Algebra I

Randall R. Holmes
Auburn University

Last revision: August 10, 2019

Contents

0 Introduction

This course is primarily a course in group theory. We will start from scratch, even giving the definition and elementary properties of a group. However, the material in the undergraduate group theory course will be reviewed rather quickly and at times using an approach more sophisticated than that used in the first course. After the review, more advanced topics in group theory will be covered. If time permits, topics from module theory will be included.

The course contains an introduction to category theory, which is playing an ever-increasing role in mathematical discourse. I do not intend to suggest that category theory is a major focus of the course (it is not), but I do want to take a moment to say a little bit about what it is in case you are unfamiliar with the idea.

Your experience with mathematical constructs to date has most likely been confined to an inspection of internal structure. For instance, in group theory you have studied things like the order of an element, the cyclic subgroup generated by an element, the cosets of a subgroup, and so forth. In topology (or analysis) you have studied things like limit points, interiors of sets, boundaries of sets, least upper bounds, and so forth.

In category theory, one looks at the bigger picture. Taking the case of groups, for example, the groups themselves become the elements as one steps back and views the collection of all groups as a new mathematical construct, a "category." Information is obtained by studying the structure preserving maps (homomorphisms) running between the groups. The standard visualization is that of a directed graph, which is roughly an array of points together with various arrows joining the points. The points represent the groups and an arrow from one point to another represents a homomorphism between the corresponding groups.

As a simple example of this new way of thinking, take the trivial group. Using the internal viewpoint, one characterizes the trivial group as the group having a single element (one has to look inside the group to see that it has only one element). In category theory, all groups look alike (they are all points) except for the array of arrows (homomorphisms) going out of them and coming into them. So how is one even to recognize the trivial group in the vast collection of all groups? A little reflection reveals that it is the only group with a single arrow going to each other group. (I have taken the liberty of identifying isomorphic groups here.)

One can form other categories as well, like the category of all rings, or the category of all topological spaces, or the category of all differentiable manifolds. Then one can step back even further and view these categories as points themselves with structure-preserving maps (functors) represented by arrows between them. For instance, each pointed topological space gives rise to a certain group, namely, its fundamental group at the distinguished point. This correspondence defines a functor from the category of all pointed topological spaces to the category of all groups.

By seeing how a mathematical construct interacts with other mathematical constructs through functors one gains insights beyond those made possible by an isolated study.

1 Definition of group and examples

1.1 Definition

A **group** is a pair $(G, *)$, where G is a set and $*$ is a binary operation on G satisfying the following:

(i) $x * (y * z) = (x * y) * z$ for all $x, y, z \in G$,

(ii) there exists $e \in G$ such that $x * e = x$ and $e * x = x$ for all $x \in G$,

(iii) for each $x \in G$ there exists $y \in G$ such that $x * y = e$ and $y * x = e$.

If $(G, *)$ is a group, we say that G is a group under $*$ (or just that G is a group, when the binary operation is clear from the context). Part (i) says that $*$ is **associative**. An element e satisfying (ii) is an **identity** element. An element y satisfying (iii) is an **inverse** of x.

1.2 Examples: Z, Q, R, C

- $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{C}$ are all groups under addition. In each case, an identity element is 0 and an inverse of x is $-x$.

- $\mathbf{Q}^\times$, $\mathbf{R}^\times$, and $\mathbf{C}^\times$ are all groups under multiplication. (The symbol $\times$ signifies that the element 0 is omitted.) In each case, an identity element is 1 and an inverse of x is $1/x$.

1.3 Nonexamples

- $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{C}$ are not groups under multiplication; the number 1 is an identity element (and the only candidate for such), but the element 0 has no inverse.

- $\mathbf{Z}^\times$, $\mathbf{Q}^\times$, $\mathbf{R}^\times$, and $\mathbf{C}^\times$ are not groups under addition. In fact, they are not even closed under addition since, for instance, they contain 1 and -1 but not the sum $1 + (-1) = 0$.

- $\mathbf{Z}^\times$ is not a group under multiplication; the number 1 is an identity element (and the only candidate for such), but 2 has no inverse.

1.4 Example: Integers modulo n

Let n be a positive integer and put $\mathbf{Z}_n = \{0, 1, \ldots, n-1\}$. Define a binary operation $+$ on this set by putting $x + y = r$, with r being the remainder upon division by n of $x + y$ (usual sum). For instance, if $n = 5$, then $4 + 3 = 2$. This binary operation is **addition modulo** n. $(\mathbf{Z}_n, +)$ is a group, the **group of integers modulo** n.

The check that $+$ is associative is not difficult, but it is tedious because it requires that one check cases. (See Section 2.13 for a proof of associativity that avoids this checking of cases.)

1.5 Example: $\mathbf{R}^n$

Let n be a positive integer. The set $\mathbf{R}^n = \{(x_1, x_2, \ldots, x_n) \,|\, x_i \in \mathbf{R}\}$ is a group under componentwise addition.

1.6 Example: General linear group

Let n be a positive integer. Denote by $\mathrm{GL}_n(\mathbf{R})$ the set of all $n \times n$ matrices with entries coming from $\mathbf{R}$ and having nonzero determinant. $(\mathrm{GL}_n(\mathbf{R}), \cdot)$ is a group, where $\cdot$ is usual matrix multiplication. It is the **general linear group** of degree n over $\mathbf{R}$.

1.7 Example: Symmetric group

Let X be a nonempty set. A bijection from X to itself is a **permutation** of X. Denote by $\mathrm{Sym}(X)$ the set of all permutations of X. $(\mathrm{Sym}(X), \circ)$ is a group, where $\circ$ is function composition. It is the **symmetric group** on X. The identity element of this group is the identity map $\varepsilon = 1_X : X \to X$ given by $\varepsilon(x) = x$ $(x \in X)$. It is customary to use juxtaposition to denote the composition of functions, so for $\sigma, \tau \in \mathrm{Sym}(X)$ one writes $\sigma\tau$ to mean $\sigma \circ \tau$.

For $n \in \mathbf{N}$, the **symmetric group of degree** n (or the **symmetric group on** n **elements**) is $\mathrm{Sym}(X)$, where $X = \{1, 2, \ldots, n\}$. This group is denoted S_n.

It is convenient to identify an element σ of S_n with the $2 \times n$ matrix that displays the numbers 1 to n in the top row and each corresponding image directly beneath:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

1.8 Example: Circle group

Let U denote the set of those complex numbers that have modulus one: $U = \{z \in \mathbf{C} \,|\, |z| = 1\}$. Writing the complex number z in the form $a+bi$, we have $|z| = \sqrt{a^2+b^2}$, so U identifies with the unit circle under the correspondence $a+bi \leftrightarrow (a,b)$ between the set of complex numbers and the points in the plane. U is closed under multiplication of complex numbers and it is a group under this binary operation, the **circle group**.

1.9 Example: Group of nth roots of unity

Let n be a positive integer. An **nth root of unity** is a complex number z satisfying $z^n = 1$. Let U_n be the set of all nth roots of unity. Under the correspondence $a+bi \leftrightarrow (a,b)$ this set is identified with n points on the unit circle spaced evenly and including $(1,0)$. More precisely,

$$\begin{aligned} U_n &= \{e^{2\pi j i/n} \,|\, 0 \leq j < n \ (j \in \mathbf{Z})\} \\ &= \{\cos(2\pi j/n) + i\sin(2\pi j/n) \,|\, 0 \leq j < n \ (j \in \mathbf{Z})\}. \end{aligned}$$

U_n is closed under multiplication and it is a group under this binary operation, the **group of nth roots of unity**.

1 – Exercises

1–1 Put $G = \mathbf{R}\backslash\{-1\}$. For $a, b \in G$, put $a * b = a + b + ab$. Prove that $(G, *)$ is a group. (Note: You need to show that $*$ is a well-defined binary operation on G.)

1–2 Let X be a set and for subsets S and T of X put $S - T = \{s \in S \,|\, s \notin T\}$ and $S + T = (S - T) \cup (T - S)$. Prove that $(P(X), +)$ is an abelian group, where $P(X)$ denotes the power set of X (i.e., the set of all subsets of X).

Hint: Associativity of $+$ can be shown using an elementary (but tedious) argument. Consider instead using the (modular) characteristic function $\chi_S : X \to \mathbf{Z}_2$ of $S \in P(X)$ given by

$$\chi_S(x) = \begin{cases} 1, & \text{if } x \in S, \\ 0, & \text{if } x \notin S. \end{cases}$$

Note that for $S, T \in P(X)$ we have $S = T$ if and only if $\chi_S = \chi_T$.

2 Elementary notions

2.1 Multiplicative notation

In the definition of a group (1.1) the symbol $*$ is used to emphasize the fact that the notation for the binary operation can be anything (e.g., $+$, $\cdot$, $\circ$). However, when referring to a group in general we will always denote the binary operation by $\cdot$, write $x \cdot y$ simply as xy, and call this the product of x and y.

2.2 Abelian group

A group G is **abelian** if its binary operation is commutative, that is, if $xy = yx$ for all $x, y \in G$. A group is **nonabelian** if it is not abelian.

The general linear group $(\mathrm{GL}_n(\mathbf{R}), \cdot)$ is nonabelian if $n \geq 2$, and the symmetric group $\mathrm{Sym}(X)$ is nonabelian if $|X| \geq 3$. All of the other groups in Section 1 are abelian.

2.3 Generalized associativity

The assumption of associativity (1.1(i)) implies that when indicating the product of three group elements it is unnecessary to use parentheses. A routine proof by induction shows that this is true of a product of more than three group elements as well.

For instance, let G be a group and let $x, y, z, u, v \in G$. The expression

$$xyzuv$$

can be computed as

$$(((xy)z)u)v$$

or as

$$((xy)(zu))v,$$

or using any of the other possible groupings, and the result will be the same (provided that the order of the factors remains the same). This is the **generalized associativity property**. Even though parentheses are unnecessary, they are often used, nonetheless, to draw the reader's attention to particular groupings.

2.4 Unique identity, unique inverse

Let G be a group.

THEOREM.

(i) *There is a unique identity element in G.*

(ii) *Each element of G has a unique inverse.*

Proof. (i) If e and e' are identity elements of G, then $e = ee' = e'$.

(ii) Let $x \in G$ and let y and z be inverses of x. Then $y = ye = y(xz) = (yx)z = ez = z$. □

It now makes sense to speak of *the* identity of the group G; it is denoted e. Similarly, if $x \in G$ it makes sense to speak of *the* inverse of x; it is denoted x^{-1}. (For some specific groups other notations are more customary. For instance, when the binary operation is $+$, one writes 0 for e and $-x$ for x^{-1}.)

2.5 Left and right cancellation

Let G be a group and let $x, y, z \in G$.

THEOREM.

(i) *If $xy = xz$, then $y = z$.*

(ii) *If $xz = yz$, then $x = y$.*

Proof. (i) If $xy = xz$, then $y = ey = x^{-1}xy = x^{-1}xz = ez = z$.

(ii) Similar. □

(i) and (ii) are the **left** and **right cancellation properties**, respectively.

2.6 Properties of inverse

Let G be a group and let $x, y \in G$.

THEOREM.

(i) *If $xy = e$, then $y = x^{-1}$ and $x = y^{-1}$.*

(ii) $(x^{-1})^{-1} = x$,

(iii) $(xy)^{-1} = y^{-1}x^{-1}$.

Proof. (i) Assume that $xy = e$. We have $xy = e = xx^{-1}$, so by left cancellation (2.5) we get $y = x^{-1}$. Similarly, $x = y^{-1}$.

(ii) Since $x^{-1}x = e$, the inverse of x^{-1} is x (by part (i)), that is, $(x^{-1})^{-1} = x$.

(iii) Since

$$(xy)(y^{-1}x^{-1}) = x(yy^{-1})x^{-1} = xex^{-1} = xx^{-1} = e,$$

the inverse of xy is $y^{-1}x^{-1}$ (by part (i)), that is $(xy)^{-1} = y^{-1}x^{-1}$. □

2.7 Solving equations

Let G be a group.

Theorem. *For any $a, b \in G$, there exist unique $x, y \in G$ such that $ax = b$ and $ya = b$.*

Proof. Let $a, b \in G$. We have $ax = b$, where $x = a^{-1}b$. Moreover, if also $x' \in G$ satisfies $ax' = b$, then

$$x' = ex' = a^{-1}ax' = a^{-1}b = x.$$

The statements involving y are proved similarly. □

(Cf. Exercise 2–2.)

2.8 Power of element

Let G be a group and let $x \in G$. Define $x^0 = e$ and for a positive integer n, define $x^n = xx\cdots x$ (n factors) and $x^{-n} = (x^{-1})^n$. This defines the nth **power** x^n of the element x for every $n \in \mathbf{Z}$.

Theorem. *For all $m, n \in \mathbf{Z}$,*

(i) $x^{-n} = (x^n)^{-1}$,

(ii) $x^m x^n = x^{m+n}$,

(iii) $(x^m)^n = x^{mn}$.

Sketch of proof. The proofs of these formulas are carried out using induction. Various cases, depending on the signs of m and n, are treated separately. As an aid to remembering the formulas, it is worthwhile to keep in mind the case of positive m and n. For instance, in this case

$$x^m x^n = (\underbrace{xx\cdots x}_{m \text{ factors}})(\underbrace{xx\cdots x}_{n \text{ factors}}) = x^{m+n},$$

which proves (ii) (albeit without the rigor of an inductive proof). □

If G is an additive group (meaning that the binary operation is denoted $+$), then the power x^n ($n \in \mathbf{N}$) really means $x + x + \cdots + x$ (n summands), and so it is written nx.

2.9 Order of group, order of group element

Let G be a group. The **order** of G, denoted $|G|$ is the cardinality of the set G. The group G is a **finite group** if $|G|$ is finite; otherwise, G is an **infinite group**. If G is finite, then the order of G is simply the number of elements in G.

Let $x \in G$. If $x^n = e$ for some positive integer n, then the least such positive integer is the **order** of x, denoted $o(x)$. If no such integer exists, then the order of x is defined to be $\aleph_0$ (the cardinality of the set of integers), and one says x has infinite order. (Cf. 3.9.)

For instance, the element $x = 3$ of the group $\mathbf{Z}_{12}$ has order 4, since $1x = 3, 2x = 6, 3x = 9, 4x = 0$ (see Section 2.8), while the element $x = 2$ of the group $\mathbf{Q}^\times$ (under multiplication) has infinite order, since $x^n = 2^n \neq 1 = e$ for all $n \in \mathbf{N}$.

Theorem. *If $x \in G$ has finite order, then $x^m = e$ if and only if $o(x) \mid m$ ($m \in \mathbf{Z}$).*

Proof. Let x be an element of G of finite order n and let $m \in \mathbf{Z}$. Assume that $x^m = e$. By the division algorithm, there exist integers q and r with $0 \leq r < n$ such that $m = qn + r$. Now

$$x^r = x^{m-qn} = x^m(x^n)^{-q} = e(e)^{-q} = e,$$

so $r = 0$ due to the minimality property of order. Thus, $m = qn$ and $o(x) = n \mid m$, as desired.

Conversely, if $o(x) \mid m$, then $m = o(x)q$ for some integer q, whence

$$x^m = x^{o(x)q} = (x^{o(x)})^q = e^q = e,$$

as desired. □

2.10 Direct product of two groups

Let G_1 and G_2 be two groups. The set $G_1 \times G_2 = \{(x_1, x_2) \mid x_i \in G_i\}$ is a group under **componentwise multiplication**:

$$(x_1, x_2)(y_1, y_2) = (x_1y_1, x_2y_2).$$

It is the **direct product** of G_1 and G_2.

If the symbol $+$ is used for the binary operations in both G_1 and G_2, then the direct product is written $G_1 \oplus G_2$ and it is called the **direct sum** of G_1 and G_2. In this case, the binary operation is **componentwise addition**:

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2).$$

A more general notion of direct product (of which $\mathbf{R}^n$ is an example) is given in Section 16.1.

2.11 Operation table

Let G be a finite group. The **operation table** of G is the table with rows and columns labeled with the elements of G (in a fixed order, usually with e coming first) and with the product xy displayed in the row labeled x and the column labeled y ($x, y \in G$).

For example, the operation table of $\mathbf{Z}_4$ is

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

2.12 Isomorphism

A **binary structure** is a pair $(S, *)$ with S a nonempty set and $*$ a binary operation on S. A group is an example of a binary structure. As with groups we refer to a binary structure $(S, *)$ by just S when no confusion can arise.

Let $(S, *)$ and $(T, \circ)$ be two binary structures. An **isomorphism** from S to T is a bijection $\varphi : S \to T$ satisfying the **homomorphism property**:

$$\varphi(x * y) = \varphi(x) \circ \varphi(y) \text{ for all } x, y \in S.$$

S and T are **isomorphic**, written $S \cong T$, if there exists an isomorphism from S to T. (This definition appears to be asymmetrical, but in fact $S \cong T$ if and only if $T \cong S$ by Exercise 2–3.)

Assume that S and T are isomorphic and let $\varphi : S \to T$ be an isomorphism. We can view φ as a “renaming function”; it takes an element x in S and renames it $\varphi(x)$ in T. Since φ is injective, no two elements of S end up with the same name after renaming, and since it is surjective, every name in T gets used. Moreover, since φ satisfies the homomorphism property, the binary operation $\circ$ in T acts on the renamed elements in exactly the same way the binary operation $*$ in S acts on the elements before renaming.

In short, the renaming function φ accomplishes nothing but a change in the notation for the elements of S and a change in the symbol for the binary operation. This means that the binary structures $(S, *)$ and $(T, \circ)$ are exactly the same aside from notational choices.

As an illustration, we have that $(\mathbf{Z}_4, +) \cong (U_4, \,\cdot\,)$. Indeed, $\varphi : \mathbf{Z}_4 \to U_4$ by

$$\begin{array}{ccr} 0 & \longmapsto & 1 \\ 1 & \longmapsto & i \\ 2 & \longmapsto & -1 \\ 3 & \longmapsto & -i \end{array}$$

is an isomorphism (Exercise 2–4). This renaming function transforms the operation table of $\mathbf{Z}_4$ to that of U_4 and therefore effects nothing more than a change of notation:

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$$\longrightarrow$$

$\cdot$	1	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

2.13 Example

Since two isomorphic binary structures are identical, except possibly for the notation used for their elements and the symbol used for their binary operations (Section 2.12), it follows that if one has a property expressible entirely in terms of its elements and its binary operation, then the other must also have that property. Here is an illustration of this principle:

Theorem. *Let $(S, *)$ and $(T, \circ)$ be binary structures and assume that $S \cong T$. If $*$ is associative, then so is $\circ$.*

Proof. Assume that $*$ is associative. Let $x', y', z' \in T$. Since $S \cong T$, there exists an isomorphism $\varphi : S \to T$. Since φ is surjective, there exist $x, y, z \in S$ such that $\varphi(x) = x'$, $\varphi(y) = y'$, and $\varphi(z) = z'$. Therefore,

$$\begin{aligned} x' \circ (y' \circ z') &= \varphi(x) \circ (\varphi(y) \circ \varphi(z)) = \varphi(x) \circ \varphi(y * z) = \varphi(x * (y * z)) \\ &= \varphi((x * y) * z) = \varphi(x * y) \circ \varphi(z) = (\varphi(x) \circ \varphi(y)) \circ \varphi(z) \\ &= (x' \circ y') \circ z', \end{aligned}$$

which proves that $\circ$ is associative. □

Recall that we refrained from proving, due to the unpleasantness of checking cases, that the associative property holds for $(\mathbf{Z}_n, +)$. We now have a way around this checking of cases. By Exercise 2–4 (and Exercise 2–3), $(U_n, \cdot\,) \cong (\mathbf{Z}_n, +)$. Now the elements of U_n are complex numbers, and it is well-known that multiplication of complex numbers is associative. Therefore, the theorem says that $+$ (addition modulo n) is associative as well.

If you know a given binary structure has a certain property and you wish to conclude that a binary structure isomorphic to that binary structure has the same property, then it is usually safe to skip this formalism of a theorem and proof and simply draw the conclusion. For instance, assuming a group G is isomorphic to a group H, it is customary just to assert: if G is abelian, then so is H; if G is finite, then so is H; if G has an element of order two, then so does H; and so on.

2 – Exercises

2–1 Let G be a group. Assume that $x^2 = e$ for every $x \in G$. Prove that G is abelian.

2–2 Let G be a nonempty set on which an associative binary operation $\cdot$ is defined. Assume that for any $a, b \in G$ there exist $x, y \in G$ such that $ax = b$ and $ya = b$. Prove that G is a group (cf. Theorem of 2.7).

2–3 Prove that the property of being isomorphic ($\cong$) is an equivalence relation (i.e., reflexive, symmetric, and transitive) on the class of all binary structures.

2–4 Prove that $(\mathbf{Z}_n, +) \cong (U_n, \cdot\,)$ $(n \in \mathbf{N})$.

2–5 Let X be a nonempty set, let S be the set of all functions $X \to X$, let G be a subset of S that is closed under composition $\circ$ of functions, and assume that $(G, \circ)$ is a group.

(a) Show by example that it is possible to have $|G| > 1$ and $G \nsubseteq \mathrm{Sym}(X)$.

(b) Prove that if G contains an injection, then $G \subseteq \mathrm{Sym}(X)$.

2–6 Prove that a group of even order has an element of order two (called an **involution**).

2–7 Let G be a group and let a and b be elements of G of finite order.

(a) Prove that if $ab = ba$, then ab has finite order.

(b) Give an example to show that it is possible for ab to have infinite order.

HINT: For (b) consider $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$.

2–8 Let G be a group, let p be a prime number, and let a be an element of G of finite order m such that $p \,|\, m$. Prove that if $b \in G$ and $b^p = a$, then b has order pm.

2–9 Let G be a group. Assume that for each three-element subset $\{x, y, z\}$ of G, either $xy = yx$, $xz = zx$, or $yz = zy$. Prove that G is abelian.

3 Subgroup

3.1 Definition

Let G be a group. A subset H of G is a **subgroup**, written $H \leq G$, if

(i) $e \in H$,

(ii) $x, y \in H$ implies $xy \in H$,

(iii) $x \in H$ implies $x^{-1} \in H$.

In (i), e denotes the identity of G. Part (ii) says that H is closed under the binary operation of G. Part (iii) says that H is closed under inversion.

If H is a subgroup of G, then H is a group in its own right with binary operation being that obtained by restricting the binary operation on G to H.

3.2 Proper subgroup, trivial subgroup

Let G be a group. G is a subgroup of itself. If H is a subgroup of G, but $H \neq G$, then H is a **proper subgroup**, written $H < G$. The singleton set $\{e\}$ is a subgroup of G, the **trivial subgroup**. (Authors tend to disagree in their use of the terms "proper" and "trivial.")

3.3 Examples

- $\mathbf{Z} \leq \mathbf{Q} \leq \mathbf{R}$ (under addition).
- For any integer n, we have $n\mathbf{Z} \leq \mathbf{Z}$, where $n\mathbf{Z} = \{nm \,|\, m \in \mathbf{Z}\}$.
- Although $\mathbf{R}^\times$ is a group under multiplication and $\mathbf{R}^\times \subset \mathbf{R}$ it is *not* the case that $\mathbf{R}^\times$ is a subgroup of the group $\mathbf{R}$ under addition.

3.4 Subgroups of the symmetric group

Let X be a nonempty set and let $Y \subseteq X$. Let

$$C(Y) = \{\sigma \in \mathrm{Sym}(X) \,|\, \sigma(y) = y \text{ for all } y \in Y\}$$

and let

$$S(Y) = \{\sigma \in \mathrm{Sym}(X) \,|\, \sigma(Y) = Y\}.$$

Then $C(Y) \leq S(Y) \leq \mathrm{Sym}(X)$.

3.5 Subgroups of the general linear group

Let n be a positive integer. Let

$$U = \{[c_{ij}] \in \mathrm{GL}_n(\mathbf{R}) \mid c_{ij} = 0 \text{ for all } i > j\},$$

the set of upper triangular nonsingular matrices. Let

$$D = \{[c_{ij}] \in \mathrm{GL}_n(\mathbf{R}) \mid c_{ij} = 0 \text{ for all } i \neq j\},$$

the set of diagonal nonsingular matrices. Then $D \leq U \leq \mathrm{GL}_n(\mathbf{R})$.

3.6 Centralizer of a set; Center

Let G be a group and let S be a subset of G. The **centralizer** of S in G is the set $C_G(S) = \{g \in G \mid gx = xg \text{ for all } x \in S\}$. $C_G(S)$ is a subgroup of G. When S is a singleton set $\{x\}$ it is customary to write $C_G(\{x\})$ simply as $C_G(x)$.

The **center** of G, denoted $Z(G)$, is the centralizer of G in G, that is, $Z(G) = C_G(G)$. Thus, $Z(G) = \{g \in G \mid gx = xg \text{ for all } x \in G\}$, the set of all elements of G that commute with every element of G. An element of the center is a **central** element.

3.7 Intersection of subgroups is a subgroup

Let G be a group and let $\{H_\alpha\}_{\alpha \in I}$ be an indexed family of subgroups of G.

THEOREM. *$\bigcap_{\alpha \in I} H_\alpha$ is a subgroup of G.*

Proof. Exercise 3–9. □

3.8 Subgroup generated by a subset

Let G be a group and let S be a subset of G. Denote by $\langle S \rangle$ the intersection of all subgroups of G containing S:

$$\langle S \rangle = \bigcap_{\substack{H \leq G \\ H \supseteq S}} H.$$

By 3.7, $\langle S \rangle$ is a subgroup of G, the **subgroup generated by** S. It is the smallest subgroup of G containing S in the sense that if H is a subgroup of G containing S, then $\langle S \rangle \subseteq H$.

It is convenient to have a description of the elements of $\langle S \rangle$. In the following theorem the empty product (i.e., the case $n = 0$) is interpreted to mean e, the identity of G, and $S^{-1} = \{x^{-1} \mid x \in S\}$.

THEOREM. $\langle S\rangle = \{x_1x_2\cdots x_n \,|\, n \in \mathbf{N}\cup\{0\}, x_i \in S\cup S^{-1}\}$

Sketch of proof. The set on the right hand side of the equality is a subgroup of G containing S, so it contains $\langle S\rangle$. On the other hand, every subgroup of G that contains S must contain this set because of the closure properties, so this set is contained in the intersection of all such subgroups, which is $\langle S\rangle$. □

So the subgroup of G generated by S consists of all products of finitely many elements of G each of which either lies in S or has inverse lying in S (with "products" including the case of no elements, meaning e, and the case of one element x_1, meaning x_1).

3.9 Cyclic subgroup generated by an element

Let G be a group and let $g \in G$. The **cyclic subgroup** of G generated by g is the subgroup of G generated by the singleton set $\{g\}$ (see 3.8). It is customary to write this subgroup as $\langle g\rangle$ instead of $\langle\{g\}\rangle$.

THEOREM.

(i) $\langle g\rangle = \{g^i \,|\, i \in \mathbf{Z}\}$.

(ii) *If $o(g) = \aleph_0$, then the elements g^i for $i \in \mathbf{Z}$ are distinct.*

(iii) *If $o(g) = n \in \mathbf{N}$, then the elements g^i for $0 \leq i < n$ are distinct and $\langle g\rangle = \{g^i \,|\, 0 \leq i < n\} = \{e, g, g^2, \ldots, g^{n-1}\}$.*

(iv) $o(g) = |\langle g\rangle|$.

Proof. (i) This follows immediately from Section 3.8.

(ii) Assume that $o(g) = \aleph_0$. If $g^i = g^j$ for integers $i \leq j$, then $j - i \geq 0$ and $g^{j-i} = g^jg^{-i} = e$, implying that $j - i = 0$, that is, $i = j$.

(iii) Assume that $o(g) = n \in \mathbf{N}$. If $g^i = g^j$ for integers $0 \leq i \leq j < n$, then $0 \leq j - i < n$ and $g^{j-i} = g^jg^{-i} = e$, implying that $j - i = 0$, that is, $i = j$. This proves that the elements g^i for $0 \leq i < n$ are distinct. Let $x \in \langle g\rangle$. According to (i), $x = g^m$ for some integer m. By the division algorithm, there exist integers q and r with $0 \leq r < n$ such that $m = nq+r$. We have $x = g^m = g^{nq+r} = (g^n)^qg^r = eg^r = g^r$. This shows that $\langle g\rangle \subseteq \{g^i \,|\, 0 \leq i < n\}$. The other inclusion follows from (i).

(iv) This follows immediately from (i), (ii), and (iii). □

3 – Exercises

3–1 Let G be a group and let H and K be subgroups of G. Prove that $H \cup K$ is a subgroup of G if and only if either $H \subseteq K$ or $K \subseteq H$.

3–2 (**Dedekind law**) Let G be a group. For subsets U and V of G, define $UV = \{uv \,|\, u \in U, v \in V\}$. Let $V \subseteq G$ and $U \subseteq W \leq G$. Prove that $UV \cap W = U(V \cap W)$.

3–3 Let G be a finite group and let C be a nonempty subset of G that is closed under the binary operation of G. Prove that C is a subgroup of G.

3–4 Let G be a group that has precisely two subgroups. Prove that the order of G is prime.

3–5 Let X be a set and let Y be a proper subset of X. Prove that $C(Y) \cong \mathrm{Sym}(X \backslash Y)$ (see Sections 3.4 and 2.12).

3–6 Let $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$, where $i^2 = -1$. Let $H = \langle \{A, B\} \rangle \leq \mathrm{GL}_2(\mathbf{C})$. Prove that H has order 8.

Hint: $BA = A^3B$.

3–7 Let G be a group that has only finitely many subgroups. Prove that G is finite.

3–8 Let n be a positive integer. Define $\mathrm{SL}_n(\mathbf{R} = \{A \in \mathrm{GL}_n(\mathbf{R}) \,|\, \det(A) = 1\}$, the **special linear group** of degree n over $\mathbf{R}$. Prove that $\mathrm{SL}_n(\mathbf{R})$ is a subgroup of $\mathrm{GL}_n(\mathbf{R})$.

3–9 Prove Theorem 3.7.

4 Generators of a group

4.1 Definition

Let G be a group. If $G = \langle S \rangle$ for some subset S of G, then we say that G is **generated by** S and we call the elements of S **generators** of G. G is **finitely generated** if $G = \langle S \rangle$ for some finite set S.

Of course, G is always generated by the set G itself, but it is often useful to have a small subset of G that generates G (see Examples 4.2, 4.3).

4.2 Example: General linear group

Let $n \in N$. Denote by I the **identity matrix**: $I = [\delta_{ij}]$ (δ_{ij} = Kronecker delta). For $1 \leq i, j \leq n$, let e_{ij} denote the $n \times n$ matrix with 1 in the (i, j) position and zeroes elsewhere and for $\lambda \in \mathbf{R}$ define $b_{ij}(\lambda) = I + \lambda e_{ij}$. For $\delta \in \mathbf{R}$ define $d_\delta = I + (\delta - 1)e_{11}$ (the matrix obtained from I by replacing the first diagonal entry by δ).

The proof of the following theorem requires some preliminaries from linear algebra, which we now discuss.

Let a be an $n \times n$ matrix. The matrix $b_{ij}(\lambda)$ corresponds to an elementary row (or column) operation. Multiplying a on the left by this matrix has the effect of adding λ times the jth row of a to the ith row of a. Similarly, multiplying a on the right by this matrix has the effect of adding λ times the ith column of a to the jth column of a.

For $1 \leq i, j \leq n$ and $\lambda \in \mathbf{R}^\times$, put

$$\begin{aligned} x_{ij}(\lambda) &= b_{ji}(-\lambda^{-1}) b_{ij}(\lambda) b_{ji}(-\lambda^{-1}) \\ &= I - e_{ii} - e_{jj} + \lambda e_{ij} - \lambda^{-1} e_{ji} \end{aligned}$$

if $i \neq j$, and put $x_{ii}(\lambda) = I$ (identity matrix). In the case $i \neq j$, $x_{ij}(\lambda)$ is the matrix obtained from the identity matrix by applying the following sequence of row operations: multiply the ith row by $-\lambda^{-1}$, multiply the jth row by λ, interchange rows i and j. Therefore, multiplying the matrix a on the left by $x_{ij}(\lambda)$ effects these row operations (in the same order) on a. Multiplying a on the right by $x_{ji}(\lambda)$ effects the same corresponding column operations.

Theorem. *The general linear group* $\mathrm{GL}_n(\mathbf{R})$ *is generated by the set*

$$\{b_{ij}(\lambda), d_\delta \mid 1 \leq i, j \leq n,\ i \neq j,\ \lambda \in \mathbf{R}, \delta \in \mathbf{R}^\times\}.$$

Proof. Let $a = [a_{ij}] \in \mathrm{GL}_n(\mathbf{R})$. Set

$$B = B_n = \langle b_{ij}(\lambda) \,|\, 1 \leq i, j \leq n, i \neq j, \lambda \in \mathbf{R} \rangle \quad \text{and} \quad D = \{d_\delta \,|\, \delta \in \mathbf{R}^\times\}.$$

We claim that there exist $b, b' \in B$ such that $bab' \in D$. The proof is by induction on n. If $n = 1$, then $a \in D$ so we can let b and b' both be the identity matrix.

Assume $n > 1$. Since a has nonzero determinant, there exist $1 \leq i, j \leq n$ with either $i \neq n$ or $j \neq n$ such that $a_{ij} \neq 0$. The sets B and D are closed under transposition of matrices so, by replacing a by its transpose if necessary, we may (and do) assume that $\lambda := a_{ij} \neq 0$ for some $1 \leq i, j \leq n$ with $j \neq n$. The matrix

$$a' := x_{ni}(1) a x_{jn}(\lambda^{-1}) \in BaB$$

has (n, n) entry equal to 1. Multiplying a' on the left and on the right by suitable elements of B (effecting row and column operations as described above) produces a matrix a'' of the form

$$a'' = \begin{bmatrix} * & \cdots & * & 0 \\ \vdots & \ddots & \vdots & \vdots \\ * & \cdots & * & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \in BaB.$$

The $(n-1) \times (n-1)$ matrix represented by the $*$'s is in $\mathrm{GL}_{n-1}(\mathbf{R})$, which we view as a subgroup of $\mathrm{GL}_n(\mathbf{R})$ by appending to matrices the indicated bottom row and right column. By the induction hypothesis, there exist matrices b_1 and b_1' in $B_{n-1} \subseteq B$ such that $b_1 a'' b_1' \in D$ (noting that these multiplications by b_1 and b_1' alter neither the bottom row nor the right column of a'') . Since $a'' \in BaB$, we conclude that there exist matrices b and b' in B such that $bab' = d$ for some $d \in D$. Hence, $a = b^{-1} d b'^{-1}$, which is in the subgroup $\langle S \rangle$, where S is the set in the statement of the theorem. Thus, $\mathrm{GL}_n(\mathbf{R}) \subseteq \langle S \rangle$ and the theorem follows. □

Note that the set of generators is closed under inversion since $b_{ij}(\lambda)^{-1} = b_{ij}(-\lambda)$ and $d_\delta^{-1} = d_{\delta^{-1}}$. Therefore every matrix in $\mathrm{GL}_n(\mathbf{R})$ can be written as a product of the indicated generators (since, in the application of 3.8, $S \cup S^{-1} = S$).

4.3 Example: Dihedral group

Let n be an integer with $n \geq 3$. Let $P_n \subset \mathbf{R}^2$ be the regular n-gon with vertices $v_j = (\cos(j-1)\alpha, \sin(j-1)\alpha)$, $j = 1, 2 \ldots, n$, where $\alpha = 2\pi/n$. Then

P_3 is an equilateral triangle, P_4 is a square, and so forth. Denote by D_{2n} the set of all symmetries of P_n, that is, the set of all distance-preserving maps $\sigma : \mathbf{R}^2 \to \mathbf{R}^2$ satisfying $\sigma(P_n) = P_n$. $(D_{2n}, \circ)$ is a group, the **dihedral group** of order $2n$. (That D_{2n} indeed has order $2n$ is shown below. We point out that some authors write D_n instead of D_{2n} and call it the dihedral group of degree n.)

Let $\sigma \in D_{2n}$. Since σ preserves distances, it maps line segments to congruent line segments, and hence maps the set of vertices of P_n onto itself sending adjacent vertices to adjacent vertices. We identify each vertex v_j with its index j and thereby view σ as an element of S_n (so $\sigma(i) = \sigma(j)$ when $\sigma(v_i) = \sigma(v_j)$). In the $2 \times n$ matrix representation of σ the entries in the bottom row either increase by one from left to right or they decrease by one from left to right (interpreting $n+1$ as 1 and 0 as n).

Put

$$\rho = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{pmatrix} \in D_{2n}$$

(counterclockwise rotation through α radians) and

$$\tau = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ 1 & n & n-1 & \cdots & 2 \end{pmatrix} \in D_{2n}$$

(reflection through x-axis).

Theorem.

(i) *The dihedral group D_{2n} is generated by the set $\{\rho, \tau\}$.*

(ii) *The elements $\rho^j, \rho^j\tau$ for $0 \leq j < n$ are distinct and*

$$D_{2n} = \{\rho^j, \rho^j\tau \,|\, 0 \leq j < n\}.$$

(iii) $|D_{2n}| = 2n$.

Sketch of proof. From the geometrical descriptions of ρ and τ, we see that these are both elements of D_{2n}. The elements ρ^j, $0 \leq j < n$, are precisely the permutations having the property that, in their $2 \times n$ matrix representations, the bottom rows increase by one from left to right. Similarly, the elements $\rho^j\tau$, $0 \leq j < n$, are the permutations having bottom rows that decrease by one from left to right. Moreover, these $2n$ elements are distinct. The theorem follows. □

4.4 Cyclic group

Let G be a group. G is **cyclic** if $G = \langle g \rangle$ for some $g \in G$. From Theorem 3.9(i) it is clear that a cyclic group is abelian.

$\mathbf{Z}$ is cyclic because $\mathbf{Z} = \langle 1 \rangle$. Similarly, $\mathbf{Z}_n$ is cyclic for each $n \in \mathbf{N}$. In fact, these are the only cyclic groups in the sense that any cyclic group is isomorphic to one of these as will be shown in Section 9.2.

Theorem. *Every subgroup of a cyclic group is cyclic.*

Proof. Let G be a cyclic group and let H be a subgroup of G. Since G is cyclic, there exists $g \in G$ such that $G = \langle g \rangle$.

If $H = \{e\}$, then $H = \langle e \rangle$ and H is cyclic. Now suppose $H \neq \{e\}$. Since every element of G, and hence H, is a power of g, it follows that $g^n \in H$ for some nonzero integer n. If $n < 0$, then $g^{-n} = (g^n)^{-1} \in H$, so we may (and do) assume that n is positive and that it is the least positive integer for which $g^n \in H$.

We claim that $H = \langle g^n \rangle$. Let $h \in H$. We have $h = g^m$ for some integer m. By the division algorithm, there exist integers q and r with $0 \leq r < n$ such that $m = nq + r$. Therefore,

$$g^r = g^{m-nq} = g^m (g^n)^{-q} \in H,$$

which implies, due to the minimality of n, that $r = 0$. Thus, $h = g^m = g^{nq} = (g^n)^q \in \langle g^n \rangle$. This proves that $H \subseteq \langle g^n \rangle$ and, since the other inclusion is immediate, the theorem follows. □

Since $\mathbf{Z}$ is cyclic, every subgroup of $\mathbf{Z}$ is cyclic and hence of the form $\langle n \rangle = n\mathbf{Z}$ for some $n \in \mathbf{Z}$.

4 – Exercises

4–1 Prove that the multiplicative group of positive rational numbers is generated by the set $\{1/p \,|\, p \text{ is a prime number}\}$.

4–2 Prove that the group $\mathbf{Q}$ is not finitely generated. (Conclude that $\mathbf{Q}$ is not cyclic.)

4–3 Let G be a group, let $x \in G$, let $C \subseteq Z(G)$, and assume $G = \langle \{x\} \cup C \rangle$. Prove that G is abelian.

4–4 Find the center of the dihedral group D_{2n}.

Hint: In order to facilitate computations, subtract one from all of the entries in both ρ and τ to obtain functions from $\mathbf{Z}_n$ to $\mathbf{Z}_n$ and find formulas to express these functions.

5 Coset

5.1 Definition

Let G be a group, let H be a subgroup of G, and let $a \in G$. The **left coset** of H determined by a is

$$aH = \{ah \,|\, h \in H\},$$

and the **right coset** of H determined by a is

$$Ha = \{ha \,|\, h \in H\}.$$

Note that, since $e \in H$, these sets each contain a.

If the notation for the binary operation of G is other than $\cdot$, then the notation for a coset is adjusted accordingly. For instance, if the binary operation is $+$, then aH is written $a + H$ $(= \{a + h \,|\, h \in H\})$.

5.2 Example

Let $G = \mathbf{Z}$, $H = \langle 3 \rangle = 3\mathbf{Z}$. Then

$$\begin{aligned} 0 + H &= \{\dots, -6, -3, 0, 3, 6, \dots\} = H, \\ 1 + H &= \{\dots, -5, -2, 1, 4, 7, \dots\}, \\ 2 + H &= \{\dots, -4, -1, 2, 5, 8, \dots\}, \end{aligned}$$

$3 + H = 0 + H$, $4 + H = 1 + H$, $5 + H = 2 + H$,....

5.3 Example

Let $G = D_8$ (the dihedral group of order 8) and let the notation be as in Example 4.3. If $H = \langle \tau \rangle = \{1, \tau\}$, then

$$\rho H = \{\rho, \rho\tau\} \neq \{\rho, \rho^3\tau\} = \{\rho, \tau\rho\} = H\rho,$$

which shows that left and right cosets determined by an element need not coincide.

5.4 Equality of cosets

Let G be a group, let H be a subgroup of G, and let $a, b \in G$. As Example 5.2 shows, it is possible to have $aH = bH$ with $a \neq b$. Here is a useful criterion for equality of cosets:

Theorem.

(i) *$aH = bH$ if and only if $a^{-1}b \in H$,*

(ii) *$Ha = Hb$ if and only if $ab^{-1} \in H$.*

Proof. (i) Assume that $aH = bH$. Then $b = be \in bH = aH$ so $b = ah$ for some $h \in H$. Thus, $a^{-1}b = h \in H$.

Now assume that $a^{-1}b \in H$. Then, for any $h \in H$ we have $bh = a(a^{-1}bh) \in aH$, implying $bH \subseteq aH$. Since $b^{-1}a = (a^{-1}b)^{-1} \in H$, this argument shows that $aH \subseteq bH$ as well, giving equality.

The proof of (ii) is similar. □

For example, it was observed in 5.2 that if $H = 3\mathbf{Z} < \mathbf{Z}$, then $4 + H = 1 + H$ and we check that $-4 + 1 = -3$ is indeed in H in agreement with (i) (noting that $a^{-1}b$ in additive notation is $-a + b$).

5.5 Congruence modulo a subgroup

Let G be a group and let H be a subgroup of G. Define a relation $\equiv_l$ on G by putting $a \equiv_l b$ if $a^{-1}b \in H$. Then $\equiv_l$ is an equivalence relation on G (Exercise 5–4), called **left congruence modulo** H. Denote by $[a]_l$ the equivalence class of $a \in G$ relative to $\equiv_l$. Thus, $[a]_l = \{b \in G \mid a \equiv_l b\}$.

Similarly, **right congruence modulo** H is the equivalence relation $\equiv_r$ on G obtained by putting $a \equiv_r b$ if $ab^{-1} \in H$. The equivalence class of $a \in G$ relative to $\equiv_r$ is denoted $[a]_r$.

$a \equiv_l b$ is written $a \equiv_l b \mod H$ if the subgroup H is not clear from the context, and similarly for $\equiv_r$.

Theorem. *For $a \in G$*

(i) *$[a]_l = aH$,*

(ii) *$[a]_r = Ha$.*

Proof. (i) Let $x \in G$. Then

$$\begin{aligned} x \in [a]_l &\Leftrightarrow a \equiv_l x \\ &\Leftrightarrow a^{-1}x \in H \\ &\Leftrightarrow x \in aH. \end{aligned}$$

Thus, $[a]_l = aH$.

The proof of (ii) is similar. □

5.6 Cosets partition the group

Let G be a group and let H be a subgroup of G.

THEOREM. *The set of left cosets of H forms a partition of G. More precisely,*

(i) $G = \bigcup_{a \in G} aH$,

(ii) *if $a, b \in G$ and $aH \cap bH \neq \emptyset$, then $aH = bH$.*

Proof. It is shown in set theory that any equivalence relation on a set gives rise to a partition of the set with the equivalence classes being the cells. Therefore, the theorem follows from Section 5.5. For the instructional value it provides, we give an elementary proof as well:

(i) If $x \in G$, then $x = xe \in xH \subseteq \bigcup_{a \in G} aH$, so $G \subseteq \bigcup_{a \in G} aH$. The other inclusion is immediate.

(ii) Let $a, b \in G$ and assume that $aH \cap bH \neq \emptyset$. Then there exists $x \in aH \cap bH$ and we have $x = ah$ and $x = bk$ for some $h, k \in H$. Therefore, $ah = x = bk$, so that $a^{-1}b = hk^{-1} \in H$. By 5.4(i), we have $aH = bH$. □

(The set of right cosets of H also forms a partition of G.) Example 5.2 provides an illustration of this theorem with $G = \mathbf{Z}$ and $H = 3\mathbf{Z}$. The set of left cosets is $\{a + H \,|\, a \in \mathbf{Z}\} = \{0 + H, 1 + H, 2 + H\}$. $\mathbf{Z}$ is the union of $0 + H$, $1 + H$, and $2 + H$, and these cosets are pairwise disjoint.

5.7 Cosets have same cardinality

Let G be a group and let H be a subgroup of G.

THEOREM. *For all $a \in G$,*

(i) $|aH| = |H|$,

(ii) $|Ha| = |H|$.

Proof. (i) Let $a \in G$. The function $f : H \to aH$ given by $f(h) = ah$ is surjective by the definition of the coset aH and it is injective by the cancellation property. Therefore, f is a bijection and we conclude that $|aH| = |H|$.

The proof of (ii) is similar. □

In particular, every left coset has the same cardinality as every right coset.

5.8 Index of subgroup

Let G be a group and let H be a subgroup of G. The set of left cosets of H and the set of right cosets of H need not be the same, but at least they have the same cardinality:

THEOREM. $|\{aH \,|\, a \in G\}| = |\{Ha \,|\, a \in G\}|$.

Proof. Let $L = \{aH, \,|\, a \in G\}$ and $R = \{Ha \,|\, a \in G\}$. Define $f : L \to R$ by $f(aH) = Ha^{-1}$. For $a, b \in G$ we have

$$aH = bH \quad\Leftrightarrow\quad a^{-1}b \in H \quad\Leftrightarrow\quad a^{-1}(b^{-1})^{-1} \in H \quad\Leftrightarrow\quad Ha^{-1} = Hb^{-1},$$

where we have used Section 5.4. This shows that f is well defined and injective. If $Ha \in R$, then $a^{-1}H \in L$ and $f(a^{-1}H) = Ha$, so f is surjective. Therefore, f is a bijection and the claim follows. □

The **index** of H in G, denoted $|G : H|$, is the cardinality of the set of left cosets of H in G, that is, $|G : H| = |\{aH \,|\, a \in G\}|$. According to the theorem, $|G : H|$ is also the cardinality of the set of right cosets of H in G.

For example, in view of 5.2, we have $|\mathbf{Z} : 3\mathbf{Z}| = 3$.

5.9 Lagrange's theorem

Let G be a group and let H be a subgroup of G.

THEOREM. $|G| = |G : H| \cdot |H|$.

Proof. Let A be a complete system of representatives for the left cosets of H in G. Thus $G = \bigcup_{a \in A} aH$, and for $a, a' \in A$, $aH \cap a'H \neq \emptyset$ if and only if $a = a'$ (such an A exists by 5.6). Define $f : A \times H \to G$ by $f((a, h)) = ah$. Since G is the union of $\{aH \,|\, a \in A\}$, f is surjective. Let $(a, h), (a', h') \in A \times H$ and assume that $f((a, h)) = f((a', h'))$. Then $ah = a'h'$, so that $ah \in aH \cap a'H$. This implies that $a = a'$ and, in turn, $h = h'$, so that $(a, h) = (a', h')$. Therefore, f is injective. Since f is bijective, we have $|G| = |A \times H| =: |A| \cdot |H| = |G : H| \cdot |H|$, as claimed. □

For instance, if $G = \mathbf{Z}_{12}$ and $H = \langle 3 \rangle = \{0, 3, 6, 9\}$, then the set of left cosets of H in G is $\{0 + H, 1 + H, 2 + H\}$, so $|G| = 12 = 3 \cdot 4 = |G : H| \cdot |H|$ in agreement with the theorem.

If G is finite, the theorem says that the order of H divides the order of G and

$$|G : H| = \frac{|G|}{|H|}.$$

5.10 Corollaries of Lagrange's theorem

Let G be a finite group.

Corollary. *If the order of G is prime, then G is cyclic.*

Proof. Assume that the order of G is prime. Since a prime is at least two, G has a nonidentity element a. The order of the subgroup $\langle a \rangle$ generated by a is greater than one and divides $|G|$ by Lagrange's theorem (5.9). Therefore, $|\langle a \rangle| = |G|$ and we have $G = \langle a \rangle$. Thus, G is cyclic. □

Corollary. *The order of each element of G divides the order of G.*

Proof. The order of an element equals the cardinality of the cyclic subgroup generated by that element (see 3.9), so the claim follows immediately from Lagrange's theorem (5.9). □

Corollary. *For each $x \in G$ we have $x^{|G|} = e$.*

Proof. Let $x \in G$. By the previous corollary, $|G| = o(x)n$ for some $n \in \mathbf{N}$. Thus $x^{|G|} = (x^{o(x)})^n = e^n = e$. □

5 – Exercises

5–1 Let G be a group and let H and K be subgroups of G. For $x \in G$, put $HxK = \{hxk \mid h \in H, k \in K\}$ (called the (H, K) **double coset** of G determined by x).

(a) Define a relation on G by putting $x \sim y$ if $x = hyk$ for some $h \in H$ and $k \in K$. Prove that $\sim$ is an equivalence relation and that the equivalence class of $x \in G$ relative to this equivalence relation is precisely HxK. (Conclude that the distinct (H, K) double cosets of G are disjoint and form a partition of G.)

(b) Prove that for each $x \in G$, the double coset HxK is a union of right cosets of H and the cardinality of the set of these cosets is $|K : K \cap H^x|$, where $H^x = x^{-1}Hx = \{x^{-1}hx \mid h \in H\}$.

Hint: For (b) find a bijection from the set of right cosets of H in HxK to the set of right cosets of $K \cap H^x$ in K.

5–2 Let G be a group and let H and K be finite subgroups of G having relatively prime orders (i.e., $\gcd(|H|, |K|) = 1$). Prove that $H \cap K = \{e\}$.

5–3 Let G be an abelian group of order $2n$ with n odd. Prove that G has precisely one element of order 2 (cf. Exercise 2–6).

5–4 Let G be a group and let H be a subgroup of G. Prove that $\equiv_l$ (that is, $\equiv_l \mod H$) is an equivalence relation relation on G. (An **equivalence relation** is a relation that is reflexive, symmetric, and transitive.)

5–5 Let $G = \langle g \rangle$ be a cyclic group and let m be an integer. Prove that g^m is a generator of G if and only if $\gcd(m, o(g)) = 1$.

Hint: If $a, b \in \mathbf{Z}$, then $\gcd(a, b) = 1$ if and only if $aj + bk = 1$ for some integers j and k.

6 Normal Subgroup

6.1 Conjugation

Let G be a group. For $x, g \in G$ define $x^g = g^{-1}xg$, the **conjugate** of x by g. The notation x^g is convenient because it obeys some familiar laws of exponents:

Theorem. *For $x, y, g, h \in G$,*

(i) $x^e = x$,

(ii) $(x^g)^h = x^{gh}$,

(iii) $(xy)^g = x^g y^g$.

Let $x, y \in G$. The element x is **conjugate** to the element y if $x^g = y$ for some $g \in G$. The relation $\sim$ on G obtained by putting $x \sim y$ if x is conjugate to y is an equivalence relation on G. The equivalence class of $x \in G$ relative to this equivalence relation is $x^G = \{x^g \,|\, g \in G\}$, called the **conjugacy class** of x.

More generally, for $S \subseteq G$ and $g \in G$ define $S^g = \{x^g \,|\, x \in S\}$, the **conjugate** of S by g. The theorem remains valid if one replaces $x, y \in G$ with $S, T \subseteq G$. If H is a subgroup of G and $g \in G$, then H^g is a subgroup of G.

6.2 Example: Conjugation and change of basis

The notion of conjugation arises in linear algebra when one considers the effect of a basis change on the matrix of a linear transformation.

Let n be a positive integer, let $T : \mathbf{C}^n \to \mathbf{C}^n$ be an invertible linear transformation, and let $B = (v_1, v_2, \dots, v_n)$ be an ordered basis of $\mathbf{C}^n$. The **matrix of T relative to B** is the matrix $A = [a_{ij}]$ defined by $T(v_i) = \sum_j a_{ij} v_j$. Invertibility of T implies that $A \in \mathrm{GL}_n(\mathbf{C})$. Let $C = (w_1, w_2, \dots, w_n)$ be another ordered basis of $\mathbf{C}^n$. Writing $v_i = \sum_j p_{ij} w_j$ $(1 \leq j \leq n)$, we get the **change of basis matrix** $P = [p_{ij}] \in \mathrm{GL}_n(\mathbf{C})$. The matrix of T relative to C is the conjugate $P^{-1}AP$.

It is natural to ask whether there is a basis of $\mathbf{C}^n$ relative to which the matrix of T is particularly simple, say, a diagonal matrix. This is the same as asking whether A is conjugate to such a matrix. The next theorem says that A is conjugate to a matrix that, if not diagonal, is at least almost diagonal.

For $\lambda \in \mathbf{C}$ and a positive integer m, the corresponding **Jordan block** matrix is the $m \times m$ matrix

$$J_{\lambda,m} = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}.$$

Theorem (Jordan Canonical Form). *There exist nonzero $\lambda_1, \lambda_2, \ldots, \lambda_s \in \mathbf{C}$ and positive integers $m_1, m_2, \ldots, m_s$ such that A is conjugate to the block diagonal matrix*

$$\operatorname{diag}(J_{\lambda_1,m_1}, J_{\lambda_2,m_2}, \ldots, J_{\lambda_s,m_s}).$$

The Jordan blocks J_{λ_i,m_i} are uniquely determined up to order.

This theorem follows from a very general theorem about finitely generated modules over a principal ideal domain.

6.3 Definition of normal subgroup

Let G be a group. A subgroup H of G is a **normal subgroup**, written $H \lhd G$, if $H^g = H$ for all $g \in G$. There are various characterizations of normality, each of which proves useful in some situations:

Theorem. *Let H be a subgroup of G. The following are equivalent:*

(i) *$H \lhd G$;*

(ii) *$Hg = gH$ for all $g \in G$;*

(iii) *$H^g \subseteq H$ for all $g \in G$;*

(iv) *$h^g \in H$ for all $h \in H, g \in G$.*

Proof. For $g \in G$,

$$H^g = H \iff g^{-1}Hg = H \iff Hg = gH,$$

so (i) and (ii) are equivalent.

That (iii) and (iv) are equivalent is immediate, as is the fact that (i) implies (iii).

Therefore, a proof that (iii) implies (i) will establish the proof. Assuming (iii), we have for every $g \in G$

$$H^g \subseteq H \quad \text{and} \quad H = (H^{g^{-1}})^g \subseteq H^g,$$

implying $H^g = H$. The proof is complete. □

Another useful characterization is given in Theorem 6.4(ii).

The equivalence of (i) and (iii) suggests that it might be the case that, for any $g \in G$, we have the equality $H^g = H$ if and only if $H^g \subseteq H$. This does not hold in general (see Exercise 6–1), but it does hold if G is finite.

Let $H \triangleleft G$ and let $g \in G$. In view of (i)⇒(ii) we can drop the terms "left" and "right" in referring to cosets of H and define the **coset** of H determined by g to be either gH or Hg.

We record a few simple observations about normal subgroups:

- G and $\{e\}$ are both normal subgroups of G.
- If H is a subgroup of G contained in the center $Z(G)$ of G, then H is normal. In particular, $Z(G)$ is a normal subgroup of G.
- If G is abelian, then every subgroup of G is normal.

6.4 Normalizer of a set

Let G be a group and let S be a subset of G. The **normalizer** of S in G is the set $N_G(S) = \{g \in G \mid S^g = S\}$. $N_G(S)$ is a subgroup of G.

THEOREM. *Let H and K be subgroups of G.*

(i) $H \triangleleft N_G(H)$.

(ii) $H \triangleleft G$ *if and only if* $N_G(H) = G$.

(iii) *If* $H \triangleleft K$, *then* $K \subseteq N_G(H)$.

Proof. Since H is a normal subgroup of itself, it follows that $H \subseteq N_G(H)$. The remaining parts of the theorem are immediate from the definitions. □

It follows from parts (i) and (iii) that $N_G(H)$ is the largest subgroup of G having H as a normal subgroup.

6.5 Product of subgroup and normal subgroup

Let G be a group. For subsets S and T of G, define the **product** $ST = \{st \mid s \in S, t \in T\}$.

Theorem. *Let $H, K \leq G$ and assume that $H \subseteq N_G(K)$.*

(i) $HK = KH$,

(ii) $HK \leq G$,

(iii) $\langle H \cup K \rangle = HK$.

In particular, if $H \triangleleft G$, then (i)–(iii) *hold.*

Proof. (i) For each $h \in H$, we have $hK = hK^h = Kh$. Thus $HK = KH$.

(ii) First, $e = ee \in HK$. Next, using (i) we have

$$(HK)(HK) = H(KH)K = H(HK)K = (HH)(KK) \subseteq HK,$$

so HK is closed under multiplication. Finally, $(HK)^{-1} = K^{-1}H^{-1} \subseteq KH = HK$, so HK is closed under inversion. Thus, $HK \leq G$.

(iii) $H = He \subseteq HK$ and $K = eK \subseteq HK$, so HK contains $H \cup K$. Since HK is a subgroup of G (by (ii)) we have $\langle H \cup K \rangle \subseteq HK$. On the other hand, every subgroup of G containing $H \cup K$ must contain HK by closure. Therefore, the other inclusion follows as well. □

6.6 Index two subgroup is normal

Let G be a group and let H be a subgroup of G.

Theorem. *If $|G : H| = 2$, then $H \triangleleft G$.*

Proof. Exercise 6–2. □

(See Exercise 13–2 for a generalization of this theorem in the case where G is finite.)

6.7 Normality is not transitive

Let G be a group. If $H \leq K \leq G$, then $H \leq G$, so the property of being a subgroup is transitive. However, it is *not* true in general that if $H \triangleleft K \triangleleft G$, then $H \triangleleft G$, as the following example shows.

- Let G be the dihedral group D_8 of order 8. Let

  $$N = \langle \rho^2 \rangle = \{\varepsilon, \rho^2\} \quad \text{and} \quad H = \langle \tau \rangle = \{\varepsilon, \tau\}$$

  with notation as in Section 4.3. One easily checks that ρ^2, and hence N, is in the center of G so that N is normal. Therefore,

  $$K = NH = \{\varepsilon, \tau, \rho^2, \rho^2\tau\}$$

  is a subgroup of G (see 6.5) and it contains H.

  The listed elements of K are distinct (see 4.3), so K has order 4. We have $|G : K| = 2$ and $|K : H| = 2$, so $H \triangleleft K \triangleleft G$ by Section 6.6.

  However, as was shown in Section 5.3, a left coset of H need not equal the corresponding right coset, so H is not a normal subgroup of G (using 6.3).

This example shows that, in general, normality is not transitive. (Cf. Exercise 6–3.)

6 – Exercises

6–1 Give an example to show that it is possible to have a subgroup H of a group G satisfying both of the following:

(a) There exists an element g of G such that $g \notin N_G(H)$ but $H^g \subseteq H$.

(b) $\{x \in G \mid H^x \subseteq H\}$ is not a subgroup.

Hint: Consider $G = \mathrm{GL}_2(\mathbf{R})$, $H = \left\{ \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix} \mid n \in \mathbf{Z} \right\}$, $g = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix}$.

6–2 Let G be a group and let H be a subgroup of G with $|G : H| = 2$. Prove that H is normal.

6–3 Let G be a group. A subgroup H of G is a **characteristic subgroup**, written $H \operatorname{char} G$, if $\varphi(H) \subseteq H$ for every isomorphism $\varphi : G \to G$. Prove that if H and K are subgroups of G with $H \operatorname{char} K \triangleleft G$, then $H \triangleleft G$.

7 Quotient Group

7.1 Definition

Let G be a group and let N be a normal subgroup of G. Denote by G/N the set of all (left) cosets of N in G:

$$G/N = \{aN \mid a \in G\}.$$

In the following theorem, $(aN)(bN)$ denotes the product of the subsets aN and bN of G (see Section 6.5).

Theorem.

(i) *For $aN, bN \in G/N$ we have $(aN)(bN) = abN \in G/N$. In particular, this product defines a binary operation $\cdot$ on G/N.*

(ii) *$(G/N, \cdot\,)$ is a group. Its identity element is eN $(= N)$ and for $aN \in G/N$ we have $(aN)^{-1} = a^{-1}N$.*

Proof. We first observe that $NN = N$. Indeed $NN \subseteq N$ by closure and, if $n \in N$, then $n = ne \in NN$ giving the other inclusion as well.

If $aN, bN \in G/N$, then $(aN)(bN) = abN^bN = abNN = abN$ using normality of N. This gives part (i). Associativity of coset multiplication now follows immediately from the associativity of multiplication in G. Also, that eN is an identity and that $a^{-1}N$ is an inverse of $aN \in G/N$ are immediate, so (ii) follows. □

$(G/N, \cdot\,)$ is the **quotient group** (or **factor group**) of G by N. The notation G/N is read "G modulo N" or "G mod N".

The normality assumption on N is essential here. In fact, if $H \leq G$ is not normal, then the set of left cosets of H in G is *never* closed under set multiplication (see Exercise 7–1).

7.2 Example: Z/3Z

Let $G = \mathbf{Z}$ and $N = \langle 3 \rangle = 3\mathbf{Z}$. Then $N \triangleleft G$ (since G is abelian). The group $G/N = \{0+N, 1+N, 2+N\}$ has operation table

$+$	$0+N$	$1+N$	$2+N$
$0+N$	$0+N$	$1+N$	$2+N$
$1+N$	$1+N$	$2+N$	$0+N$
$2+N$	$2+N$	$0+N$	$1+N$

7.3 Quotient by commutator subgroup

Let G be a group. For $a, b \in G$ define $[a, b] = a^{-1}b^{-1}ab$, the **commutator** of a and b. The **commutator subgroup** (or **derived subgroup**) of G, denoted $G^{(1)}$, is the subgroup of G generated by the set of all commutators. In symbols

$$G^{(1)} = \langle [G, G] \rangle,$$

where $[G, G] = \{[a, b] \mid a, b \in G\}$.

According to Theorem 3.8, every element of $G^{(1)}$ is a product of elements each of which is either a commutator or the inverse of a commutator. Since $[a, b]^{-1} = [b, a]$, the inverse of a commutator is also a commutator. We conclude that every element of $G^{(1)}$ can be expressed as a product of commutators.

Let N be a normal subgroup of G.

Theorem. *G/N is abelian if and only if $N \supseteq G^{(1)}$.*

Proof. For a and b in G,

$$\begin{aligned}(aN)(bN) = (bN)(aN) &\iff abN = baN \\ &\iff (ba)^{-1}(ab) \in N \\ &\iff [a, b] = a^{-1}b^{-1}ab \in N.\end{aligned}$$

The claim follows. □

Since $G^{(1)}$ is normal (Exercise 7–3), the theorem says that it is the smallest normal subgroup of G for which the corresponding quotient is abelian.

7.4 Simple group

Let G be a group. Since G and $\{e\}$ are both normal subgroups of G one can form the corresponding quotients:

- G/G is the one element group {G},
- $G/\{e\}$ is the group $\{a\{e\} \mid a \in G\} = \{\{a\} \mid a \in G\}$ with binary operation given by $\{a\}\{b\} = \{ab\}$, so it is isomorphic to G with $\{a\} \mapsto a$ defining an isomorphism.

G is **simple** if it is nontrivial and G and $\{e\}$ are its only normal subgroups.

For example, $\mathbf{Z}_p$ is simple for each prime number p. Indeed, since the order of a subgroup must divide the order of the group (5.9), $\{0\}$ and $\mathbf{Z}_p$

are the only subgroups of $\mathbf{Z}_p$ (and hence the only normal subgroups). On the other hand, if n is not prime, say $n = n_1 n_2$ with $1 < n_i < n$, then $\mathbf{Z}_n$ is not simple, since $o(n_1) = n_2$ implying $\langle n_1 \rangle$ is a proper, nontrivial (normal) subgroup of $\mathbf{Z}_n$.

The reason for the term "simple" is the fact that a simple group has only the trivial group and itself as quotients, as seen above. It is also due to the fact that finite simple groups are the building blocks of all finite groups (in a sense to be made precise in Section 10.5).

7 – Exercises

7–1 Let G be a group and let H be a subgroup of G. Prove that if the set of left cosets of H in G is closed under set multiplication, then H is normal.

7–2 Let G be a group and let H be a subgroup of the center $Z(G)$ of G so that, in particular, $H \triangleleft G$. Prove that if G/H is cyclic, then G is abelian.

7–3 Let G be a group. Prove that the commutator subgroup $G^{(1)}$ of G is normal.

7–4 Let G be an abelian group and let n be a positive integer. Let N be the subgroup of G generated by the set $\{a^n \,|\, a \in G\}$ and put $\bar{G} = G/N$. Prove that the order of every element of $\bar{G}$ divides n.

7–5 Let G be a finite abelian group and let p be a prime number. Prove that if p divides the order of G, then G has an element of order p.

Hint: Use induction on the order of G as well as Exercise 3–4.

8 Homomorphism

8.1 Definitions

Let G and G' be groups. A **homomorphism** from G to G' is a function $\varphi : G \to G'$ satisfying $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in G$.

- A **monomorphism** is an injective homomorphism.
- An **epimorphism** is a surjective homomorphism.
- An **isomorphism** is a bijective homomorphism.
- An **automorphism** is an isomorphism from a group to itself.

"Isomorphism" was defined earlier in Section 2.12. Recall that two groups G and G' are **isomorphic**, written $G \cong G'$, if there exists an isomorphism from one to the other. In this case, the groups G and G' are identical as far as their group properties are concerned.

8.2 Examples

- The exponential function $\varphi : (\mathbf{R}, +) \to (\mathbf{R}^+, \cdot\,)$ given by $\varphi(x) = e^x$ is a homomorphism. In fact, it is an isomorphism, so $(\mathbf{R}, +) \cong (\mathbf{R}^+, \cdot\,)$.
- The determinant function $\det : \mathrm{GL}_n(\mathbf{R}) \to \mathbf{R}^\times$ given by $A \mapsto \det(A)$ is an epimorphism.
- Let n be a positive integer. The function $\varphi : \mathbf{Z} \to \mathbf{Z}_n$ given by $\varphi(m) = r$, where r is the remainder of m upon division by n, is an epimorphism called **reduction modulo** n.
- Let G be a group and let N be a normal subgroup of G. The function $\pi : G \to G/N$ given by $\pi(x) = xN$ is an epimorphism, the **canonical epimorphism**.
- Let G be a group and let $g \in G$. The function $\iota_g : G \to G$ given by $\iota_g(x) = x^g$ is an automorphism of G, the **inner automorphism** of G determined by g.

8.3 Elementary properties

Let $\varphi : G \to G'$ be a group homomorphism and let e' denote the identity element of G'.

Theorem.

(i) $\varphi(e) = e'$.

(ii) $\varphi(x^{-1}) = \varphi(x)^{-1}$ *for each* $x \in G$.

(iii) *If* $H \leq G$, *then* $\varphi(H) \leq G'$, *where* $\varphi(H) = \{\varphi(h) \,|\, h \in H\}$.

(iv) *If* $H' \leq G'$, *then* $\varphi^{-1}(H') \leq G$, *where* $\varphi^{-1}(H') = \{x \in G \,|\, \varphi(x) \in H'\}$.

Proof. (i) We have

$$\varphi(e)\varphi(e) = \varphi(ee) = \varphi(e),$$

so cancellation gives $\varphi(e) = e'$.

(ii) Let $x \in G$. Using part (i) we have

$$\varphi(x)\varphi(x^{-1}) = \varphi(xx^{-1}) = \varphi(e) = e',$$

so $\varphi(x^{-1}) = \varphi(x)^{-1}$.

(iii) Let $H \leq G$. By (i), $e' = \varphi(e) \in \varphi(H)$. For $h, k \in H$, we have $\varphi(h)\varphi(k) = \varphi(hk) \in \varphi(H)$ and $\varphi(h)^{-1} = \varphi(h^{-1}) \in \varphi(H)$ (using (ii)), so $\varphi(H)$ is closed under both multiplication and inversion. Therefore, $\varphi(H) \leq G'$.

(iv) Let $H' \leq G'$. By (i), $\varphi(e) = e' \in H'$, so $e \in \varphi^{-1}(H')$. Let $x, y \in \varphi^{-1}(H')$, so that $\varphi(x), \varphi(y) \in H'$. We have $\varphi(xy) = \varphi(x)\varphi(y) \in H'$ and $\varphi(x^{-1}) = \varphi(x)^{-1} \in H'$ (using (ii)), so $xy, x^{-1} \in \varphi^{-1}(H')$. This shows that $\varphi^{-1}(H')$ is closed under both multiplication and inversion. Therefore, $\varphi^{-1}(H') \leq G$. □

8.4 Kernel and image

Let $\varphi : G \to G'$ be a homomorphism. Associated with φ are two important subgroups:

- $\ker\varphi = \varphi^{-1}(\{e'\}) = \{x \in G \,|\, \varphi(x) = e'\}$, the **kernel** of φ,
- $\operatorname{im}\varphi = \varphi(G) = \{\varphi(x) \,|\, x \in G\}$, the **image** of φ.

The kernel of φ is a subgroup of G by Theorem 8.3(iv) with $H' = \{e'\}$. The image of φ is a subgroup of G' by Theorem 8.3(iii) with $H = G$.

8.5 Kernel same thing as normal subgroup

Let G be a group. The following theorem says that the notions "kernel of a homomorphism from G" and "normal subgroup of G" amount to the same thing.

THEOREM. *If $\varphi : G \to G'$ is a homomorphism, then $\ker \varphi$ is a normal subgroup of G. Conversely, if N is a normal subgroup of G, then N is the kernel of a homomorphism from G, namely the canonical epimorphism $\pi : G \to G/N$.*

Proof. Let $\varphi : G \to G'$ be a homomorphism. It was observed in 8.4 that $\ker \varphi$ is a subgroup of G, so it suffices to check normality, for which we use the condition given in (iv) of 6.3. Let $k \in \ker \varphi$ and $g \in G$. We have

$$\varphi(k^g) = \varphi(g^{-1}kg) = \varphi(g^{-1})\varphi(k)\varphi(g) = \varphi(g)^{-1}e'\varphi(g) = e',$$

which says that $k^g \in \ker \varphi$. Thus, $\ker \varphi$ is a normal subgroup of G.

Now let N be a normal subgroup of G and let $\pi : G \to G/N$ be the canonical epimorphism (see 8.2). If $g \in G$, then

$$\begin{aligned} g \in \ker \pi &\iff \pi(g) = N \\ &\iff gN = N \\ &\iff g \in N, \end{aligned}$$

so $\ker \pi = N$, as claimed. □

8.6 Homomorphism is injective iff kernel is trivial

Let $\varphi : G \to G'$ be a homomorphism.

THEOREM. *φ is injective if and only if* $\ker \varphi = \{e\}$.

Proof. Assume that φ is injective. Let $k \in \ker \varphi$. Then $\varphi(k) = e'$. But also, $\varphi(e) = e'$ by Section 8.3. So $\varphi(k) = \varphi(e)$ and injectivity of φ gives $k = e$. This shows that $\ker \varphi \subseteq \{e\}$. Since a kernel is a subgroup, the other inclusion is immediate.

Now assume that $\ker \varphi = \{e\}$. Let $x, y \in G$ and assume that $\varphi(x) = \varphi(y)$. Then $\varphi(xy^{-1}) = \varphi(x)\varphi(y)^{-1} = e'$, implying that $xy^{-1} \in \ker \varphi = \{e\}$. Thus, $xy^{-1} = e$, that is, $x = y$. Therefore, φ is injective. □

As a practical matter, we observe that, in order to show that a homomorphism φ is injective, it suffices to show that $\ker \varphi \subseteq \{e\}$, since the other inclusion always holds ($\ker \varphi$ is a subgroup).

8.7 Fundamental Homomorphism Theorem

Let $\varphi : G \to G'$ be a homomorphism.

Fundamental Homomorphism Theorem. *If N is a normal subgroup of G with $N \subseteq \ker\varphi$, then there exists a unique homomorphism $\overline{\varphi} : G/N \to G'$ such that $\overline{\varphi}\pi = \varphi$, where $\pi : G \to G/N$ is the canonical epimorphism. The function $\overline{\varphi}$ is given by $\overline{\varphi}(aN) = \varphi(a)$.*

Proof. Let N be a normal subgroup of G with $N \subseteq \ker\varphi$. As in the statement of the theorem, let $\overline{\varphi} : G/N \to G'$ be the function given by $\overline{\varphi}(aN) = \varphi(a)$.

If $aN = bN$ $(a, b \in G)$, then $a^{-1}b \in N \subseteq \ker\varphi$, so that $\varphi(a)^{-1}\varphi(b) = \varphi(a^{-1}b) = e'$, implying $\varphi(a) = \varphi(b)$. Thus, $\overline{\varphi}$ is well defined.

For $aN, bN \in G/N$, we have

$$\begin{aligned}\overline{\varphi}((aN)(bN)) &= \overline{\varphi}((ab)N) = \varphi(ab)\\ &= \varphi(a)\varphi(b) = \overline{\varphi}(aN)\overline{\varphi}(bN),\end{aligned}$$

so $\overline{\varphi}$ is a homomorphism.

For $a \in G$, we have

$$(\overline{\varphi}\pi)(a) = \overline{\varphi}(\pi(a)) = \overline{\varphi}(aN) = \varphi(a),$$

giving $\overline{\varphi}\pi = \varphi$.

Finally, let $\psi : G/N \to G'$ be a homomorphism such that $\psi\pi = \varphi$. Then for any $aN \in G/N$ we have

$$\psi(aN) = \psi(\pi(a)) = (\psi\pi)(a) = \varphi(a) = \overline{\varphi}(aN),$$

so that $\psi = \overline{\varphi}$, thus establishing uniqueness. □

The condition $N \subseteq \ker\varphi$ is a necessary condition for the function $\overline{\varphi}$ given in the theorem to be well defined, as the reader can check.

8 – Exercises

8–1 Let $G = \left\{ \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \mid a \in \mathbf{R} \right\}$ viewed as a group under matrix multiplication. Prove that $G \cong \mathbf{R}$.

8–2 Let G and G' be groups, let $\varphi : G \to G'$ be an epimorphism, let N be a normal subgroup of G, and assume that $N \cap \ker \varphi = \{e\}$. Prove that $\varphi(C_G(x)) = C_{G'}(\varphi(x))$ for all $x \in N$. (See Section 3.6 for notation.)

8–3 Let G and G' be groups and let $\varphi : G \to G'$ be a homomorphism.

(a) Prove that if N' is a normal subgroup of G', then $\varphi^{-1}(N')$ is a normal subgroup of G.

(b) Prove that if N is a normal subgroup of G and φ is surjective, then $\varphi(N)$ is a normal subgroup of G'.

(c) Give an example to show that without the assumption of surjectivity in the previous part, $\varphi(N)$ need not be a normal subgroup of G'.

8–4 Let G and G' be groups with G finite, let $\varphi : G \to G'$ be an epimorphism, let p be a prime number, and let g' be an element of G' of order p^n for some nonnegative integer n. Prove that there exists an element g of G of order p^m for some nonnegative integer m such that $\varphi(g) = g'$.

9 Isomorphism Theorems

9.1 First Isomorphism Theorem

Let $\varphi : G \to G'$ be a homomorphism. By Theorem 8.5, $\ker \varphi$ is a normal subgroup of G so the quotient group $G/\ker \varphi$ is defined.

Theorem (First Isomorphism Theorem).

$$G/\ker \varphi \cong \operatorname{im} \varphi.$$

Proof. Put $N = \ker \varphi$. By Section 8.7, the function $\overline{\varphi} : G/N \to G'$ given by $\overline{\varphi}(aN) = \varphi(a)$ is a well-defined homomorphism. By restricting the codomain to $\operatorname{im}\varphi$ we obtain an epimorphism $G/N \to \operatorname{im}\varphi$, which we continue to denote by $\overline{\varphi}$.

Let $aN \in \ker \overline{\varphi}$. Then $\varphi(a) = \overline{\varphi}(aN) = e'$, so that $a \in \ker \varphi = N$. Thus, $aN = N$. This shows that $\ker \overline{\varphi} \subseteq \{N\}$ so that $\overline{\varphi}$ is injective (see 8.6).

Therefore, $\overline{\varphi} : G/N \to \operatorname{im}\varphi$ is an isomorphism and $G/\ker \varphi = G/N \cong \operatorname{im}\varphi$. □

9.2 Example: Classification of cyclic groups

Let G be a cyclic group.

Theorem. *If G is infinite, then $G \cong \mathbf{Z}$. Otherwise, $G \cong \mathbf{Z}/n\mathbf{Z}$, where $n = |G|$.*

Proof. Since G is cyclic, we have $G = \langle a \rangle$ for some $a \in G$. Define $\varphi : \mathbf{Z} \to G$ by $\varphi(m) = a^m$. For $m, m' \in \mathbf{Z}$, we have

$$\varphi(m + m') = a^{m+m'} = a^m a^{m'} = \varphi(m)\varphi(m'),$$

so φ is a homomorphism. By Section 3.9, $G = \{a^m \mid m \in \mathbf{Z}\}$ so φ is surjective.

Since $\mathbf{Z} = \langle 1 \rangle$ is cyclic, so is $\ker \varphi$ by 4.4, and therefore, $\ker \varphi = \langle n \rangle = n\mathbf{Z}$ for some integer n, which we can (and do) take to be nonnegative. By the First Isomorphism Theorem (9.1), we have $\mathbf{Z}/n\mathbf{Z} \cong G$.

If $n = 0$, then $\mathbf{Z}/n\mathbf{Z} = \mathbf{Z}/\{0\} \cong \mathbf{Z}$. If $n > 0$, then $\mathbf{Z}/n\mathbf{Z} = \{0 + n\mathbf{Z}, 1 + n\mathbf{Z}, \dots, (n-1) + n\mathbf{Z}\}$ and, since these cosets are distinct, $\mathbf{Z}/n\mathbf{Z}$ has order n. The theorem follows. □

9.3 Quotient same thing as homomorphic image

Let G be a group. The following theorem says that the notions "quotient of G" and "homomorphic image of G" amount to the same thing.

THEOREM. *If G/N $(N \triangleleft G)$ is a quotient of G, then G/N is isomorphic to a homomorphic image of G, namely the image under the canonical homomorphism $\pi : G \to G/N$. Conversely, the image $\varphi(G)$ of G under a homomorphism $\varphi : G \to G'$ is isomorphic to a quotient of G, namely $G/\ker\varphi$.*

Proof. There is nothing to prove in the first statement once it has been checked that the canonical map $\pi : G \to G/N$ is indeed an epimorphism, and this was observed in Section 8.2. The second statement is given by the First Isomorphism Theorem (9.1). □

9.4 Second Isomorphism Theorem

Let G be a group and let H and N be subgroups of G with $H \subseteq N_G(N)$ (this latter being the case if N is normal, for instance). By Theorem 6.5, HN is a subgroup of G. It contains N as a normal subgroup so the quotient group HN/N is defined. Also, $H \cap N$ is a subgroup of H and it is normal (which is easily checked, although it is a consequence of the proof below), so the quotient group $H/H \cap N$ is defined.

THEOREM (Second Isomorphism Theorem).

$$H/H \cap N \cong HN/N.$$

Proof. Define $\varphi : H \to HN/N$ by $\varphi(h) = hN$. Then φ is a homomorphism (it is simply the restriction to H of the canonical epimorphism $HN \to HN/N$).

For $h \in H$ we have

$$h \in \ker\varphi \iff \varphi(h) = N \iff hN = N \iff h \in H \cap N,$$

so $\ker\varphi = H \cap N$.

Let $x \in HN/N$. Then $x = hnN = hN$ for some $h \in H$ and $n \in N$, and we have $\varphi(h) = hN = x$, so φ is surjective.

By the First Isomorphism Theorem (9.1),

$$H/(H \cap N) = H/\ker\varphi \cong \operatorname{im}\varphi = HN/N,$$

and the proof is complete. □

9.5 Third Isomorphism Theorem

Let G be a group and let M and N be normal subgroups of G with $M \supseteq N$. Then M/N is a normal subgroup of G/N (as is easily checked), so the quotient group $(G/N)/(M/N)$ is defined.

Theorem (Third Isomorphism Theorem).

$$(G/N)/(M/N) \cong G/M.$$

Proof. Let $\psi : G \to G/M$ be the canonical epimorphism. Since $N \subseteq M = \ker \psi$, the Fundamental Homomorphism Theorem (8.7) says that the function $\varphi : G/N \to G/M$ given by $\varphi(aN) = \psi(a) = aM$ is a well-defined homomorphism. It follows from the indicated formula that φ is surjective.

We claim that $\ker \varphi = M/N$. Let $aN \in G/N$. We first observe that if $aN \in M/N$, then $aN = bN$ for some $b \in M$, implying $a = bn \in M$ for some $n \in N$. Using this observation to supply the direction $\Leftarrow$ of the final step, we have

$$\begin{aligned} aN \in \ker \varphi &\iff \varphi(aN) = M \iff aM = M \\ &\iff a \in M \iff aN \in M/N, \end{aligned}$$

so the claim is established.

By the First Isomorphism Theorem (9.1),

$$(G/N)/(M/N) = (G/N)/\ker \varphi \cong \operatorname{im} \varphi = G/M,$$

and the proof is complete. □

9.6 Correspondence Theorem

Let $\varphi : G \to G'$ be an epimorphism. Let

$$\begin{aligned} \mathbf{S} &= \{H \mid \ker \varphi \subseteq H \leq G\}, \\ \mathbf{S}' &= \{H' \mid H' \leq G'\}. \end{aligned}$$

Theorem (Correspondence Theorem).

(i) *The map $\mathbf{S} \to \mathbf{S}'$ given by $H \mapsto \varphi(H)$ is a bijection. Its inverse map $\mathbf{S}' \to \mathbf{S}$ is given by $H' \mapsto \varphi^{-1}(H')$.*

(ii) *For $H, K \in \mathbf{S}$, $\varphi(H) \subseteq \varphi(K)$ if and only if $H \subseteq K$, and in this case $|\varphi(K) : \varphi(H)| = |K : H|$.*

(iii) *For $H, K \in \mathbf{S}$, $\varphi(H) \triangleleft \varphi(K)$ if and only if $H \triangleleft K$, and in this case $\varphi(K)/\varphi(H) \cong K/H$.*

Proof. (i) By 8.3, if H is a subgroup of G, then $\varphi(H)$ is a subgroup of G' so the map is well defined. By this same section, if H' is a subgroup of G', then $\varphi^{-1}(H')$ is a subgroup of G, and this latter subgroup contains $\ker\varphi$ since $\varphi(k) = e' \in H'$ for all $k \in \ker\varphi$. Therefore, the indicated inverse map is also well defined. It suffices to show that both compositions of these two functions yield the respective identity maps on $\mathbf{S}$ and $\mathbf{S}'$.

Let $H \in \mathbf{S}$. We need to show that $\varphi^{-1}(\varphi(H)) = H$. Let $g \in \varphi^{-1}(\varphi(H))$. Then $\varphi(g) \in \varphi(H)$, implying that $\varphi(g) = \varphi(h)$ for some $h \in H$. Therefore,

$$\varphi(gh^{-1}) = \varphi(g)\varphi(h)^{-1} = e'$$

so that $gh^{-1} \in \ker\varphi \subseteq H$. It follows that $g \in H$. This gives $\varphi^{-1}(\varphi(H)) \subseteq H$. The other inclusion is immediate.

Let $H' \in \mathbf{S}'$. We need to show that $\varphi(\varphi^{-1}(H')) = H'$. Let $h' \in H'$. Since φ is surjective, there exists $g \in G$ such that $\varphi(g) = h'$. But this last equation says that $g \in \varphi^{-1}(H')$, so $h' \in \varphi(\varphi^{-1}(H'))$. This gives $H' \subseteq \varphi(\varphi^{-1}(H'))$. The other inclusion is immediate.

(ii) Let $H, K \in \mathbf{S}$. If $\varphi(H) \subseteq \varphi(K)$, then, using (i), we have

$$H = \varphi^{-1}(\varphi(H)) \subseteq \varphi^{-1}(\varphi(K)) = K,$$

and the other implication is immediate.

Assume that $H \subseteq K$, so that $\varphi(H) \subseteq \varphi(K)$. We claim that the map $f : \{kH \mid k \in K\} \to \{\varphi(k)\varphi(H) \mid k \in K\}$ given by $f(kH) = \varphi(k)\varphi(H)$ is a well-defined bijection. For $k, k' \in K$, we have

$$\begin{aligned} kH = k'H \Rightarrow k^{-1}k' \in H &\Rightarrow \varphi(k)^{-1}\varphi(k') = \varphi(k^{-1}k') \in \varphi(H) \\ &\Rightarrow \varphi(k)\varphi(H) = \varphi(k')\varphi(H), \end{aligned}$$

so f is well-defined. Let $k, k' \in K$ and suppose that $f(kH) = f(k'H)$. Then $\varphi(k)\varphi(H) = \varphi(k')\varphi(H)$, implying that $\varphi(k^{-1}k') = \varphi(k)^{-1}\varphi(k') \in \varphi(H)$. Therefore, $k^{-1}k' \in \varphi^{-1}(\varphi(H)) = H$, and so $kH = k'H$. This shows that f is injective. That f is surjective is immediate, so the claim that f is bijective is established. We conclude that the domain and the codomain of f have the same cardinality, that is, $|K : H| = |\varphi(K) : \varphi(H)|$.

(iii) Let $H, K \in \mathbf{S}$. Denote by φ_K the restriction of φ to K, so that $\varphi_K : K \to \varphi(K)$ is a well-defined epimorphism. Assume that $\varphi(H) \triangleleft \varphi(K)$. Applying part (i) of this theorem to φ_K and then using Exercise 8–3(a) we get

$$H = \varphi_K^{-1}(\varphi_K(H)) = \varphi_K^{-1}(\varphi(H)) \triangleleft K.$$

Conversely, if $H \lhd K$ we get $\varphi(H) = \varphi_K(H) \lhd \varphi(K)$ by Exercise 8–3(b).

Assume that $H \lhd K$, so that $\varphi(H) \lhd \varphi(K)$. Put

$$\psi = \pi\varphi_K : K \overset{\varphi_K}{\to} \varphi(K) \overset{\pi}{\to} \varphi(K)/\varphi(H),$$

where π is the canonical epimorphism, and note that ψ is an epimorphism. We have

$$\ker\psi = \varphi_K^{-1}(\ker\pi) = \varphi_K^{-1}(\varphi(H)) = \varphi_K^{-1}(\varphi_K(H)) = H,$$

the last equality due to part (i) applied to φ_K. By the First Isomorphism Theorem (9.1)

$$K/H = K/\ker\psi \cong \operatorname{im}\psi = \varphi(K)/\varphi(H)$$

and the proof is complete. □

9.7 Quotient is simple iff normal subgroup is maximal

Let G be a group. A **maximal normal subgroup** of G is a normal subgroup N of G satisfying

(i) $N \neq G$,

(ii) If $N \subsetneqq N' \lhd G$, then $N' = G$.

Let N be a normal subgroup of G.

Theorem. *The quotient G/N is simple if and only if N is a maximal normal subgroup of G.*

Proof. Let $\pi : G \to G/N$ be the canonical epimorphism. According to the Correspondence theorem (9.6), the map $N' \mapsto \pi(N') = N'/N$ sets up a one-to-one correspondence between the normal subgroups of G containing N and the normal subgroups of G/N.

The quotient G/N is simple if and only if it is nontrivial and its only normal subgroups are G/N and $\{N\}$ (which equals N/N). So, by the Correspondence theorem, G/N is simple if and only if $N \neq G$ and the only normal subgroups of G containing N are G and N, which holds if and only if N is maximal. □

9.8 Group with operator domain

Let Ω be a set. An **Ω-group** (or **group with operator domain** Ω) is a group G together with a function $G \times \Omega \to G$ denoted $(g, \omega) \mapsto g^\omega$ satisfying

$$(gh)^\omega = g^\omega h^\omega \text{ for each } g, h \in G,\ \omega \in \Omega.$$

Note that any group can be viewed as an Ω-group with $\Omega = \emptyset$, so the notion of "Ω-group" generalizes the notion of "group."

An example of an Ω-group is a vector space V over $\mathbf{R}$ (or, more generally, an R-module where R is a ring). Indeed, V is a group under vector addition, and if we let $\Omega = \mathbf{R}$ and define $v^\alpha = \alpha v$ ($v \in V, \alpha \in \mathbf{R}$), then V is an Ω-group, since

$$(v+w)^\alpha = \alpha(v+w) = \alpha v + \alpha w = v^\alpha + w^\alpha$$

($v, w \in V$, $\alpha \in \mathbf{R}$).

Let G be an Ω-group. An **Ω-subgroup** of G is a subgroup H of G satisfying $h^\omega \in H$ for all $h \in H$, $\omega \in \Omega$.

Let G' be another Ω-group. An **Ω-homomorphism** from G to G' is a group homomorphism $\varphi : G \to G'$ satisfying $\varphi(g^\omega) = (\varphi(g))^\omega$ for each $g \in G$, $\omega \in \Omega$.

Note that in the vector space example an Ω-subgroup is a subspace and an Ω-homomorphism is a linear transformation.

Let N be a normal Ω-subgroup of G (i.e., a normal subgroup that is also an Ω-subgroup). The quotient group G/N is an Ω-group with the definition $(aN)^\omega = a^\omega N$ ($a \in G$, $\omega \in \Omega$).

The general results proved so far involving the notions of group, subgroup, and homomorphism remain valid if one replaces these terms with Ω-group, Ω-subgroup, and Ω-homomorphism, respectively. In particular, the Fundamental homomorphism theorem, the three Isomorphism theorems, and the Correspondence theorem are all valid in the Ω-group setting. In particular, these results apply to vector spaces over $\mathbf{R}$ (and, more generally, to modules over any ring).

9 – Exercises

9–1 In the additive group $\mathbf{R}^3$, let

$$D = \{(t,t,t) \mid t \in \mathbf{R}\} \quad \text{and} \quad P = \{(u, v, -u) \mid u, v \in \mathbf{R}\},$$

both normal subgroups of $\mathbf{R}^3$. Use the second isomorphism theorem to prove that $\mathbf{R}^3/D \cong P$.

9–2 Let $\varphi : G \to G'$ be a group homomorphism with G' abelian. Prove that if H is a subgroup of G with $H \supseteq \ker\varphi$, then H is normal.

Hint: Consider the Correspondence theorem (9.6).

9–3 Let n be a positive integer. Prove that

$$\mathrm{SL}_n(\mathbf{R}) \triangleleft \mathrm{GL}_n(\mathbf{R}) \quad \text{and} \quad \mathrm{GL}_n(\mathbf{R})/\,\mathrm{SL}_n(\mathbf{R}) \cong \mathbf{R}^\times.$$

(See Exercise 3–8.)

9–4 Let G be a group and let H and K be normal subgroups of G such that $G = HK$. Prove that

$$G/(H \cap K) \cong (G/H) \times (G/K).$$

9–5 Let G be a group and let H be a subgroup of G. Prove that

$$N_G(H)/C_G(H)$$

is isomorphic to a subgroup of the group $\mathrm{Aut}(H) \leq \mathrm{Sym}(H)$ of all automorphisms of H. (See Sections 3.6 and 6.4 for notation.)

10 Composition series

10.1 Definition

Let G be a group. A **subnormal series** of G is a tuple

$$(G_i) = (G_0, G_1, \dots G_r),$$

where

- $G_0 = G$,
- $G_r = \{e\}$,
- $G_{i+1} \lhd G_i$ for all $0 \leq i < r$.

Thus, a subnormal series of G is a finite descending sequence of subgroups of G starting at G and ending at $\{e\}$ with each term normal in the preceding term.

Let $(G_i) = (G_0, G_1, \dots, G_r)$ be a subnormal series of G. The **factors** of (G_i) are the quotient groups G_i/G_{i+1}, $0 \leq i < r$ (or any list of r groups isomorphic to these quotient groups). The **length** of (G_i) is the number of its nontrivial factors.

If every factor of the subnormal series (G_i) of G is simple (7.4), it is a **composition series**.

In view of 9.7, (G_i) is a composition series of G if and only if G_{i+1} is a maximal normal subgroup of G_i for each i. A group need not have a composition series ($\mathbf{Z}$ has none for example). However, if G is finite and nontrivial, then it has a composition series since, for instance, one can be constructed by starting with $G_0 = G$ and recursively choosing G_{i+1} to be a maximal normal subgroup of G_i (this process necessarily ending with $G_r = \{e\}$ for some r since the orders of the subgroups strictly decrease).

10.2 Example: $\mathbf{Z}_6$

The group $\mathbf{Z}_6$ has as composition series $(\mathbf{Z}_6, \langle 3\rangle, \{0\})$. The factors are $\mathbf{Z}_6/\langle 3\rangle$ and $\langle 3\rangle/\{0\}$, which are isomorphic to $\mathbf{Z}_3$ and $\mathbf{Z}_2$, respectively.

This group also has as composition series $(\mathbf{Z}_6, \langle 2\rangle, \{0\})$. Its factors are $\mathbf{Z}_6/\langle 2\rangle$ and $\langle 2\rangle/\{0\}$, which are isomorphic to $\mathbf{Z}_2$ and $\mathbf{Z}_3$, respectively.

For one thing, this example shows that a group can have more than one composition series. But is also suggests that any two composition series of a group have the same factors (except possibly for the order in which they occur). This is indeed the case as will be shown in Section 10.5.

10.3 Zassenhaus butterfly lemma

Let G be a group and let A_1, A_2, B_1, B_2 be subgroups of G with $A_2 \triangleleft A_1$ and $B_2 \triangleleft B_1$. By 6.5, each of the products

$$\begin{aligned} A_{11} &= (A_1 \cap B_1)A_2 \\ A_{12} &= (A_1 \cap B_2)A_2 \\ B_{11} &= (A_1 \cap B_1)B_2 \\ B_{21} &= (A_2 \cap B_1)B_2 \end{aligned}$$

is a subgroup of G. Moreover, $A_{12} \triangleleft A_{11}$ and $B_{21} \triangleleft B_{11}$, as is revealed in the proof of the next result.

Lemma (Zassenhaus butterfly lemma).

$$A_{11}/A_{12} \cong B_{11}/B_{21}.$$

Proof. The lemma is proved by establishing the isomorphisms

$$A_{11}/A_{12} \cong (A_1 \cap B_1)/D \cong B_{11}/B_{21},$$

where $D = (A_1 \cap B_2)(A_2 \cap B_1)$.

Since $A_1 \cap B_1 \subseteq N_G(A_{12})$, the Second Isomorphism Theorem (9.4) gives

$$\begin{aligned} A_{11}/A_{12} &= (A_1 \cap B_1)A_{12}/A_{12} \cong (A_1 \cap B_1)/[(A_1 \cap B_1) \cap A_{12}] \\ &= (A_1 \cap B_1)/D, \end{aligned}$$

the last equality from the Dedekind law (see Exercise 3–2). The second isomorphism above is proved similarly. □

10.4 Schreier refinement theorem

Let G be a group and let (G_i) be a subnormal series of G. A subnormal series (R_j) of G is a **refinement** of (G_i) if for each i we have $G_i = R_{j(i)}$ for some $j(i)$ (so (R_j) can be thought of as being obtained from (G_i) by inserting terms).

Two subnormal series of G are **equivalent** if there exists a one-to-one correspondence between their nontrivial factors such that corresponding factors are isomorphic.

Let (A_i) and (B_i) be two subnormal series of G.

Theorem (Schreier refinement theorem). *There exist refinements of* (A_i) *and* (B_i)*, respectively, such that the refinements are equivalent.*

Proof. For each i and j (for which the indicated subgroup is defined), put

$$A_{ij} = (A_i \cap B_j)A_{i+1} \quad \text{and} \quad B_{ij} = (A_i \cap B_j)B_{j+1}.$$

Then $A_{i0} = A_i$ and $B_{0j} = B_j$ for each i and j, so (A_{ij}) (ordered lexicographically according to the double index) is a refinement of (A_i) and (B_{ij}) (ordered reverse lexicographically according to the double index) is a refinement of (B_j). Moreover, by the lemma of Zassenhaus (10.3),

$$A_{ij}/A_{i,j+1} \cong B_{ij}/B_{i+1,j}$$

for each i and j. Therefore, these refinements are equivalent. □

10.5 Jordan-Hölder theorem

Let G be a group.

Let (G_i) be a composition series of G and let (R_j) be a refinement of (G_i). Since G_{i+1} is a maximal normal subgroup of G_i for each i (see 10.1), it must be the case that each term in the series (R_j) equals some term in the series (G_i). In particular, the nontrivial factors of (R_j) are precisely the factors of (G_i), so that (R_j) is equivalent to (G_i).

This observation, together with the Schreier refinement theorem (10.4), gives the following theorem.

Theorem (Jordan-Hölder theorem). *Any two composition series of* G *are equivalent.*

Proof. By the Schreier refinement theorem, any two composition series of G have equivalent refinements, each of which is equivalent to the series it refines by the preceding remarks. □

Assume that G has a composition series (which is the case if G is finite and nontrivial, for instance). The **composition factors** of G are the factors of any composition series of G. This notion is well-defined due to the theorem.

For example, the composition factors of $\mathbf{Z}_6$ are $\mathbf{Z}_2$ and $\mathbf{Z}_3$.

10.6 Example: Fundamental theorem of arithmetic

Let n be a natural number with $n \geq 2$.

Theorem (Fundamental theorem of arithmetic). *The number n is a product of prime numbers and these prime numbers are uniquely determined up to order.*

Here is a proof of this well-known theorem using the Jordan-Hölder theorem. The group $\mathbf{Z}_n$ is finite, so it has a composition series $(G_0, G_1, \ldots, G_r)$. Each G_i is cyclic (4.4), and a quotient of a cyclic group is cyclic (easy proof), so each factor G_i/G_{i+1} of the series is cyclic and simple and hence isomorphic to $\mathbf{Z}_{p_i}$ for some prime number p_i (see 7.4). By repeated use of Lagrange's theorem we have $n = p_0 p_1 \cdots p_{r-1}$. Finally, given a prime factorization of n, a proof by induction using Exercise 7–5 shows that $\mathbf{Z}_n$ has a composition series with composition factors having the prime factors as orders, so the uniqueness statement follows from the uniqueness statement in the Jordan-Hölder theorem.

10.7 Classification of finite simple groups

Let G be a nontrivial finite group. According to the Jordan-Hölder theorem, G determines a unique (up to permutations of terms and isomorphisms) list of simple groups $S_1, S_2, \ldots, S_r$, namely, its composition factors. By Lagrange's theorem, its order is the product of the orders of these simple groups:

$$|G| = |S_1||S_2| \cdots |S_r|.$$

In the special case $G = \mathbf{Z}_n$, these observations yield the Fundamental theorem of arithmetic (10.6) where the orders of the simple groups turn out to be the prime factors of n.

Therefore, this way of associating to a given finite group a list of simple groups can be viewed as a generalization of the associating to a given natural number its list of prime factors. Just as we regard prime numbers as building blocks of all numbers, we can regard simple groups as building blocks of all finite groups. Due to the pivotal role finite simple groups play, much effort has been expended in their determination, and this is now complete:

Theorem (Classification of finite simple groups). *A finite simple group is isomorphic to one of the following:*

(i) *a cyclic group $\mathbf{Z}_p$ with p prime,*

(ii) *an alternating group A_n with $n \geq 5$,*

(iii) *a simple group of Lie type,*

(iv) *a sporadic group.*

The cyclic groups $\mathbf{Z}_p$ with p prime are simple due to Lagrange's theorem. They are the only abelian finite simple groups.

The alternating group A_n is an index two subgroup of the symmetric group S_n defined in 11.6. Its simplicity for $n \geq 5$ is discussed in 14.6.

The simple groups of Lie type are certain groups constructed from matrix groups over finite fields. An example of one is the projective special linear group $\mathrm{PSL}_n(F)$, which is the quotient by its center of the group of determinant one $n \times n$ matrices over the finite field F.

The sporadic groups are 26 simple groups not accounted for among the infinite families in (i)-(iii). The largest is the **Monster**. Its order is

$$2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71 \approx 8 \cdot 10^{53}.$$

10.8 Extension problem

In 10.7 an analogy was drawn between the composition factors of a finite group and the prime factors of a natural number. The analogy is not perfect, for, given a list of prime numbers, there is only one natural number having those primes as factors, but given a list of finite simple groups, it is possible to have two nonisomorphic groups both having those simple groups as composition factors.

Indeed, $\mathbf{Z}_4$ is not isomorphic to $\mathbf{Z}_2 \oplus \mathbf{Z}_2$ since the first is cyclic and the second is not, but these groups have composition series $(\mathbf{Z}_4, \langle 2 \rangle, \{0\})$ and $(\mathbf{Z}_2 \oplus \mathbf{Z}_2, \langle (1,0) \rangle, \{(0,0)\})$, respectively, both of which have factors $\mathbf{Z}_2$ and $\mathbf{Z}_2$. Intuitively speaking, $\mathbf{Z}_4$ and $\mathbf{Z}_2 \oplus \mathbf{Z}_2$ are composed of the same building blocks, but the blocks are not stacked the same way in the one group as in the other. This example leads one to consider the "extension problem."

Let N and Q be groups. A group G is an **extension** of N by Q if it has a normal subgroup isomorphic to N with corresponding quotient isomorphic to Q.

Extension Problem. *For finite groups N and Q with Q simple, determine all* (*isomorphism classes of*) *extensions of N by Q.*

A solution to this problem would allow for a recursive determination of all finite groups: The groups with one composition factor are known–they are the simple groups. Assuming the groups with n composition factors are known, they can play the role of N in the extension problem and a solution would yield all groups with $n + 1$ composition factors.

The extension problem has not been solved, nor does it appear that a solution is on the horizon. However, many results involving special cases have been worked out. We mention one such result.

Assume that N is abelian. It has been shown that there is an abelian group $H^2(Q, N)$ (called a "cohomology group") having elements in natural one-to-one correspondence with (equivalence classes of) extensions of N by Q with N contained in the center. There is always one such extension, namely $G = N \times Q$. It corresponds to the identity of $H^2(Q, N)$. Taking the case $N = \mathbf{Z}_2$ and $Q = \mathbf{Z}_2$, it turns out that $H^2(Q, N) \cong \mathbf{Z}_2 = \{0, 1\}$. The group $\mathbf{Z}_2 \oplus \mathbf{Z}_2$ corresponds to 0, while the group $\mathbf{Z}_4$ corresponds to 1.

10 – Exercises

10–1 Let G be a group and assume that G has a composition series. Prove that if H is a normal subgroup of G, then G has a composition series (G_i) with $G_i = H$ for some i.

10–2 Let G be a nontrivial abelian group. Prove that G has a composition series if and only if the order of G is finite.

10–3 Let G be a finite group and let N be a normal subgroup of G. Prove that if N has composition factors $S_1, \dots, S_m$ and G/N has composition factors $T_1, \dots, T_n$, then G has composition factors

$$S_1, \dots, S_m, T_1, \dots, T_n.$$

10–4 Determine the composition factors of the group $D_8 \times \mathbf{Z}_{12}$. (Completely support your claims by, for instance, exhibiting a composition series, or invoking theorems and/or exercises.)

11 Symmetric group of degree n

11.1 Definition

Let n be a positive integer. Recall (Section 1.7) that S_n denotes the **symmetric group of degree** n, which is the group of all permutations of the set $\{1, 2, \ldots, n\}$, with binary operation being function composition (written using juxtaposition).

An element σ of S_n has a $2 \times n$ matrix representation

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

Since each such representation is completely determined by its bottom row (and different bottom rows correspond to different permutations), the number of permutations of $\{1, 2, \ldots, n\}$ is simply the number of arrangements of these numbers. Thus $|S_n| = n!$.

11.2 Cycle

Let n be a positive integer and let $i_1, i_2, \ldots, i_r$ be distinct integers with $1 \leq i_j \leq n$. Let σ be the element of S_n that satisfies

$$\sigma(i_j) = \begin{cases} i_{j+1} & j < r, \\ i_1 & j = r, \end{cases}$$

and $\sigma(k) = k$ for all $k \notin \{i_1, i_2, \ldots, i_r\}$. This permutation is written $\sigma = (i_1, i_2, \ldots, i_r)$. It is called an r**-cycle** (or a **cycle** of **length** r) and we write $\operatorname{length}(\sigma) = r$.

For example, $(1, 5, 2, 4) \in S_5$ is a 4-cycle. Using the $2 \times n$ matrix representation of a permutation, we have

$$(1, 5, 2, 4) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 1 & 2 \end{pmatrix}.$$

Note that a cycle is invariant under (unchanged by) any cyclic permutation of its entries. (A **cyclic permutation** of a list is the moving of any number of its elements from the end to the beginning without changing their order.) For instance:

$$(1, 5, 2, 4) = (4, 1, 5, 2) = (2, 4, 1, 5) = (5, 2, 4, 1).$$

It is easy to check that the inverse of a cycle is obtained by writing the entries in reverse order. For example,

$$(1,5,2,4)^{-1} = (4,2,5,1).$$

A **transposition** is a 2-cycle. The transposition (i_1, i_2) transposes (interchanges) the two numbers i_1 and i_2 and fixes every other number.

A 1-cycle (a_1) is the identity since it fixes a_1 as well as every other number.

11.3 Permutation is product of disjoint cycles

Two cycles $(i_1, i_2, \ldots, i_r)$ and $(k_1, k_2, \ldots, k_s)$ are **disjoint** if $\{i_1, i_2, \ldots, i_r\} \cap \{k_1, k_2, \ldots, k_s\} = \emptyset$. Since a cycle moves only those numbers appearing, disjoint cycles commute.

Let

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 7 & 8 & 9 & 5 & 4 & 2 & 1 & 6 \end{pmatrix} \in S_9$$

We can write σ as a product (meaning composition) of (pairwise) disjoint cycles as follows:

- Start with 1. We have $\sigma : 1 \mapsto 3 \mapsto 8 \mapsto 1$ (back to where we started). This completes the cycle $(1,3,8)$.

- Pick the smallest number not yet appearing, namely, 2. We have $\sigma : 2 \mapsto 7 \mapsto 2$. This completes the cycle $(2,7)$, which we compose with the cycle above to get $(1,3,8)(2,7)$.

- Again, pick the smallest number not yet appearing, namely, 4. We have $\sigma : 4 \mapsto 9 \mapsto 6 \mapsto 4$ giving the cycle $(4,9,6)$, so we now have $(1,3,8)(2,7)(4,9,6)$.

- Only 5 remains, and since $\sigma : 5 \mapsto 5$ we get the cycle (5) producing $(1,3,8)(2,7)(4,9,6)(5)$.

The proof in the following theorem formalizes this algorithm and shows that σ equals the resulting composition of cycles:

$$\sigma = (1,3,8)(2,7)(4,9,6)$$

(we usually suppress 1-cycles as we have done here with (5) since a 1-cycle is the identity and composing with the identity has no effect).

Both sides of this equation should have the same effect on each number $1, 2, \ldots, 9$. Let's use 7 as a test. Looking at the definition of σ we see that $\sigma(7) = 2$. On the other hand, the composition applied to 7 is:

$$\begin{aligned}
[(1,3,8)(2,7)(4,9,6)]\,(7) &= [(1,3,8)(2,7)]\,(7) && (4,9,6) \text{ fixes } 7\\
&= [(1,3,8)](2) && (2,7) \text{ sends } 7 \text{ to } 2\\
&= 2 && (1,3,8) \text{ fixes } 2,
\end{aligned}$$

as desired. (Note that the cycles in the composition were constructed left to right, but they are applied right to left. Since the cycles are disjoint, the order of the factors is irrelevant.)

In the statement of the theorem the word "product" has the usual broad meaning (any number of factors, with the case of a product of one factor meaning that element itself).

Theorem. *Any element of S_n can be written as a product of disjoint cycles. Moreover, any two such factorizations are the same except possibly for the order of the factors* (*provided cycles of length one are omitted*).

Proof. Let $\sigma \in S_n$. Put $F_\sigma = \{i \,|\, \sigma(i) = i\}$, the set of fixed points of σ. We use reverse induction on $|F_\sigma|$ to show the existence of a factorization of σ as a product of disjoint cycles. If $|F_\sigma| = n$, then σ is the identity map and can therefore be written $\sigma = (1)$, which is a product of disjoint cycles according to our convention.

Assume that $|F_\sigma| < n$. Then there exists i_1 such that $\sigma(i_1) \neq i_1$. Recursively define $i_{j+1} = \sigma(i_j)$. There is a least positive integer r for which $\sigma(i_r) = i_j \in \{i_1, i_2, \ldots, i_r\}$. If $1 < j \leq r$, then i_{j-1} is defined and $\sigma(i_{j-1}) = i_j = \sigma(i_r)$, which, since $i_{j-1} \neq i_r$, violates injectivity of σ. Therefore, we conclude that $\sigma(i_r) = i_1$.

Let σ_1 denote the cycle $(i_1, i_2, \ldots, i_r)$ (defined since the i_j are distinct by construction) and put $\sigma' = \sigma_1^{-1}\sigma$.

We claim that $F_{\sigma'} \supseteq F_\sigma$. Let $i \in F_\sigma$. Then $i \neq i_j$ for all j, so that $\sigma_1(i) = i$. Therefore, $\sigma'(i) = \sigma_1^{-1}\sigma(i) = \sigma_1^{-1}(i) = i$, showing that $i \in F_{\sigma'}$ and establishing the claim.

For each j, $\sigma'(i_j) = \sigma_1^{-1}\sigma(i_j) = \sigma_1^{-1}(i_{j+1}) = i_j$ (interpreting the sum $j+1$ modulo r), so $i_j \in F_{\sigma'}$. On the other hand, $\sigma(i_1) \neq i_1$ (by the definition of i_1), so $i_1 \in F_{\sigma'} \backslash F_\sigma$. We conclude that $|F_{\sigma'}| > |F_\sigma|$. Therefore, the induction hypothesis applies to σ' and $\sigma' = \sigma_2\sigma_3 \cdots \sigma_t$, a product of disjoint cycles.

Fix $1 \leq j \leq r$ and assume that i_j appears in one of the cycles $\sigma_2, \sigma_3, \ldots, \sigma_t$. Then it appears in precisely one of the cycles and, since disjoint cycles commute, we may (and do) assume that it appears in σ_2. As we have seen, σ' fixes i_j, so $i_j = \sigma'(i_j) = \sigma_2(i_j)$. Thus σ_2 has length one.

Therefore, $\sigma = \sigma_1\sigma' = \sigma_1\sigma_2\cdots\sigma_t$ and, after removing any cycles of length one in this product (a harmless act, since each such is the identity), we are left with a product of disjoint cycles.

Next we turn to the uniqueness statement. Let

$$\sigma = \sigma_1\sigma_2\cdots\sigma_t \quad \text{and} \quad \sigma = \rho_1\rho_2\cdots\rho_s$$

be two factorizations of σ as a product of disjoint cycles. We argue again by reverse induction on $|F_\sigma|$. If $|F_\sigma| = n$, then σ is the identity so that the cycles σ_i and ρ_i each have length one, and the uniqueness statement follows (in light of the agreement to ignore cycles of length one).

Assume that $|F_\sigma| < n$. Then $\sigma(k) \neq k$ for some k. Now k must appear in one of the σ_i and also in one of the ρ_i. Since disjoint cycles commute, we may (and do) assume that k appears in σ_1 and also in ρ_1. For each positive integer m we have

$$\sigma_1^m(k) = \sigma^m(k) = \rho_1^m(k).$$

In particular, since the length of σ_1 is characterized as the least positive integer r for which $\sigma_1^r(k) = k$, the cycle ρ_1 must also have this same length r and

$$\sigma_1 = (k, \sigma_1(k), \sigma_1^2(k), \ldots, \sigma_1^{r-1}(k)) = (k, \rho_1(k), \rho_1^2(k), \ldots, \rho_1^{r-1}(k)) = \rho_1.$$

Put $\sigma' = \sigma_1^{-1}\sigma$ and note that $\sigma' = \rho_1^{-1}\sigma$ as well. Then

$$\sigma' = \sigma_2\sigma_3\cdots\sigma_t \quad \text{and} \quad \sigma' = \rho_2\rho_3\cdots\rho_s.$$

By an argument similar to that given in the first part of the proof, $|F_{\sigma'}| > |F_\sigma|$, so the induction hypothesis applies to σ' and, after a relabeling to reflect any rearrangement of factors or deletions of cycles of length one, we have $t = s$ (with the possibility that this common number is 1, implying no factors at all) and $\sigma_i = \rho_i$ for all i. □

Although it is customary to suppress cycles of length one when expressing an element of S_n as a product of cycles, there are times when it is convenient not to do so. A **complete factorization** of $\sigma \in S_n$ is a factorization $\sigma = \sigma_1\sigma_2\cdots\sigma_t$, where the σ_i are disjoint cycles and each $1 \leq k \leq n$ appears in at least one (and hence precisely one) cycle σ_i.

For example, $\sigma = (1,3,8)(2,7)(4,9,6)(5)$ is a complete factorization of the permutation σ defined at the first of the section.

It follows from the theorem that every element of S_n has a complete factorization.

11.4 Permutation is product of transpositions

Let n be a positive integer. In the statement of the theorem, the case $n = 1$ is included by allowing the possibility of a product with no factors, which is interpreted to be the identity. (This is necessary since $S_1 = \{\varepsilon\}$ has no transpositions at all.)

THEOREM. *Any element of S_n can be written as a product of transpositions.*

Proof. By Section 11.2 it suffices to show that every cycle can be written as a product of transpositions. Let $\sigma = (i_1, i_2, \ldots, i_r)$ be a cycle. We proceed by induction on r. If $r = 1$, then σ is the identity, which, according to our convention, is a product of transpositions with no factors.

Assume that $r > 1$. We claim that $(i_1, i_2)\sigma = \sigma'$, where $\sigma' = (i_2, i_3, \ldots, i_r)$. We have,

$$\begin{aligned}
(i_1, i_2)\sigma(i_1) &= i_1 = \sigma'(i_1),\\
(i_1, i_2)\sigma(i_r) &= i_2 = \sigma'(i_r),\\
(i_1, i_2)\sigma(i_j) &= i_{j+1} = \sigma'(i_j), \quad 1 < j < r,\\
(i_1, i_2)\sigma(k) &= k = \sigma(k), \quad k \neq i_1, i_2, \ldots, i_r,
\end{aligned}$$

so the claim is established. Now the cycle σ' has length $r - 1$, so the induction hypothesis says that it is a product of transpositions. Therefore, $\sigma = (i_1, i_2)\sigma'$ is a product of transpositions as well. □

The proof of the theorem provides an algorithm for writing a given element of S_n as a product of transpositions: write as a product of disjoint cycles and then write each cycle $(i_1, i_2, \ldots, i_r)$ as the product

$$(i_1, i_2)(i_2, i_3)(i_3, i_4) \cdots (i_{r-1}, i_r)$$

of transpositions. For example,

$$\begin{aligned}
\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 1 & 3 & 6 & 9 & 2 & 5 & 8 & 7 \end{pmatrix} &= (1, 4, 6, 2)(5, 9, 7)\\
&= (1, 4)(4, 6)(6, 2)(5, 9)(9, 7).
\end{aligned}$$

It is important to note that an element of S_n can be written as a product of transpositions in more than one way. For instance, if $\sigma = (1, 2, 3)$, then

$$\begin{aligned}
\sigma &= (1, 2)(2, 3),\\
\sigma &= (2, 3)(1, 3),\\
\sigma &= (1, 3)(2, 3)(1, 2)(1, 3),\\
\sigma &= (1, 3)(2, 3)(1, 2)(1, 3)(1, 2)(1, 2),
\end{aligned}$$

as is easily checked. However, in any two such factorizations the **parity** (i.e., even or odd) of the number of factors will always be the same (see Section 11.5).

11.5 Even permutation, odd permutation

Let n be a positive integer. An element of S_n is **even** if it can be written as a product of an even number of transpositions. An element of S_n is **odd** if it can be written as a product of an odd number of transpositions.

By Theorem 11.4 an element of S_n is either even or odd, but possibly *both* for all we know at this point. Choosing to apply these terms to permutations would be a bad idea if a permutation could be both even and odd. However, this is not the case:

Theorem. *An element of S_n is not both even and odd.*

Proof. Let $\sigma \in S_n$. Let $\sigma = \sigma_1\sigma_2\cdots\sigma_t$ be a complete factorization of σ (11.2). Define $N(\sigma) = \sum_i(\text{length}(\sigma_i) - 1)$ (well defined by the uniqueness statement of 11.2). Let $1 \leq a, b \leq n$ with $a \neq b$. Any cycle in which a and b appear can be written $(a, c_1, \dots, c_h, b, d_1, \dots, d_k)$ (after applying a cyclic permutation to the elements, if necessary). For such a cycle, a routine check verifies the equation

$$(a, b)(a, c_1, \dots, c_h, b, d_1, \dots, d_k) = (b, d_1, \dots, d_k)(a, c_1, \dots, c_h).$$

Multiplying both sides of this equation by $(a, b)^{-1} = (a, b)$ gives

$$(a, b)(b, d_1, \dots, d_k)(a, c_1, \dots, c_h) = (a, c_1, \dots, c_h, b, d_1, \dots, d_k).$$

It follows from these equations that

$$N((a,b)\sigma) = \begin{cases} N(\sigma) - 1, & \text{if } a \text{ and } b \text{ appear in the same } \sigma_i, \\ N(\sigma) + 1, & \text{otherwise.} \end{cases}$$

Indeed, assuming that a and b both appear in σ_i for some i, we may (and do) assume that $i = 1$ (since disjoint cycles commute) and $\sigma_1 = (a, c_1, \dots, c_h, b, d_1, \dots, d_k)$, so that, writing $\sigma' = \sigma_2\cdots\sigma_t$,

$$\begin{aligned} N((a,b)\sigma) &= N((a,b)\sigma_1) + N(\sigma') \\ &= k + h + N(\sigma') \\ &= N(\sigma_1) - 1 + N(\sigma') \\ &= N(\sigma) - 1, \end{aligned}$$

and similarly for the other case. In particular, $N((a,b)\sigma)$ and $N(\sigma)$ always have opposite parities (i.e., if one is even, then the other is odd).

Let $\sigma = \tau_1\tau_2\cdots\tau_s$ be a factorization of σ with each τ_i a transposition. Then $\tau_s\tau_{s-1}\cdots\tau_1\sigma = \varepsilon$, so $N(\tau_s\tau_{s-1}\cdots\tau_1\sigma) = N(\varepsilon) = 0$, which is an even number. By repeated application of the observation above, we find that $N(\sigma)$ is even or odd according as s is even or odd. This is to say that the parity of the number s of factors in our factorization $\sigma = \tau_1\tau_2\cdots\tau_s$ is the same as the parity of the number $N(\sigma)$. Since this latter depends only on σ the proof is complete. □

11.6 Alternating group

Let n be an integer greater than 1. Let A_n be the set of all even permutations in S_n:

$$A_n = \{\sigma \in S_n \,|\, \sigma \text{ is even}\}.$$

Theorem.

(i) $A_n \lhd S_n$,

(ii) $|S_n : A_n| = 2$,

(iii) $|A_n| = n!/2$.

Proof. The function $\varphi : S_n \to \mathbf{Z}_2$ given by

$$\varphi(\sigma) = \begin{cases} 0, & \sigma \text{ even} \\ 1, & \sigma \text{ odd} \end{cases}$$

is well defined by 11.5, it is an epimorphism (as is easily checked), and its kernel is A_n. Since kernels are normal subgroups, this proves (i). By the first isomorphism theorem,

$$S_n/A_n = S_n/\ker\varphi \cong \operatorname{im}\varphi = \mathbf{Z}_2.$$

Therefore, $|S_n : A_n| = |S_n/A_n| = |\mathbf{Z}_2| = 2$, giving (ii). By Lagrange's theorem, $|A_n| = |S_n|/|S_n : A_n| = n!/2$, which gives (iii) and completes the proof. □

A_n is the **alternating group** of degree n. We show in 14.6 that this group is simple if $n \geq 5$.

11.7 Conjugacy classes in the symmetric group

Let n be a positive integer and let $\sigma \in S_n$. Let $\sigma = \sigma_1\sigma_2\cdots\sigma_t$ be a complete factorization of σ (11.2). Assume that the σ_i are ordered in nonincreasing order according to length, so that $i < j \Rightarrow \operatorname{length}(\sigma_i) \geq \operatorname{length}(\sigma_j)$. The **cycle structure** of σ is the tuple

$$[\operatorname{length}(\sigma_1), \operatorname{length}(\sigma_2), \ldots, \operatorname{length}(\sigma_t)].$$

For example, the cycle structure of

$$\sigma = (1,4)(2,3,8)(5,9) = (2,3,8)(1,4)(5,9)(6)(7) \in S_9$$

is $[3,2,2,1,1]$. For the sake of brevity, exponential notation is often used to indicate repeated entries, so this cycle structure is also written $[3, 2^2, 1^2]$.

A **partition** of n is a nonincreasing sequence $[a_1, a_2, \ldots, a_t]$ of positive integers with $\sum_i a_i = n$. The cycle structure of an element of S_n is an example of a partition.

Theorem.

(i) *Two elements of S_n are conjugate if and only if they have the same cycle structure.*

(ii) *There is a one-to-one correspondence between the set of conjugacy classes of S_n and the set of partitions of n defined by mapping a conjugacy class to the cycle structure of one (and hence all) of its elements.*

Proof. (i) Let σ and σ' be two elements of S_n. Assume that σ' is conjugate to σ. Then $\sigma' = \sigma^\rho$ for some $\rho \in S_n$. Let $\sigma = \sigma_1\sigma_2\cdots\sigma_t$ be a factorization of σ as a product of disjoint cycles. We have

$$\sigma' = \sigma^\rho = \sigma_1^\rho\sigma_2^\rho\cdots\sigma_t^\rho.$$

By Exercise 11–2, $\sigma_i^\rho = {}^{\rho^{-1}}\sigma_i$ is a cycle having the same length as σ_i. Therefore, σ' has the same cycle structure as σ.

Now assume that σ and σ' have the same cycle structure. Let $\sigma = \sigma_1\sigma_2\cdots\sigma_t$ be a complete factorization σ with factors in nonincreasing order according to length and write $\sigma_i = (a_{i1}, a_{i2}, \ldots, a_{ir_i})$. Then the cycle structure of σ is $[r_1, r_2, \ldots, r_t]$. Since σ' has this cycle structure as well, there is a complete factorization of σ' of the form $\sigma' = \sigma_1'\sigma_2'\cdots\sigma_t'$, with $\sigma_i' = (a_{i1}', a_{i2}', \ldots, a_{ir_i}')$. The formula $\rho(a_{ij}) = a_{ij}'$ defines an element ρ of S_n.

Using Exercise 11–2 again, we see that $\sigma' = {}^{\rho}\sigma = \sigma^{\rho^{-1}}$, so σ' is conjugate to σ.

(ii) By part (i) the indicated map is well defined and injective. If $a = [a_1, a_2, \dots, a_t]$ is a partition of n, then $\sigma = \sigma_1\sigma_2 \cdots \sigma_t$, where σ_i is the a_i-cycle with jth entry $j + \sum_{k=1}^{i-1} a_k$, is an element of S_n with cycle structure a, so the map is surjective as well. □

For example, the group S_4 has 5 conjugacy classes since the partitions of 4 are $[1^4]$, $[2, 1^2]$, $[2, 2]$, $[3, 1]$, and $[4]$.

11 – Exercises

11–1 Let σ be an element of S_n ($n \in \mathbf{N}$) and let $\sigma = \sigma_1\sigma_2 \cdots \sigma_t$ be a decomposition of σ as a product of disjoint cycles. Prove that the order of σ is the least common multiple of the lengths of the cycles σ_i, $1 \leq i \leq t$.

11–2 Let ρ be an element of S_n ($n \in \mathbf{N}$) and let $\sigma = (i_1, i_2, \dots, i_r) \in S_n$ be an r-cycle. Prove that ${}^{\rho}\sigma = (\rho(i_1), \rho(i_2), \dots, \rho(i_r))$, where ${}^{\rho}\sigma := \rho\sigma\rho^{-1}$.

11–3 Let n be an integer greater than one. Prove that

$$S_n = \langle (1, 2), (1, 2, \dots, n) \rangle.$$

Hint: Exercise 11–2.

12 Group action

12.1 Definition

Let G be a group. A (left) G-**set** is a pair $(S, \cdot\,)$, where S a set and $\cdot$ is a function $G \times S \to S$, denoted by $(g, s) \mapsto g \cdot s$ (or just gs) satisfying

(i) $e \cdot s = s$ for all $s \in S$,

(ii) $(gh) \cdot s = g \cdot (h \cdot s)$ for all $g, h \in G$, $s \in S$.

There is a corresponding notion of a right G-set.

Let $(S, \cdot\,)$ be a G-set. We say that G acts on S and call $\cdot$ the **action**. When the action is clear from the context we just say S is a G-set.

12.2 Example: Natural action of symmetric group

Let X be a nonempty set. The symmetric group $\mathrm{Sym}(X)$ acts on X with action given by $\sigma \cdot x = \sigma(x)$ ($\sigma \in \mathrm{Sym}(X)$, $x \in X$). This is the **natural action** of $\mathrm{Sym}(X)$ on X.

12.3 Example: Left translation

Let G be a group. G acts on itself with action given by $g \cdot x = gx$ ($g, x \in G$), where the product on the right is the product of group elements. This action is **left translation**.

Let H be a subgroup of G. G acts on the set $\{aH \mid a \in G\}$ of left cosets of H with action given by $g \cdot aH = (ga)H$ ($g, a \in G$). This action is also called left translation.

12.4 Example: Conjugation

Let G be a group. G acts on itself with action given by $g \cdot x = {}^{g}x$ ($g, x \in G$), where ${}^{g}x = gxg^{-1}$. This action is **conjugation**. Conjugation of x by g as defined in Section 6.1 defines a *right* action of G on itself: $x \cdot g = x^g = g^{-1}xg$ ($g, x \in G$). The two exponential notations for conjugation are related by the formula $x^g = {}^{g^{-1}}x$.

G acts on the set $\{H \mid H \leq G\}$ of all subgroups of G with action given by $g \cdot H = {}^{g}H$ ($g \in G$). This action is also called conjugation.

12.5 Permutation representation

Let G be a group. A **permutation representation** of G is a homomorphism from G to a symmetric group $\mathrm{Sym}(S)$ (S, a set). The following theorem says that G-sets and permutation representations of G are essentially the same things.

Theorem.

(i) *If S is a G-set, then the map $\rho : G \to \mathrm{Sym}(S)$ given by $\rho(g)(s) = g \cdot s$ is a permutation representation of G.*

(ii) *If $\rho : G \to \mathrm{Sym}(S)$ is a permutation representation of G, then S is a G-set with the action given by $g \cdot s = \rho(g)(s)$ $(g \in G, s \in S)$.*

Proof. We begin with a general observation. Let S be a set and let S^S denote the set of all functions from S to S viewed as a binary structure under composition of functions. Let $\rho : G \to S^S$ and $G \times S \to S$, $(g, s) \mapsto g \cdot s$, be maps with $\rho(g)(s) = g \cdot s$ for all $g \in G$ and $s \in S$. For any $g, h \in G$,

$$\begin{aligned} \rho(gh) = \rho(g)\rho(h) &\iff \rho(gh)(s) = [\rho(g)\rho(h)](s) \quad \forall s \in S \\ &\iff \rho(gh)(s) = \rho(g)(\rho(h)(s)) \quad \forall s \in S \\ &\iff (gh) \cdot s = g \cdot (h \cdot s) \quad \forall s \in S. \end{aligned}$$

Assume that S is a G-set under the given map $(g, s) \mapsto g \cdot s$. For all $s \in S$, we have $\rho(e)(s) = e \cdot s = s$, so $\rho(e)$ is the identity map ε on S. Let $g \in G$. Using the above observation, we have $\rho(g)\rho(g^{-1}) = \rho(gg^{-1}) = \rho(e) = \varepsilon$ and similarly $\rho(g^{-1})\rho(g) = \varepsilon$. Therefore, $\rho(g)$ is a bijection and hence an element of $\mathrm{Sym}(S)$. This shows that $\rho : G \to \mathrm{Sym}(S)$ is well defined. By the above observation, ρ is a homomorphism, and hence a permutation representation of G. This proves (i).

Now assume that the given ρ is a permutation representation. Then, for all $s \in S$ we have $e \cdot s = \rho(e)(s) = \varepsilon(s) = s$, so the first property of a G-set is satisfied. By the above observation, the second property is satisfied as well. This proves (ii). □

In (i), ρ is the permutation representation of G **afforded by** the G-set S. In (ii), S is the G-set **affording** the permutation representation ρ.

12.6 Cayley's Theorem

Theorem (Cayley). *Every group is isomorphic to a subgroup of a symmetric group.*

Proof. Let G be a group. View the set G as a G-set with the action being left translation and let $\rho : G \to \mathrm{Sym}(G)$ be the permutation representation afforded by this G-set. If g is in the kernel of ρ, then $e = \varepsilon(e) = \rho(g)(e) = ge = g$. Therefore, the kernel of ρ is trivial, which implies that ρ is injective. We conclude that G is isomorphic to $\operatorname{im}\rho$, which is a subgroup of $\mathrm{Sym}(G)$. □

12.7 Orbit

Let G be a group and let S be a G-set. For s and t in S put

$$s \sim t \iff s = g \cdot t \text{ for some } g \in G.$$

THEOREM.

(i) *$\sim$ is an equivalence relation on S.*

(ii) *For $s \in S$, we have $\bar{s} = G \cdot s = \{g \cdot s \,|\, g \in G\}$, where $\bar{s}$ is the equivalence class of s relative to $\sim$.*

(iii) *$\{G \cdot s \,|\, s \in S\}$ is a partition of S.*

Proof. (i) If $s \in S$, then $s = e \cdot s$, so $s \sim s$ and $\sim$ is reflexive. Let $s, t \in S$ and assume that $s \sim t$. Then $s = g \cdot t$ for some $g \in G$. We have

$$t = e \cdot t = (g^{-1}g) \cdot t = g^{-1} \cdot (g \cdot t) = g^{-1} \cdot s,$$

so that $t \sim s$. Therefore, $\sim$ is symmetric. Let $s, t, u \in S$ and assume that $s \sim t$ and $t \sim u$. Then $s = g \cdot t$ and $t = h \cdot u$ for some $g, h \in G$. We have

$$s = g \cdot (h \cdot u) = (gh) \cdot u,$$

so that $s \sim u$. Therefore, $\sim$ is transitive. This proves that $\sim$ is an equivalence relation.

(ii) Let $s, t \in S$. We have

$$t \in \bar{s} \iff t \sim s \iff t = g \cdot s \text{ for some } g \in G \iff t \in G \cdot s,$$

and the claim follows.

(iii) This is immediate from (ii) and the fact that given any equivalence relation on a set, the collection of equivalence classes forms a partition of the set. □

For $s \in S$, the set $\bar{s} = G \cdot s$ is the **orbit** of s under the action of G.

12.8 Example: Orbits of a permutation

Let n be a positive integer and let $\sigma \in S_n$. Put $G = \langle \sigma \rangle$ and $S = \{1, 2, \ldots, n\}$. Then S is a G-set relative to the natural action. Let

$$\sigma = (i_{11}, i_{12}, \ldots, i_{1r_1})(i_{21}, i_{22}, \ldots, i_{2r_2}) \cdots (i_{t1}, i_{t2}, \ldots, i_{tr_t})$$

be a complete factorization of σ (see 11.2). The orbits in S under the action of G are the sets

$$\{i_{11}, i_{12}, \ldots, i_{1r_1}\}, \{i_{21}, i_{22}, \ldots, i_{2r_2}\}, \ldots, \{i_{t1}, i_{t2}, \ldots, i_{tr_t}\}.$$

For example, if $\sigma = (1, 4, 6, 2)(5, 9, 7) \in S_9$, then $\sigma = (1, 4, 6, 2)(5, 9, 7)(3)(8)$ is a complete factorization of σ and the orbits are

$$\{1, 4, 6, 2\}, \{5, 9, 7\}, \{3\}, \{8\}.$$

12.9 Stabilizer

Let G be a group and let S be a G-set. For $s \in S$ define

$$G_s = \{g \in G \,|\, g \cdot s = s\},$$

the **stabilizer** of s.

Theorem.

(i) *G_s is a subgroup of G for each $s \in S$.*

(ii) *$|\bar{s}| = |G : G_s|$ for each $s \in S$.*

(iii) *If S is finite, then $|S| = \sum_{i=1}^{n} |G : G_{s_i}|$, where $\bar{s}_1, \bar{s}_2, \ldots, \bar{s}_n$ are the distinct orbits in S.*

Proof. (i) Let $s \in S$. Since $e \cdot s = s$, we have $e \in G_s$. Let $g, h \in G_s$. Then $(gh) \cdot s = g \cdot (h \cdot s) = g \cdot s = s$, so $gh \in G_s$. Also $g^{-1} \cdot s = g^{-1} \cdot (g \cdot s) = (g^{-1}g) \cdot s = e \cdot s = s$, so $g^{-1} \in G_s$. Therefore, G_s is a subgroup of G.

(ii) Let $s \in S$ and let C be the set of left cosets of G_s in G. Define $f : C \to G \cdot s$ by $f(gG_s) = g \cdot s$. For $g, h \in G$,

$$gG_s = hG_s \iff g^{-1}h \in G_s \iff g^{-1}h \cdot s = s \iff g \cdot s = h \cdot s,$$

so f is well defined and injective. It is immediate that f is surjective. Therefore, f is bijective, whence $|\bar{s}| = |G \cdot s| = |C| = |G : G_s|$.

(iii) This is immediate from (ii) and the fact that S is the disjoint union of the orbits $\bar{s}_i$. □

12.10 Example: Coset is orbit

Let G be a finite group and let H be a subgroup of G. View G as an H-set relative to left translation. The orbit of $a \in G$ is $\bar{a} = Ha$, the right coset of H determined by a. So the distinct orbits are the distinct right cosets, say, $Ha_1, Ha_2, \ldots, Ha_n$. Note that for each i, the stabilizer H_{a_i} is the trivial subgroup $\{e\}$. By Theorem 12.9(iii), we have

$$|G| = \sum_{i=1}^{n} |H : H_{a_i}| = \sum_{i=1}^{n} |H| = n|H| = |G : H||H|,$$

which is Lagrange's Theorem (5.9) in the special case of finite G.

12.11 Class equation

Let G be a finite group. Recall that the center of G is $Z(G) = \{x \in G \,|\, gx = xg \text{ for all } g \in G\}$. Let $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ be the distinct conjugacy classes of G *not* contained in the center.

Theorem (Class equation).

$$|G| = |Z(G)| + \sum_{i=1}^{n} |G : C_G(x_i)|.$$

Proof. View the set G as a G-set with the action being conjugation. The orbit of $x \in G$ under this action is $G \cdot x = {}^{G}x$, which is precisely the conjugacy class x. For each $x, g \in G$,

$$\begin{aligned} g \in G_x &\iff g \cdot x = x \iff {}^{g}x = x \iff gxg^{-1} = x \\ &\iff xg = gx \iff g \in C_G(x), \end{aligned}$$

so that $G_x = C_G(x)$. If $x \in Z(G)$, then the orbit of x is the singleton set $\{x\}$ and $G_x = C_G(x) = G$. Since the orbits in G are $\{x\}$, $x \in Z(G)$, and $\bar{x}_i$, $1 \leq i \leq n$, Section 12.9 gives

$$|G| = \sum_{x \in Z(G)} |G : G_x| + \sum_{i=1}^{n} |G : G_{x_i}| = |Z(G)| + \sum_{i=1}^{n} |G : C_G(x_i)|,$$

as claimed. □

12 – Exercises

12–1 Let G be a group of odd order and let $g \in G$. Prove that if g is conjugate to g^{-1}, then $g = e$.

HINT: Consider an orbit under the action of conjugation.

12–2 Let X be a set and let n be a positive integer. For $\sigma \in S_n$ and $x = (x_1, x_2, \ldots, x_n) \in X^n$ define $\sigma x = (x_{\sigma^{-1}(1)}, x_{\sigma^{-1}(2)}, \ldots, x_{\sigma^{-1}(n)})$. Prove that X^n is an S_n-set with this action. (S_n is said to act on X^n by **place permutation**.)

12–3 Let G be a group and assume that G has an element that has precisely one conjugate other than itself. Prove that G is not simple.

12–4 Let G be a group and let S be a G-set. Assume that for each pair $s, t \in S$ there exists $g \in G$ such that $g \cdot s = t$ (the action of G on S is said to be **transitive**). Prove that the one-point stabilizers form a single conjugacy class. More precisely, prove that for any $s \in S$, one has $\bar{G}_s = \{G_t \mid t \in S\}$, where $\bar{G}_s$ denotes the orbit of G_s under the (left) conjugacy action of G on the set of its subsets.

13 p-Group

13.1 Definition

Let p be a prime number. A finite group G is a p-**group** if its order is a power of p, that is, if there exists a nonnegative integer n such that $|G| = p^n$.

(There is a more general notion of p-group that does not require the group to be finite (see Section 13.5) and, since some of the statements about p-groups made below fail to hold if the group is infinite, we have considered it prudent to insert the word "finite" in those cases.)

13.2 Fixed points of p-group action

Let p be a prime number, let G be a (finite) p-group, and let S be a finite G-set. Define

$$S_0 = \{s \in S \,|\, g \cdot s = s \text{ for all } g \in G\},$$

the set of fixed points of S under the action of G.

Theorem. $|S_0| \equiv |S| \mod p$.

Proof. By assumption, $|G| = p^n$ for some nonnegative integer n. The set S_0 is precisely the union of the singleton orbits of S under the action of G. Let $\bar{s}_1, \bar{s}_2, \ldots, \bar{s}_n$ be the orbits of S not contained in S_0. For each i, $|G : G_{s_i}| = |\bar{s}_i| > 1$ and, since $|G : G_{s_i}|$ is a divisor of $|G| = p^n$, we conclude that $p \,|\, |G : G_{s_i}|$. By Section 12.9,

$$|S| = |S_0| + \sum_i |G : G_{s_i}| \equiv |S_0| \mod p,$$

and the proof is complete. □

13.3 Center of nontrivial p-group is nontrivial

Let p be a prime number and let G be a (finite) p-group. Recall that $Z(G)$ denotes the center of G.

Theorem. *If H is a nontrivial normal subgroup of G, then $H \cap Z(G)$ is nontrivial. In particular, if G is nontrivial, then $Z(G)$ is nontrivial.*

Proof. Let H be a nontrivial normal subgroup of G. By Lagrange's theorem and the fact that G is a p-group, we have $|H| = p^n$ for some positive integer n. Now G acts on $S = H$ by conjugation and, in the notation of Section 13.2, $S_0 = H \cap Z(G)$. By the theorem of that section,

$$|H \cap Z(G)| = |S_0| \equiv |S| = |H| \equiv 0 \mod p,$$

so $H \cap Z(G)$ has at least p elements and is therefore nontrivial. Putting $H = G$ gives the second statement. □

The second statement also follows directly from the class equation (12.11).

13.4 Cauchy's theorem

Let G be a finite group and let p be a prime number.

Theorem (Cauchy). *If p divides the order of G, then G has an element of order p.*

Proof. (This proof is due to J. H. McKay.) Assume that p divides the order of G. By Exercise 12–2, the symmetric group S_p acts on the set G^p by place permutation and hence so does the subgroup $H = \langle \sigma \rangle$ of S_p, where $\sigma = (1, 2, \ldots, p)$.

We claim that the subset

$$S = \{(a_1, a_2, \ldots, a_p) \in G^p \mid a_1 a_2 \cdots a_p = e\}$$

of G^p is closed under the action of H. If $a = (a_1, a_2, \ldots, a_p) \in S$, then

$$a_p a_1 a_2 \cdots a_{p-1} = a_p(a_1 a_2 \cdots a_{p-1} a_p) a_p^{-1} = a_p e a_p^{-1} = e,$$

so $\sigma a = (a_{\sigma^{-1}(1)}, \ldots, a_{\sigma^{-1}(p)}) = (a_p, a_1, a_2, \ldots, a_{p-1}) \in S$. Since H is generated by σ, the claim follows and S is an H-set.

The tuple $(a_1, a_2, \ldots, a_p)$ is in S if and only if $a_p = (a_1 a_2 \cdots a_{p-1})^{-1}$, so the map $S \to G^{p-1}$ given by $(a_1, a_2, \ldots, a_p) \mapsto (a_1, a_2, \ldots, a_{p-1})$ is a bijection.

Applying Section 13.2 to the H-set S (valid since $|H| = p$), we get

$$|S_0| \equiv |S| = |G|^{p-1} \equiv 0 \mod p.$$

Now $S_0 = \{(x, x, \ldots, x) \in G^p \mid x^p = e\}$. Since this set contains $(e, e, \ldots, e)$, it is nonempty, and it therefore has at least p elements by the above congruence. In particular, S_0 contains some tuple $(x, x, \ldots, x)$ with $a \neq e$, and x is the desired element of order p. □

Recall that Lagrange's theorem says that if H is a subgroup of G, then its order divides the order of G. The (full) converse of Lagrange's theorem would read "If the natural number n divides the order of G, then G has a subgroup of order n." This statement does not hold in general. However, according to Cauchy's theorem, if n is *prime*, then the statement is true (since the element of order n guaranteed by the theorem generates a subgroup of order n). Because of this, Cauchy's theorem can be viewed as a partial converse to Lagrange's theorem.

13.5 Element characterization of p-group

Let G be a finite group and let p be a prime number. To say an element g of G has order a power of p is to say that there exists a nonnegative integer n such that $o(g) = p^n$.

THEOREM. *The group G is a p-group if and only if every element of G has order a power of p.*

Proof. Assume that G is a p-group so that $|G| = p^n$ for some nonnegative integer n. Let $g \in G$. By a corollary of Lagrange's theorem (5.10), $o(g) \,|\, |G| = p^n$, so g has order a power of p.

Assume that every element of G has order a power of p. Let q be an arbitrary prime divisor of $|G|$. By Cauchy's theorem (13.4), G has an element of order q. But, by our assumption, this order must be a power of p. Therefore, $q = p$. It follows that $|G|$ is a power of p. □

In the literature a (not necessarily finite) group is called a p-group if each of its elements has order a power of p. According to the present theorem, this is in agreement with our definition of p-group in the case where the group is finite. It should be emphasized that several properties of finite p-groups, including Theorems 13.3 and 13.6, fail to hold for infinite p-groups.

13.6 Normalizer of p-subgroup

Let G be a finite group, let p be a prime number, and let H be a **p-subgroup** of G (meaning, H is a subgroup of G and H is a p-group).

THEOREM.

(i) $|N_G(H) : H| \equiv |G : H| \mod p$.

(ii) *If p divides $|G : H|$, then $N_G(H) \neq H$.*

(iii) *If G is a p-group and $H \neq G$, then $N_G(H) \neq H$.*

There is an argument in the proof that we will need in the next section as well, so we separate it out in the form of a lemma (and state it in general enough terms to handle both applications). In the statement, S_0 is as defined in Section 13.2. The proof is Exercise 13–4.

LEMMA. *Let G be a group and let H and K be subgroups of G. Let $S = \{gK \,|\, g \in G\}$, the set of all left cosets of K in G. Regard S as an H-set with the action being left translation: $h \cdot gK = hgK$ ($h \in H$, $gK \in S$). Then $S_0 = \{gK \,|\, {}^gK \supseteq H\}$.*

Now we prove the theorem.

Proof. (i) View the set $S = \{gH \,|\, g \in G\}$ of all left cosets of H in G as an H-set by left translation. By the above lemma (with $K = H$), we have

$$S_0 = \{gH \in S \,|\, {}^gH \supseteq H\} = \{gH \in S \,|\, g \in N_G(H)\} = N_G(H)/H,$$

where S_0 is as in Section 13.2. By that same section,

$$|N_G(H) : H| = |N_G(H)/H| = |S_0| \equiv |S| = |G : H| \mod p,$$

as claimed.

(ii) Assume that p divides $|G : H|$. By part (i), p divides $|N_G(H) : H|$. Since $|N_G(H) : H|$ is not zero, it must be at least p. Therefore, $N_G(H) \neq H$.

(iii) Assume that G is a p-group and $H \neq G$. By Lagrange's theorem, $|G : H|$ is a divisor of $|G|$ and is therefore a power of p. Since $H \neq G$, this power of p is not p^0. Hence, p divides $|G : H|$ and part (ii) gives $N_G(H) \neq H$. □

13 – Exercises

13–1 Prove that a group of order p^2, with p prime, is abelian.

HINT: Exercise 7–2.

13–2 Let G be a nontrivial finite group and let p be the smallest prime divisor of the order of G. Prove that if H is a subgroup of G of index p, then H is normal.

HINT: Assume otherwise and let G and H be a counterexample with $|G|$ as small as possible. Since H is not normal, it has a conjugate K with $K \neq H$. Use Exercise 5–1 to prove that $|K : K \cap H| = p$ and $|H : H \cap K| = p$. Conclude that $H \cap K \triangleleft G$ and $|G : H \cap K| = p^2$ and derive a contradiction.

13–3 Use the proof of Cauchy's theorem (13.4) to establish Fermat's (little) theorem: If p is a prime number, then $n^p \equiv n \pmod{p}$ for every integer n.

Hint: First assume that p does not divide n and prove that $n^{p-1} \equiv 1 \pmod{p}$ by letting G be $\mathbf{Z}_n$ (or indeed any group of order n) in the proof of Cauchy's theorem.

13–4 Let G be a group and let H and K be subgroups of G. Let $S = \{gK \mid g \in G\}$, the set of all left cosets of K in G. Regard S as an H-set with the action being left translation: $h \cdot gK = hgK$ $(h \in H,\ gK \in S)$. Prove that $S_0 = \{gK \in S \mid {}^{g}K \supseteq H\}$, where S_0 is as defined in Section 13.2.

14 Sylow theorems

14.1 Definition

Let G be a finite group and let p be a prime number. A **Sylow p-subgroup** of G is a subgroup of G of order p^n, where p^n is the greatest power of p that divides the order of G.

- Let $G = \mathbf{Z}_{12}$. Then $|G| = 2^2 \cdot 3$, so
$$\begin{aligned} \langle 3 \rangle &= \{0, 3, 6, 9\} \text{ is a Sylow 2-subgroup,} \\ \langle 4 \rangle &= \{0, 4, 8\} \text{ is a Sylow 3-subgroup,} \end{aligned}$$
and $\{0\}$ is a Sylow p-subgroup if $p \neq 2, 3$.

By Lagrange's theorem, a Sylow p-subgroup of G is necessarily a maximal p-subgroup of G (meaning, not properly contained in another p-subgroup of G).

Denote by $\mathrm{Syl}_p(G)$ the set of all Sylow p-subgroups of G. Although it is not immediate that $\mathrm{Syl}_p(G)$ is nonempty, this is in fact the case, as is shown in Section 14.2.

14.2 Sylow existence theorem

Let G be a finite group and let p be a prime number. Let p^n be the greatest power of p that divides the order of G. Then $|G| = p^n m$ for some $m \in \mathbf{N}$ with $p \nmid m$.

Theorem (Sylow existence theorem). *G has a Sylow p-subgroup. In fact, for each $0 \leq i \leq n$, there exists a subgroup of G of order p^i, and each subgroup of order p^{i-1} is normal in some subgroup of order p^i.*

Proof. The case $i = n$ of the second statement shows that G has a Sylow p-subgroup, so it suffices to prove the second statement, and this we do by induction on i. If $i = 0$, then $\{e\}$ is a subgroup of order p^i, and there is no subgroup of order p^{i-1}, so the second part is vacuously true.

Let $0 < i \leq n$. By the induction hypothesis, G has a subgroup of order p^{i-1}. Let H be an arbitrary such subgroup. We have $H \lhd N_G(H)$ and
$$|N_G(H)/H| = |N_G(H) : H| \equiv |G : H| = p^{n-i+1} m \equiv 0 \mod p,$$

by Section 13.6. Therefore, by Cauchy's theorem (13.4), the group $N_G(H)/H$ contains an element of order p and hence a subgroup A of order p.

Put $H_1 = \pi^{-1}(A)$, where $\pi : N_G(H) \to N_G(H)/H$ is the canonical epimorphism. Then H_1 is a subgroup of $N_G(H)$ (and hence a subgroup of G) and it contains H. Since H is normal in $N_G(H)$ it is normal in H_1 as well. Moreover, $H_1/H \cong A$ by the Correspondence theorem (9.6), so

$$|H_1| = |H_1 : H||H| = p \cdot p^{i-1} = p^i,$$

and the proof is complete. □

14.3 Sylow conjugacy theorem

Let G be a finite group and let p be a prime number. G acts on the set of its subsets by conjugation:

$$g \cdot S = {}^gS \quad (g \in G, S \subseteq G).$$

Let $P \in \mathrm{Syl}_p(G)$. Denote by $\bar{P}$ the orbit of P with respect to this action, so that $\bar{P} = \{{}^gP \mid g \in G\}$.

Theorem (Sylow conjugacy theorem).

(i) *If H is a p-subgroup of G, then ${}^gP \supseteq H$ for some $g \in G$.*

(ii) $\bar{P} = \mathrm{Syl}_p(G)$.

Proof. (i) Let H be a p-subgroup of G. Let $S = \{gP \mid g \in G\}$, the set of all left cosets of P in G. This set is an H-set with the action being left translation. By Lemma 13.6, (with $K = P$), we have $S_0 = \{gP \in S \mid {}^gP \supseteq H\}$, where S_0 is as in Section 13.2. By that same section, $|S_0| \equiv |S| \not\equiv 0 \mod p$, so S_0 is nonempty. Hence, there exists $g \in G$ such that ${}^gP \supseteq H$.

(ii) Every conjugate of P has the same order as P and is therefore a Sylow p-subgroup of G. This gives the inclusion $\bar{P} \subseteq \mathrm{Syl}_p(G)$.

If $Q \in \mathrm{Syl}_p(G)$, then part (i) with $H = Q$ gives $g \in G$ with ${}^gP \supseteq Q$. Since both of these sets have the same cardinality, we get $Q = {}^gP \in \bar{P}$. Therefore, $\mathrm{Syl}_p(G) \subseteq \bar{P}$. With the earlier inclusion we get the desired equality. □

Part (ii) of the theorem says that a conjugate of a Sylow p-subgroup is a Sylow p-subgroup and that any Sylow p-subgroup is conjugate to every other Sylow p-subgroup.

Since a conjugate of a Sylow p-subgroup is a Sylow p-subgroup, part (i) implies that every p-subgroup of G is contained in a Sylow p-subgroup.

If there is only one Sylow p-subgroup P, then $\bar{P} = \mathrm{Syl}_p(G) = \{P\}$ implying that P is normal.

14.4 Sylow number theorem

Let G be a finite group and let p be a prime number. Let p^n be the greatest power of p that divides the order of G so that $|G| = p^n m$ with $p \nmid m$. The factor p^n is the p-**part** of $|G|$ and the factor m is the p'-**part** of $|G|$.

The following theorem says that the number of Sylow p-subgroups of G is a divisor of the p'-part of $|G|$ and that it is congruent to one modulo p.

Theorem (Sylow number theorem).

(i) $|\operatorname{Syl}_p(G)|$ *divides* $|G|/p^n$.

(ii) $|\operatorname{Syl}_p(G)| \equiv 1 \mod p$.

Proof. (i) According to the Sylow conjugacy theorem (14.3) the set $\operatorname{Syl}_p(G)$ of Sylow p-subgroups is an orbit, say $\operatorname{Syl}_p(G) = \bar{P}$, relative to the action of G on its subsets by conjugation. The stabilizer G_P of P contains P, so

$$|\operatorname{Syl}_p(G)| = |\bar{P}| = |G : G_P| = \frac{|G : P|}{|G_P : P|} \mid |G : P| = |G|/p^n,$$

as claimed.

(ii) Retaining the notation of the last paragraph, view $S = \bar{P}$ as a P-set by restricting the action.

We claim that $S_0 = \{P\}$, where S_0 is as in Section 13.2. Since $a \cdot P = {}^aP = P$ for all $a \in P$, we have $\{P\} \subseteq S_0$. Let $Q \in S_0$. We have ${}^aQ = a \cdot Q = Q$ for all $a \in P$, so $P \subseteq N_G(Q)$. As a consequence of Lagrange's theorem, p^n is the greatest power of p that divides the order of $N_G(Q)$, and, since P and Q both have order p^n, they are both Sylow p-subgroups of $N_G(Q)$. By the Sylow conjugacy theorem (14.3) $Q = {}^gQ = P \in \{P\}$ for some $g \in N_G(Q)$. Therefore, $S_0 \subseteq \{P\}$ and the claim is established.

By Section 13.2,

$$|\operatorname{Syl}_p(G)| = |\bar{P}| = |S| \equiv |S_0| = |\{P\}| = 1 \mod p,$$

as claimed. □

14.5 Example: Group of order 28 not simple

Let G be a group of order 28. By Theorem 14.4(i), the number $|\operatorname{Syl}_7(G)|$ of Sylow 7-subgroups of G divides $|G|/7 = 4$ and is therefore 1, 2, or 4. But by (ii) of the same theorem, this number is also congruent to 1 modulo 7 eliminating the possibilities 2 and 4. Therefore, G has exactly one Sylow 7-subgroup. This subgroup is normal by Section 14.3, and it is nontrivial and proper since its order is 7. Therefore, G is not simple.

14.6 Alternating group is simple

The proof we give here of the simplicity of the alternating group relies heavily on the theory of group actions and on Sylow theory; it provides a nice illustration of the power of these theories.

Theorem. *The alternating group A_n is simple for $n \geq 5$.*

Proof. The proof is by induction on $n \geq 5$. Put $G = A_n$ and first assume that $n = 5$. Let N be a nontrivial normal subgroup of G. We need to show that $N = G$.

We first argue that $|N|$ is divisible by either 3 or 5. Suppose otherwise. Then, since $|G| = 5!/2 = 60 = 2^2 \cdot 3 \cdot 5$, it follows from Lagrange's theorem that $|N|$ is either 2 or 4.

Suppose that $|N| = 2$. Then $N = \{\varepsilon, \sigma\}$ with σ an element of order 2. The cycle structure of σ is $[2^2, 1]$, so σ fixes some $s \in \{1, 2, 3, 4, 5\} =: S$, whence $N \subseteq G_s$, where G_s is the stabilizer of s relative to the natural action of G on S. Since N is normal, it is contained in each conjugate of G_s and therefore in each one-point stabilizer G_t ($t \in S$) by Exercise 12–4. (This exercise applies since G contains the element $(1, 2, 3, 4, 5)$ and hence each power of this element so the action of G on S is transitive.) But this implies that σ fixes each element of S, so that $\sigma = \varepsilon$, a contradiction.

Therefore, $|N| = 4$ and N is a Sylow 2-subgroup of G. Since N is normal, it contains every 2-subgroup of G (see 14.3) and hence every element with cycle structure $[2^2, 1]$. There are more than 4 such elements (in fact, 15 such), so this is a contradiction.

This establishes the claim that $|N|$ is divisible by either 3 or 5. Suppose that $|N|$ is divisible by 3. Then N contains a Sylow 3-subgroup of G. By the Sylow conjugacy theorem (14.3), N contains every Sylow 3-subgroup of G and hence every 3-cycle, of which there are 20. Now N also contains the identity, so its order is at least 21. Since the order of N divides the order of G, this forces $|N|$ to be either 30 or 60. In either case, $|N|$ is divisible by 5. Thus, N contains a Sylow 5-subgroup of G and hence every 5-cycle, of which there are 24. We conclude that $|N| = 60$, so that $N = G$ as desired. The case $|N|$ divisible by 5 is handled similarly.

Now assume that $n > 5$ and let N be a nontrivial normal subgroup of G. There exists $\sigma \in N$ with $\sigma \neq \varepsilon$. By relabeling the elements of $S = \{1, 2, \ldots, n\}$ if necessary we may (and do) assume that the complete factorization of σ begins $(1, 2)(3, 4) \cdots$ or $(1, 2, 3, \ldots) \cdots$. Put $\rho = (3, 5, 6) \in G$. Using Exercise 11–2, we see that ${}^{\rho}\sigma$ equals $(1, 2)(5, 4) \cdots$ or $(1, 2, 5, \ldots) \cdots$, respectively. In either case, ${}^{\rho}\sigma \neq \sigma$ and $\sigma^{-1} \cdot {}^{\rho}\sigma \in N$ fixes the element 1.

By the preceding paragraph, $N \cap G_1 \neq \{\varepsilon\}$. Now G_1 is isomorphic to A_{n-1}, which is simple by the induction hypothesis. Since $N \cap G_1$ is a nontrivial normal subgroup of G_1, it follows that N contains G_1 and hence every conjugate of G_1. By Exercise 12–4, N contains every one-point stabilizer G_s $(s \in S)$. (The exercise applies since A_n contains $(1, 2, \ldots, n)$ if n is odd and both $(1, 2, \ldots, n-1)$ and $(2, 3, \ldots, n)$ if n is even.) The product of any two transpositions moves at most four elements and hence lies in G_s for some s. Since the set of all such products, a subset of N, generates G, we conclude that $N = G$. This completes the proof. □

14 – Exercises

14–1 Let G be a group, let H be a finite normal subgroup of G, let p be a prime number, and let P be a Sylow p-subgroup of H. Prove that $G = HN_G(P)$.

Hint: Let $g \in G$. Consider ${}^gP$ in light of the Sylow conjugacy theorem (14.3).

14–2 Let P be a Sylow 2-subgroup of the symmetric group S_6. Prove that $P \cong D_8 \times \mathbf{Z}_2$.

Hint: $\{1, 2, 3, 4, 5, 6\} = \{1, 2, 3, 4\} \cup \{5, 6\}$.

14–3 Let G be a finite group, let p be a prime number, and let P be a Sylow p-subgroup of G. Prove that $N_G(N_G(P)) = N_G(P)$.

14–4 Prove that a group of order 12 is not simple.

14–5 Let G be a simple group of order 168. Find the number of elements of G that have order 7.

15 Category

15.1 Definition

A **category $\mathbf{C}$** consists of the following:

- a class $\text{obj}(\mathbf{C})$ of **objects**;
- a class $\text{mor}(\mathbf{C})$ of **morphisms**;
- two functions $s, t : \text{mor}(\mathbf{C}) \to \text{obj}(\mathbf{C})$, such that for every pair (A, B) of objects, the class $\text{mor}(A, B)$ of all morphisms f with **source** $s(f)$ equal to A and **target** $t(f)$ equal to B is a set (for such an f we write $f : A \to B$ or $A \xrightarrow{f} B$ and say f is a morphism from A to B);
- a function $\text{mor}(A, B) \times \text{mor}(B, C) \to \text{mor}(A, C)$ for each triple (A, B, C) of objects, denoted by $(f, g) \mapsto g \circ f$, the **composition** of f and g;

satisfying two axioms:

(i) if $A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$ are morphisms, then $h \circ (g \circ f) = (h \circ g) \circ f$,

(ii) for each object A there exists a morphism $1_A : A \to A$ such that for any morphisms $f : A \to B$ and $g : B \to A$ we have $f \circ 1_A = f$ and $1_A \circ g = g$.

Part (i) is the **associative property** of morphisms. A morphism 1_A as in (ii) is unique (same proof as that for uniqueness of group identity). It is the **identity morphism** of the object A.

15.2 Examples

- **Set** denotes the category of sets. The object class is the class of all sets. The morphisms are functions (maps) between sets and composition of morphisms is usual function composition. The identity morphism of an object is the identity function on that set which sends each element to itself. (In the remaining examples, composition and identity morphisms are as defined here unless stated otherwise.)
- **Grp** denotes the category of groups. The object class is the class of all groups. The morphisms are homomorphisms between groups.

- **Ab** denotes the category of abelian groups. The object class is the class of all abelian groups. The morphisms are homomorphisms between abelian groups.

- $\mathbf{Vec}_F$ denotes the category of vector spaces over the field F (for example, $F = \mathbf{R}$). The object class is the class of all vector spaces over F. The morphisms are linear transformations between vector spaces.

- **Top** denotes the category of topological spaces. The object class is the class of all topological spaces. The morphisms are continuous maps between spaces.

- Ω-**Grp** denotes the category of Ω-groups with Ω a set. The object class is the class of all Ω-groups. The morphisms are Ω-homomorphisms between Ω-groups.

- G-**Set** denotes the category of G-sets with G a group. The object class is the class of all G-sets. The morphisms are G-maps between G-sets. (A ***G*-map** from a G-set S to a G-set T is a function $f : S \to T$ satisfying $f(g \cdot s) = g \cdot f(s)$ for all $g \in G$, $s \in S$.)

- $\mathbf{PO}(S)$ denotes the category associated with the partially ordered set S (S has an order $\preceq$ that is reflexive, transitive and antisymmetric, this last term meaning $s \preceq t, t \preceq s \Rightarrow s = t$). The object class is the set of elements of S (so $\mathrm{obj}(\mathbf{PO}(S)) = S$). For objects s and t

  $$\mathrm{mor}(s,t) = \begin{cases} \{s \preceq t\} & \text{if } s \preceq t, \\ \emptyset & \text{if } s \not\preceq t, \end{cases}$$

  where $\{s \preceq t\}$ is interpreted as a singleton set with the indicated string of characters as its sole element. For morphisms $f = s \preceq t$ and $g = t \preceq u$ the composition $g \circ f$ is $s \preceq u$ (well-defined by transitivity of $\preceq$). The identity morphism of the object s is $s \preceq s$ (well-defined by reflexivity of $\preceq$). (Note that antisymmetry is not required for this construction.)

- $\mathbf{C}(G)$ denotes the category associated with the group G. The object class is a singleton set $\{\cdot\}$. The morphisms are the elements of the group G (so $\mathrm{mor}(\mathbf{C}(G)) = \mathrm{mor}(\,\cdot\,,\cdot) = G\,$). For morphisms $g, h : \cdot \to \cdot$ the composition $h \circ g$ is the product hg of the group elements. The identity morphism (for the sole object $\cdot\,$) is the identity element e of the group G. (Note that the existence of inverses in the group is not required for this construction.)

15.3 Equivalence

Let $\varphi : G \to G'$ be a group homomorphism, that is, a morphism in the category **Grp**. A check to see whether φ is an isomorphism requires–using the current definition (8.1)–an inspection of group elements to see whether φ is bijective (i.e., injective and surjective). This makes the definition unsuitable for the category setting since in **Grp** one has available only objects and morphisms. The following characterization of group isomorphism remedies this problem since it uses only category-theoretic notions. It also provides a model for an appropriate generalization of isomorphism to an arbitrary category.

Theorem. *The homomorphism $\varphi : G \to G'$ is an isomorphism if and only if there exists a homomorphism $\psi : G' \to G$ such that $\psi \circ \varphi = 1_G$ and $\varphi \circ \psi = 1_{G'}$.*

Proof. Assume that φ is an isomorphism. Define $\psi : G' \to G$ by $\psi(g') = g$, where $g \in G$ satisfies $\varphi(g) = g'$. (Since φ is surjective, there exists at least one such g; since φ is injective, there exists at most one such g. Therefore, ψ is well defined.) It is immediate from the definition of ψ that $\psi \circ \varphi = 1_G$ and $\varphi \circ \psi = 1_{G'}$. For $g', h' \in G'$ we have

$$\varphi(\psi(g'h')) = g'h' = \varphi(\psi(g'))\varphi(\psi(h')) = \varphi(\psi(g')\psi(h')),$$

and, since φ is injective, we get $\psi(g'h') = \psi(g')\psi(h')$, implying that ψ is a homomorphism.

Now assume that there exists a homomorphism $\psi : G' \to G$ such that $\psi \circ \varphi = 1_G$ and $\varphi \circ \psi = 1_{G'}$. If $\varphi(g) = \varphi(h)$ $(g, h \in G)$, then

$$g = 1_G(g) = \psi(\varphi(g)) = \psi(\varphi(h)) = 1_G(h) = h,$$

implying that φ is injective. If $g' \in G'$, then $\psi(g') \in G$ and $\varphi(\psi(g')) = 1_{G'}(g') = g'$, so φ is surjective as well. Therefore φ is an isomorphism. This completes the proof. □

Let $\mathbf{C}$ be a category. A morphism $f : A \to B$ in $\mathbf{C}$ is an **equivalence** if there exists a morphism $g : B \to A$ such that $g \circ f = 1_A$ and $f \circ g = 1_B$. In this case, such a g is an **inverse** of f.

Two objects A and B of a category are **equivalent**, written $A \cong B$, if there exists an equivalence from one to the other. The property of being equivalent is (not surprisingly) an equivalence relation on the object class of the category and each cell of the corresponding partition consists of objects that are identical to each other in regard to how they relate through morphisms to other objects in the category.

In the category **Set**, an equivalence is the same as a bijection. In the categories **Grp**, $\mathbf{Vec}_F$, Ω-**Grp**, and G-**Set**, an equivalence is the same as a bijective morphism (called an isomorphism in each case). In **Top**, an equivalence is a homeomorphism, which is *not* the same as a bijective morphism (e.g., the identity function $\mathbf{R} \to \mathbf{R}$ is bijective and continuous if the first $\mathbf{R}$ has the discrete topology and the second has the usual topology, but it is not a homeomorphism). The equivalences in $\mathbf{PO}(S)$ are precisely the identity morphisms $s \preceq s$ (due to antisymmetry) and in $\mathbf{C}(G)$ every morphism is an equivalence.

15.4 Motivation for definition of product

If category theory is to be of use in the study of groups there needs to be a way to identify standard constructions by just looking at objects and morphisms. For instance, given two groups, G_1 and G_2, we seek a way of singling out among all of the objects of **Grp** the one we called the direct product P of G_1 and G_2 (i.e., $P = G_1 \times G_2$) by just looking at how objects and morphisms relate to G_1 and G_2.

The first thing we note is that there are two very natural homomorphisms $\pi_1 : P \to G_1$ and $\pi_2 : P \to G_2$ (namely, the ones given by $\pi_1((g_1, g_2)) = g_1$ and $\pi_2((g_1, g_2)) = g_2$). Next, we note that if G is any group that also admits homomorphisms $f_i : G \to G_i$ ($i = 1, 2$), then there exists one and only one way to define a homomorphism $f : G \to P$ such that $\pi_i \circ f = f_i$ ($i = 1, 2$). (Suppose such an f does exist. Let $g \in G$ and write $f(g) = (g_1, g_2)$. Then

$$g_1 = \pi_1((g_1, g_2)) = \pi_1(f(g)) = (\pi_1 \circ f)(g) = f_1(g),$$

and similarly $g_2 = f_2(g)$, so f is forced to satisfy $f(g) = (f_1(g), f_2(g))$. One easily checks that f so defined is indeed a homomorphism.)

So we have identified features of the direct product that can be expressed entirely in terms of morphisms and objects. It turns out that these features uniquely determine P (up to isomorphism) as we will see once we generalize what we have so far to an arbitrary category and an arbitrary family of objects.

15.5 Product

Let $\mathbf{C}$ be a category and let $\{A_i\}_{i \in I}$ be a family of objects of $\mathbf{C}$. A **product** of the family is a pair $(P, \{\pi_i\})$, where P is an object and $\pi_i : P \to A_i$, $i \in I$, are morphisms having the property that if A is any object and $f_i : A \to A_i$, $i \in I$, are any morphisms, then there exists a unique morphism $f : A \to P$ such that $\pi_i \circ f = f_i$ for all $i \in I$.

Products are unique up to equivalence in the following strong sense.

Theorem. *Let $(P, \{\pi_i\})$ and $(P', \{\pi'_i\})$ be products of the family $\{A_i\}_{i \in I}$. There exists a unique equivalence $f : P \to P'$ such that $\pi'_i \circ f = \pi_i$ for all i. In particular, $P \cong P'$.*

Proof. Since $(P, \{\pi_i\})$ is a product of the family, putting $(A, \{f_i\}) = (P', \{\pi'_i\})$ in the definition, we get a unique morphism $g : P' \to P$ such that $\pi_i \circ g = \pi'_i$ for all i. Similarly, since $(P', \{\pi'_i\})$ is a product of the family, putting $(A, \{f_i\}) = (P, \{\pi_i\})$ in the definition, we get a unique morphism $f : P \to P'$ such that $\pi'_i \circ f = \pi_i$ for all i. Combining, we find that $g \circ f : P \to P$ satisfies

$$\pi_i \circ (g \circ f) = (\pi_i \circ g) \circ f = \pi'_i \circ f = \pi_i$$

for all i. But also $1_P : P \to P$ satisfies $\pi_i \circ 1_P = \pi_i$ for all i. Using the uniqueness statement in the definition of product, this time with $(P, \{\pi_i\})$ as the product and $(A, \{f_i\}) = (P, \{\pi_i\})$, we conclude that $g \circ f = 1_P$.

A similar argument yields $f \circ g = 1_{P'}$. Thus f is an equivalence and $P \cong P'$. The uniqueness claim follows from the fact that $f : P \to P'$ is the unique morphism satisfying $\pi'_i \circ f = \pi_i$ for all i. □

Warning: The theorem says only that if two products exist, then they must be equivalent; it does not address the issue of existence. In fact, there are categories and families of objects for which no product exists. For instance, if $S = \{s, t\}$ and one defines a relation $\preceq$ on S by declaring only $s \preceq s$ and $t \preceq t$, then the family $\{s, t\}$ in the category $\mathbf{PO}(S)$ has no product (s cannot be the object part of a product since there is no morphism $s \to t$ to play the role of π_2, and similarly for t). There are less trivial examples as well.

If a product of the family $\{A_i\}_{i \in I}$ exists, its object part (or rather the equivalence class of its object part) is denoted $\prod_{i \in I} A_i$ (or $A_1 \sqcap A_2 \sqcap \cdots \sqcap A_n$ if $I = \{1, 2, \ldots, n\}$).

In the categories **Set**, $\mathbf{Vec}_F$, Ω-**Grp**, G-**Set**, and **Top**, a product of two objects is obtained, much as we did for **Grp**, by forming their Cartesian product (endowed with the relevant structure in a natural way) and using for $\{\pi_i\}$ the maps defined above. In 16.1, by generalizing the construction of a direct product of two groups, it is shown that a product exists for any family of objects in the category **Grp**.

15.6 Coproduct

A nice thing about the category point of view is that every construction involving morphisms, which we think of as arrows, can be dualized simply by reversing the direction of the arrows. If one construction is useful, then one expects its dual construction to be useful as well.

The dual of a product is called a "coproduct." (This follows the custom in the naming of dual constructions of prefixing co- to the name of the original construction.)

Let $\mathbf{C}$ be a category and let $\{A_i\}_{i\in I}$ be a family of objects of $\mathbf{C}$. A **coproduct** of the family is a pair $(C, \{\iota_i\})$, where C is an object and $\iota_i : A_i \to C$, $i \in I$, are morphisms having the property that if A is any object and $f_i : A_i \to A$, $i \in I$, are any morphisms, then there exists a unique morphism $f : C \to A$ such that $f \circ \iota_i = f_i$ for all $i \in I$.

One obtains a proof of the uniqueness of a coproduct by "turning the arrows around" in the proof of the corresponding theorem about products.

Theorem. *Let $(C, \{\iota_i\})$ and $(C', \{\iota'_i\})$ be coproducts of the family $\{A_i\}_{i\in I}$. There exists a unique equivalence $f : C \to C'$ such that $f \circ \iota_i = f_i$ for all i. In particular, $C \cong C'$.*

If a coproduct of the family $\{A_i\}_{i\in I}$ exists, its object part (or rather the equivalence class of its object part) is denoted $\bigsqcup_{i\in I} A_i$ (or $A_1 \sqcup A_2 \sqcup \cdots \sqcup A_n$ if $I = \{1, 2, \ldots, n\}$).

In the categories **Set**, G-**Set**, and **Top**, a coproduct of two objects is obtained by forming their disjoint union (see 15.7), with the maps ι_1 and ι_2 taken to be the inclusion maps. In $\mathbf{Vec}_F$ and **Ab** the direct product (together with the natural injection maps) is a coproduct of two objects; in general the direct sum is a coproduct of a family of objects (see 16.2).

In the category **Grp** (and similarly for Ω-**Grp**), a coproduct of two objects G_1 and G_2 exists and is called their **free product**, denoted $G_1 * G_2$. Here is a rough description of this group. First, it is assumed that the sets G_1 and G_2 are disjoint (which can be arranged for by a simple renaming of elements if necessary). As a set, $G_1 * G_2$ consists of all "reduced words," which are formal strings $x_1x_2\cdots x_n$ ($n \geq 0$, $x_i \in G_1 \cup G_2$) where no factor is an identity and no adjacent factors lie in the same group. The product of two reduced words is defined to be juxtaposition followed by reduction: if the factors on the joined ends lie in the same group, they are multiplied; if that product is the identity, then it is removed; if this leaves adjacent factors in the same group then they are multiplied; if that product is the identity, then it is removed; this process is repeated until a reduced word is obtained.

For $i \in \{1, 2\}$ the map $\iota_i : G_i \to G_1 * G_2$ sends an element of G_i to itself regarded as a word with a single factor.

15.7 Example: Coproduct of sets is disjoint union

Let S_1 and S_2 be *disjoint* sets. Put $C = S_1 \cup S_2$ and define maps $\iota_1 : S_1 \to C$ and $\iota_2 : S_2 \to C$ by $\iota_1(s_1) = s_1$ and $\iota_2(s_2) = s_2$ (the **inclusion maps**). We claim that $(C, \{\iota_i\})$ is a coproduct of the family $\{S_1, S_2\}$ in the category **Set**. Let A be a set and let $f_i : S_i \to A$ $(i = 1, 2)$ be maps. Define $f : C \to A$ by

$$f(c) = \begin{cases} f_1(c) & \text{if } c \in S_1, \\ f_2(c) & \text{if } c \in S_2. \end{cases}$$

Then for s_i in S_i we have

$$(f \circ \iota_i)(s_i) = f(\iota_i(s_i)) = f(s_i) = f_i(s_i),$$

so $f \circ \iota_i = f_i$ $(i = 1, 2)$. Also, f as defined is easily seen to be the unique map satisfying these two identities. Therefore, $(C, \{\iota_i\})$ is a coproduct of the family as claimed and we can write $S_1 \sqcup S_2 = S_1 \cup S_2$.

If the sets S_1 and S_2 are not disjoint and the functions f_1 and f_2 do not agree on the intersection, then clearly we cannot define f as above. Nevertheless, in this case we can make disjoint copies of S_1 and S_2 and then argue that *their* union is (the object part of) a coproduct of S_1 and S_2, the so-called "disjoint union" (see Exercise 15–1).

15 – Exercises

15–1 Let $\{S_i\}_{i \in I}$ be a family of sets. For each i put $S_i' = \{(s, i) \mid s \in S_i\}$. Prove that there exists a coproduct of the family $\{S_i\}_{i \in I}$ in the category **Set** having object part $\bigcup_{i \in I} S_i'$ (called the **disjoint union** of the family $\{S_i\}_{i \in I}$).

15–2 For groups G_1 and G_2 let $\iota_i : G_i \to G_1 \times G_2$ be the injections defined by $\iota_1(g_1) = (g_1, e_2)$ and $\iota_2(g_2) = (e_1, g_2)$. Prove that $(G_1 \times G_2, \{\iota_i\})$ is *not*, in general, a coproduct of the family $\{G_i\}$ in the category **Grp**.

Hint: Let G be a nonabelian group and take $G_i = G$ $(i = 1, 2)$. In the definition of coproduct, let A be G and let $f_i = 1_G$ $(i = 1, 2)$.

16 Direct product and direct sum

16.1 Definition: Direct product

Let $\{G_i\}_{i\in I}$ be an indexed family of groups. Denote by $\prod_{i\in I} G_i$ (or just $\prod G_i$) the set of all functions $a : I \to \bigcup_i G_i$ such that $a_i = a(i) \in G_i$ for each $i \in I$. For $a, b \in \prod G_i$ define $ab \in \prod G_i$ by $(ab)_i = a_i b_i$. With this binary operation $\prod G_i$ is a group, the **direct product** of the family $\{G_i\}_{i\in I}$. (See the theorem below for the category interpretation.)

Consider the special case $I = \{1, 2\}$. Here, an element a of $\prod G_i$ can be represented by a pair (a_1, a_2) with the image of 1 occupying the first position and the image of 2 occupying the second position. Note that for a and b in $\prod G_i$ the product ab is represented by the pair (a_1b_1, a_2b_2). Thus, the group $\prod G_i$ identifies with the direct product $G_1 \times G_2$ of G_1 and G_2 as defined in 2.10.

Similarly, if $I = \{1, 2, \ldots, r\}$ for some positive integer r, then $\prod G_i$ identifies with the group

$$G_1 \times G_2 \times \cdots \times G_r = \{(a_1, a_2, \ldots, a_r) \,|\, a_i \in G_i\}$$

with componentwise binary operation. The general definition of direct product given above is needed to handle the case of an arbitrary (possibly infinite) indexing set I. However, when the indexing set is $\{1, 2, \ldots, r\}$, it is customary to use the tuple notation for elements.

For each $i \in I$, define $\pi_i : \prod G_j :\to G_i$ by $\pi_i(a) = a_i$.

Theorem. *$(\prod G_i, \{\pi_i\})$ is a product of the family $\{G_i\}_{i\in I}$ in the category **Grp**.*

Proof. Put $P = \prod G_i$. For $i \in I$, we have, for $a, b \in P$, $\pi_i(ab) = (ab)_i = a_i b_i = \pi_i(a)\pi_i(b)$, so π_i is a homomorphism.

Let G be a group and let $f_i : G \to G_i$ ($i \in I$) be homomorphisms. Define $f : G \to P$ by $f(g)_i = f_i(g)$. If $g, h \in G$, then, for each $i \in I$, we have

$$f(gh)_i = f_i(gh) = f_i(g)f_i(h) = f(g)_i f(h)_i = (f(g)f(h))_i,$$

so $f(gh) = f(g)f(h)$. Therefore, f is a homomorphism.

For each $g \in G$ and $i \in I$, we have $(\pi_i \circ f)(g) = \pi_i(f(g)) = f(g)_i = f_i(g)$, so $\pi_i \circ f = f_i$.

Finally, if $f' : G \to P$ is a homomorphism that satisfies $\pi_i \circ f' = f_i$ for all $i \in I$, then for each $g \in G$ and $i \in I$, we have $f'(g)_i = \pi_i(f'(g)) = (\pi_i \circ f')(g) = f_i(g) = f(g)_i$, so that $f' = f$, demonstrating uniqueness.

This completes the proof. □

16.2 Definition: Direct sum

Let $\{G_i\}_{i\in I}$ be an indexed family of *abelian* groups (with additive notation). Denote by $\sum_{i\in I} G_i$ (or just $\sum G_i$) the subset of $\prod G_i$ consisting of all a for which $|\{i \in I \mid a_i \neq 0\}| < \infty$. Thus, $\sum G_i$ is the set of all elements of $\prod G_i$ that map all but finitely many elements of I to the identity element of the corresponding group. It is easily checked that $\sum G_i$ is a subgroup of $\prod G_i$. It is the **direct sum** of the family $\{G_i\}_{i\in I}$. If $I = \{1, 2, \dots, r\}$, this group is also denoted $G_1 \oplus G_2 \oplus \cdots \oplus G_r$.

Note that $\prod G_i = \sum G_i$ if I is finite.

For each $i \in I$, define $\iota_i : G_i \to \sum G_j$ by $\iota_i(g)_j = \delta_{ij}(g)$, where $\delta_{ij} : G_i \to G_j$ is the **Kronecker delta** function, which is the identity map if $i = j$ and the zero map if $i \neq j$.

THEOREM. *($\sum G_i, \{\iota_i\}$) is a coproduct of the family $\{G_i\}_{i\in I}$ in the category **Ab**.*

Proof. Put $C = \sum G_i$. Let $i \in I$. For $g, h \in G_i$, we have

$$\begin{aligned}\iota_i(g+h)_j &= \delta_{ij}(g+h) = \delta_{ij}(g) + \delta_{ij}(h) = \iota_i(g)_j + \iota_i(h)_j \\ &= (\iota_i(g) + \iota_i(h))_j\end{aligned}$$

for each j, so $\iota_i(g+h) = \iota_i(g) + \iota_i(h)$. Therefore, ι_i is a homomorphism.

Let A be an abelian group and for each $i \in I$ let $f_i : G_i \to A$ be a homomorphism. Define $f : C \to A$ by $f(c) = \sum_i f_i(c_i)$ (this possibly infinite sum being regarded as a finite sum by ignoring each term with $c_i = 0$). For $c, d \in C$, we have

$$\begin{aligned}f(c+d) &= \sum_i f_i((c+d)_i) = \sum_i f_i(c_i + d_i) = \sum_i [f_i(c_i) + f_i(d_i)] \\ &= \sum_i f_i(c_i) + \sum_i f_i(d_i) = f(c) + f(d),\end{aligned}$$

so f is a homomorphism.

Let $i \in I$. For each $g \in G_i$, we have

$$f \circ \iota_i(g) = f(\iota_i(g)) = \sum_j f_j(\iota_i(g)_j) = \sum_j f_j(\delta_{ij}(g)) = f_i(g),$$

so $f \circ \iota_i = f_i$.

Finally, let $f' : C \to A$ be a homomorphism that satisfies $f' \circ \iota_i = f_i$ for each i. Let $c \in C$. For each j, we have $c_j = \sum_i \delta_{ij}(c_i) = \sum_i \iota_i(c_i)_j = (\sum_i \iota_i(c_i))_j$, so that $c = \sum_i \iota_i(c_i)$. Therefore,

$$\begin{aligned} f'(c) &= f'(\sum_i \iota_i(c_i)) = \sum_i f'(\iota_i(c_i)) = \sum_i f' \circ \iota_i(c_i) \\ &= \sum_i f_i(c_i) = f(c), \end{aligned}$$

so that $f' = f$, demonstrating uniqueness.

This completes the proof.

□

16.3 Internal direct product/sum

Let G be a group and let $N_1, N_2, \ldots, N_r$ be normal subgroups of G. The group G is the **internal direct product** of $N_1, N_2, \ldots, N_r$, written $G = \dot{\prod} N_i$ if

(i) $G = N_1 N_2 \cdots N_r = \{n_1 n_2 \cdots n_r \,|\, n_i \in N_i, 1 \leq i \leq r\}$,

(ii) $N_i \cap (N_1 N_2 \cdots \hat{N_i} \cdots N_r) = \{e\}$ for each $1 \leq i \leq r$,

where $\hat{N_i}$ signifies that the ith factor is omitted. If G is an additive group and these conditions are met (with the appropriate translation to additive notation), it is the **internal direct sum** of $N_1, N_2, \ldots, N_r$, written $G = \dot{\sum} N_i$ (or $G = N_1 \dot{+} N_2 \dot{+} \cdots \dot{+} N_r$). In the notations for direct product and direct sum, the dot can be thought of as indicating the trivial intersection property (ii).

Lemma. *Let $1 \leq i, j \leq r$.*

(i) *If $N_i \cap N_j = \{e\}$, then the elements of N_i commute with those of N_j.*

(ii) *If G satisfies* (ii) *in the definition above and $i \neq j$, then the elements of N_i commute with those of N_j.*

Proof. (i) Assume that $N_i \cap N_j = \{e\}$. For $n_i \in N_i$ and $n_j \in N_j$, we have

$$n_i n_j n_i^{-1} n_j^{-1} \in N_i \cap N_j = \{e\},$$

as can be seen by grouping the first three factors and using the normality assumption, and then doing the same with the last three factors. Therefore, $n_i n_j = n_j n_i$ and the claim follows.

(ii) Assume that G satisfies (ii) in the above definition and that $i \neq j$. Then $N_i \cap N_j \subseteq N_i \cap (N_1 N_2 \cdots \hat{N}_i \cdots N_r) = \{e\}$, so the claim follows from part (i). $\square$

Theorem. *$G = \dot{\prod} N_i$ if and only if each a in G can be expressed uniquely as a product $a = n_1 n_2 \cdots n_r$ with $n_i \in N_i$ for each i.*

Proof. Assume that $G = \dot{\prod} N_i$. Let $a \in G$. By (i) of the definition, $a = n_1 n_2 \cdots n_r$ with $n_i \in N_i$.

Now suppose that also $a = m_1 m_2 \cdots m_r$ with $m_i \in N_i$. Then we have $n_1 n_2 \cdots n_r = m_1 m_2 \cdots m_r$, and for each i,

$$m_i^{-1} n_i = \prod_{j \neq i} n_j^{-1} m_j \in N_i \cap (N_1 N_2 \cdots \hat{N}_i \cdots N_r) = \{e\}$$

by (ii) of the definition (and the lemma). Therefore, $n_i = m_i$. This shows uniqueness of the expression.

Now assume that each a in G can be expressed uniquely as a product $a = n_1 n_2 \cdots n_r$ with $n_i \in N_i$. Then (i) of the definition is satisfied. Fix i and let $n_i \in N_i \cap (N_1 N_2 \cdots \hat{N}_i \cdots N_r)$. Then $n_i = n_1 n_2 \cdots n_{i-1} n_{i+1} \cdots n_r$ for some $n_j \in N_j$. This equation can be written

$$ee \cdots e n_i e \cdots e = n_1 n_2 \cdots n_{i-1} e n_{i+1} \cdots n_r,$$

so the uniqueness assumption says that $n_i = e$. This shows that $N_i \cap (N_1 N_2 \cdots \hat{N}_i \cdots N_r) = \{e\}$, which is (ii) of the definition. The proof is complete. $\square$

16.4 Internal direct product is isomorphic to direct product

Let G be a group and let $N_1, N_2, \ldots, N_r$ be normal subgroups of G.

Theorem. *If $G = \dot{\prod} N_i$, then $G \cong \prod N_i$.*

Proof. Assume that $G = \dot{\prod} N_i$. Define $\varphi : \prod N_i \to G$ by

$$\varphi((n_1, n_2, \ldots, n_r)) = n_1 n_2 \cdots n_r.$$

This function is bijective by 16.3.

For $n, m \in \prod N_i$, we have

$$\begin{aligned}\varphi(nm) = \varphi((n_1 m_1, n_2 m_2, \ldots, n_r m_r)) &= \prod_i n_i m_i = \prod_i n_i \prod_i m_i \\ &= \varphi(n)\varphi(m),\end{aligned}$$

where we have used the lemma in Section 16.3. Therefore, φ is a homomorphism and hence an isomorphism. This proves that $G \cong \prod N_i$. $\square$

16.5 Example: Vector space

Let V be a finite dimensional vector space over $\mathbf{R}$ and let $v_1, v_2, \ldots, v_r$ be a basis of V. For each i, put

$$N_i = \mathbf{R}v_i = \{\alpha v_i \mid \alpha \in \mathbf{R}\}.$$

Then each N_i is a normal subgroup of the additive group V. The spanning property of a basis implies that $V = N_1 + N_2 + \cdots + N_r$, which is (i) in the definition of internal direct sum (16.3).

Let $v \in N_i \cap (N_1 + \cdots + \hat{N_i} + \cdots N_r)$. Then $v = \alpha_i v_i$ and also

$$v = \alpha_1 v_1 + \cdots + \alpha_{i-1}v_{i-1} + \alpha_{i+1}v_{i+1} + \cdots \alpha_r v_r$$

for some $\alpha_j \in \mathbf{R}$ $(1 \leq j \leq r)$. The linear independence property of basis then implies that $\alpha_j = 0$ for each j so that $v = 0$. This gives (ii) in the definition of internal direct sum. Thus $V = \dot{\sum} N_i$.

By 16.4, $V \cong \sum N_i$. Since each N_i is isomorphic to $\mathbf{R}$ as groups we get the well-known isomorphism $V \cong \sum_{i=1}^{r} \mathbf{R} = \mathbf{R}^r$ (shown here to be an isomorphism of groups, but in fact this is an isomorphism of vector spaces as well).

16 – Exercises

16–1 Let A, B, and C be (additive) abelian groups, let $f : A \to C$ be a monomorphism, and let $g : C \to B$ be an epimorphism, and assume that $\operatorname{im} f = \ker g$.

(a) Assume that there exists a homomorphism $h : C \to A$ such that $h \circ f = 1_A$. Prove that $C \cong A \oplus B$.

(b) Assume that there exists a homomorphism $h : B \to C$ such that $g \circ h = 1_B$. Prove that $C \cong A \oplus B$.

17 Group decomposition

17.1 Definition

A group is **decomposable** if it is the internal direct product of two (or more) of its *proper* normal subgroups. A group is **indecomposable** if it is nontrivial and not decomposable.

In view of 16.4, a group is decomposable if and only if it is isomorphic to a direct product (or sum if additive) of two groups neither of which is trivial.

17.2 Example: Normal Sylow p-subgroups

Let G be a finite group and let $p_1, p_2, \ldots, p_r$ be the (distinct) prime divisors of the order of G. Assume that for each i, G has a normal Sylow p_i-subgroup P_i. By Section 14.3, this assumption is equivalent to the assumption that G has a unique Sylow p_i-subgroup P_i for each i.

Theorem. $G = \dot{\prod}_i P_i$.

Proof. By Section 6.5, $H := P_1 P_2 \cdots P_r$ is a subgroup of G. We claim that H is the internal direct product of its normal subgroups P_i $(1 \leq i \leq r)$. Part (i) of the definition is immediate.

Let $1 \leq i, j \leq r$ and assume that $i \neq j$. It follows from Lagrange's theorem that $P_i \cap P_j = \{e\}$ since this intersection is a subgroup of both P_i and P_j, which have relatively prime orders. By the lemma of 16.3, the elements of P_i commute with those of P_j.

Fix i and let $g \in P_i \cap (P_1 P_2 \cdots \hat{P_i} \cdots P_r)$. We have $g = \prod_{j \neq i} g_j$ for some $g_j \in P_j$. Let $n = \prod_{j \neq i} n_j$, where $n_j = |P_j|$. Using the previous paragraph, we get

$$g^n = (\prod_{j \neq i} g_j)^n = \prod_{j \neq i} g_j^n = e,$$

which implies that the order of g divides n. But g is an element of P_i, so its order divides $|P_i|$ as well. Since n and $|P_i|$ are relatively prime, we conclude that $g = e$. Therefore, part (ii) of the definition of internal direct product holds and the claim is established.

By 16.4, H is isomorphic to the direct product $\prod_i P_i$, so $|H| = \prod_i |P_i| = |G|$. We conclude that $G = H = \dot{\prod}_i P_i$. □

17.3 Example: Finite cyclic group

Let m and n be positive integers and denote by $\gcd(m,n)$ the greatest common divisor of m and n.

Theorem.

(i) $\mathbf{Z}_{mn} \cong \mathbf{Z}_m \oplus \mathbf{Z}_n$ *if and only if* $\gcd(m,n) = 1$.

(ii) $\mathbf{Z}_n$ *is indecomposable if and only if* $n = p^k$ *for some prime number* p *and some* $k \in \mathbf{N}$.

Proof. (i) Assume that $\mathbf{Z}_{mn} \cong \mathbf{Z}_m \oplus \mathbf{Z}_n$. Let $(a,b) \in \mathbf{Z}_m \oplus \mathbf{Z}_n$ correspond to $1 \in \mathbf{Z}_{mn}$ under an isomorphism. Let k be an arbitrary multiple of both m and n, so that $k = m'm$ and $k = n'n$ for some integers m' and n'. We have

$$k(a,b) = (ka,kb) = (m'ma, n'nb) = (m'e, n'e) = (e,e),$$

so that $mn = o((a,b)) \,|\, k$. This shows that mn is the least common multiple of m and n, which in turn implies that $\gcd(m,n) = 1$.

Now assume that $\gcd(m,n) = 1$. Since $\mathbf{Z}_{mn}$ is abelian, its Sylow p-subgroups are normal, so Section 17.2 applies. By reason of order, the sum P_m of the Sylow p-subgroups of $\mathbf{Z}_{mn}$ with p dividing m has order m, and it is cyclic (4.4) and therefore isomorphic to $\mathbf{Z}_m$. Similarly, the sum P_n of the Sylow p-subgroups with p dividing n is isomorphic to $\mathbf{Z}_n$. By 17.2, $\mathbf{Z}_{mn}$ is the internal direct sum of P_m and P_n and is therefore isomorphic to $P_m \oplus P_n \cong \mathbf{Z}_m \oplus \mathbf{Z}_n$.

(ii) Assume that $\mathbf{Z}_n$ is indecomposable. Since an indecomposable group is nontrivial, n is not 1 and is therefore divisible by some prime number p. By part (i), n is divisible by at most one prime number. We conclude that $n = p^k$ for some $k \in \mathbf{N}$.

Now assume that $n = p^k$ for some prime number p and some $k \in \mathbf{N}$. Let A and B be subgroups of $\mathbf{Z}_n$ and assume that $\mathbf{Z}_n$ is the internal direct sum of A and B. By Section 4.4, A and B are both cyclic, so that $A \cong \mathbf{Z}_l$ and $B \cong \mathbf{Z}_m$ for some positive integers l and m. Then $\mathbf{Z}_n \cong \mathbf{Z}_l \oplus \mathbf{Z}_m$ by Section 16.4, implying that $\gcd(l,m) = 1$ by part (i). But l and m are divisors of n by Lagrange's theorem and hence powers of p. Therefore, either $l = 1$ or $m = 1$, that is, either A is trivial or B is trivial. We conclude that $\mathbf{Z}_n$ is indecomposable. □

17.4 Krull-Remak-Schmidt theorem

Let G be a group and assume that G has a composition series. The following theorem says that G can be decomposed into a direct product of indecomposable groups and done so in an essentially unique way. In the statement, the term "direct product" has the usual broad meaning, which includes the possibility of only one factor.

Theorem (Krull-Remak-Schmidt).

(i) *G is isomorphic to a direct product of finitely many indecomposable groups.*

(ii) *If $G \cong H_1 \times \cdots \times H_r$ and also $G \cong K_1 \times \cdots \times K_s$ with each H_i and K_j indecomposable groups, then $r = s$ and there exists a permutation $\sigma \in S_r$ such that $H_i \cong K_{\sigma(i)}$ for each $1 \leq i \leq r$.*

Proof. Omitted. □

17.5 Fundamental theorem of finite abelian groups

Let G be a nontrivial finite abelian group.

Theorem (Fundamental theorem of finite abelian groups).

(i) *There exist unique prime numbers $p_1 < p_2 < \cdots < p_n$ and unique positive integers $k_{i1} \leq k_{i2} \leq \cdots \leq k_{im_i}$, $1 \leq i \leq n$, such that*

$$G \cong \sum_{i,j} \mathbf{Z}_{p_i^{k_{ij}}}.$$

(ii) *There exist unique positive integers $m_1 | m_2 | \dots | m_t$ such that*

$$G \cong \sum_{i} \mathbf{Z}_{m_i}.$$

Proof. (i) The uniqueness statement follows immediately from the Krull-Remak-Schmidt theorem (17.4) and (ii) of Section 17.3, so we turn to the existence statement. By Section 17.2, we may (and do) assume that G is a p-group for some prime number p.

Let $C = \langle c \rangle$ be a cyclic subgroup of G of greatest possible order. We claim that there exists a subgroup H of G such that G is the internal direct product of C and H. Once this is established, the desired decomposition will follow, since a proof by induction on the order of G will show that G is isomorphic to a direct product of cyclic groups, each of order a power of p.

We prove the claim by induction on $n = |G|/|C|$. If $n = 1$, then $C = G$ and we can let $H = \{e\}$. Assume that $n > 1$, so that $C \neq G$. There exists an element $a \in G$ with $a \notin C$. Assume that a has been chosen with minimal order. Now $o(a^p) < o(a)$ (since $a \neq e$), so that $a^p \in C$. If a^p were to generate C, we would have $|\langle a \rangle| = p|C| > |C|$, contrary to the choice of C. Therefore, a^p does not generate C, so, by Exercise 5–5, we have $a^p = c^{pk}$ for some integer k. Hence, ac^{-k} is an element of order p not in C. By the choice of a, a too has order p.

Put $A = \langle a \rangle$ and $G' = G/A$, and let $\pi : G \to G'$ be the canonical epimorphism. The kernel of π is A, which intersects C trivially, so π restricts to an isomorphism $C \to \pi(C) =: C'$. Since a homomorphic image of an element has order at most the order of the element, it follows that C' is a cyclic subgroup of G' of greatest possible order.

Now $|G'|/|C'| = |G|/(p|C|) < n$ so the induction hypothesis applies to guarantee the existence of a subgroup H' of G' such that $G' = C'H'$ and $C' \cap H' = \{e'\}$. By the correspondence theorem (9.6), $H' = \pi(H)$ for some subgroup H of G with $H \supseteq A$. Now $\pi(G) = G' = C'H' = \pi(C)\pi(H) = \pi(CH)$, so, again by the correspondence theorem, $G = CH$. Also,

$$\pi(C \cap H) \subseteq \pi(C) \cap \pi(H) = C' \cap H' = \{e'\},$$

so $C \cap H \subseteq \ker \pi = A$. Therefore,

$$C \cap H \subseteq A \cap (C \cap H) = (A \cap C) \cap H = \{e\} \cap H = \{e\}.$$

Thus, G is the internal direct product of C and H. In light of the earlier remarks, we see that this completes the proof of part (i).

(ii) This claim is proved by exhibiting an algorithm for computing such $m_1, m_2, \dots, m_t$ from the decomposition in part (i) (establishing existence), as well as an algorithm for going the other way (establishing uniqueness). Rather than presenting the general algorithms, we illustrate in Section 17.6 how they are carried out for a specific group. □

The prime powers $p_i^{k_{ij}}$ $(1 \leq i \leq n, 1 \leq j \leq m_i)$ appearing in (i) are the **elementary divisors** of G. The positive integers m_i $(1 \leq i \leq t)$ appearing in (ii) are the **invariant factors** of G.

17.6 Example

Let $G = \mathbf{Z}_{10} \oplus \mathbf{Z}_{35} \oplus \mathbf{Z}_8 \oplus \mathbf{Z}_{98}$. Using 17.3(i) as well as associativity and commutativity of direct sum, we get

$$\begin{aligned} G &= \mathbf{Z}_{10} \oplus \mathbf{Z}_{35} \oplus \mathbf{Z}_8 \oplus \mathbf{Z}_{98} \\ &\cong (\mathbf{Z}_2 \oplus \mathbf{Z}_5) \oplus (\mathbf{Z}_5 \oplus \mathbf{Z}_7) \oplus \mathbf{Z}_{2^3} \oplus (\mathbf{Z}_2 \oplus \mathbf{Z}_{7^2}) \\ &\cong \mathbf{Z}_2 \oplus \mathbf{Z}_2 \oplus \mathbf{Z}_{2^3} \oplus \mathbf{Z}_5 \oplus \mathbf{Z}_5 \oplus \mathbf{Z}_7 \oplus \mathbf{Z}_{7^2}, \end{aligned}$$

so the elementary divisors of G are $2, 2, 2^3, 5, 5, 7, 7^2$.

Note that if these prime powers are arranged in rows with the greatest powers forming the bottom row,

$$\begin{array}{lll} 2 & & \\ 2 & 5 & 7 \\ 2^3 & 5 & 7^2, \end{array}$$

then each row product divides the next: $2|70|1960$. Using 17.3(i) again, we get

$$\begin{aligned} G &\cong \mathbf{Z}_2 \oplus (\mathbf{Z}_2 \oplus \mathbf{Z}_5 \oplus \mathbf{Z}_7) \oplus (\mathbf{Z}_{2^3} \oplus \mathbf{Z}_5 \oplus \mathbf{Z}_{7^2}) \\ &\cong \mathbf{Z}_2 \oplus \mathbf{Z}_{70} \oplus \mathbf{Z}_{1960}, \end{aligned}$$

so the invariant factors of G are $2, 70, 1960$. Conversely, if one begins with these invariant factors, then, by displaying their prime power factorizations as above, one obtains the elementary divisors.

18 Solvable group

18.1 Definition

Let G be a group. A subnormal series $(G_i) = (G_0, G_1, \ldots, G_r)$ of G is a **solvable series** if each of its factors G_i/G_{i+1} is abelian. The group G is **solvable** if it has a solvable series.

For example, an abelian group is solvable. (Indeed, if G is an abelian group, then $(G, \{e\})$ is a subnormal series of G with a single factor (isomorphic to) the abelian group G.) Therefore, the notion of "solvable group" generalizes the notion of "abelian group."

The term "solvable" is due to a theorem involving solutions of polynomial equations. This theorem says that if $p(x)$ is a polynomial over $\mathbf{Q}$ (or more generally over any field of characteristic zero), then the equation $p(x) = 0$ is solvable by radicals over $\mathbf{Q}$ (meaning roughly that the (complex) solutions can all be expressed using only $\mathbf{Q}$ and the operations $+, -, \times, \div, \sqrt[n]{\ }$) if and only if a certain group connected to $p(x)$, its Galois group over $\mathbf{Q}$, is solvable as defined above.

18.2 Example: Dihedral group is solvable

Let n be a positive integer. Recall (4.3) that the dihedral group D_{2n} is generated by the set $\{\rho, \tau\}$. In cycle notation we have $\rho = (1, 2, \ldots, n)$, so $\langle\rho\rangle$ is a cyclic subgroup of D_{2n} of order n. Since D_{2n} has order $2n$, this subgroup has index 2 and is therefore normal. Thus, $(D_{2n}, \langle\rho\rangle, \{\varepsilon\})$ is a subnormal series of D_{2n}. Since the factors of this series are (isomorphic to) $\mathbf{Z}_2$ and $\mathbf{Z}_n$, both of which are abelian, the dihedral group is solvable.

18.3 Derived series

Let G be a group. Recall that the commutator of two elements a and b of G is $[a, b] = a^{-1}b^{-1}ab$ and the commutator subgroup of G is

$$G^{(1)} = \langle [G, G] \rangle,$$

where $[G, G] = \{[a, b] \mid a, b \in G\}$.

The **derived series** of G is the sequence of subgroups $G^{(0)}, G^{(1)}, G^{(2)}, \ldots$ defined recursively by

- $G^{(0)} = G$,

- $G^{(i)} = (G^{(i-1)})^{(1)} \quad (i > 0)$.

So the derived series of G starts with G and each remaining term is the commutator subgroup of the preceding term.

Theorem. *G is solvable if and only if $G^{(n)} = \{e\}$ for some n.*

Proof. Assume that G is solvable. There exists a solvable series $(G_0, G_1, \dots, G_r)$ of G.

We claim that $G_i \supseteq G^{(i)}$ for each $0 \leq i \leq r$. The proof is by induction on i. The case $i = 0$ is $G_0 = G = G^{(0)}$, so it holds. Assume that $i > 0$. Using the induction hypothesis, we have

$$G^{(i)} = \langle [G^{(i-1)}, G^{(i-1)}] \rangle \subseteq \langle [G_{i-1}, G_{i-1}] \rangle = G_{i-1}^{(1)}.$$

Now G_{i-1}/G_i is abelian, so Section 7.3 and this inclusion give $G_i \supseteq G_{i-1}^{(1)} \supseteq G^{(i)}$, as desired.

Since $\{e\} = G_r \supseteq G^{(r)}$, we have $G^{(r)} = \{e\}$, so the condition is satisfied with $n = r$.

Now assume that $G^{(n)} = \{e\}$ for some n. For each $0 \leq i \leq n$, put $G_i = G^{(i)}$. We have $G_i = (G^{(i-1)})^{(1)} = G_{i-1}^{(1)}$ for each $0 < i \leq n$. By Section 7.3, $(G_0, G_1, \dots, G_n)$ is a solvable series of G. Therefore, G is solvable. □

18.4 Subgroup and homomorphic image

Let G be a group.

Theorem.

(i) *If G is solvable and $H \leq G$, then H is solvable.*

(ii) *If G is solvable and $\varphi : G \to G'$ is a homomorphism, then $\operatorname{im} \varphi$ is solvable.*

(iii) *If $N \lhd G$, then G is solvable if and only if N and G/N are both solvable.*

Proof. The proof continually uses, with no further indication, the characterization of solvable group given in 18.3.

(i) We begin by making the observation that for any subgroups H and K of G with $H \subseteq K$,

$$H^{(1)} = \langle [H, H] \rangle \subseteq \langle [K, K] \rangle = K^{(1)}.$$

Let $H \leq G$. We claim that $H^{(i)} \subseteq G^{(i)}$ for each i. The proof is by induction on i. Since $H^{(0)} = H \subseteq G = G^{(0)}$, the claim holds for $i = 0$.

Assume that $i > 0$. By the induction hypothesis and the observation made above, we have

$$H^{(i)} = (H^{(i-1)})^{(1)} \subseteq (G^{(i-1)})^{(1)} = G^{(i)},$$

so the claim is established.

Assume that G is solvable. We have $G^{(n)} = \{e\}$ for some n, so that $H^{(n)} \subseteq G^{(n)} = \{e\}$. Therefore, $H^{(n)} = \{e\}$ and H is solvable.

(ii) Let $\varphi : G \to G'$ be a homomorphism. We begin by making the observation that for any subgroup H of G,

$$\varphi(H)^{(1)} = \langle[\varphi(H), \varphi(H)]\rangle = \langle\varphi([H, H])\rangle = \varphi(\langle[H, H]\rangle) = \varphi(H^{(1)}).$$

We claim that $\varphi(G)^{(i)} = \varphi(G^{(i)})$ for each i. The proof is by induction on i. Since $\varphi(G)^{(0)} = \varphi(G) = \varphi(G^{(0)})$, the claim holds for $i = 0$.

Assume that $i > 0$. By the induction hypothesis and the observation made above, we have

$$\varphi(G)^{(i)} = (\varphi(G)^{(i-1)})^{(1)} = \varphi(G^{(i-1)})^{(1)} = \varphi((G^{(i-1)})^{(1)}) = \varphi(G^{(i)}),$$

so the claim is established.

Assume that G is solvable. We have $G^{(n)} = \{e\}$ for some n, so that $\varphi(G)^{(n)} = \varphi(G^{(n)}) = \varphi(\{e\}) = \{e'\}$. Therefore, $\operatorname{im}\varphi = \varphi(G)$ is solvable.

(iii) Let $N \triangleleft G$. If G is solvable, then N is solvable by part (i) and G/N is solvable by part (ii), using the canonical epimorphism $\pi : G \to G/N$.

Assume that N and G/N are both solvable. We have $(G/N)^{(m)} = \{N\}$ for some m. By the established claim in the proof of part (ii),

$$\{N\} = (G/N)^{(m)} = \pi(G)^{(m)} = \pi(G^{(m)}),$$

implying that $G^{(m)} \subseteq N$. By part (i), $G^{(m)}$ is solvable, so that $G^{(m+n)} = (G^{(m)})^{(n)} = \{e\}$ for some n. Therefore, G is solvable. □

18.5 S_n not solvable for $n \geq 5$

Theorem. *The symmetric group S_n is not solvable for $n \geq 5$.*

Proof. Let $n \geq 5$. Since a subgroup of a solvable group is solvable (18.4) it is enough to prove that the alternating group $G = A_n$ is not solvable. Since G is simple (14.6), $G^{(1)}$ is either G or $\{e\}$.

Suppose that $G^{(1)} = \{e\}$. Then the quotient $G/G^{(1)}$ is isomorphic to G, which is nonabelian since, for instance, the cycles $(1, 2, 3)$ and $(3, 4, 5)$ are elements of G and

$$(1, 2, 3)(3, 4, 5) = (1, 2, 3, 4, 5) \neq (1, 2, 4, 5, 3) = (3, 4, 5)(1, 2, 3).$$

But this contradicts Section 7.3.

Therefore, $G^{(1)} = G$, and it follows that $G^{(n)} = G \neq \{\epsilon\}$ for all n. We conclude from Section 18.3 that G is not solvable. This completes the proof. □

It is shown in Galois theory that the Galois group over $\mathbf{Q}$ of the quintic polynomial $x^5 - 4x + 2$ is S_5. In view of the remarks in 18.1 and the present theorem, the equation $x^5 - 4x + 2 = 0$ is not solvable by radicals over $\mathbf{Q}$. In particular, there is no generalization of the quadratic formula that gives the solutions of a general fifth degree polynomial equation.

18.6 A group of odd order is solvable

The following result, known as the Odd Order theorem (also the Feit-Thompson theorem), was a major step toward the eventual proof of the classification theorem of finite simple groups (see 10.7). The proof, published in 1963 and occupying 255 journal pages, built on pioneering work of Michio Suzuki.

Theorem (Feit-Thompson). *A group of odd order is solvable.*

18 – Exercises

18–1 Let $H_1, H_2, \ldots, H_n$ be groups and put $G = H_1 \times \cdots \times H_n$. Prove that G is solvable if and only if H_i is solvable for each $1 \leq i \leq n$.

18–2 Let p and q be prime numbers. Prove that a group of order pq is solvable.

Hint: Reduce to the case $p \neq q$ and then use Sylow theory.

19 Nilpotent group

19.1 Definition

Let G be a group. A **central series** of G is a tuple $(N_0, N_1, \dots, N_r)$, where

(i) $N_i \lhd G$ for all i,

(ii) $N_{i-1} \subseteq N_i$ for all $0 < i \leq r$,

(iii) $N_i/N_{i-1} \subseteq Z(G/N_{i-1})$ for all $0 < i \leq r$.

G is **nilpotent** if it has a central series $(N_0, N_1, \dots, N_r)$ such that $N_0 = \{e\}$ and $N_r = G$.

- If G is abelian, then it is nilpotent since $(\{e\}, G)$ is a central series.

19.2 Upper central series

Let G be a group. The **upper central series** of G is the sequence $Z_0, Z_1, Z_2, \dots$ of subgroups of G with $Z_i = Z_i(G)$ defined recursively by

(i) $Z_0 = \{e\}$,

(ii) $Z_i = \pi^{-1}(Z(G/Z_{i-1}))$ for $i > 0$,

where $\pi : G \to G/Z_{i-1}$ is the canonical epimorphism. Since π is surjective we have for each $i > 0$

$$Z_i/Z_{i-1} = \pi(Z_i) = \pi(\pi^{-1}(Z(G/Z_{i-1}))) = Z(G/Z_{i-1}).$$

This formula shows that $(Z_0, Z_1, \dots, Z_r)$ is a central series of G for each nonnegative integer r.

19.3 Lower central series

Let G be a group. The **lower central series** of G is the sequence $L_0, L_1, L_2, \dots$ of subgroups of G with $L_i = L_i(G)$ defined recursively by

(i) $L_0 = G$,

(ii) $L_i = \langle [G, L_{i-1}] \rangle$ for $i > 0$, where $[G, L_{i-1}] = \{[g, x] \mid g \in G, x \in L_{i-1}\}$.

For each positive integer r, $(L_r, L_{r-1}, \dots, L_0)$ is a central series of G (see Exercise 19–1).

There is a transparent relationship between the lower central series of G and the derived series $G^{(0)}, G^{(1)}, G^{(2)}, \dots$ of G (see 18.3 for the definition):

Theorem. *$G^{(i)} \subseteq L_i$ for each $i \geq 0$.*

Proof. The proof is by induction on i. Since $G^{(0)} = G = L_0$, the case $i = 0$ holds.

Assume that $i > 0$. We have

$$G^{(i)} = \langle [G^{(i-1)}, G^{(i-1)}] \rangle \subseteq \langle [L_{i-1}, L_{i-1}] \rangle \subseteq \langle [G, L_{i-1}] \rangle = L_i$$

where the first inclusion is due to the induction hypothesis. □

19.4 Upper/lower central series characterization of nilpotent

Let G be a group.

Theorem. *The following are equivalent:*

(i) *G is nilpotent;*

(ii) *$L_r = \{e\}$ for some r;*

(iii) *$Z_r = G$ for some r.*

Proof. (i) implies (ii): Assume that (i) holds, so that there exists a central series $(N_0, N_1, \dots, N_r)$ of G with $N_0 = \{e\}$ and $N_r = G$.

It is enough to prove that $L_i \subseteq N_{r-i}$ for each i, for then $L_r = N_{r-r} = N_0 = \{e\}$. We proceed by induction on i. Since $L_0 = G = N_{r-0}$, the case $i = 0$ holds.

Assume that $i > 0$. For $n \in N_{r-i+1}$ and $g \in G$, we have, using the definition of central series, $[g, n]N_{r-i} = [gN_{r-i}, nN_{r-i}] = N_{r-i}$, so that $[g, n] \in N_{r-i}$. This shows that $[G, N_{r-i+1}] \subseteq N_{r-i}$. Therefore,

$$L_i = \langle [G, L_{i-1}] \rangle \subseteq \langle [G, N_{r-(i-1)}] \rangle \subseteq N_{r-i},$$

where the first inclusion is from the induction hypothesis. This completes the proof of this implication.

(ii) implies (iii): Assume that (ii) holds, so that $L_r = \{e\}$ for some r.

It is enough to prove that $Z_i \supseteq L_{r-i}$ for each i, for then $Z_r \supseteq L_{r-r} = L_0 = G$. We proceed by induction on i. Since $Z_0 = \{e\} = L_{r-0}$, the case $i = 0$ holds.

Assume that $i > 0$. Let $l \in L_{r-i}$ and $g \in G$. Using the definition of the lower central series and the induction hypothesis, we have

$$[g, l] \in L_{r-i+1} = L_{r-(i-1)} \subseteq Z_{i-1},$$

so that $[gZ_{i-1}, lZ_{i-1}] = [g, l]Z_{i-1} = Z_{i-1}$. It follows that $lZ_{i-1} \in Z(G/Z_{i-1})$. Therefore, $Z_i = \pi^{-1}(Z(G/Z_{i-1}) \supseteq L_{r-i}$, where $\pi : G \to G/Z_{i-1}$ is the canonical epimorphism. This completes the proof of this implication.

(iii) implies (i): Assume that (iii) holds, so that $Z_r = G$ for some r. From the equation $Z_i/Z_{i-1} = Z(G/Z_{i-1})$ (see 19.2), it follows that (Z_i) is an upper central series of G. Since $Z_0 = \{e\}$ and $Z_r = G$, G is nilpotent. This completes the proof. □

19.5 Nilpotent group is solvable

Let G be a group.

Theorem. *If G is nilpotent, then G is solvable.*

Proof. Assume that G is nilpotent so that $L_r = \{e\}$ for some r, where (L_i) is the lower central series of G (see 19.4). By Section 19.3, $G^{(r)} \subseteq L_r = \{e\}$, so that G is solvable by Section 18.3. □

The converse of this theorem does not hold (see Section 19.9 below).

19.6 Finite p-group is nilpotent

Let p be a prime number.

Theorem. *A finite p-group is nilpotent.*

Proof. Let G be a finite p-group. By Section 19.4, it is enough to show that $Z_r = G$ for some r, where (Z_i) is the upper central series of G. If $i \geq 0$ and $Z_i \neq G$, then the quotient G/Z_i is nontrivial and it is a p-group (using Lagrange's theorem), so Section 13.3 gives $Z_{i+1} = \pi^{-1}(Z(G/Z_i)) \supsetneqq Z_i$, where $\pi : G \to G/Z_i$ is the canonical epimorphism. Since G is finite, we must have $Z_r = G$ for some r (else, the upper central series would be a strictly increasing sequence of subgroups). □

19.7 Subgroup and homomorphic image

Let G be a nilpotent group.

THEOREM.

(i) *If H is a subgroup of G, then H is nilpotent.*

(ii) *If $\varphi : G \to G'$ is a homomorphism, then $\operatorname{im}\varphi$ is nilpotent.*

Proof. (i) Let H be a subgroup of G. By Section 19.4, it is enough to prove that $L_i(H) \subseteq L_i(G)$ for each i, for then, since $L_r(G) = \{e\}$ for some r, we would have $L_r(H) = \{e\}$ as well. We proceed by induction on i. Since $L_0(H) = H \subseteq G = L_0(G)$, the case $i = 0$ holds. Assume that $i > 0$. We have

$$L_i(H) = \langle [H, L_{i-1}(H)] \rangle \subseteq \langle [G, L_{i-1}(G)] \rangle = L_i(G),$$

where the inclusion is from the induction hypothesis. This completes the proof of this part.

(ii) Let $\varphi : G \to G'$ be a homomorphism. By Section 19.4, it is enough to prove that $L_i(\varphi(G)) \subseteq \varphi(L_i(G))$ for each i, for then, since $L_r(G) = \{e\}$ for some r, we would have

$$L_r(\operatorname{im}\varphi) = L_r(\varphi(G)) \subseteq \varphi(L_r(G)) = \varphi(\{e\}) = \{e'\}$$

as well. We proceed by induction on i. Since $L_0(\varphi(G)) = \varphi(G) = \varphi(L_0(G))$, the case $i = 0$ holds. Assume that $i > 0$. We have

$$\begin{aligned} L_i(\varphi(G)) &= \langle [\varphi(G), L_{i-1}(\varphi(G))] \rangle \subseteq \langle [\varphi(G), \varphi(L_{i-1}(G))] \rangle \\ &\subseteq \varphi\langle [G, L_{i-1}(G)] \rangle = \varphi(L_i(G)), \end{aligned}$$

where the first inclusion is from the induction hypothesis. This completes the proof. □

19.8 Product is nilpotent iff factors are

Let $G_1, G_2, \dots, G_n$ be groups and put $G = G_1 \times G_2 \times \cdots \times G_n$.

THEOREM. *G is nilpotent if and only if G_i is nilpotent for each $1 \leq i \leq n$.*

Proof. Assume that G is nilpotent. For each $1 \leq i \leq n$, the map $\pi_i : G \to G_i$ given by $\pi_i((a_1, a_2, \dots, a_n)) = a_i$ is an epimorphism, so G_i is nilpotent by part (ii) of 19.7.

Assume that G_i is nilpotent for each $1 \leq i \leq n$. By Section 19.4, it is sufficient to prove that $L_j(G) \subseteq L_j(G_1) \times \cdots \times L_j(G_n)$ for each j, for then, since, for each i, there exists r_i such that $L_{r_i}(G_i) = \{e_i\}$, we would have

$$\begin{aligned} L_r(G) &\subseteq L_r(G_1) \times \cdots \times L_r(G_n) \subseteq L_{r_1}(G_1) \times \cdots \times L_{r_n}(G_n) \\ &= \{e_1\} \times \cdots \times \{e_n\} = \{e\}, \end{aligned}$$

where r is the largest of the r_i. We proceed by induction on j. Since $L_0(G) = G = G_1 \times \cdots \times G_n = L_0(G_1) \times \cdots \times L_0(G_n)$, the case $j = 0$ holds. Assume that $j > 0$. We have

$$\begin{aligned} L_j(G) &= \langle [G, L_{j-1}(G)] \rangle \\ &\subseteq \langle [G_1 \times \cdots \times G_n, L_{j-1}(G_1) \times \cdots \times L_{j-1}(G_n)] \rangle \\ &= \langle [G_1, L_{j-1}(G_1)] \rangle \times \cdots \times \langle [G_n, L_{j-1}(G_n)] \rangle \\ &= L_j(G_1) \times \cdots \times L_j(G_n), \end{aligned}$$

where the inclusion is from the induction hypothesis. The proof is complete.

□

19.9 Finite group nilpotent iff product of p-groups

Let G be a finite group.

THEOREM. *The following are equivalent:*

(i) *G is nilpotent;*

(ii) *If H is a proper subgroup of G, then $N_G(H) \supsetneq H$;*

(iii) *G has a normal Sylow p-subgroup for each prime number p;*

(iv) *G is isomorphic to a direct product of p-groups for various prime numbers p.*

Proof. (i) implies (ii): Assume that (i) holds, so that $Z_r = G$ for some r, where (Z_i) is the upper central series of G (see 19.4). Let H be a proper subgroup of G. Since $Z_0 = \{e\} \subseteq H$ and $Z_r = G \not\subseteq H$, there exists a largest m for which $Z_m \subseteq H$. By this choice of m, there exists some element a of Z_{m+1} that is not in H.

Let $h \in H$. From the definition of Z_{m+1}, we have $[h, a]Z_m = [hZ_m, aZ_m]$ $= Z_m$, so that $[h, a] \in Z_m$. Therefore, $h^a = h(h^{-1}a^{-1}ha) = h[h, a] \in H$, which implies that $a \in N_G(H) \backslash H$ and (ii) follows.

(ii) implies (iii): Assume that (ii) holds. Let p be a prime number and let P be a Sylow p-subgroup of G (such exists by the Sylow existence theorem 14.2). Suppose that $N_G(P) \neq G$. Since G is finite, $N_G(P)$ is contained in some maximal (proper) subgroup H of G. By our assumption, $N_G(H) \supsetneq H$, so that $N_G(H) = G$. Therefore, H is normal. By the Frattini argument (14–1), we have $G = HN_G(P) \subseteq H$, contradicting that H is proper. Therefore, $N_G(P) = G$, so that P is normal giving (iii).

(iii) implies (iv): Assume that (iii) holds. By Section 17.2, G is the internal direct product of its Sylow p-subgroups for various primes numbers p, so (iv) now follows from Section 16.4.

(iv) implies (i): This implication is immediate from Sections 19.8 and 19.6. □

A nilpotent group is solvable (19.5), but a solvable group need not be nilpotent. Indeed, S_3 is solvable since $(S_3, A_3, \{\varepsilon\})$ is a subnormal series of S_3 having abelian factors $\mathbf{Z}_2$ and $\mathbf{Z}_3$, but it is not nilpotent because its Sylow 2- and 3-subgroups are isomorphic to $\mathbf{Z}_2$ and $\mathbf{Z}_3$, respectively, and the direct product of these groups is abelian (while S_3 is not).

19 – Exercises

19–1 Let G be a group and let (L_i) be the lower central series of G. Prove that, for each nonnegative integer r, $(L_r, L_{r-1}, \ldots, L_0)$ is a central series of G.

19–2 Let G be a group, let $(N_0, N_1, \ldots, N_r)$ be a central series of G with $N_0 = \{e\}$, and let (Z_i) be the upper central series of G. Prove that $Z_i \supseteq N_i$ for each $0 \leq i \leq r$.

20 Free object of concrete category

20.1 Motivation for definition of free object

Let V be a vector space over a field F. A subset $B = \{v_1, v_2, \dots, v_n\}$ of V is a **basis** for V if

- B generates V (meaning, $V = \langle B \rangle =$ intersection of all subspaces of V containing B),
- B is linearly independent (meaning $\sum_i \alpha_i v_i = 0$, with $\alpha_i \in F$, implies $\alpha_i = 0$ for all i).

Note that $\langle B \rangle = \{\sum_i \alpha_i v_i \,|\, \alpha_i \in F\}$, since this latter set is a subspace containing B (as is easily checked) and it is contained in every subspace of V that contains B by the closure properties of a subspace. Therefore, to say that B generates V is the same as saying that every vector in V can be written as a linear combination of the vectors $v_1, v_2, \dots, v_n$ (which is the spanning property of a basis one learns in elementary linear algebra).

The two defining properties of a basis given above are convenient for checking whether a given subset of V is a basis, but the following characterization is more useful.

- B is a basis for V if and only if every vector v in V can be written *uniquely* in the form $v = \sum_i \alpha_i v_i$.

($\Rightarrow$) If B is a basis, then every vector $v \in V$ can be written $v = \sum_i \alpha_i v_i$ in at least one way by the generating (spanning) property, and if $v = \sum_i \beta_i v_i$ as well, then the sums must be equal giving $\sum_i (\alpha_i - \beta_i) v_i = 0$, whence $\alpha_i = \beta_i$ for all i since the coefficients must be 0 by linear independence.

($\Leftarrow$) If the unique expression property holds, then B clearly spans (generates) V and if $\sum_i \alpha_i v_i = 0$ then, since $\sum_i 0 v_i = 0$ as well, uniqueness says $\alpha_i = 0$ for all i.

The idea of a basis is so useful in the study of vector spaces (as one knows from linear algebra) that it makes one wonder whether an analogous concept might be available in categories besides $\mathbf{Vec}_F$. Such a concept for a general category would have to be defined entirely in terms of objects and morphisms since these are all we have in an arbitrary category. As an aid to our search for a suitable generalization, we look for a characterization of vector space basis that avoids the inspection of elements required in the characterizations given above. The following well-known characterization is a first attempt.

- B is a basis for V if and only if given any vector space W and any vectors $w_1, w_2, \ldots, w_n$ in W there exists a unique linear transformation $T : V \to W$ such that $T(v_i) = w_i$ for all i.

The direction ($\Rightarrow$) says that for any vector space W, we can arrange for a linear transformation $T : V \to W$ that has any effect we wish on the basis vectors, that is, we are *free* to pick the images $T(v_i)$ (since the vectors w_i are arbitrary) and, once these images are chosen there is a (unique) extension of T to a linear transformation.

This characterization is a step in the right direction since it involves the linear transformation T, which is a morphism in the category $\mathbf{Vec}_F$. However, it falls short of our goal due to the continued dependence on elements. Here is a slight reworking that will serve as a second attempt:

- B is a basis for V if and only if given any vector space W and any function $f : B \to W$, there exists a unique linear transformation $T : V \to W$ such that $T(v_i) = f(v_i)$ for all i.

In this version, we have expressed the choice of basis vector images using the more categorical function $f : B \to W$. Although this is not a morphism in the category $\mathbf{Vec}_F$, it *is* a morphism in the category **Set** provided we forget that W has the structure of a vector space and just think of it as a set.

This leads us to the idea of characterizing the notion of a basis by using *two* categories, $\mathbf{Vec}_F$ and **Set**, and the fact that one can associate to any vector space W its underlying set, which we denote $\sigma(W)$. In order to clearly delineate statements made in **Set** from those made in $\mathbf{Vec}_F$ we introduce a new set $X = \{x_1, x_2, \ldots, x_n\}$ and put $\iota(x_i) = v_i$, thus defining an injection $\iota : X \to \sigma(V)$, which is a morphism in **Set**. The preceding characterization now reads

- $\iota(X)$ is a basis for V if and only if given any object W of $\mathbf{Vec}_F$ and any morphism $f : X \to \sigma(W)$ in **Set**, there exists a unique morphism $T : V \to W$ in $\mathbf{Vec}_F$ such that $\sigma(T) \circ \iota = f$.

The notation $\sigma(T)$ signifies that T is to be viewed simply as a map of sets, $\sigma(T) : \sigma(V) \to \sigma(W)$, which is what is required for $\sigma(T) \circ \iota = f$ to make sense as an equation involving morphisms in **Set**.

In 20.4, the category **Vec** will be replaced by an arbitrary category **C** for which there is a suitable way to associate to each object A of **C** a set $\sigma(A)$, and then the fact that V and $\iota : X \to \sigma(V)$ satisfy the above condition will be expressed by saying that the pair (V, ι) is "free" on X. The association $\sigma : \mathbf{C} \to \mathbf{Set}$ we need is an example of a "functor," the definition of which is given in 20.2.

20.2 Functor

Let $\mathbf{C}$ and $\mathbf{D}$ be categories. A **functor** from $\mathbf{C}$ to $\mathbf{D}$, written $F : \mathbf{C} \to \mathbf{D}$, is a function

$$F : \begin{cases} \mathrm{obj}(\mathbf{C}) \to \mathrm{obj}(\mathbf{D}) \\ \mathrm{mor}(\mathbf{C}) \to \mathrm{mor}(\mathbf{D}) \end{cases}$$

satisfying the following:

- $A \xrightarrow{f} B$ in $\mathbf{C}$ implies $F(A) \xrightarrow{F(f)} F(B)$ in $\mathbf{D}$,
- $F(1_A) = 1_{F(A)}$ for each $A \in \mathrm{obj}(\mathbf{C})$,
- $F(g \circ f) = F(g) \circ F(f)$ for all $f, g \in \mathrm{mor}(\mathbf{C})$ with $t(f) = s(g)$.

Another way of saying the first condition is that if $f \in \mathrm{mor}(\mathbf{C})$, then $F(s(f)) = s(F(f))$ and $F(t(f)) = t(F(f))$, so that F respects the source and target maps. Intuitively, F maps all arrows (morphisms) from one vertex (object) to another to arrows from the image of the first vertex to the image of the second vertex. There is another useful notion of functor that reverses the direction of the arrows (see Exercise 2).

A functor $F : \mathbf{C} \to \mathbf{D}$ is **faithful** if for each pair (A, B) of objects of $\mathbf{C}$ the restriction of F to $\mathrm{mor}(A, B)$ is injective. A **full** functor is defined similarly with "surjective" replacing "injective."

A **concrete category** is a pair $(\mathbf{C}, \sigma)$, where $\mathbf{C}$ is a category and $\sigma : \mathbf{C} \to \mathbf{Set}$ is a faithful functor. If $(\mathbf{C}, \sigma)$ is a concrete category and it is clear what σ is intended to be, we say that the category $\mathbf{C}$ is concrete.

20.3 Examples of functors

- The **forgetful functor** $\sigma : \mathbf{Grp} \to \mathbf{Set}$ sends a group $(G, *)$ to its underlying set G and a homomorphism to itself (just viewed as a map of sets). There are forgetful functors for the categories $\mathbf{Ab}$, $\mathbf{Vec}_F$, Ω-$\mathbf{Grp}$, G-$\mathbf{Set}$, and $\mathbf{Top}$ as well. These categories paired with their forgetful functors are all concrete.

- The **power set functor** $P : \mathbf{Set} \to \mathbf{Set}$ sends a set to its power set (set of subsets) and sends a morphism $f : X \to Y$ to the map $P(f) : P(X) \to P(Y)$ defined by $P(f)(S) = f(S)$ for $S \subseteq X$.
- Let $\mathbf{C}$ be a category and let A be an object of $\mathbf{C}$. The (covariant) **mor functor** induced by A is the functor $F_A : \mathbf{C} \to \mathbf{Set}$ that sends an object B to the set $\mathrm{mor}(A, B)$ and a morphism $f : B \to B'$ to the map $f_* : \mathrm{mor}(A, B) \to \mathrm{mor}(A, B')$ defined by $f_*(g) = f \circ g$.
- The **fundamental group functor** is the functor $\pi : \mathbf{PTop} \to \mathbf{Grp}$ from the category of pointed topological spaces to the category of groups that sends a pointed space to the fundamental group at the distinguished point and a morphism to the induced homomorphism.
- If $\varphi : G \to G'$ is a group homomorphism, then a functor $\mathbf{C}(\varphi) : \mathbf{C}(G) \to \mathbf{C}(G')$ is obtained by defining $\mathbf{C}(\varphi)(\cdot) = \cdot$ and $\mathbf{C}(\varphi)(g) = \varphi(g)$ $(g \in G)$.

20.4 Free object

Let $(\mathbf{C}, \sigma)$ be a concrete category and let X be a set. A **free object** on X in $\mathbf{C}$ is a pair (F, ι) with F an object of $\mathbf{C}$ and $\iota : X \to \sigma(F)$ a map with the property that given any object A of $\mathbf{C}$ and map $f : X \to \sigma(A)$ there exists a unique morphism $\bar{f} : F \to A$ in $\mathbf{C}$ such that $\sigma(\bar{f}) \circ \iota = f$.

If (F, ι) is a free object on X, one sometimes just says that F is free on X.

A free object on X, if it exists, is unique up to equivalence in the following strong sense.

Theorem. *Let (F, ι) and (F', ι') be free on X in $\mathbf{C}$. There exists a unique equivalence $f : F \to F'$ such that $\sigma(f) \circ \iota = \iota'$. In particular, $F \cong F'$.*

Proof. Letting (F', ι') play the role of (A, f) in the definition of free object, we get a unique morphism $\bar{\iota'} : F \to F'$ satisfying $\sigma(\bar{\iota'}) \circ \iota = \iota'$. Similarly, with (F, ι) now playing the role of (A, f), we get a unique morphism $\bar{\iota} : F' \to F$ satisfying $\sigma(\bar{\iota}) \circ \iota' = \iota$. Using the fact that a functor respects compositions of morphisms, we have

$$\sigma(\bar{\iota} \circ \bar{\iota'}) \circ \iota = \sigma(\bar{\iota}) \circ \sigma(\bar{\iota'}) \circ \iota = \sigma(\bar{\iota}) \circ \iota' = \iota.$$

But also $\sigma(1_F) \circ \iota = 1_{\sigma(F)} \circ \iota = \iota$. From the uniqueness statement in the definition of free object, we conclude that $\bar{\iota} \circ \bar{\iota'} = 1_F$. Similarly, $\bar{\iota'} \circ \bar{\iota} = 1_{F'}$. Therefore, $f = \bar{\iota'} : F \to F'$ is an equivalence, and it is the unique morphism such that $\sigma(f) \circ \iota = \iota'$ as noted above. □

20.5 Free group

Let X be a set.

THEOREM. *There exists a free object on X in the category* **Grp**.

Proof. Omitted. □

In lieu of a proof, we give a rough description of a construction of such a free object (F, ι). A first attempt is to let F be the set of all formal products $x_1 x_2 \cdots x_n$ ($x_i \in X$) with multiplication defined to be concatenation. The empty product (i.e., the indicated product with $n = 0$) serves as an identity. However, there are no inverses, so this does not quite work. We fix this problem by inventing inverses: Let X^{-1} be a set in one-to-one correspondence with X and with $X \cap X^{-1} = \emptyset$. For $x \in X$, denote the corresponding element of X^{-1} by x^{-1} and call it the inverse of x. Let F be the set of all formal products $u_1 u_2 \cdots u_n$ ($u_i \in X \cup X^{-1}$), in which an element of X and its inverse do not appear side by side. Such a formal product is a **reduced word** on X. Define multiplication in F to be concatenation followed by reduction: if the factors on the joined ends are an element of X and its inverse, then they are removed; if this produces a juxtaposition of an element of X and its inverse, then these factors are removed as well; this process is continued until a reduced word is obtained. Finally, $\iota : X \to \sigma(F)$ is defined by sending $x \in X$ to the word x.

We need to check that (F, ι) satisfies the definition of free object. Let A be a group and let $f : X \to \sigma(A)$ be a map. For $x \in X$, put $\bar{f}(\iota(x)) = f(x)$ and $\bar{f}(\iota(x)^{-1}) = f(x)^{-1}$ and for a reduced word $w = u_1 u_2 \cdots u_n \in F$ put $\bar{f}(w) = \bar{f}(u_1)\bar{f}(u_2) \cdots \bar{f}(u_n)$. Then $\bar{f} \circ \iota = f$ and it follows easily that $\bar{f} : F \to A$ is a homomorphism. Moreover $F = \langle \iota(X) \rangle$, so if $\bar{f}$ is to satisfy these conditions it has to be defined as above, and uniqueness follows.

The group F is called the **free group** on X. It is sometimes written $F(X)$.

COROLLARY. *Every group G is a homomorphic image of a free group on a set X, and this set can be chosen to be any generating set for G.*

Proof. Let G be a group and let $X \subseteq G$ be a set of generators of G. By the theorem, there exists a free object (F, ι) on X in the category **Grp**. Taking $f : X \to \sigma(G)$ to be the inclusion map in the definition of free object, we get a homomorphism $\bar{f} : F \to G$ such that $\sigma(\bar{f}) \circ \iota = f$. We have

$$\operatorname{im} \bar{f} \supseteq \bar{f}(\iota(X)) = (\sigma(\bar{f}) \circ \iota)(X) = f(X) = X,$$

and since $\operatorname{im} \bar{f}$ is a subgroup of G, it follows that $\operatorname{im} \bar{f} \supseteq \langle X \rangle = G$, that is, $\bar{f}$ is surjective, as desired. □

20.6 Initial/Terminal object

The reader might have been struck by the similarity among the statements (and proofs) of uniqueness up to equivalence of products, coproducts, and free objects (see 15.5, 15.6, 20.4). In this section, we discuss the unifying notions of initial and terminal objects, of which these constructions are examples.

Let $\mathbf{D}$ be a category. An object I of $\mathbf{D}$ is an **initial object** if for each object A of $\mathbf{D}$ there exists a unique morphism $I \to A$. An object T of $\mathbf{D}$ is a **terminal object** if for each object A of $\mathbf{D}$ there exists a unique morphism $A \to T$. (Initial objects and terminal objects are sometimes referred to collectively as "universal objects" with the former being called universally repelling and the latter being called universally attracting.)

Theorem.

(i) *If I and I' are two initial objects of $\mathbf{D}$, then there exists a unique equivalence $I \to I'$.*

(ii) *If T and T' are two terminal objects of $\mathbf{D}$, then there exists a unique equivalence $T \to T'$.*

Proof. (i) Let I and I' be two initial objects of $\mathbf{D}$. Since I is an initial object, there exists a unique morphism $f : I \to I'$, and since I' is an initial object, there exists a unique morphism $g : I' \to I$. This gives a morphism $g \circ f : I \to I$. But $1_I : I \to I$ is also a morphism. By the uniqueness statement in the definition of initial object, we get $g \circ f = 1_I$. Similarly, $f \circ g = 1_{I'}$. Therefore, $f : I \to I'$ is an equivalence. As noted earlier, f is the unique morphism from I to I', so it must be the unique equivalence as well.

The proof of (ii) is similar. □

Here we show the connection with free objects. Let $(\mathbf{C}, \sigma)$ be a concrete category and let X be a set. Form a new category $\mathbf{D}$:

- objects are pairs (A, α) with $A \in \operatorname{obj}(\mathbf{C})$ and $\alpha : X \to \sigma(A)$ a map,
- morphisms from the object (A, α) to the object (B, β) are morphisms $f : A \to B$ in $\mathbf{C}$ such that $\sigma(f) \circ \alpha = \beta$,

and composition is the same as the composition in $\mathbf{C}$. It is trivial to check that this is indeed a category.

Let (F, ι) and (F', ι') be free objects on X in $\mathbf{C}$. By the definition of free object, these are both initial objects of $\mathbf{D}$. Therefore, the theorem says that there exists a unique equivalence $(F, \iota) \to (F', \iota')$ in $\mathbf{D}$, which is the same as saying that there exists a unique equivalence $f : F \to F'$ in $\mathbf{C}$ such that $\sigma(f) \circ \iota = \iota'$ (cf. 20.4).

This point of view has the advantage of consolidating the two things that make up a free object, an object F and a map ι, into a single object (F, ι). Of course, we pay the price of having to work in the category $\mathbf{D}$, which is more complicated than the category $\mathbf{C}$. However, one can argue that, as far as the notion of freeness on the set X is concerned, $\mathbf{D}$ is the correct category to work in. For instance, when we said earlier that free objects were unique up to equivalence we had to say "in a strong sense" to refer to the fact that there is a unique equivalence that is compatible with the related maps. However, just ordinary equivalence (of initial objects) in the category $\mathbf{D}$ entails this strong sense automatically.

One can mimic the construction of the category $\mathbf{D}$ above to create a category in which a coproduct $(C, \{\iota_i\})$ is an initial object, and then appeal to the theorem on uniqueness of initial objects to get the strong sense of uniqueness of coproducts stated in 15.6 (see Exercise 6). A similar construction can be done for products; they end up being terminal objects.

Initial and terminal objects are ubiquitous throughout mathematics. They can be discovered almost anywhere it makes sense to talk about an optimal thing. For instance, the greatest common divisor of a finite set of natural numbers is an optimal choice from among all divisors of every number in the set. A greatest common divisor is a terminal object in an appropriate category. Here is a short list of other notions that can be characterized as either initial or terminal objects in appropriate categories:

- least common multiple of a finite set of natural numbers,
- supremum of a subset of $\mathbf{R}$,
- infimum of a subset of $\mathbf{R}$,
- empty set,
- trivial group,
- quotient group,

- closure of a subset of $\mathbf{R}^n$,
- interior of a subset of $\mathbf{R}^n$,
- universal covering space of a locally connected topological space,
- universal enveloping algebra of a Lie algebra,
- tensor product of modules,
- completion of a metric space.

20 – Exercises

1. For a group G, put $F(G) = G/G^{(1)}$, where $G^{(1)}$ is the commutator subgroup of G. Define $F(\varphi)$ for φ a homomorphism in such a way that $F : \mathbf{Grp} \to \mathbf{Ab}$ becomes a functor.

2. A **contravariant functor** is defined just like a functor, except with conditions

   - $A \xrightarrow{f} B$ in $\mathbf{C}$ implies $F(B) \xrightarrow{F(f)} F(A)$ in $\mathbf{D}$,
   - $F(1_A) = 1_{F(A)}$ for each $A \in \operatorname{obj}(\mathbf{C})$,
   - $F(g \circ f) = F(f) \circ F(g)$ for all $f, g \in \operatorname{mor}(\mathbf{C})$ with $t(f) = s(g)$,

   so a contravariant functor reverses the direction of arrows. An ordinary functor is called a **covariant functor** if it needs to be distinguished from a contravariant functor.

   Modify the definition of the power set functor (see 20.3) to obtain a contravariant functor $\mathbf{Set} \to \mathbf{Set}$.

3. Let $\mathbf{C}$ be a category and let B be an object of $\mathbf{C}$. By analogy with the covariant mor functor of 20.3, define a contravariant mor functor $F^B : \mathbf{C} \to \mathbf{Set}$ induced by B and verify the axioms.

4. For a vector space V over a field F, put $V^* = \operatorname{mor}(V, F)$ (= set of morphisms in $\mathbf{Vec}_F$). Define a contravariant functor D:$\mathbf{Vec}_F \to \mathbf{Vec}_F$ with object map $V \mapsto V^*$. (Hint: See Exercise 3.)

5. Let $(\mathbf{C}, \sigma)$ be a concrete category, let X be a set, and let (F, ι) be free on X in $\mathbf{C}$. Assume that there exists an object A of $\mathbf{C}$ with $|\sigma(A)| > 1$.

   (a) Prove that ι is injective.

(b) Give an example to show that ι need not be injective if no such object A exists.

6. Let $\{A_i\}_{i\in I}$ be a family of objects of a category $\mathbf{C}$. Construct a category $\mathbf{D}$ in which a coproduct of the family is an initial object.

7. Let S be a finite set of natural numbers. A greatest common divisor (gcd) of S is a natural number g having the properties:

   - $g \,|\, s$ for all $s \in S$,
   - if $d \in \mathbf{N}$ has the property that $d \,|\, s$ for all $s \in S$, then $d \,|\, g$.

   Construct a category in which a gcd of S is a terminal object and use uniqueness up to equivalence of terminal objects to prove uniqueness of a gcd.

8. Let G be a group and let N be a normal subgroup of G. Construct a category in which G/N (or rather G/N paired with a certain homomorphism) is an initial object. (Hint: Fundamental Homomorphism Theorem.)

21 Group presentation

21.1 Motivation

We begin with a simple example. Let $F = \langle g \rangle$ be an infinite (multiplicative) cyclic group. It is easy to see that F is free on the set $\{g\}$ (with ι taken to be the inclusion map). Let $N = \langle g^6 \rangle$ and for $x \in F$, let $\overline{x}$ denote the image of x under the canonical epimorphism $F \to F/N =: \overline{F}$. Then

$$\overline{g}^6 = g^6 N = N = \overline{e},$$

so, intuitively, the effect of passing to the quotient $F/\langle g^6 \rangle$ is to make g^6 equal to the identity. We express this by writing

$$\overline{F} = \langle g \,|\, g^6 = e \rangle.$$

and by saying that $\overline{F}$ has the presentation $(g \,|\, g^6 = e)$.

The group $\overline{F}$ is the most general group with one generator having sixth power the identity in the sense that any other such group is a homomorphic image of $\overline{F}$. (This follows from von Dyck's theorem (21.4), which follows almost immediately from the Fundamental Homomorphism theorem.) Since $\overline{F}$ is isomorphic to $\mathbf{Z}_6$ this fact provides one way (albeit too fancy) of seeing that there are epimorphisms $\mathbf{Z}_6 \to A$ for $A = \mathbf{Z}_6$, $\mathbf{Z}_3$, $\mathbf{Z}_2$, and $\{e\}$.

21.2 Generators and relations

Let X be a set and let $F = F(X)$ be the free group on X. Let R be a set of reduced words relative to X as in 20.5 and let N be the *normal* subgroup of F generated by R (so N is the intersection of all normal subgroups of F containing the set R). We define

$$\langle X \,|\, R = e \rangle := F/N,$$

where $R = e$ denotes the set of all expressions $r = e$ $(r \in R)$. This is the group with **generators** X and **relations** $R = e$.

For example, given a natural number n,

$$\langle a, b \,|\, a^n = e, b^2 = e, abab = e \rangle$$

denotes the quotient of the free group on $\{a, b\}$ by the normal subgroup generated by $\{a^n, b^2, abab\}$. It turns out that this group is isomorphic to the dihedral group D_{2n} of order $2n$ (see Exercise 1).

This group is also sometimes written

$$\langle a, b \,|\, a^n, b^2, aba = b^{-1} \rangle.$$

The convention for notations appearing to the right of the bar ($|$) is that all equations are to be viewed in the form $r = e$ by applying standard group operations (so $aba = b^{-1}$ means $abab = e$), and any expression not involving an equality sign is understood to equal e (so a^n means $a^n = e$ and b^2 means $b^2 = e$).

It should be pointed out that considerable collapsing can occur. For instance, in the group $\langle a, b \,|\, ab^{-1} = e \rangle$ we have $a = b$ (meaning $aN = bN$), so the two distinct (reduced) words a and b collapse to the same group element after passing to the quotient F/N. The "word problem" is the problem of finding a general algorithm for deciding when two words collapse to the same group element. It was shown by Novikov in 1955 that the word problem is unsolvable–even when both X and R are assumed to be finite.

21.3 Presentation of a group

Let X be a set and let R be a set of reduced words on X. A group G has **presentation** $(X \,|\, R = e)$ if $G \cong \langle X \,|\, R = e \rangle$.

(We draw a distinction between the ordered pair $(X \,|\, R = e)$ and the group $\langle X \,|\, R = e \rangle$ while most authors use one or the other of these notations to mean simultaneously the ordered pair and the group.)

Theorem. *Every group has a presentation.*

Proof. Let G be a group. By Section 20.5, if F is the free group on any generating set X (e.g., $X = G$) of G, then there exists an epimorphism $\varphi : F \to G$. For any set R of generators of $\ker \varphi$ (e.g., $R = \ker \varphi$) we have

$$G = \operatorname{im} \varphi \cong F/\ker \varphi = \langle X \,|\, R = e \rangle,$$

so that G has presentation $(X \,|\, R = e)$. □

The generating sets X and R are generally chosen to be as small and natural as possible.

21.4 Von Dyck's theorem

Let X be a set, let R be a set of reduced words on X, let H be a group, and let $f : X \to H$ be a map. We say that the relations $R = e$ are **satisfied** in H relative to f if $\bar{f}(r) = e$ for all $r \in R$, where $\bar{f} : F(X) \to H$ is the homomorphism induced by f. More simply put, the relations $R = e$ are satisfied in H if they are valid after replacing each $x \in X$ by its image $f(x) \in H$ (and replacing the identity e of $F(X)$ by the identity e of H).

Theorem (von Dyck). *Let G be a group with presentation $(X \mid R = e)$. If the relations $R = e$ are satisfied in H relative to $f : X \to H$, and $H = \langle f(X) \rangle$, then there exists an epimorphism $G \to H$.*

Proof. Assume that the relations $R = e$ are satisfied in H relative to $f : X \to H$, and that $H = \langle f(X) \rangle$. With $\iota : X \to F(X) =: F$ denoting the natural inclusion, we have $\bar{f} \circ \iota = f$. Therefore, $\operatorname{im} \bar{f} \supseteq \bar{f}(\iota(X)) = f(X)$, and, since $f(X)$ generates H, it follows that $\bar{f}$ is surjective.

Now $\ker \bar{f}$ is a normal subgroup of F containing R, so $\ker \bar{f} \supseteq N$, where N is the normal subgroup of F generated by R. By the fundamental homomorphism theorem (8.7), there exists a homomorphism $\varphi : F/N \to H$ such that $\varphi \circ \pi = \bar{f}$, where $\pi : F \to F/N$ is the canonical epimorphism. Since $\bar{f}$ is surjective, φ is as well. Finally, since G has presentation $(X \mid R = e)$, we have $G \cong \langle X \mid R = e \rangle = F/N$, so φ can be composed with an isomorphism to get an epimorphism $G \to H$, as desired. □

21 – Exercises

1. Prove that the dihedral group D_{2n} has the presentation $(a, b \mid a^n = e, b^2 = e, abab = e)$. (Hint: Prove that every element in this presentation can be written in the form $a^i b^j$ with $0 \leq i < n, 0 \leq j < 2$. Use the theorem of von Dyck.)

Appendix

A Writing proofs

A.1 Strings of relations

In a string of relations, the main news value should appear at the ends of the string and all of the intermediate steps should be easily verifiable.

- *If* $r > 2$, *then* $r^2 + r - 6 = (r + 3)(r - 2) > 0$ $(r \in \mathbf{R})$.

  The point being made is that if r is greater than 2, then $r^2 + r - 6$ is positive. The equality $r^2+r-6 = (r+3)(r-2)$ is verified by multiplying out the right hand side; the inequality $(r + 3)(r - 2) > 0$ follows from the fact that both factors are positive under the assumption $r > 2$.

- $(2 + 3)^2 = 5^2 = 25 \neq 13 = 4 + 9 = 2^2 + 3^2$.

  This says that $(2 + 3)^2 \neq 2^2 + 3^2$.

- $\frac{1}{2} + \frac{2}{3} - \frac{1}{4} = \frac{6}{12} + \frac{8}{12} - \frac{3}{12} = \frac{11}{12} \notin \mathbf{Z}$.

  This says that $\frac{1}{2} + \frac{2}{3} - \frac{1}{4}$ is not an integer. It is confusing to the reader if this point is made by writing $\frac{11}{12} = \frac{6}{12} + \frac{8}{12} - \frac{3}{12} = \frac{1}{2} + \frac{2}{3} - \frac{1}{4} \notin \mathbf{Z}$. In working from left to right, he can easily check each step except for the last, $\frac{1}{2} + \frac{2}{3} - \frac{1}{4} \notin \mathbf{Z}$. For this, he has to work backwards to see that $\frac{1}{2} + \frac{2}{3} - \frac{1}{4}$ equals $\frac{11}{12}$ which is not an integer.

A.2 If P, then Q.

To prove a statement of the form “If P, then Q” (which is the same as “P implies Q”), assume that P is true and show that Q is true.

- *Let* $a, b, c \in \mathbf{R}$. *If* $a < b$ *and* $c < 0$, *then* $ca > cb$.

  Proof: Assume that $a < b$ and $c < 0$. Since $a < b$, we have $a - b < 0$. Therefore, $ca - cb = c(a - b) > 0$. Hence, $ca > cb$, as desired. □

A.3 P if and only if Q

A statement of the form “P if and only if Q” is a combination of the two statements “If P, then Q” and “If Q, then P,” so it is often written with a double implication symbol: “$P \Leftrightarrow Q$.” To prove such a statement, take each implication separately and proceed as in A.2.

- *For $r \in \mathbf{R}$, $r^2 - 2r = -1$ if and only if $r = 1$.*

  Proof: Let $r \in \mathbf{R}$.

  ($\Rightarrow$) Assume $r^2 - 2r = -1$. Then $(r-1)^2 = r^2 - 2r + 1 = 0$, which implies $r - 1 = 0$. Hence, $r = 1$.

  ($\Leftarrow$) Assume $r = 1$. Then $r^2 - 2r = 1^2 - 2(1) = -1$. $\square$

It is common to use ($\Rightarrow$) and ($\Leftarrow$) as above to introduce the particular implication being proved. Incidentally, you should convince yourself that ($\Leftarrow$) corresponds to the statement "P if Q" while ($\Rightarrow$) corresponds to the statement "P only if Q."

A.4 Counterexample

To show that a statement involving "for every" is false, provide a single, explicit counterexample.

- *For every positive real number r, we have $r^3 > r^2$.*

  This statement is false, for if $r = \frac{1}{2}$, then $r^3 = \frac{1}{8} \not> \frac{1}{4} = r^2$.

I could also have said that the statement is false, for if r is any real number less than 1, then $r^3 - r^2 = r^2(r-1) < 0$, whence $r^3 < r^2$. However, the explicit counterexample above is preferable to this argument in that it is easier to understand and it says just what needs to be said.

A.5 Showing "there exists"

To prove a statement involving "there exists," just exhibit a single such object and show that it satisfies the stated property.

- *There exists an $r \in \mathbf{R}$ satisfying $r^2 + r - 12 = 0$.*

  Proof: Let $r = 3$. Then, $r^2 + r - 12 = 3^2 + 3 - 12 = 0$.

Note that I did not tell the reader how I came up with an r that works. There is no obligation to reveal the thought process that leads to the insight. In fact, doing so risks confusing the reader since it is unexpected. Also, I did not include that $r = -4$ also works since exhibiting a single r sufficed.

A.6 Showing "for every"

To prove a statement involving "for every," start with an arbitrary such object and show that it satisfies the given property.

- *For every* $r \in \mathbf{R}$ *with* $r \geq 3$, *we have* $r^2 - 2r + 1 \geq 4$.

  Proof: Let $r \in \mathbf{R}$ with $r \geq 3$. Then $r^2 - 2r + 1 = (r-1)^2 \geq (3-1)^2 = 4$. □

The first sentence of the proof means "Let r denote an arbitrary (i.e., any old) real number greater than or equal to 3."

A.7 Proof by contradiction

There is a method for proving a statement called "Proof by contradiction" which is sometimes useful. To use this method, one assumes that the given statement is false and then proceeds to derive a contradiction. The contradiction signals the presence somewhere of an invalid step. Therefore, provided all the other steps are valid, one can conclude that the initial assumption was not correct, which is to say that the given statement is in fact true.

- *There are infinitely many prime numbers.* (A *prime number* is an integer greater than 1 that is evenly divisible by no positive integers except 1 and itself (e.g., 2, 3, 5, 7, 11, ...).)

  Proof: Suppose the statement is false. In other words, suppose there are only finitely many primes. We may enumerate them: $p_1, p_2, \ldots, p_n$. Consider the number $s := p_1 p_2 \cdots p_n + 1$. Now s is an integer greater than 1, so it must be divisible by some prime, say p_i. This means that $s = p_i m$ for some integer m. But then, $1 = s - p_1 p_2 \cdots p_n = p_i(m - p_1 p_2 \cdots \hat{p_i} \cdots p_n)$ where the symbol $\hat{p_i}$ means "delete p_i." The expression in the parentheses is just some integer and, since it is not possible to multiply the prime p_i by another integer and get 1, this is an obvious contradiction. Hence, our original assumption is wrong, that is, there are infinitely many prime numbers. □

This is essentially Euclid's famous proof of the infinitude of primes.

A.8 Contrapositive

A statement of the form "If P, then Q" is logically equivalent to the statement "If not Q, then not P" meaning that the first statement is true if and only if the second statement is true (you should be able to convince yourself that this is the case). This second statement is called the *contrapositive* of the first. Sometimes, proving the contrapositive of a statement is easier than proving the statement itself.

- *If* $r \neq s$*, then* $2r + 3 \neq 2s + 3$ $(r, s \in \mathbf{R})$.

  Proof: We prove the contrapositive: If $2r + 3 = 2s + 3$, then $r = s$. Assume $2r + 3 = 2s + 3$. Subtracting 3 from both sides and dividing through by 2 gives $r = s$, as desired. □

Occasionally, people give a proof by contradiction (see A.7) of a statement that can be established more directly by proving its contrapositive. For example, to prove the above statement by contradiction, we would start off assuming that there exist $r, s \in \mathbf{R}$ such that $r \neq s$ and $2r + 3 = 2s + 3$. Then, as above, we would obtain $r = s$, contradicting that $r \neq s$. This proof is valid, but it is not as direct as the first proof. When a proof by contradiction ends up contradicting one of the initial assumptions, as in this case, it can usually be recast using the contrapositive. (Note that this was not the case in the example worked for A.7.)

A.9 Negation

In order to formulate the contrapositives of statements or to give proofs by contradiction, one needs to be able to negate statements. Usually, this is easy; for instance, the negative of $a = b$ is $a \neq b$. However, more complicated statements require some thought. Logicians have formal rules that can be used to accurately negate extremely complex statements, but since most statements occurring in mathematics have very simple logical structures, mathematicians tend not to use the formulas relying instead on their own reasoning. Statements involving "for every" sometimes cause problems, so here is an example.

- $ab = ba$ *for every* $a, b \in G$.

  The negative is "There exist $a, b \in G$ such that $ab \neq ba$" (not "$ab \neq ba$ for every $a, b \in G$").

A.10 Variable scope

The "scope" of a variable in a proof refers to the portion of the proof that starts where the variable is introduced and ends where the variable no longer has meaning.

Generally, if a variable x is introduced with "If $x\dots$" or "For every $x\dots$," then that variable (and every variable that depends on it), ceases to have meaning at the end of the sentence. Such a variable x is said to have "local scope."

On the other hand, a variable x introduced using "Let $x\dots$" or "There exists $x\dots$" has meaning all the way to the end of the proof. Such a variable is said to have "global scope."

- If n is an even integer, then $n = 2m$ for some integer m. Therefore, $m = n/2$.

  (Incorrect. Due to the conditional "If $\dots$" the variable n has no meaning past the first sentence. Since m depends on this n, it too has no meaning past the first sentence.)

- Let n be an even integer. Then $n = 2m$ for some integer m. Therefore, $m = n/2$.

  (Correct. The phrase "Let n be an even integer" fixes an arbitrary even integer, and from that point on n refers to that fixed even integer. The m in the next sentence is chosen to satisfy $n = 2m$, so it too continues to have meaning from that point on.)

- For every odd integer n, the integer $n+1$ is even. Therefore, $n+1 = 2m$ for some $m \in \mathbf{Z}$.

  (Incorrect. Due to the quantifier "For every," n ceases to have meaning past the first sentence.)

- Let n be an odd integer. Then $n + 1$ is even, so $n + 1 = 2m$ for some integer m. Therefore, $m = (n + 1)/2$.

  (Correct. Both n and m have the indicated meaning to the end of the proof, unless the meaning is overwritten by a new statement, such as "Let n be an even integer.")