups is then proved by explicit construction.

Lemma 1. If F is free on X, then X generates F.

Proof. Let $H = \langle X \rangle := \cap \{K \leq F \mid K \supseteq X\}$, and let $\theta : X \to H$ denote inclusion, with $\theta' : F \to H$ the corresponding extension. Letting $\iota : H \to F$ denote inclusion, we see from the picture that $\theta' \iota$ extends $\theta \iota = \text{inc}$. But so does the identity map 1_F. By uniqueness, $\theta' \iota = 1_F$, whence $F = \text{Im } 1_F = \text{Im } \theta' \iota = \text{Im } \theta' \subseteq H$.

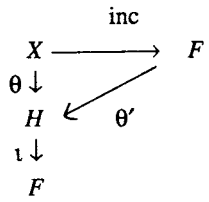

Fig. 2

Proposition 1. If F_i is free on X_i $(i = 1,2)$ and $F_1 \cong F_2$, then $|X_1| = |X_2|$.

Proof. Apply Remark 4 above with $G = Z_2$ the group of order 2. Since $F_1 \cong F_2$, $|\text{Hom}(F_1, Z_2)| = |\text{Hom}(F_2, Z_2)|$, whence $|\text{Map}(X_1, Z_2)| = \text{Map}(X_2, Z_2)|$. But for any sets B, C of cardinalities b, c, respectively, $|\text{Map}(B,C)| = c^b$. So in this case, $2^{|X_1|} = 2^{|X_2|}$, and the result follows by taking logs to the base 2.

Proposition 2. If F_i is free on X_i $(1 = 1,2)$ and $|X_1| = |X_2|$, then $F_1 \cong F_2$.

Proof. Assume $|X_1| = |X_2|$, so that there is a bijection $\kappa : X_1 \to X_2$. Let α, β be the extensions given by:

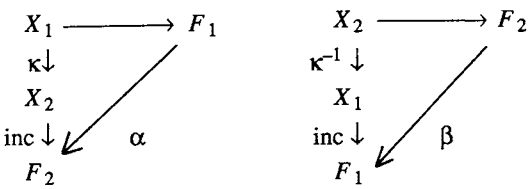

Fig. 3

Now for all $x_1 \in X_1$, $x_1 \alpha\beta = (x_1\kappa)\beta = (x_1\kappa)\kappa^{-1} = x_1$, so that $\alpha\beta : F_1 \to F_1$ extends $\text{inc} : X_1 \to F_1$. But so does the identity map 1_{F_1} on F_1, whence $\alpha\beta = 1_{F_1}$ by uniqueness. Similarly, $\beta\alpha = 1_{F_2}$, and so α is an isomorphism.

2. Existence of $F(X)$

Let X be any abstract set. There follows a recipe for constructing a free group $F(X)$ containing X as a basis.

Step 1. First form another copy of X, $\hat{X} = \{\hat{x} \mid x \in X$(whose elements will later become the inverses of the elements of X), and consider their union $X^{\pm} := X \cup \hat{X}$.

Next form the *words* $W_n = (X^{\pm})^{\times n}$ of length $n \geq 0$ in $X^{\pm}$, which are just n-tuples of elements of $X^{\pm}$. Thus.

W_0 consists only of ( ), the *empty* word, usually written e,

W_1 consists of (x), $(\hat{x})$, $x \in X$, and so looks like $X^{\pm}$,

W_2 consists of pairs (x,y), $x,y \in X^{\pm}$, and so on.

Now discard all words containing a pair $x,\hat{x}$ (for the *same* $x \in X$) in adjacent positions, in either order. The remaining words, namely those without this property, are called *reduced*: let $\tilde{W}_n$ denote the set of reduced words of *length* n.

Finally, define $F(X) = \bigcup_{n \geq 0} \tilde{W}_n$.

Step 2. For $F(X)$ to be a group, we need a binary operation on $F(X)$, defined roughly as "juxtaposition plus cancellation", and precisely as follows: given

$$a = (x_1, \ldots, x_l) \in \tilde{W}_l, \quad b = (y_1, \ldots, y_m) \in \tilde{W}_m \ , \tag{2}$$

put

$$ab = (x_1, \ldots, x_{l-r}, \ y_{r+1}, \ldots, y_m) \ , \tag{3}$$

where r is the largest value of $k \geq 0$ for which none of (x_l, y_1), (x_{l-1}, y_2), ..., (x_{l-k+1}, y_k) is reduced. This condition guarantees that $ab \in \tilde{W}_{l+m-2r}$. Of course $r \leq \min(l,m)$ and equality may occur.

Closure being assured, it remains to check the other axioms for a group. Clearly, e is the identity, and

$$(x_1, \ldots, x_l)^{-1} = (\hat{x}_l, \ldots, \hat{x}_1) \ ,$$

interpreting $\hat{\hat{x}} = x$ when $x \in X$. For the associative law, let a, b, ab be as in (2) and (3), and let

$$c = (z_1, \ldots, z_n) \in \tilde{W}_n, \quad bc = (y_1, \ldots, y_{m-s}, z_{s+1}, \ldots, z_n) \in \tilde{W}_{m+n-2s} \ .$$

Thus we have to show that

$$(ab)c = a(bc) \ . \tag{4}$$

If any of a, b or c is e, this is obvious, and so we can assume that l, m, $n \geq 1$. There are three cases to consider.

Case 1: $r + s < m$. Here, both sides of (4) are equal to

$$(x_1, \ldots, x_{l-r}, y_{r+1}, \ldots, y_{m-s}, z_{s+1}, \ldots, z_n) \in \tilde{W}_{l+m+n-2r-2s} \ .$$

Case 2: $r + s = m$. In this case, both sides of (4) are equal to the product

$$(x_1, \ldots, x_{l-r})(z_{s+1}, \ldots, z_n) \ .$$

Case 3: $r + s > m$. Put

$$\beta = (y_1, \ldots, y_{m-s}), \quad \gamma = (y_{m-s+1}, \ldots, y_r), \quad \delta = (y_{r+1}, \ldots, y_m) \ ,$$

so that γ has length ≥ 1 by case hypothesis, and

$b = \beta\gamma\delta$, which is unambiguous by Case 1, as are

$a = \alpha\gamma^{-1}\beta^{-1}$, where $\alpha = (x_1, \ldots, x_{l-r})$, and

$c = \delta^{-1}\gamma^{-1}\varepsilon$, where $\varepsilon = (z_{s+1}, \ldots, z_n)$.

Then

$$(ab)c = (\alpha\delta)(\delta^{-1}\gamma^{-1}\varepsilon) = \alpha(\gamma^{-1}\varepsilon) \ ,$$

and

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

Since α and γ^{-1} are adjacent in the reduced word a, there is no cancellation in forming their product, and similarly with γ^{-1} and ε (in c), whence $\alpha(\gamma^{-1}\varepsilon) = (\alpha\gamma^{-1})\varepsilon$ by Case 1,

showing that (4) holds in every case.

Step 3. It is now appropriate to use the bijection

$$\left.\begin{array}{rcl} W_1 = \tilde{W}_1 & \longrightarrow & W^{\pm} \\ (x) & \longmapsto & x \end{array}\right\}$$

to simplify notation by removing brackets and commas. Then $\hat{x}$ naturally becomes identified with x^{-1}, and in view of such rules as $x^{-1}xy = y$ and $(xy)^{-1} = y^{-1}x^{-1}$, the above definitions of product and inverse are in accordance with standard notational conventions. Reduced words are those in which no cancellation in the usual sense is possible. We retain the notion of length, and write $l(w) = n$ if $w \in \tilde{W}_n$. Note that $X \subseteq F(X)$, and that $<X>$ contains $X^{-1} = \{x^{-1} \mid x \in X\}$ (by inversion) and thus every $\tilde{W}_n$ (by closure), whence X generates $F(X)$.

Step 4. Finally, it must be shown that $F(X)$ is free on X. For a given G, and a given $\theta : X \to G$, define

$$e\theta' = e, \quad x\theta' = x\theta \text{ and } (x^{-1})\theta' = (x\,\theta)^{-1} \text{ for } x \in X \ ,$$

and

$$(x_1 \ldots x_l)\,\theta' = (x_1\theta') \ldots (x_l\theta'), \text{ for } x_1 \ldots x_l \in \tilde{W}_l \ .$$

It is clear that θ' extends θ and that there can be at most one homomorphism with this property, as X generates $F(X)$. It remains only to prove that $(ab)\theta' = (a\theta')\,(b\theta')$, with a, b, ab as in (2) and (3), above. By the definition of ab, no (x_{l-i+1}, y_i) is reduced for $1 \le i \le r$, and so $y_i = x_{l-i+1}^{-1}$, and $y_i\theta' = (x_{l-i+1}\theta')^{-1}$ for all such i, by the definition of θ' on words of length 1. It follows from the definition of θ' on longer words that

$$\begin{aligned} (a\theta')^{-1}(ab)\theta'(b\theta')^{-1} &= (x_l\theta')^{-1} \ldots (x_{l-r+1}\theta')^{-1}\,(y_r\theta')^{-1} \ldots (y_1\theta')^{-1} \\ &= (y_1\theta') \ldots (y_r\theta')\,(y_r\theta')^{-1} \ldots (y_1\theta')^{-1} \\ &= e \ , \end{aligned}$$

as required.

Theorem 1. The group $F(X)$ of reduced words in $X^{\pm}$ is free on X.

This result has two important consequences which provide an internal characterisation of free groups and a starting point for the theory of presentations, respectively.

Proposition 3. A group F is free on a subset X if and only if

(i) X generates F, and

(ii) no reduced word in $X^{\pm}$ of positive length is equal to e.

Proof. Let $\theta' : F(X) \to F$ be the homomorphism extending the inclusion $\theta : X \to F$. Then (i) and (ii) are respectively equivalent to the assertions that θ' is surjective and injective. On the other hand, if F is free on X, the extension $\phi' : F \to F(X)$ of the inclusion $\phi : X \to F(X)$ is clearly an inverse of θ' (using Lemma 1), while if θ' has an inverse, it must be an isomorphism, and freeness is an isomorphism invariant (see Exercise 4). Thus,

$$F \text{ is free on } X \Leftrightarrow \theta' \text{ is a bijection} \Leftrightarrow \text{(i) and (ii) hold.}$$

Proposition 4. Every group is isomorphic to a factor group of some free group.

Proof. Given a group G, let X be a set of generators for G (which always exists: take $X = G$, for example). Then let $\theta' : F(X) \to G$ be the extension of the inclusion $\theta : X \to G$. Now $\operatorname{Im} \theta' = G$, since $\langle X \rangle = G$, so that if $K = \operatorname{Ker} \theta'$, then

$$G = \operatorname{Im} \theta' \cong F(X) / K \ ,$$

by an Isomorphism Theorem.

3. Further properties of $F(X)$

Several important properties of free groups depend only on elementary arguments involving the lengths of words. The examples which follow are fairly typical and are included for three reasons. a) They form a rather pleasing array of interrelated results; b) their culmination (Theorem 2) is in several works left as an exercise; c) Theorem 2 provides a useful illustration of the power of the Nielsen-Schreier theorem (proved in Chapter 2), of which it is an easy consequence (see Exercise 2).

Let $F = F(X)$ denote the free group on a fixed set X. The following definition will be useful.

Definition 2. A reduced word $a = x_1 x_2 \dots x_l$, $x_i \in X^{\pm}$, $1 \le i \le l$, is called *cyclically reduced* if $x_l \neq x_1^{-1}$.

Now let $a = x_1 \dots x_l$ be any reduced word in $X^{\pm}$, and let $a^2 = x_1 \dots x_{l-r} x_{r+l} \dots x_l$, reduced, so that $l(a^2) = l(a) - 2r$. How big can r be? Clearly, $r = 0$ if and only if a is cyclically reduced. To answer this in general, first let $l = 2k + 1$ be odd. Then it is clear that $r \le k$, for otherwise, $x_{k+1} = x_{k+1}^{-1}$, that is, $x_{k+1}^2 = e$, which is impossible by condition (ii) in Proposition 3. On the other hand, when $l = 2k$ is even, we must have $r < k$, for otherwise $x_k = x_{k+1}^{-1}$, contrary to the fact that a is reduced. It follows that $r < l/2$, whence $a = u^{-1} \overset{\vee}{a} u$, where

$$u^{-1} = x_1 \dots x_r = x_l^{-1} \dots x_{l-r+1}^{-1}, \quad \overset{\vee}{a} = x_{r+1} \dots x_{l-r} \ , \tag{5}$$

with $\overset{\vee}{a} \neq e$ and $x_{r+1} \neq x_{l-r}^{-1}$, so that $\overset{\vee}{a}$ is cyclically reduced. Note that

$$l(a^2) = 2l - 2r > 2l - l = l = l(a) \ .$$

More generally, for $n \in \mathbb{N}$, $a^n = u^{-1} \overset{\vee}{a}{}^n u$, and since it is clear that $\overset{\vee}{a}{}^n$ is cyclically reduced, it follows that

$$l(a^n) = nl(\overset{\vee}{a}) + 2r > (n-1) l(\overset{\vee}{a}) + 2r = l(a^{n-1}) \ . \tag{6}$$

Thus, no elements of $F(X)$ other than e can have finite order. Groups with this property are called *torsion-free*.

Proposition 5. $F(X)$ is torsion-free.

The next point is that free groups are as non-commutative as they can possibly be, in the following sense. In any group, if two elements are powers of a common element, then they must commute. The key lemma asserts that in $F(X)$, the converse holds.

Lemma 2. Let $a, b \in F(X)$ satisfy $ab = ba$. Then there is a $c \in F(X)$ such that $a = c^k$, $b = c^h$, $h, k \in \mathbb{Z}$.

Proof. Proceed by induction on $l(a) + l(b)$. Since the result is clear when either a or b is e, the basis for the induction is established, and we may assume that $a \neq e \neq b$. Letting $a = x_1 \dots x_l$, $b = y_1 \dots y_m$, assume that $l = l(a) \le l(b) = m$, by symmetry. Now consider the equation $ab = ba$ in reduced form:

$$x_1 \dots x_{l-r} y_{r+1} \dots y_m = y_1 \dots y_{m-r} x_{r+1} \dots x_l \ , \tag{7}$$

where $0 \le r \le \min(l,m) = l$, by assumption. Distinguish three cases.

Case (i): $r = 0$. Then it follows from (7) that $x_i = y_i$, $1 \le i \le l$, by comparing initial segments. Then, $b = au$, with $l(u) = m - l < m$, so that $l(a) + l(u) < l(a) + l(b)$. Then $au = b = a^{-1}ba = a^{-1}aua = ua$, and the inductive hypothesis yields that a and u are both powers of some $c \in F$, whence so is $b = au$.

Case (ii): $r = l$. Here, $y_i = x_{l-i+1}^{-1}$, $1 \le i \le l$, and $b = a^{-1}u$, with $l(u) = m - l < m$. It follows as in case (i) that a^{-1} and u commute and so are again powers of a common c, and $b = a^{-1}u$ is too.

Case (iii): $0 < r < l$. In this case,

$$x_1 = y_1,\ x_l = y_m,\ x_l = y_1^{-1},\ y_m = x_1^{-1}\ ,$$

whence

$$a = x_1 a' x_1^{-1},\ b = x_1 b' x_1^{-1}\ ,$$

where $l(a') = l - 2$, $l(b') = m - 2$. Conjugation of $ab = ba$ by x_1 yields $a'b' = b'a'$, and a', b' are powers of a common c', by induction. But then a, b are both powers of $c = x_1 c' x_1^{-1}$.

The desired conclusion holds in every case, and the proof is complete.

Proposition 6. (i) In a free group F, nth roots, when they exist, are unique, that is, if $a,b \in F$ satisfy $a^n = b^n$, $n \in \mathbb{N}$, then $a = b$.

(ii) Any element $w \in F$ has only finitely many roots, that is, the set $\{a \in F \mid a^n = w, \text{ some } n \in \mathbb{N}\}$ is finite.

Proof. (i) Write $a = u^{-1}\overset{\vee}{a}u$, $b = v^{-1}\overset{\vee}{b}v$ as in (5), with $\overset{\vee}{a}$, $\overset{\vee}{b}$ cyclically reduced and $l(u) = r$, $l(v) = s$, say. Now apply (6) to the equations $a^n = b^n$ and $a^{2n} = b^{2n}$ to obtain

$$nl(\overset{\vee}{a}) + 2r = nl(\overset{\vee}{b}) + 2s\ ,$$

$$2nl(\overset{\vee}{a}) + 2r = 2nl(\overset{\vee}{b}) + 2s\ .$$

Thus, $l(\overset{\vee}{a}) = l(\overset{\vee}{b})$ and $r = s$, and since the equation

$$u^{-1}\overset{\vee}{a}^n u = v^{-1}\overset{\vee}{b}^n v$$

involves no cancellation, it follows that $u = v$ and $\overset{\vee}{a} = \overset{\vee}{b}$, whence $a = b$.

(ii) Let a be a root of w, say $a^n = w$, $n \in \mathbb{N}$. If $w = e$, the result follows from Proposition 5. If $w \neq e$, neither is a nor $\overset{\vee}{a}$, and it follows from (6) that $n \leq l(w)$. Then, w is an nth power for at most finitely many n and for each such n, w is the nth power of at most one element by part (i). Hence the total number of roots is finite.

Proposition 6(i) can be used to strengthen Lemma 2, as follows.

Lemma 3. If $a^h b^k = b^k a^h$ for $a,b \in F$ and $h,k \in \mathbb{Z} \setminus \{0\}$, then a and b are powers of a common element.

Proof. By simple manipulation, we may assume that $h,k \in \mathbb{N}$. Then

$$a^h = b^k a^h b^{-k} = (b^k a b^{-k})^h \ ,$$

whence $a = b^k a b^{-k}$ by Proposition 6(i). Thus,

$$b^k = a b^k a^{-1} = (a b a^{-1})^k \ ,$$

and $b = a b a^{-1}$ for the same reason. Hence $ab = ba$, and the result follows from Lemma 2.

Proposition 7. Commutation is an equivalence relation on $F \setminus \{e\}$, that is, the centralizer $C(w) := \{w \in F \mid aw = wa\}$ of any $a \in F \setminus \{e\}$ is abelian.

Proof. We must show that if $u,v \in F$ both commute with some $a \in F \setminus \{e\}$, then they commute with each other. Assume that $ua = au$ and $va = av$, and that $u \neq e \neq v$ to avoid triviality. Thus, by Lemma 2, there exist $b,d \in F$ and $p,q,r,s \in \mathbb{Z} \setminus \{O\}$ such that

$$u = b^p, \ a = b^q, \ a = d^r, \ v = d^s \ .$$

But then b^q and d^r commute, as they are equal, and it follows from Lemma 3 that there is a $c \in F$ and $h,k \in \mathbb{Z}$ such that

$$b = c^h, \ d = c^k \ .$$

Thus, $u = c^{hp}$ and $v = c^{ks}$ must commute.

Finally, Proposition 7 can be strengthened as follows.