Section I.9. Free Groups, Free Products, and Generators and Relations

Note. This section includes material covered in Fraleigh's Sections VII.39 and VII.40. We define a free group on a set and show (in Theorem I.9.2) that this idea of "free" is consistent with the idea of "free on a set" in the setting of a concrete category (see Definition I.7.7). We also define generators and relations in a group presentation.

Note. To define a free group F on a set X, we will first define "words" on the set, have a way to reduce these words, define a method of combining words (this combination will be the binary operation in the free group), and then give a reduction of the combined words. The free group will have the reduced words as its elements and the combination as the binary operation. If set $X = \varnothing$ then the free group on X is $F = \langle e \rangle$.

Definition. Let X be a nonempty set. Define set X^{-1} to be disjoint from X such that $|X| = |X^{-1}|$. Choose a bijection from X to X^{-1} and denote the image of $x \in X$ as x^{-1}. Introduce the symbol "1" (with X and X^{-1} not containing 1). A *word* on X is a sequence $(a_1, a_2, \ldots)$ with $a_i \in X \cup X^{-1} \cup \{1\}$ for $i \in \mathbb{N}$ such that for some $n \in \mathbb{N}$ we have $a_k = 1$ for all $k \geq n$. The sequence $(1, 1, \ldots)$ is the *empty word* which we will also sometimes denote as 1.

Example 1. Let $X = \{x, y, z\}$. Then $X^{-1} = \{x^{-1}, y^{-1}, z^{-1}\}$. An example of a word is the sequence

$$w = (x, x, 1, x, 1, x, x, x^{-1}, x^{-1}, x^{-1}, x^{-1}, y, y, y, x^{-1}, x^{-1}, x, z, z, z^{-1}, z^{-1}, 1, 1, 1, \ldots).$$

Definition. A word $(a_1, a_2, \ldots)$ on set X is *reduced* provided

(i) for all $x \in X$, x and x^{-1} are not adjacent (that is, $a_i = x$ implies $a_{i+1} \neq x^{-1}$, and $a_i = x^{-1}$ implies $a_{i+1} \neq x$ for all $i \in \mathbb{N}$ and $x \in X$); and

(ii) $a_k = 1$ implies that $a_i = 1$ for all $i > k$.

Note. The word w in Example 1 is not reduced. If we interpret the sequence w as a product of its entries, it would be tempting to eliminate products giving 1, and to drop the 1's to get: $xyyyx^{-1}$. This is the motivation for the following.

Note. Every nonempty reduced word is of the form $(x_1^{\lambda_1}, x_2^{\lambda_2}, \ldots, x_n^{\lambda_n}, 1, 1, 1, \ldots)$ where $n \in \mathbb{N}$, $x_i \in X$ (the x_i's may not be distinct), and $\lambda_i = \pm 1$ (here we represent x as x^1). From now on we represent reduced words by dropping the 1's and eliminating the parentheses to get $x_1^{\lambda_1}, x_2^{\lambda_2}, \ldots, x_n^{\lambda_n}$. This differs from Fraleigh's notation in that Fraleigh allows integer exponents and so $x^1x^1x^1 = x^3$ and $y^{-1}y^{-1} = y^{-2}$. We denote the set of all reduced words on X as $F(X)$.

Note. In this new notation, two words $x_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n}$ and $y_1^{\delta_1}y_2^{\delta_2}\cdots y_m^{\delta_m}$ are *equal* (based on the definition of equal sequences) if and only if both are 1, or $m = n$, $x_i = y_i$, and $\lambda_i = \delta_i$ for all $i = 1, 2, \ldots, n$. In this way, each element of X can be thought of as an element of $F(X)$, and so X can be thought of as a subset of $F(X)$. Now we define a binary operation on $F(X)$. To do so, we must insure that the product of two reduced words is itself a reduced word.

Definition. Let X be a set and $F(X)$ the set of reduced words on X. For words $1, w \in F(X)$ define $1w = w1 = w$. For words $x_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n}$ and $y_1^{\delta_1}y_2^{\delta_2}\cdots y_m^{\delta_m}$ in $F(X)$, first consider the word formed by "juxtaposition" (that is, by placing one word after another) to get

$$(x_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n})(y_1^{\delta_1}y_2^{\delta_2}\cdots y_m^{\delta_m}) = x_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n}y_1^{\delta_1}y_2^{\delta_2}\cdots y_m^{\delta_m}.$$

However this new word may not be reduced.

(1) Suppose $n \leq m$ and let k be the largest integer such that $x_{n-j}^{\lambda_{n-j}} = y_{j+1}^{-\delta_{j+1}}$ for $j = 0, 1, \ldots, k-1$ (so $0 \leq k \leq n$). Then reduce the juxtaposition to

$$\begin{cases} x_1^{\lambda_1}x_2^{\lambda_2}\cdots x_{n-k}^{\lambda_{n-k}}y_{k+1}^{\delta_{k+1}}y_{k+2}^{\delta_{k+2}}\cdots y_m^{\delta_m} \text{ if } k < n \\ y_{n+1}^{\delta_{n+1}}y_{n+2}^{\delta_{n+2}}\cdots y_m^{\delta_m} \text{ if } k = n < m \\ 1 \text{ if } k = m = n. \end{cases}$$

(2) Suppose $m \leq n$ and let k be the largest integer such that $x_{n-j}^{\lambda_{n-j}} = y_{j+1}^{-\delta_{j+1}}$ for $j = 0, 1, \ldots, k-1$ (so $0 \leq k \leq m$). Then reduce the juxtaposition to

$$\begin{cases} x_1^{\lambda_1}x_2^{\lambda_2}\cdots x_{n-k}^{\lambda_{n-k}}y_{k+1}^{\delta_{k+1}}y_{k+2}^{\delta_{k+2}}\cdots y_m^{\delta_m} \text{ if } k < m \\ x_1^{\delta_1}x_2^{\delta_2}\cdots x_{n-m}^{\delta_{n-m}} \text{ if } k = m < n \\ 1 \text{ if } k = m = n. \end{cases}$$

Note. Since $x_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n}$ is a reduced word, $y_1^{\delta_1}y_2^{\delta_2}\cdots y_m^{\delta_m}$ is a reduced word, and $x_{n-k}^{\lambda_{n-k}} \neq y_{k+1}^{-\delta_{k+1}}$ then this new word is reduced.

Example. Let $X = \{x, y, z\}$. Then

$$(x, x, y, y, z, z, x, y, z, 1, 1, 1, \ldots)(z^{-1}, y^{-1}, x^{-1}, z^{-1}, z^{-1}, y^{-1}, z, 1, 1, 1, \ldots)$$

becomes the word (in the shorthand notation)

$$xxyyzzxyzz^{-1}y^{-1}x^{-1}z^{-1}z^{-1}y^{-1}z$$

which by the previous definition reduces to $xxyz$. Notice that $k = 6$ here.

Theorem I.9.1. If X is a nonempty set and $F = F(X)$ is the set of all reduced words on X, then F is a group under the binary operation defined in the previous definition. Also, $F = \langle X \rangle$ (where $\langle X \rangle$ represents the group generated by set X). The group $F = F(X)$ is called the *free group on set* X.

Note. The following properties hold for free groups:

(1) If $|X| \geq 2$ then the free group on X is nonabelian.

(2) Every element of a free group (except 1) has infinite order.

(3) If $X = \{a\}$, then the free group on X is the infinite cyclic group $\langle a \rangle \cong \mathbb{Z}$.

(4) Every subgroup of a free group is itself a free group on some set.

Property 1 holds by considering $x, y \in X$ with $x \neq y$. Then word $x^{-1}y^{-1}xy$ is reduced and so $x^{-1}y^{-1}xy \neq 1$ and hence $xy \neq yx$. Property 2 is Exercise I.9.1

and Property 3 is Exercise I.9.2. Hungerford refers to Property 4 as "a decidedly nontrivial fact" and refers to Rotman's *The Theory of Groups* (1973); this result appears as the "Nielsen-Schreier Theorem," Theorem 11.23 on page 258 (you can probably find Rotman's book in PDF online).

Theorem I.9.2. Let F be the free group on set X and $\iota : X \to F$ the inclusion map (see page 4). If G is a group and $f : X \to G$ is a map of sets, then there exists a unique homomorphism of groups $\overline{f} : F \to G$ such that $\overline{f}\iota = f$. In other words, F is a free object on the set X in the category of groups.

Note. Theorem I.7.8 tells us that if, in the concrete category $\mathcal{C}$, object F is free on set X and object F' is free on set X' where $|X| = |X'|$, then F and F' are equivalent (that is, there is morphism $f : F \to F'$ and morphism $g : F' \to F$ such that $f \circ g = 1_{F'}$ and $g \circ f = 1_F$). So by combining this result with Theorem I.9.2 (where the category is the category of all groups and the morphisms are group homomorphisms) we have that if F and F' are free groups on set X then $F \cong F'$.

Corollary I.9.3. Every group G is the homomorphic image of a free group.

Theorem. Gallian's "Universal Quotient Group Property."
Every group G is isomorphic to a quotient group of a free group.

Proof. Let G be a group. By Corollary I.9.3 there is a free group F and a homomorphism $\overline{f}$ such that $\overline{f}(F) = G$. Let $N = \text{Ker}(\overline{f})$. Then by the First Isomorphism Theorem (Corollary I.5.7), $G = \text{Im}(\overline{f}) \cong F/\text{Ker}(\overline{f}) = F/N$. ∎

Note I.9.1. By the Universal Quotient Group Property, group G is determined (up to isomorphism) if we know set X, group F, and kernel N. But group F is determined (up to isomorphism) by set X by Theorem I.7.8. Kernel N is determined by a subset of group F which generates N as a subgroup of F. If $x_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n} \in F$ is such a generator of N, then $\overline{f}(x_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n}) = x_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n} = e$ where $e \in G$. This equation is called a *relation* on the generators x_i of G. So if we know set X of generators of group G and a sufficient number of relations in order to determine kernel N, then group G is determined. It is not surprising that the choice of set X and set R of relations is not unique for a given group G (as is shown in Exercises I.9.6 and I.9.9).

Note I.9.2. By the previous note, we see that for any group G, there is a "complete description" of G in terms of a set X and a set of relations R (in which case $G \cong F/N$ where F is free on set X and the set of relations determine kernel N in terms of some of the generators of N). We now want to consider the converse. That is, given a set X and a set Y of reduced words on the elements of X, is there a group G such that G is generated by set X and such that all relations in the set $R = \{w = e \mid w \in Y\}$ are valid? The answer is "yes" as seen in the next note, provided we allow for the possibility that the elements of X may not be distinct. If, for example, $a, b \in X$ and $a^1b^{-1} \in Y$ then $ab^{-1} = e$ is an equation in R and it must be that in fact $a = b$ (in which case a^1b^{-1} turns out NOT to really be a reduced word).

Note I.9.3. Let X be a set of "generators" and Y a set of reduced words on the elements of X. Group G generated by set X which satisfies the relations in set $R = \{w = e \mid w \in Y\}$ is constructed as follows. Let group F be the free group on X (unique up to isomorphism by the Note following Theorem I.9.2 in these class notes) and let N be the normal subgroup of F generated by set Y (that is, the intersection of all normal subgroups of F that contain Y—see Exercise I.5.2). Let G be the quotient group F/N. "Identify" X with its image in F/N under the canonical epimorphism (see page 43); as noted above, this may involve identifying some elements of X with one another (in particular, all elements of X in the kernel of the canonical epimorphism are identified together; more generally, elements of X in the same coset of N under the canonical epimorphism are identified together). Then G is a group generated by X (by definition). Any relation $w = e$ in set R is satisfied because, for $w = x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_n^{\lambda_n} \in Y$ we have $w = e$ and so under the canonical epimorphism w is mapped to N. So, as an element of F, we have $w \in N$. Then $wN = (x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_n^{\lambda_n})N = N$. Since N is the identity in $G = F/N$, we have in group G that $w = x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_n^{\lambda_n} = e$ where e denotes the identity in $G = F/N$.

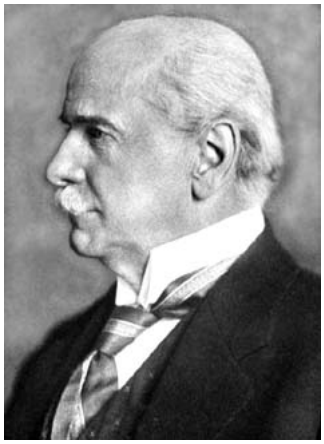
Definition I.9.4. Let X be a set and Y a set of reduced words on X. A group G is the *group defined by the generators* $x \in X$ *and relations* $w = e$ for $w \in Y$ provided $G \cong F/N$, where F is the free group on X and N is the normal subgroup of F generated by Y. We say that $(X \mid Y)$ is a *presentation* of group G.

Note. By Notes I.9.1, I.9.2, and I.9.3 above we know that any sets X and Y result in a group G. Hungerford says that G "is the largest possible such group in the sense" given by the following theorem. The theorem involves an onto homomorphism (i.e., an epimorphism) which has a cardinality implication (namely, $|G| \geq |H|$).

Theorem I.9.5. von Dyck's Theorem.

Let X be a set, Y a set of reduced words on X and G the group defined by the generators $x \in X$ and relations $w = e$ for $w \in Y$. If H is any group such that $H = \langle X \rangle$ and H satisfies all the relations $w = e$ for $w \in Y$, then there is an epimorphism mapping $G \to H$.

Note. Theorem I.9.5 is named for Walther von Dyck (1856–1934; not "Van Dyck" as Hungerford states).



Walther Franz Anton von Dyck (December 6, 1856 to November 5, 1934)

He was a student of Felix Klein's in Munich. In addition to his result on group presentations, he made contributions to function theory, topology and potential

theory. He was involved in the creation of the German Museum of Natural Science and Technology which had interactive displays of various scientific principles; these displays were the first of their kind and soon were copied around the world. von Dyck was also part of a project to publish the complete works of Johannes Kepler. For more information, check out the University of Saint Andrews biographical webpage.

Example. To illustrate von Dyck's Theorem, consider $X = \{b\}$ and $Y = \{b^m\}$ (here we use exponential notation instead of, as above, insisting that all exponents be ± 1). Then $(X \mid Y)$ is a presentation of $G \cong \mathbb{Z}_m$ (by Exercise I.9.9). When m is even, another group generated by X and which also satisfies the relation $b^m = e$ is the group $H = \mathbb{Z}_{m/2}$. The epimorphism mapping $G \to H$ (or $\mathbb{Z}_m \to \mathbb{Z}_{m/2}$ here) maps the equivalence class containing x to the equivalence class containing $x \pmod{m/2}$.

Note. As in the previous example, we use exponential notation and allow integer-valued exponents (as Fraleigh does). Hungerford mentions "the sort of *ad hoc* arguments that are often the only way of investigating a given presentation."

Example I.9.A. Let $X = \{a, b\}$ and $R = \{a^4 = e, a^2b^{-2} = e, abab^{-1} = e\}$. Exercise I.4.14 shows that the quaternions, Q_8, is a group of order 8 generated by elements a, b which satisfy the three relations. So by Theorem I.9.5 there is an epimorphism $\varphi : G \to Q_8$. Hence $|G| \geq |Q_8| = 8$. Since G is generated by a, b then

by Theorem I.2.8 each element of G is of the form $a_1^{n_1}a_2^{n_2}\cdots a_k^{n_k}$ where $a_i \in \{a, b\}$ and $n_i \in \mathbb{Z}$. Since $a^4 = e$ and $a^2b^{-2} = e$ (or $a^2 = b^2$ or $a^4 = b^4 = e$) then we can take all $n_i \in \{0, 1, 2, 3\}$. Since $abab^{-1} = e$ then $ba = a^{-1}b = a^3b$. Therefore in any product of powers of a's and b's, we can one-by-one move the b's to the right of the a's (picking up factors of a^2 along the way). So every element of G is of the form a^ib^j where $i, j \in \{0, 1, 2, 3\}$. Since $a^2 = b^2$, if $j = 2$ or 3 then we can replace b^j by a^2b^{j-2}, thus giving all elements of G of the form a^ib^j where $i \in \{0, 1, 2, 3\}$ and $j \in \{0, 1\}$. Hence $|G| \leq 8$. So epimorphism φ is one-to-one and φ is in fact an isomorphism. So $G \cong Q_8$.

Example. By Exercise I.9.8, $(X \mid Y)$ where $X = \{a, b\}$ and $Y = \{a^n, b^2, abab\}$ (for $n \geq 3$) is a presentation of the dihedral group D_n. Geometrically, a is a rotation of a regular n-gon and b is a mirror reflection.

Example. Let X be a set and $Y = \varnothing$. Then $(X \mid Y)$ is a presentation of the free group F on set X (the group is "free" in the sense that there are no relations imposed on the generators of F). Notice that since $Y = \varnothing$ then $N = \langle e \rangle$ and so $F/N = F/\langle e \rangle \cong F$.

Note. A relative of the Conway, Curtis, Norton, Parker, and Wilson $\mathbb{ATLAS}$ *of Finite Groups* (Oxford: Clarendon Press, 1985) is "ATLAS of Finite Groups—Version 3" and is available online: $\mathbb{ATLAS}$ *of Finite groups* (accessed 2/21/2021). This gives a presentation of the group M_{11} (a "Mathieu group"—see my "Finite Simple

Groups", a supplement to Introduction to Modern Algebra [MATH 4127/5127]). The group is of order $|M_{11}| = 7920$ and a Cayley table for the group would consist of $(7920)^2 = 62{,}726{,}400$ entries. The online ATLAS gives a presentation of M_{11} as $X = \{a, b\}$ and $Y = \{a^2, b^4, (ab)^{11}, (abababbababbabb)^4\}$. So one advantage of a group presentation is that it gives a complete description of the group without actually listing the Cayley table for the group.

Note. We now look at coproducts in the category of groups. We go light on the details, which are similar to those in the construction of free groups.

Note. Given a family of groups $\{G_i \mid i \in I\}$ we may assume (by relabeling of the elements of the G_i) that the G_i are mutually disjoint sets (so, for example, we use a different symbol for each identity element).

Definition. Let $\{G_i \mid i \in I\}$ be a disjoint collection of groups (as described above). Let $X = \cup_{i \in I} G_i$ and let $\{1\}$ be a one-element set disjoint from X. A *word* on X is a sequence $(a_1, a_2, \ldots)$ such that $a_i \in X \cup \{1\}$ and for some $n \in \mathbb{N}$ we have $a_i = 1$ for all $i > n$. Word $(a_1, a_2, \ldots)$ is *reduced* provided

(i) no $a_i \in X$ is the identity element in its group G_j;

(ii) for all $i, j \geq 1$ we have that a_i and a_{i+1} are not in the same group G_j;

(iii) $a_k = 1$ implies that $a_i = 1$ for all $i \geq k$.

Note. The word $1 = (1, 1, 1, \ldots)$ is reduced. Every reduced word (other than 1) may be written uniquely as $a_1 a_2 \cdots a_n = (a_1, a_2, \ldots, a_n, 1, 1, \ldots)$ where $a_i \in X$.

"Definition." Let $\prod_{i \in I}^{*} G_i$ be the set of all reduced words on X. If I is finite we denote this as $G_1 * G_2 * \cdots * G_n$. Define a binary operation on $\prod_{i \in I}^{*} G_i$ as follows. $1 = (1, 1, 1, \ldots)$ is the identity element. The product of two reduced words (where neither is 1) is the reduced word which results from juxtaposition and the following reduction.

- If juxtaposition puts an element of G_i next to its inverse, then eliminate these two elements.
- If juxtaposition leads to two elements of a given G_i next to each other, then replace this pair with its product in G_i.

$\prod_{i \in I}^{*} G_i$ is the *free product* of family $\{G_i \mid i \in I\}$ under this binary operation.

Example. Hungerford illustrates the product as follows. Let $a_i, b_i \in G_i$ for $i = 1, 2, 3$. Then $(a_1 a_2 a_3)(a_3^{-1} b_2 b_1 b_3) = a_1 c_2 b_1 b_3 = (a_1, c_2, b_1, b_3, 1, 1, 1, \ldots)$ where $c_2 = a_2 b_2$.

Note. We claim the following without proof.

Theorem. Let $\{G_i \mid i \in I\}$ be a family of groups. Then the free product $\prod_{i \in I}^{*} G_i$ is a group under the binary operation described above.

Note. The following result shows that the free product $\prod^* G_i$ is a coproduct in the category of groups.

Theorem I.9.6. Let $\{G_i \mid i \in I\}$ be a family of groups and $\prod^*_{i\in I} G_i$ their free product. If $\{\psi_i : G_i \to H \mid i \in I\}$ is a family of group homomorphisms, then there exists a unique homomorphism $\psi : \prod^* G_i \to H$ such that $\psi\iota_i = \psi_i$ for all $i \in I$ (where $\iota_i : G_i \to \prod^* G_i$ is the inclusion map) and this property determines $\prod^* G_i$ uniquely up to isomorphism. In other words, $\prod^*_{i\in I} G_i$ is a coproduct in the category of groups.

Note. For a final comment on group presentations, we briefly address the "group isomorphism problem." The problem is to determine if two finite group presentations represent isomorphic groups. The problem was introduced by Max Dehn in "Über unendliche diskontinuierlich Gruppen [On Infinite Discontinuous Groups]," *Mathematische Annalen* 71 (1911), 116–144. Dehn also introduced the "word problem" and the "conjugacy problem" in this paper. The word problem is to decide whether two words written in terms of the generators of a group (given by a presentation) represent the same element of the group. The conjugacy problem is to determine if for two words x and y in a group (given by a presentation) there is z in the group such that $y = zxz^{-1}$. All three problems are undecidable in the following sense. "There does not exist a computer algorithm that correctly solves every instance of [these problems], regardless of how much time is allowed for the algorithm to run." (This information is from the Wikipedia page on the Group Isomorphism

Problem.) The undecidability of the isomorphism problem was given in 1958 by S.I. Adjan (in "On Algorithmic Problems in Effectively Complete Classes of Groups," Dokl. Akad. Nuak SSSR **123** (1958), 13–16, in Russian) and M.O. Rabin (in "Recursive Unsolvability of Group Theoretic Problems," *Annals of Mathematics* **67**(1) (1958), 172–194; a preview of this article is online through JSTOR). The result is called the Adjan-Rabin Theorem. A reference on this and related problems is George Sacerdote's "Some Undecidable Problems in Group Theory" *Proceedings of American Mathematical Society* **36**(1) (1972), 231–238; a preview of this article is online through JSTOR. A more detailed discussion of group presentations can be found in Chapter IV "Presentation of Groups" in *Introduction to Knot Theory*, Richard Crowell and Ralph Fox, Ginn and Company, Boston (1963) (also, Dover Publications (2008)).

Revised: 5/1/2021