Preliminaries

These notes are derived primarily from *Abstract Algebra, Theory and Applications* by Thomas Judson (16ed). Portions are also drawn from Keith Conrad's notes on the dihedral groups.

This first set of notes covers material that students should be familiar with from MTH331.

1. Basic Set Theory

Definition: Set, elements

A *set* X is a well-defined collection of objects, called *elements*. One should be able to determine membership in a set. We write $a \in X$ to say an element is in the set.

Example. Important sets to know are

(1) $\mathbb{N}$, the natural numbers $1, 2, 3, \ldots$
(2) $\mathbb{Z}$, the integers
(3) $\mathbb{Q}$, the rationals
(4) $\mathbb{R}$, the reals
(5) $\mathbb{C}$, the complex numbers
(6) $\emptyset$, the empty set.

We now briefly discuss basic operations on sets.

Definition: Subset

A *subset* of a set X is a set Y such that for all $y \in Y$, $y \in X$. We write $Y \subset X$. We say sets X and Y are *equal* and write $X = Y$ if $X \subset Y$ and $Y \subset X$. We say Y is a *proper subset* of X if $Y \subset X$ and $Y \neq X$.

Example. $\emptyset \subset \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

Operations on sets: Let A and B be subsets of a (universal) set U.

- Union of A and B: $A \cup B = \{x : x \in A \text{ or } x \in B\}$
- Intersection of A and B: $A \cap B = \{x : x \in A \text{ and } x \in B\}$
- Complement of A in U: $A' = \{x : x \in U \text{ and } x \notin A\}$
- Difference: $A \backslash B = \{x : x \in A \text{ and } x \notin B\}$
- Cartesian Product: $A \times B = \{(a, b) : a \in A, b \in B\}$

Proposition 1: Set Laws

Let A, B, C be sets.

(1) $A \cup A = A$, $A \cap A = A$, $A \backslash A = \emptyset$.

(2) $A \cup \emptyset = A$, $A \cap \emptyset = \emptyset$.

(3) $A \cup (B \cup C) = (A \cup B) \cup C$, $A \cap (B \cap C) = (A \cap B) \cap C$.

(4) $A \cup B = B \cup A$, $A \cap B = B \cap A$.

(5) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

(6) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Proof. We will prove the first statement in (3). The remainder are left as an exercise.

Let $x \in A \cup (B \cup C)$. Then $x \in A$ or $x \in B \cup C$. In the first case $x \in A$ so $x \in A \cup B$. Thus, $x \in (A \cup B) \cup C)$. In the second case, either $x \in B$ or $x \in C$. If $x \in B$ then $x \in A \cup B$. Now in either of these cases, $x \in (A \cup B) \cup C$. Hence, $A \cup (B \cup C) \subset (A \cup B) \cup C$.

Now suppose $x \in (A \cup B) \cup C$. Then either $x \in A \cup B$ or $x \in C$. In the second case $x \in B \cup C$. In the first case, either $x \in A$, or else $x \in B$, whence $x \in B \cup C$. Thus, in all cases either $x \in A$ or $x \in B \cup C$. Thus $(A \cup B) \cup C \subset A \cup (B \cup C)$. This implies that $A \cup (B \cup C) = (A \cup B) \cup C$, as desired. □

Theorem 2: DeMorgan's Laws

Let A and B be subsets of a (universal) set U.

(1) $(A \cup B)' = A' \cap B'$.

(2) $(A \cap B)' = A' \cup B'$.

Proof. We prove (1) and leave (2) as an exercise.

Suppose $x \in (A \cup B)'$. Then $x \in U$ and $x \notin A \cup B$. If $x \in A$, then $x \in A \cup B$, a contradiction. Thus, $x \notin A$. Said otherwise, $x \in A'$. By similar logic, $x \in B'$. Thus, $x \in A' \cap B'$. Thus, $(A \cup B)' \subset A' \cap B'$.

Suppose $x \in A' \cap B'$. Then $x \in A'$ and $x \in B'$. If $x \in A \cup B$, then either $x \in A$ or $x \in B$. The first case contradicts $x \in A'$ and the second contradicts $x \in B'$. We conclude that $x \notin A \cup B$, so $x \in (A \cup B)'$. Hence, $A' \cap B' \subset (A \cup B)'$, so combined with the first paragraph we have $(A \cup B)' = A' \cap B'$. □

Formally, a *relation* on sets X and Y is a subset of the Cartesian product $X \times Y$. A function is actually a special type of relation. Formally, a function is a relation $f \subset X \times Y$ satisfying if $(a, b), (a, c) \in f$ then $b = c$. (Here, X is the domain and Y is the codomain.)

For a set X, a *binary relation* is a subset $R \subset X \times X$. However, we often use different notation for binary relations. Instead of $(a, b) \in R$, we say $\sim$ is a binary relation and write $a \sim b$.

Definition: Equivalence relation, equivalence class

A binary relation $\sim$ on X is an *equivalence relation* on a set X if it satisfies the following:

- (reflexive property) $x \sim x$ for all $x \in X$;
- (symmetric property) if $x \sim y$, then $y \sim x$;
- (transitive property) if $x \sim y$ and $y \sim z$, then $x \sim z$.

The *equivalence class* of $x \in X$ is the set $[x] = \{y \in X : x \sim y\}$.

We will often write $x \sim y$ in place of $(x, y) \in R$.

Example (Congruence mod n)**.** Fix a positive integer n. We define a binary relation on $\mathbb{Z}$ by the rule $x \sim y$ if and only if $x - y$ is divisible by n. We claim that $\sim$ is an equivalence relation.

Let $x \in \mathbb{Z}$, then $x - x = 0$ and 0 is divisible by n. Thus $x \sim x$ for all $x \in X$, so the reflexive property holds.

Now suppose that $x, y \in \mathbb{Z}$ such that $x \sim y$. Then $x - y = nk$ for some integer k. Thus, $y - x = n(-k)$ and so $y - x$ is divisible by n and so $y \sim x$. Thus, the symmetric property holds.

Finally, suppose $x, y, z \in \mathbb{Z}$ such that $x \sim y$ and $y \sim z$. Then $x - y = nk$ and $y - z = n\ell$ for some integers k and ℓ. Now $x - z = (x - y) + (y - z) = nk + n\ell = n(k + \ell)$. Since $k + \ell$ is another integer (by closure under addition) then $x - z$ is divisible by n. Therefore $x \sim z$ so $\sim$ is transitive.

Thus, $\sim$ is an equivalence relation on $\mathbb{Z}$. We will often write $x \equiv y \mod n$ if $x \sim y$. (Note that by symmetry we can also write $y \equiv x \mod n$. If $x \in \mathbb{Z}$, then the equivalence class of $[x]$ is the set of integers that differ from x by a multiple of n.

Example (Congruence mod 5)**.** We have already checked that this is an equivalence relation. A complete set of equivalence classes are $[0], [1], [2], [3], [4]$. Clearly, none of these sets are the same since none of the *representatives* differ from one another by a multiple of 5. Moreover, *any other number* differs from exactly one of the above by a multiple of 5.

We can perform arithmetic on the equivalence classes above just as we would on the integers. For example, $3 + 4 = 7$ but $7 \equiv 2 \mod 5$ and so we write $3 + 4 \equiv 2 \mod 5$. Thus, $[3] + [4] = [2]$. Now observe that if we take any elements from either equivalence class, this equality still holds. For example, $13 \in [3]$ and $24 \in [4]$ but $13 + 24 = 37 \in [2]$. Moreover, this operation is associative. The element $[0]$ acts as an identity ($[0] + [k] = [k]$) and every element has an inverse ($[1] + [4] = [0]$ and $[2] + [3] = [0]$). Thus, the equivalence classes of the integers mod 5 is another example of a *group*.

Everything we've just done works perfectly well with 5 replaced by any positive number n.

Definition: Partition

A *partition* P of a set X is a collection of nonempty sets $X_1, X_2, \ldots$ such that $X_i \cap X_j = \emptyset$ for all $i \neq j$ and $\bigcup_k X_k = X$.

Theorem 3: Equivalence classes are partitions

Let $\sim$ be an equivalence relation on a set X.

(1) If $y \sim x$, then $[x] = [y]$.

(2) Given $x, y \in X$, $[x] = [y]$ or $[x] \cap [y] = \emptyset$ (equivalence classes are either equal or disjoint).

(3) The equivalence classes of X form a partition of X.

Proof. (1) Let $z \in [x]$, then $x \sim z$, so $y \sim z$ by transitivity. Thus, $z \in [y]$ and so $[x] \subset [y]$. Similarly, $[y] \subset [x]$ (exercise), so $[x] = [y]$.

(2) Choose $x, y \in X$ and suppose $[x] \cap [y] \neq \emptyset$. Then there exists $z \in [x] \cap [y]$. Thus, $z \in [x]$ and $z \in [y]$ so $x \sim z$ and $y \sim z$. By symmetry, $z \sim y$ and by transitivity, $x \sim y$. By (1), $[x] = [y]$.

(3) Since $x \sim x$, then $x \in [x]$ and so every element of X belongs to (at least) one equivalence class. By (2), we can choose $[x_1], [x_2], \ldots$, a complete set of (disjoint) equivalence classes such that $\bigcup_{k \in X} [x_k] = X$. □

We have shown that every equivalence relation determines a partition. But, in fact, the opposite is also true: every partition determines an equivalence relation. (Thus, there is a bijection between the equivalence relations and partitions of a given set X.)

Exercise. Given a partition $P = \{X_i\}$ of a set X, define a binary relation on X by the rule that $x \sim y$ if $x, y \in X_i$. Check that this rule indeed defines an equivalence relation.

2. Functions

We have seen one definition of a function in terms of relations. The following definition is independent of that.

Definition: Function, domain, codomain

Let A and B be sets. A *function* $f : A \to B$ is a rule that assigns each $a \in A$ a unique output, denoted $f(a) \in A$. The set A is called the *domain of* f and B the *codomain of* f.

So, there are two requirements for a *rule* $A \to B$ to be a *function*. First, every input must land in B. Secondly, each input must give a *unique* output. When a rule fails to follow either of these rules, we say it is not *well-defined*.

Example. (1) Define a rule $f : \mathbb{Z} \to \mathbb{Z}$ by $f(x) = 2x$. Then every output has a unique output, so f is a (well-defined) function.

(2) Define a rule $f : \mathbb{Z} \to \mathbb{Z}$ by $f(x) = x/2$. Then $f(1) \notin \mathbb{Z}$, so f is not well-defined.

(3) Define a rule $f : \mathbb{Q} \to QQ$ by $f(p/q) = p + q$. Then $f(1/2) = 3$ while $f(2/4) = 6$. So, while $1/2 = 2/4$, $f(1/2) \neq f(2/4)$. Hence, f is not well-defined.

Definition: Surjective, injective, bijective, permutation

Let $f : A \to B$ be a function. If $f(A) = B$, then f is said to be *surjective* (or *onto*). If for all $a_1, a_2 \in A$ such that $a_1 \neq a_2$ we have $f(a_1) \neq f(a_2)$, the f is said to be *injective* (or *one-to-one*). A function that is both injective and surjective is said to be *bijective*. A bijective function from a set to itself is a *permutation*.

Example. (1) The map $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(x) = x + 1$ is bijective.

Suppose $x, y \in \mathbb{Z}$ such that $f(x) = f(y)$. Then $x+1 = y+1$, so $x = y$ and thus f is injective. Now suppose $z \in \mathbb{Z}$. Then $f(z-1) = (z-1)+1 = z$, so f is surjective.

(2) The map $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(x) = 2x$ is injective but not surjective.

Suppose $x, y \in \mathbb{Z}$ such that $f(x) = f(y)$. Then $x+1 = y+1$, so $x = y$ and thus f is injective. But note that if $f(n) = 1$ then $2n = 1$, so $n = 1/2 \notin \mathbb{Z}$. Thus, f is not surjective.

(3) The map $f : \mathbb{Z} \to \mathbb{Z}$ defined by

$$f(x) = \begin{cases} x & \text{if } x \text{ is odd} \\ x/2 & \text{if } x \text{ is even} \end{cases}$$

is surjective but not injective.

Let $z \in \mathbb{Z}$. Then $2z$ is even and so $f(2z) = (2z)/2 = z$. Thus, f is surjective. However, $f(1) = 1 = f(2)$ and $1 \neq 2$, so f is not injective.

Recall that if $f : A \to B$ and $g : B \to C$ are functions, then the *composition* $g \circ f : A \to C$ is defined by the rule

$$(g \circ f)(a) = g(f(a)) \quad \text{for all } a \in A.$$

Theorem 4: Properties of composition

Let $f : A \to B$, $g : B \to C$, and $h : C \to D$ be functions.

(1) We have $(h \circ g) \circ f = h \circ (g \circ f)$. That is, function composition is associative.

(2) If f and g are injective, then $g \circ f$ is injective.

(3) If f and g are surjective, then $g \circ f$ is surjective.

(4) If f and g are bijective, then $g \circ f$ is bijective.

Proof. (1) By definition, two functions are equal if they agree at every element of their domain. Let $a \in A$. Then

$$((h \circ g) \circ f)(a) = (h \circ g)(f(a)) = h(g(f(a)) = h((g \circ f)(a)) = (h \circ (g \circ f))(a).$$

Hence, (1) holds.

(2) Suppose f and g are injective. We claim $g \circ f : A \to C$ is injective. Suppose there are $a, a' \in A$ such that $(g \circ f)(a) = (g \circ f)(a')$. Then $g(f(a)) = g(f(a'))$. By injectivity of g, $f(a) = f(a')$. By injectivity of f, $a = a'$. Hence, $g \circ f$ is injective.

(3) Suppose f and g are surjective. We claim $g \circ f : A \to C$ is surjective. Choose $c \in C$. Because g is surjective there exists $b \in B$ of c, so $g(b) = c$. Because f is surjective, there exists a preimage $a \in A$ of b, so $f(a) = b$. Then $(g \circ f)(a) = g(f(a)) = g(b) = c$. Thus, $g \circ f$ is surjective.

(4) This follows from properties (2) and (3). □

Property (1) from Theorem 4 says that composition of functions is associative while Property (4) says that the composition of two bijections is another bijection.

If $f : A \to B$ is a function of sets, then for any subset $C \subset B$, we define the *preimage of* C to be

$$f^{-1}(C) = \{a \in A : f(a) \in C\}.$$

Note in general that f^{-1} *is not* a function. For example, consider our previous example $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(x) = x$ if x is odd and $f(x) = x/2$ if x is even. Then $f^{-1}(1) = \{1, 2\}$ since $f(1) = f(2) = 1$. Hence, f^{-1} is not well-defined.

For every set A, there is a function $\text{id}_A : A \to A$, called the *identity map on* A, defined by

$$\text{id}_A(a) = a \text{ for all } a \in A.$$

Note that if $f : A \to B$ is any function (with domain A), then $f \circ \text{id}_A = f$. Similarly, if $g : B \to A$ is any function (with codomain A), then $\text{id}_A \circ g = g$.

Definition: Invertible function, inverse

A function $f : A \to B$ is *invertible* if there exists another function $g : B \to A$ such that $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$. In this case, the map g is called an *inverse of* f.

This next argument is one that we will see many times in this course. Recall that, generally, to prove to objects are unique we assume there are two objects with the requisite property and show that they are the same.

Theorem 5: Uniqueness of the inverse

The inverse of an invertible function is unique.

Proof. Suppose $f : A \to B$ is an invertible function with inverses g and h. Let $a \in A$. Then by the definition of an inverse function and associativity of composition,

$$h(a) = h(\mathrm{id}_A(a)) = h((f \circ g)(a)) = (h \circ (f \circ g))(a) = ((h \circ f) \circ g)(a) = \mathrm{id}_B(g(a)) = g(a).$$

Thus, $h = g$. □

Hence, if f is an invertible function, then there is no ambiguity in referring to *the* inverse of f, which we denote by f^{-1}.

Theorem 6: Invertibility is equivalent to bijectivitiy

A function $f : A \to B$ is invertible if and only if is bijective.

Proof. First assume f is invertible. By definition, there exists and inverse map f^{-1}. Let $a_1, a_2 \in A$ such that $f(a_1) = f(a_2)$. Applying f^{-1} to both sides gives $a_1 = a_2$, so f is injective. Let $b \in B$ and set $a = f^{-1}(b)$. Applying f to both sides gives $f(a) = b$, so f is surjective and hence bijective.

Now assume f is bijective. Let $b \in B$, then by surjectivity there is some $a \in A$ such that $f(a) = b$. By injectivity, a is the only element in A such that $f(a) = b$. Hence, the function $g : B \to A$ defined by $g(b) = a$ (the preimage of b) is well-defined. Then $g(f(a)) = g(b) = a$ and $f(g(b)) = f(a) = b$, so $g = f^{-1}$. □

The First Principle of Mathematical Induction

Let $S(n)$ be a statement about the integers for $n \in \mathbb{N}$ and suppose $S(n_0)$ is true for some integer n_0. If for all integers k with $k \geq n_0$, $S(k)$ true implies $S(k+1)$ is true, then $S(n)$ is true for all integers $n \geq n_0$. The statement $S(k)$ is referred to as the *inductive hypothesis*.

Example. Prove $10^{n+1} + 10^n + 1$ is divisible by 3.

Here we set $n_0 = 0$, so $S(0)$ is the statement that 12 is divisible by 3. That is true so assume $S(k)$ is true for some $k \geq 0$. That is, $10^{n+1} + 10^n + 1 = 3m$ for some integer m. Now $S(k+1)$ is the statement that $10^{k+2} + 10^{k+1} + 1$ is divisible by 3. By algebra we have

$$\begin{aligned} 10^{k+2} + 10^{k+1} + 1 &= 10(10^{k+1} + 10^k) + 1 \\ &= 10(3m - 1) + 1 \quad \text{by the inductive hypothesis} \\ &= 30m - 9 \\ &= 3(10m - 3). \end{aligned}$$

Thus, $S(k+1)$ is true and so by the (First) Principle of Mathematical Induction, $S(n)$ is true for all integers $n \geq 0$.

The Well-Ordering Principle

Every nonempty subset of $\mathbb{N}$ contains a least element.

Note that the First Principle of Mathematical Induction implies the Well-Ordering Principle.

The next theorem is an example of an *existence and uniqueness proof.* While this theorem is stated for integers but applies equally well to many other sets with an almost identical proof, as we will see later in the course.

Theorem 7: The Division Algorithm

Let $a, b \in \mathbb{Z}$ with $b > 0$. Then there exist unique integers q and r such that $a = bq + r$ with $0 \leq r < b$.

Proof. (Existence) Let $S = \{a - bk : k \in \mathbb{Z} \text{ and } a - bk \geq 0\} \subset \mathbb{N}$. If $0 \in S$, then b divides a so choose $q = a/b$ and $r = 0$. Assume $0 \notin S$. If $a \geq 0$, then $a = a - b \cdot 0 \in S$. If $a < 0$, then $a - b(2a) = a(1 - 2b) \in S$. In either case, $S \neq \emptyset$ and so we may apply the Well-Ordering Principle to find a least member of S, say r. By definition of S, there exists an integer q such that $r = a - bq$. That is, $a = bq + r$ and $r \geq 0$ by definition of S. We claim $r < b$. Suppose otherwise. Then

$$a - b(q+1) = a - bq - b = r - b > 0.$$

But then $a - b(q+1) \in S$ and $a - b(q+1) < r$, contradicting the choice of r.

(Uniqueness) Suppose there exist r, r', q, q' such that $a = bq + r$ and $a = bq' + r'$ with $0 \leq r, r' < b$. Then $bq+r = bq'+r'$. Assume $r' \geq r$. Then $b(q-q') = r-r'$ so b divides $r'-r$ and $0 \leq r'-r \leq r' < b$. Thus, $r' - r = 0$ so $r = r'$ and $q = q'$. □

We say a (nonzero) integer b divides an integer a if the (unique) r appearing the Division Algorithm is zero. We will often use the notation $b \mid a$ in place of b divides a. An integer d is a *common divisor* of integers a, b if $d \mid a$ and $d \mid b$. The *greatest common divisor* of $a, b \in \mathbb{Z}$ is a positive integer d that is a common divisor of a and b and for any other common divisor d' of a and b, $d' \mid d$. We write $d = \gcd(a, b)$.

The proof of the next theorem uses a similar idea as the Division Algorithm.

Theorem 8

Let a and b be nonzero integers. There exist integers r, s such that $\gcd(a, b) = ar + bs$. Furthermore, the gcd of a and b is unique.

Proof. Let $S = \{ar + bs : r, s \in \mathbb{Z} \text{ and } ar + bs > 0\}$. If $a, b > 0$, then $a(1) + b(1) > 0$. The other cases for a and b are similar. Thus, $S \neq 0$ and so by the Well-Ordering Principle S contains a least element, say $d = ar + bs$. We claim that $d = \gcd(a, b)$.

Write $a = dq + r'$ according to the Division Algorithm, where $0 \leq r' < d$. Suppose $r' > 0$. Then

$$r' = a - dq = a - (ar + bs)q = a(1 - rq) + b(-sq) \in S.$$

Since $r' < d$, this contradicts the minimality of d, so we must have $r' = 0$. That is $d \mid a$. A similar argument shows that $d \mid b$, so d is a common divisor of a and b.

Suppose d' is any common divisor of a and b. then $a = d'h$ and $b = d'k$ for integers h and k. Then

$$d = ar + bs = (d'h)r + (d'k)s = d'(hr + ks).$$

Thus, $d' \mid d$. It follows that d is the unique gcd of a and b. □

Note that the theorem only claims uniqueness of the gcd itself, and not the integers r and s. In fact, there are *infinitely* many integers that produce the gcd. The Division Algorithm does, however, give a methodology for producing a pair of such integers, as illustrated in the next example.

Example (Euclidean Algorithm)**.** Calculate $d = \gcd(471, 562)$ and find integers r and s such that $d = 471r + 562s$.

We repeatedly apply the division algorithm.

$$\begin{aligned}
562 &= 471 \cdot 1 + 91 \\
471 &= 91 \cdot 5 + 16 \\
91 &= 16 \cdot 5 + 11 \\
16 &= 11 \cdot 1 + 5 \\
11 &= 5 \cdot 2 + 1 \\
5 &= 1 \cdot 5 + 0.
\end{aligned}$$

Thus, $d = 1$. That is, 471 and 562 are relatively prime. Now by reversing:

$$\begin{aligned}
1 &= 11 + (-2) \cdot 5 = 11 + (-2)[16 + (-1) \cdot 11] \\
&= (3) \cdot 11 + (-2) \cdot 16 = (3) \cdot [91 + (-5) \cdot 16] + (-2) \cdot 16 \\
&= (3) \cdot 91 + (-17) \cdot 16 = (3) \cdot 91 + (-17) \cdot [471 + (-5) \cdot 91] \\
&= (88) \cdot 91 + (-17) \cdot 471 = (88) \cdot [562 + (-1) \cdot 471] + (-17) \cdot 471 \\
&= (88) \cdot 562 + (-105) \cdot 471
\end{aligned}$$

Hence, $r = -105$ and $s = 88$.

An integer $p > 1$ is *prime* if $x \mid p$ for $x > 1$ implies that $x = p$. The reason for omitting 1 as a prime becomes clear in light of the Fundamental Theorem of Arithmetic, proved below. In particular, we wish for every positive integer to have a *unique* decomposition in terms of prime numbers (e.g., $30 = 2 \cdot 3 \cdot 5$). But if we allow 1 to be prime then we lose uniqueness.

Lemma 9

Let $a, b \in \mathbb{Z}$ and p a prime number. If $p \mid ab$, then $p \mid a$ or $p \mid b$.

Proof. Suppose $p \nmid a$. We must show that $p \mid b$. Since p does not divide a, then $\gcd(a, p) = 1$. Thus, there exist integers r, s such that $ar + ps = 1$. Thus,

$$b = b \cdot 1 = b(ar + ps) = (ba)r + p(bs).$$

Since $p \mid ab$ and $p \mid p$, then p divides the right-hand sides. Consequently, p must divide the left-hand side, so $p \mid b$. □

Theorem 10: The Fundamental Theorem of Arithmetic

Let n be an integer such that $n > 1$. Then $n = p_1 p_2 \cdots p_k$ where the p_i are prime. Furthermore, if $n = q_1 q_2 \cdots q_\ell$ where the q_i are prime, then $k = \ell$ and the q_i are a rearrangement of the p_i.

Proof. (Existence) Let S be the set of all integers greater than 1 that cannot be written as the product of primes and suppose $S \neq \emptyset$. By the Well-Ordering principle, S has a least element, say a. If the only positive factors of a are a and 1, then a is prime, a contradiction. Hence we may assume $a = a_1 a_2$ where $1 < a_1, a_2 < a$. Since a is the minimal element of S, then $a_1, a_2 \notin S$. Hence, both a_1 and a_2 can be written as the product of primes:

$$a_1 = p_1 \cdots p_r$$

$$a_2 = q_1 \cdots q_s.$$

But then

$$a = a_1 a_2 = p_1 \cdots p_r q_1 \cdots q_s.$$

Thus, $a \notin S$, a contradiction.

(Uniqueness) The theorem is certainly true for $n = 2$ as 2 is prime. Suppose the theorem holds for all integers[1] m such that $1 \leq m < n$. Write

$$n = p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_\ell,$$

where $p_1 \leq p_2 \leq \cdots \leq p_k$ and $q_1 \leq q_2 \leq \cdots \leq q_\ell$. By the previous lemma, $p_1 \mid q_i$ for some q_i, $i = 1, \ldots, \ell$, and $q_1 \mid p_j$ for some $j = 1, \ldots, k$. The p_i and q_i are prime. so $p_1 = q_i$ and $q_1 = p_j$. Hence, $p_1 \leq p_j = q_1 \leq q_i = p_1$, so $p_1 = q_1$. Hence, we can divide that on both sides and obtain

$$n' = p_2 \cdots p_k = q_2 \cdots q_\ell.$$

Since $n' < n$, then by the (strong) inductive hypothesis, n' has a unique factorization. Hence, $k = \ell$ and $q_i = p_i$ for $i = 1, \ldots, k$. □

[1] In this theorem we are actually making use of *strong* induction.

Introduction to Groups

At its heart, Group Theory is the study of "symmetries" of objects. What symmetry means, and what sort of objects we will study, is a major focus of this course. Though everything we will do is directly related to symmetry in some way, the precise connection will sometimes be obscured by abstractness. This is intentional so that these ideas will apply as broadly as possible.

Most of this course is devoted to the study of *groups*. Though you may not know the definition of a group yet, you have seen them throughout most of your life. Consider the set of integers ($\mathbb{Z}$) and the operation of addition. Some observations we can make, that you know well already, are

- The sum of two integers is another integer.
- The operation of addition is associative.
- The number 0 acts as an identity, so that $0 + n = n$ for all $n \in \mathbb{Z}$.
- For every integer n, there is another integer $-n$, such that $n + (-n) = 0$.

These properties collectively make $\mathbb{Z}$ with this operation into a group[1]. It would be natural to ask whether we could choose a different operation for the integers, say multiplication. You should think about whether the above properties still hold if we replace addition by multiplication,

There are lots of other examples of groups which you have seen, including:

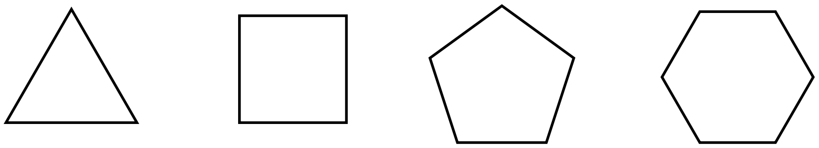
- the rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, or the complex numbers $\mathbb{C}$ under addition;
- $n \times n$ matrices (over any of the fields mentioned above) under matrix addition;
- functions $f : \mathbb{R} \to \mathbb{R}$ under pointwise addition.

We will study all of these examples, and more, but this is not where we will begin the course. Instead, we will make the connection between groups and symmetries very clear through the study of dihedral groups.

[1] You may also note that the operation of addition on the integers is commutative, that is, $m + n = n + m$ for all $m, n \in \mathbb{Z}$. This is sort of a bonus, called the *abelian* property, which we will not require for groups

1. Symmetries of an equilateral triangle

A regular polygon is one whose sides all have equal length[2]. We will refer to such an object as a regular n-gon, where n denotes the number of sides. (Throughout, $n \geq 3$ for somewhat obvious reasons.) Here are some examples of regular n-gons:



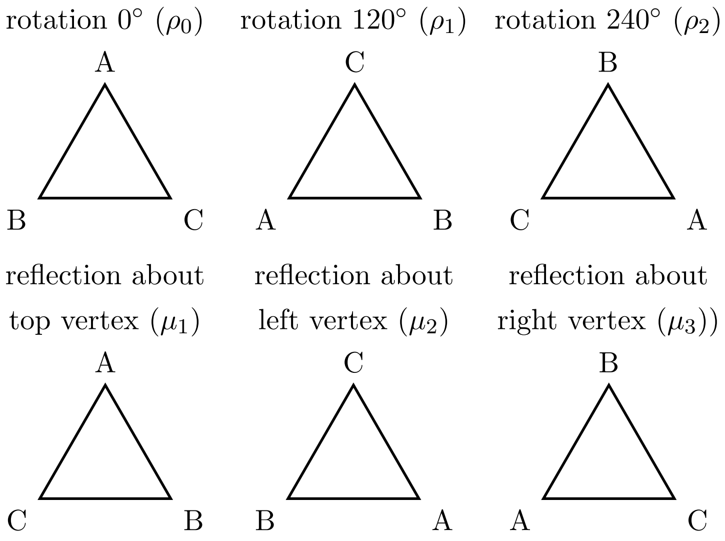
equilateral triangle   square   regular pentagon   regular hexagon

Definition: Symmetry

A *symmetry* of a regular n-gon is a permutation of the vertices preserving adjacency.

A symmetry is an example of a *rigid motion* in the plane ($\mathbb{R}^2$). However, there are other types of rigid motions, including translations, that we do not consider. To cut down on the terminology a little bit, we will first study symmetries of an equilateral triangle.

There are a total of six symmetries: three reflections and three rotations (counterclockwise). (Note that a rotation by 360° is the same as a rotation by 0°.) These symmetries do not change the object itself, but they do permute the vertices. We record the results of these below.



rotation 0° (ρ_0)   rotation 120° (ρ_1)   rotation 240° (ρ_2)

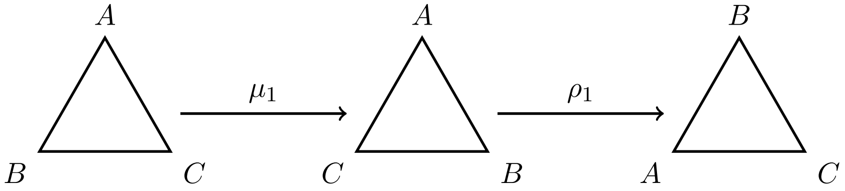
reflection about top vertex (μ_1)   reflection about left vertex (μ_2)   reflection about right vertex (μ_3))

[2] The definition of a regular polygon implies that all angles within the regular polygon are also congruent. You can prove this using congruent triangles.

The set of these symmetries is called[3] D_3. We can think of each one as a function and compose them just as we would compose two functions. Remember that function composition is from right-to-left.

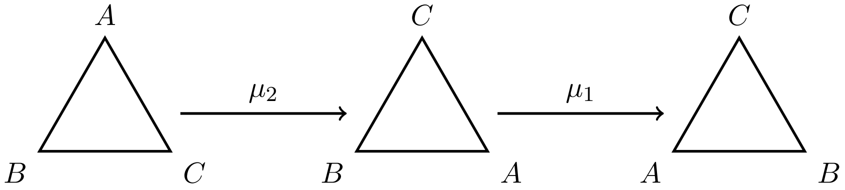
Example. Compute each of the following compositions.

(1) $\rho_1 \circ \mu_1$



Hence, $\rho_1 \circ \mu_1 = \mu_3$.

(2) $\mu_1 \circ \mu_2$



Hence, $\mu_1 \circ \mu_2 = \rho_1$.

We can check all 36 compositions and record them in a multiplication-like table. More formally, this is called a *Cayley table*. We need to be careful, though, because the element in row a and column b is $a \circ b$, but this means we perform b *first* and a *second*.

	ρ_0	ρ_1	ρ_2	μ_1	μ_2	μ_3
ρ_0	ρ_0	ρ_1	ρ_2	μ_1	μ_2	μ_3
ρ_1	ρ_1	ρ_2	ρ_0	μ_3	μ_1	μ_2
ρ_2	ρ_2	ρ_0	ρ_1	μ_2	μ_3	μ_1
μ_1	μ_1	μ_2	μ_3	ρ_0	ρ_1	ρ_2
μ_2	μ_2	μ_3	μ_1	ρ_2	ρ_0	ρ_1
μ_3	μ_3	μ_1	μ_2	ρ_1	ρ_2	ρ_0

[3] This group is sometimes referred to as D_6 (because it has 6 elements). Always pay attention to conventions when consulting a new text.

We make a few observations about the table above that reveal the defining properties of a group.

- The system is closed (the composition of two symmetries is another symmetry).
- The operation is associative (this is true for function composition in general).
- There is an identity element (the "do nothing" symmetry ρ_0).
- Every symmetry has an inverse symmetry that reverses it.

There are some other observations we can make that will be true for symmetries in general.

- The composition of two rotations is a rotation.
- The composition of two reflections is a rotation.
- The composition of a reflection and a rotation is a reflection. This is true in either order, but the order determines the resulting reflection.

Example. (1) Identify the inverse of each element in D_3.

- The inverse of ρ_0 (the identity) is itself.
- Each reflection is its own inverse.
- The rotations ρ_1 and ρ_2 are inverses of one another.

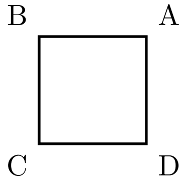
(2) For each element in D_3, determine how many times one would need to compose it with itself in order to obtain the identity. That is, for a in D_3, how many a's do we need to get $a \circ a \circ \cdots \circ a = \rho_0$.

- The identity need only be composed with itself once.
- For any reflection μ, we have $\mu \circ \mu = \rho_0$.
- The rotations ρ_1 and ρ_2 need three: $\rho_1 \circ \rho_1 \circ \rho_1 = \rho_3$.

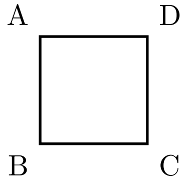
We call this the *order* of the element.

Our analysis and methods above extend easily to the symmetry group of a square, denoted D_4. There are a total of eight symmetries of a square: four rotations (including the trivial rotation) and four reflections.

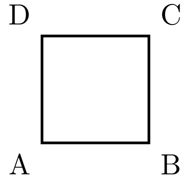
rotation 0° (ρ_0)



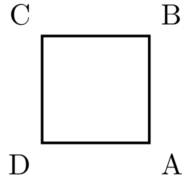
rotation 90° (ρ_1)



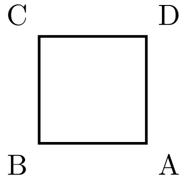
rotation 180° (ρ_2)



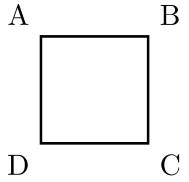
rotation 270° (ρ_3)



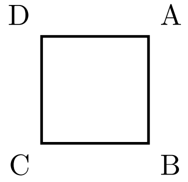
reflection about horizontal axis (μ_1)



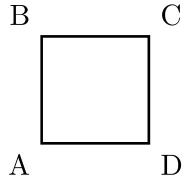
reflection about vertical axis (μ_2)



reflection about positive diagonal (μ_3)



reflection about negative diagonal (μ_4)



We will construct the Cayley table for this group.

	ρ_0	ρ_1	ρ_2	ρ_3	μ_1	μ_2	μ_3	μ_4
ρ_0	ρ_0	ρ_1	ρ_2	ρ_3	μ_1	μ_2	μ_3	μ_4
ρ_1	ρ_1	ρ_2	ρ_3	ρ_4	μ_3	μ_4	μ_2	μ_1
ρ_2	ρ_2	ρ_3	ρ_0	ρ_1	μ_2	μ_1	μ_4	μ_3
ρ_3	ρ_3	ρ_0	ρ_1	ρ_2	μ_4	μ_3	μ_1	μ_2
μ_1	μ_1	μ_4	μ_2	μ_3	ρ_0	ρ_2	ρ_3	ρ_1
μ_2	μ_2	μ_3	μ_1	μ_4	ρ_2	ρ_0	ρ_1	ρ_3
μ_3	μ_3	μ_1	μ_4	μ_2	ρ_1	ρ_3	ρ_0	ρ_2
μ_4	μ_4	μ_2	μ_3	μ_1	ρ_3	ρ_1	ρ_2	ρ_0

We notice that our observations from above for D_3 all seem to hold for D_4 as well. That is, D_4 satisfies the properties of a group.

Example. (1) Find the inverse of each element in D_4.

The inverse of a reflection μ_i is itself. The inverse of ρ_0 is itself, as is the inverse of ρ_2. Finally, ρ_1 and ρ_3 are inverses of each other.

(2) Find the order of each element in D_4.

The identity has order 1. The rotation ρ_1 and ρ_3 have order 4. The rotation ρ_2 has order 2. Each reflection has order 2.

(3) Find the elements of D_4 that commute with all other elements. (That is, find those $a \in D_4$ such that $a \circ b = b \circ a$ for all $b \in D_4$).

We can see from the Cayley table that the only such elements are ρ_0 and ρ_2.

2. The group D_n

Here we generalize our discussion on D_3 to arbitrary regular polygons. Throughout this section, $n \geq 3$. The set of symmetries of a regular n-gon, along with the operation of composition of symmetries, is known as the *dihedral group* and is denoted D_n. Our goal here will be to develop some consistent notation for D_n, along with basic facts about this group.

Theorem 1: Order of D_n

There are $2n$ symmetries of a regular n-gon.

Proof. First, we show that there are *at least* $2n$ symmetries ($|D_n| \geq 2n$). There are n rotations (counterclockwise) by $\left(\frac{360 \cdot k}{n}\right)^\circ$, $k = 0, \ldots, n-1$. These are all clearly distinct from each other. Also note that there are no fixed vertices by these rotations. In fact, the only fixed point on the n-gon is the center.

For reflections, we have to split into two cases. If there are an odd number of vertices, then for each vertex there is a reflection through a point and the midpoint of the side opposite that vertex. If there are even number of vertices, the there are $n/2$ reflections through opposite vertices and $n/2$ reflections through opposite midpoints on sides. In either case there are two points fixed by the reflection (the intersection points between the lines of reflection and the n-gon). Thus, no reflection is a rotation.

It is left to show that there are *at most* $2n$ symmetries ($|D_n| \leq 2n$). Choose two adjacent vertices, A and B in the regular n-gon. Let $g \in D_n$. There are n choices for the position of $g(A)$. Since g must preserve adjacency of vertices, then $g(A)$ and $g(B)$ are adjacent. Thus, given the position of $g(A)$, there are 2 choices for the position of $g(B)$. Also recall that g must preserve distances between points. Thus, given any other point P, the position of $g(P)$ is determined by the choice of location for $g(A)$ and $g(B)$. This gives $2n$ *rearrangements of the vertices.* By the Lemma **??**, each such rearrangement is a rigid motion and hence $|D_n| = 2n$. □

Next we show how to express D_n in more conventional group-theoretic notation. That is, we will replace composition with multiplication. The ability to switch between different notations will be an important skill to develop through this course.

Let $r \in D_n$ denote the rotation by $(360/n)^\circ$ and let $s \in D_n$ denote *any reflection through a vertex*. Let 1 be the identity in D_n. Note that $r^n = 1$ and $s^2 = 1$. Then it follows that the $2n$ elements of D_n are

$$\{1, r, r^2, \ldots, r^{n-1}, s, rs, r^2s, \ldots, r^{n-1}s\}$$

The elements $1, r, r^2, \ldots, r^{n-1}$ are the rotations while the remaining elements $s, rs, r^2s, \ldots, r^{n-1}s$ are reflections.

Note that these elements are all distinct from each other.

- Clearly the rotations are all distinct.
- Suppose $r^k s = r^\ell s$ for some k, ℓ. Then multiplying both sides by s gives $r^k = r^\ell$.
- Suppose that $r^k s = r^\ell$, then $s = r^{\ell - k}$, which implies that s is a rotation, a contradiction.

The fact that D_n has n reflections and n rotations, and that the r^i are clearly rotations, now implies that the $r^k s$ are reflections for all k.

We will now prove a critical defining relation in D_n.

Theorem 2: Relation on D_n

In D_n, $srs = r^{-1}$.

Proof. Let A be a vertex fixed by s with adjacent vertices B, B' where B is counterclockwise of A. Then

$$(srs)(A) = (sr)(s(A)) = s(r(A)) = s(B) = B' \quad \text{and} \quad r^{-1}(A) = B.$$
$$(srs)(B) = (sr(s(B)) = s(r(B')) = s(A) = A \quad \text{and} \quad r^{-1}(B) = A.$$

Since the relation holds for an arbitrary pair of adjacent vertices, it holds for all vertices. □

Example. (1) Write the Cayley table of D_3 using the r, s notation above.

	1	r	r^2	s	rs	r^2s
1	1	r	r^2	s	rs	r^2s
r	r	r^2	1	rs	r^2s	s
r^2	r^2	1	r	r^2s	s	rs
s	s	r^2s	rs	1	r^2	r
rs	rs	s	r^2s	r	1	r^2
r^2s	r^2s	rs	s	r^2	r	1

(2) Write the Cayley table of D_4 using the r, s notation above.

	1	r	r^2	r^3	s	rs	r^2s	r^3s
1	1	r	r^2	r^3	s	rs	r^2s	r^3s
r	r	r^2	r^3	1	rs	r^2s	r^3s	s
r^2	r^2	r^3	1	r	r^2s	r^3s	s	rs
r^3	r^3	1	r	r^2	r^3s	s	rs	r^2s
s	s	r^3s	r^2s	rs	1	r^3	r^2	r
rs	rs	s	r^3s	r^2s	r	1	r^3	r^2
r^2s	r^2s	rs	s	r^3s	r^2	r	1	r^3
r^3s	r^3s	r^2s	rs	s	r^3	r^2	r	r^4

3. The definition of a group

We now have sufficient setup to define the notion of a group. We will consider a few examples before returning to the dihedral group.

Recall that function f between sets X and Y, written $f : X \to Y$, is a rule such that for each $x \in X$, there is a unique output $f(x) \in Y$. This means that if $x_1 = x_2$ in X, then $f(x_1) = f(x_2)$.

Definition: Binary operation

A *binary operation* on a set S is a function $f : S \times S \to S$.

Often times we ignore the function itself, we will use operation notation. So instead of $f(s, s')$ we might write $s \cdot s'$. If $\cdot$ is a binary operation on S, then we say that S is *closed under* $\cdot$.

Example. The following are examples (or non-examples) of binary operations.

- Composition of symmetries (more generally, function composition) is a binary operation.
- Addition and multiplication of integers are both binary operations.
- Subtraction of integers is a binary operation. However, subtraction on $\mathbb{N}$ is *not* a binary operation because $1 - 2 \notin \mathbb{N}$.
- Addition and multiplication of $n \times n$ matrices are both binary operations.
- Division of real numbers is not a binary operation since $r/0$ is not defined.

Definition: Group

A *group* is a pair $(G, \cdot)$ with G a set and $\cdot$ a binary operation on G satisfying
(1) for any $a, b, c \in G$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (associativity);
(2) there exists $e \in G$ such that $a \cdot e = a = e \cdot a$ for all $a \in G$ (existence of an identity element);
(3) for each $a \in G$ there exists an element $b \in G$ such that $a \cdot b = e = b \cdot a$ (closure under inverses).

To verify that a set and operation are a group actually requires checking four axioms, since we must verify that $\cdot$ is a binary operation.

Definition: Abelian group

A group $(G, \cdot)$ such that $a \cdot b = b \cdot a$ for all $a, b \in G$ is said to be *abelian.*

Example. The following are examples of groups with which you are already familiar.

- $(\mathbb{Z}, +)$ is a group (see the introduction). Similarly, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, and $(\mathbb{C}, +)$ are groups.
- $(\mathbb{N}, +)$ is *not* a group because the additive inverse of a positive integer is a negative integer.
- $(\mathbb{Z}, \cdot)$ is *not* a group because the multiplicative inverse of an integer is typically not an integer.
- $(\mathbb{R}^\times, \cdot)$ is a group where $\mathbb{R}^\times = \mathbb{R}\backslash\{0\}$ and $\cdot$ is usual multiplication[4]. Similarly, we can define $\mathbb{Q}^\times$ and $\mathbb{C}^\times$.

  The product of two nonzero real numbers is a nonzero real number, so multiplication is a binary operation on $\mathbb{R}^\times$. Multiplication is associative (this is well-known). The elements $1 \in \mathbb{R}^\times$ is an identity element. Finally, the (multiplicative) inverse of nonzero real number r is a nonzero real number $1/r$.
- $(M_2(\mathbb{R}), +)$ is a group where $M_2(\mathbb{R})$ denotes the set of 2×2 matrices with entries in $\mathbb{R}$ and $+$ is the operation of matrix addition. (Group properties established in linear algebra.)
- $(\mathrm{GL}_2(\mathbb{R}), \cdot)$ where $\mathrm{GL}_2(\mathbb{R}) \subset M_2(\mathbb{R})$ denotes the set of 2×2 *invertible* matrices and $\cdot$ is matrix multiplication.

  Note that invertibility of a (square) matrix A is equivalent to $\det(A) \neq 0$. If $A, B \in \mathrm{GL}_2(\mathbb{R})$, then $\det(AB) = \det(A)\det(B) \neq 0$. Hence, $\mathrm{GL}_2(\mathbb{R})$ is closed under multiplication. Matrix multiplication is associative (linear algebra). The identity matrix is invertible and acts as an identity on $\mathrm{GL}_2(\mathbb{R})$. Finally, the inverse of an invertible matrix is invertible.

  In this example, one can replace 2 with any positive integer (note that $\mathrm{GL}_1(\mathbb{R}) = \mathbb{C}^\times$) and one can replace $\mathbb{R}$ with any field (e.g., $\mathbb{Q}$ or $\mathbb{C}$).

[4] This is sometimes denoted $\mathbb{R}^*$.

Example (The group $\mathcal{S}_3$)**.** A bijection from a set to itself is called a *permutation*. Let $X = \{1, 2, 3\}$. The function $f : X \to X$ given by $f(1) = 2$, $f(2) = 1$, and $f(3) = 3$ is bijective, and hence a permutation of X. In total, there are six permutations of X and we denote the set of these by $\mathcal{S}_3$:

$$\begin{array}{cccccc} \begin{pmatrix}1 & 2 & 3\\ 1 & 2 & 3\end{pmatrix} & \begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{pmatrix} & \begin{pmatrix}1 & 2 & 3\\ 3 & 1 & 2\end{pmatrix} & \begin{pmatrix}1 & 2 & 3\\ 1 & 3 & 2\end{pmatrix} & \begin{pmatrix}1 & 2 & 3\\ 3 & 2 & 1\end{pmatrix} & \begin{pmatrix}2 & 1 & 3\\ 1 & 2 & 3\end{pmatrix} \\ (1) & (123) & (132) & (23) & (13) & (12) \end{array}$$

The composition of two bijective functions is another bijective functions, so $\mathcal{S}_3$ is closed under function composition, and function composition is also associative. There is an identity element (the identity function id_X) and every bijective function is invertible. (And the inverse of a bijection is a bijection). Thus, $\mathcal{S}_3$ is a group under function composition. Now we construct a Cayley table for $\mathcal{S}_3$. This should be compared to the Cayley table for D_3.

	(1)	(123)	(132)	(23)	(13)	(12)
(1)	(1)	(123)	(132)	(23)	(13)	(12)
(123)	(123)	(132)	(1)	(12)	(23)	(13)
(132)	(132)	(1)	(123)	(13)	(12)	(23)
(23)	(23)	(13)	(12)	(1)	(123)	(132)
(13)	(13)	(12)	(23)	(132)	(1)	(123)
(12)	(12)	(23)	(13)	(123)	(132)	(1)

Example (The group $\mathcal{F}$)**.** Let $\mathcal{F}$ denote the set of functions $f : \mathbb{R} \to \mathbb{R}$ (that is, functions whose domain and codomain are both $\mathbb{R}$). If $f, g \in \mathcal{F}$, then we define the function $f + g : \mathbb{R} \to \mathbb{R}$ by

$$(f + g)(x) = f(x) + g(x) \text{ for all } x \in \mathbb{R}.$$

(This is called *pointwise addition* of functions.) It is clear that this is a binary operation since f and g are both defined at all points in $\mathbb{R}$ and the output must be in $\mathbb{R}$.

The identity element is the *zero function* $z \in \mathcal{F}$ defined by $z(x) = 0$ for all $x \in \mathbb{R}$. To see this, let $f \in \mathcal{F}$ and $x \in \mathbb{R}$. Then

$$(z + f)(x) = z(x) + f(x) = 0 + f(x) = f(x) = f(x) + 0 = f(x) + z(x) = (f + z)(x).$$

Similarly, for $f \in \mathcal{F}$ define $(-f) \in \mathcal{F}$ by $(-f)(x) = -(f(x))$ for all $x \in \mathbb{R}$. Then for $x \in \mathbb{R}$,

$$(f + (-f))(x) = f(x) + (-f)(x) = f(x) - (f(x)) = 0 = -(f(x)) + f(x) = ((-f) + f)(x).$$

Hence, $-f$ is the inverse of f.

It is left to verify that pointwise addition is associative. Let $f, g, h \in \mathcal{F}$ and let $x \in \mathbb{R}$. Then

$$\begin{aligned} ((f + g) + h)(x) &= (f + g)(x) + h(x) = (f(x) + g(x)) + h(x) \\ &= f(x) + (g(x) + h(x)) \quad \text{(by associativity of addition on } \mathbb{R}) \\ &= f(x) + (g + h)(x) = (f + (g + h))(x). \end{aligned}$$

4. The group of rotations, addition mod n, and roots of unity

Let R be the subset of D_3 consisting of rotations (though this also works for D_4 and in fact any D_n as we will see). Since the composition of two rotations is another rotation, then R is closed under composition. Furthermore,

- composition is associative on R (because it is associative on D_3),
- the identity element is a rotation (and hence in R),
- the inverse of any rotation is a rotation (R is closed under inverses).

Consequently, R is a group itself under composition of symmetries. We have now just seen our first example of a *group within a group*.

Definition: Subgroup

A *subgroup* of a group G is a subset H that is a group with respect to the operation associated to G.

Example. The following are further examples of subgroups.

(1) Let G be a group. Then G is a subgroup of itself. If $e \in G$ is the identity element, then $\{e\}$ is a subgroup called the *trivial subgroup*. A subgroup of G that is not G and not the trivial subgroup is called *proper*.

(2) $2\mathbb{Z}$, the set of even numbers, is a subgroup of $\mathbb{Z}$ (under addition). The set of odd numbers is not a subgroup.

(3) $\mathbb{R}^\times$ is a group under multiplication and $\mathbb{Q}^\times$ is a subgroup.

(4) $\mathrm{GL}_2(\mathbb{R})$ is a subset of $M_2(\mathbb{R})$ but not a subgroup because $M_2(\mathbb{R})$ is a group under matrix addition and $\mathrm{GL}_2(\mathbb{R})$ a group under matrix multiplication.

Let $R_n = \{1, r, \ldots, r^{n-1}\}$ denote the subgroup of rotations of D_n (this is a subgroup by the same logic as above). Recall that $r^n = 1$. To compute a product of rotations, so that the result is an element of R_n, we must use *modular arithmetic*. For example, in D_4, $r^2 \cdot r^3 = r$ because

$$r^5 = r^4 \cdot r = 1 \cdot r = r.$$

Example. Let R_{12} be the subgroup of rotations of D_{12} (symmetries of a dodecagon)[5]. each of the following.

(1) Compute $r^4 \cdot r^3$

$$r^4 \cdot r^3 = r^7$$

(2) Compute $r^8 \cdot r^7$

$$r^8 \cdot r^7 = r^{15} = r^{15-12} = r^3.$$

(3) Compute $r^3 \cdot r^9$

$$r^3 \cdot r^9 = r^{12} = r^{12-12} = r^0.$$

(4) Give an explicit formula for the inverse of a rotation r^k in R_{12}.

The inverse of a rotation r^k is $r^{-k} = r^{12-k}$. In general, in R_n, the inverse of r^k is r^{n-k}.

You should notice that doing computations in R_{12} is not altogether different than doing arithmetic on a clock[6].

The (sub)group R_n goes under many guises and has many different names. We will introduce two of them here.

Fix a positive integer n. Recall that congruence mod n is the equivalence relation[7] $\sim$ on $\mathbb{Z}$ defined by $x \sim y$ if $x - y$ is divisible by n. We typically write $x \equiv y \mod n$ in place of $x \sim y$. Let $\mathbb{Z}_n$ denote the collection of distinct equivalence classes under congruence mod n. So

$$\mathbb{Z}_n = \{[0], [1], \ldots, [n-1]\}.$$

(Recall that, since the equivalence classes partition $\mathbb{Z}$, every integer belongs to exactly one of the above equivalence classes and two given equivalence classes are either equal or disjoint.) We will show that $\mathbb{Z}_n$ is a group under addition of equivalence classes (addition mod n).

[5] Not to be confused with the hip-hop group D12 featuring Eminem and Proof, amongst others.

[6] My daughter can do this, and she's six, so you should certainly be able to do it.

[7] See preliminaries

Example. Let n be a positive integer. Define a rule $f : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n$ by $f([a],[b]) = [a+b]$. We claim that f is a function (i.e., is well-defined). This shows that addition mod n is a binary operation on $\mathbb{Z}_n$.

By definition, the output of f is an equivalence class in $\mathbb{Z}_n$. We must show that that each input has a *unique* output. In other words, we need to show that the output is *independent of the choice of equivalence class representative.*

Let $a, a', b, b' \in \mathbb{Z}$ such that $[a] = [a']$ and $[b] = [b']$. We claim $[a] + [b] = [a'] + [b']$. Since $[a] = [a']$, then $a' = a + kn$. Similarly, $b' = b + \ell n$. Thus,

$$a' + b' = (a + kn) + (b + \ell n) = (a + b) + (k + \ell)n.$$

Therefore, $a' + b' \in [a + b]$ so $[a + b] = [a' + b']$.

Example (The group $\mathbb{Z}_n$)**.** By the previous example, addition mod n is a binary operation on $\mathbb{Z}_n$. We now verify the remaining properties of a group.

Let $[a], [b], [c] \in \mathbb{Z}_n$. Then by definition of addition mod n (and because addition of integers is associative),

$$([a] + [b]) + [c] = [a + b] + [c] = [(a + b) + c] = [a + (b + c)] = [a] + [b + c] = [a] + ([b] + [c]).$$

Hence, addition mod n is associative.

We claim that $[0]$ (the equivalence class of 0) is the identity element in $\mathbb{Z}_n$ (under addition mod n). Let $[x] \in \mathbb{Z}_n$. Then

$$[x] + [0] = [x + 0] = [x] = [0 + x] = [0] + [x].$$

This proves the claim.

Finally, for $[z] \in \mathbb{Z}_n$, we claim that $[-z]$ is the (additive) inverse of $[z]$. We have

$$[z] + [-z] = [z + (-z)] = [0] = [(-z) + z] = [-z] + [z],$$

as claimed. Thus, $\mathbb{Z}_n$ is a group under addition mod n.

Finally, note that for $[x], [y] \in \mathbb{Z}_n$, then by commutative of addition of integers,

$$[x] + [y] = [x + y] = [y + x] = [y] + [x].$$

Hence, $\mathbb{Z}_n$ is an *abelian* group under addition mod n.

Exercise. Show that multiplication mod n is a binary operation but that $\mathbb{Z}_n$ is not a group under this operation.

It should (hopefully) be clear at this point the connection between addition mod n and the rotation subgroup of D_n. Just as [3] and [7] represent the same element in $\mathbb{Z}_4$, so do the rotations r^3 and r^7 in R_4. In fact, the relationship between these two groups can be made even more precise. First, we make a Cayley table for both groups.

R_4	1	r	r^2	r^3
1	1	r	r^2	r^3
r	r	r^2	r^3	1
r^2	r^2	r^3	1	r
r^3	r^3	1	r	r^2

$\mathbb{Z}_4$	$[0]$	$[1]$	$[2]$	$[3]$
$[0]$	$[0]$	$[1]$	$[2]$	$[3]$
$[1]$	$[1]$	$[2]$	$[3]$	$[0]$
$[2]$	$[2]$	$[3]$	$[0]$	$[1]$
$[3]$	$[3]$	$[0]$	$[1]$	$[2]$

Define a map $\phi : R_4 \to \mathbb{Z}_4$ by $r^k \mapsto [k]$. Using this map to substitute elements in the R_4 table (for example, r^3 is replaced by $[3]$) produces the $\mathbb{Z}_4$ table. When such a map exists between two groups and their corresponding Cayley tables, we say the groups are *isomorphic*. Essentially, they are the same group but represented in different ways. Next we will see another instance of this.

To start, note that $\mathbb{C}$ *is not* a group under multiplication. However, we can *fix* $\mathbb{C}$ so that it is a group under multiplication.

Example (The group $\mathbb{C}^\times$)**.** The elements of $\mathbb{C}^\times = \mathbb{C}\backslash\{0\}$ are complex numbers $a + bi$, $a, b \in \mathbb{R}$, such that a and b are not both zero (recall that $i = \sqrt{-1}$). Unlike with $\mathbb{Q}$ and $\mathbb{R}$, it is somewhat nontrivial to verify that $\mathbb{C}^\times$ is closed under multiplication[8].

Let $a + bi, c + di \in \mathbb{C}^\times$. Then

$$(a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i.$$

Suppose $d = 0$. Then $c \neq 0$ and $(a + bi)(c + di) = (ac) + (bc)i$. Since a and b cannot both be zero, then one of ac, bc is nonzero and hence the product is nonzero. Now we may assume $d \neq 0$. If $a \neq 0$, then $c = bd/a$ and so

$$ad + bc = ad + b(bd/a) = (d/a)(a^2 + b^2).$$

Since $a^2 + b^2 > 0$, then $ad + bc \neq 0$.

Associativity of multiplication is easy but tedious. The identity is $1 = 1 + 0i$. The inverse of $a + bi \in \mathbb{C}^\times$ is

$$(a + bi)^{-1} = \frac{a - bi}{a^2 + b^2}.$$

The element $a - bi$ is called the *complex conjugate* of $a + bi$.

[8] Notation note: some people write $\mathbb{C}^*$ instead of $\mathbb{C}^\times$ and it means the same thing.

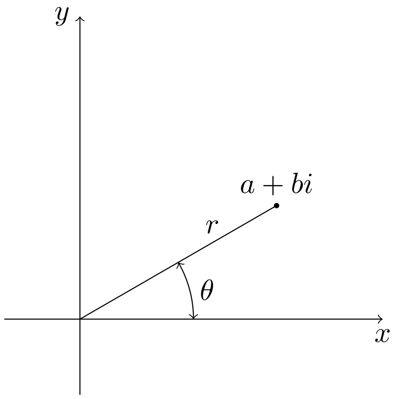
The group $\mathbb{C}^\times$ has lots of subgroups. One you might be familiar with is $H = \{1, -1, i, -i\}$ (also known as the *group of fourth roots of unity*). We should compare this group to R_4 by constructing their Cayley tables and writing down a map between them:

R_4	1	r	r^2	r^3
1	1	r	r^2	r^3
r	r	r^2	r^3	1
r^2	r^2	r^3	1	r
r^3	r^3	1	r	r^2

H	1	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

Again there is a map between the two groups so that, when applied to the Cayley table for R_4, produces the Cayley table for H. Specifically, that map is

We will explore roots of unity more, along with their connection the groups $\mathbb{Z}_n$. Given a complex number $z = a + bi$, the *complex conjugate* of z is $\overline{z} = a - bi$. The *modulus* of z is $|z| = \sqrt{a^2 + b^2}$. The complex number z can be represented on the Cartesian plane in polar form (r, θ) where:



$$a = r\cos\theta \quad \text{and} \quad b = r\sin\theta.$$

Note that we require $0 \leq \theta < 2\pi$ to ensure that this representation is well-defined. Then

$$z = r(\cos\theta + i\sin\theta).$$

(Your book uses the abbreviation cisθ for $\cos\theta + i\sin\theta$ but I will avoid this.)

Theorem 3: DeMoivre's Theorem

Let $z = r(\cos\theta + i\sin\theta)$. Then for any positive integer n,

$$z^n = r^n\left(\cos(n\theta) + i\sin(n\theta)\right).$$

Definition: Roots of unity

Let n be a positive integer. The *nth roots of unity* are the complex numbers of the form

$$z = \cos(2k\pi/n) + i\sin(2k\pi/n)$$

for $k = 0, 1, \ldots, n-1$. An nth root of unity z is *primitive* if $z^m \neq 1$ for $m = 1, \ldots, n-1$.

Example. The complex number i is a fourth root of unity since $i^4 = 1$. Moreover, it is a *primitive* fourth root of unity since $i^1 = i, i^2 = -1$, and $i^3 = -i$. We see from this that i in fact *generates* all of the fourth roots of unity. In the form above, we can identify the fourth roots of unity as

$$\begin{aligned} 1 &= \cos(0) + i\sin(0) & -1 &= \cos(\pi) + i\sin(\pi) \\ i &= \cos(\pi/2) + i\sin(\pi/2) & -i &= \cos(3\pi/2) + i\sin(3\pi/2) \end{aligned}$$

This idea of *generating* a group will be very important later when studying cyclic groups. Note that the group R_n is generated by a single rotation r (of degree $2\pi/n$). Similarly, $\mathbb{Z}_n$ is generated by $[1]$ since added that element to itself enough times produces all the elements[9] in $\mathbb{Z}_n$.

Example (The group C_n)**.** Let n be a positive integer and let ω be a primitive nth root of unity.

We first claim that the set of *all* nth roots of unity is $C_n = \{1, \omega, \omega^2, \ldots, \omega^{n-1}\}$. By definition, there are n nth roots of unity, so to justify this claim we need only show that no two of these are the same. Suppose $\omega^k = \omega^\ell$ for distinct powers k and ℓ. Without loss of generality, assume $k > \ell$. Then $\omega^{k-\ell} = 1$. But $1 < k - \ell < n$, contradicting the definition of primitive. This proves the claim.

Clearly C_n is a subset of $\mathbb{C}^\times$. It is closed under multiplication (since $\omega^i \cdot \omega^j = \omega^{i+j} \in C_n$). The operation (multiplication of complex numbers) is associative because it is associative on $\mathbb{C}^\times$. Obviously, C_n contains the the identity element 1. Furthermore, it is closed under inverses. In particular, the inverse of ω^k is $\omega^{n-k} \in C_n$. Hence, C_n is a subgroup of $\mathbb{C}^\times$ (and hence a group itself).

[9] This is not the actual definition of a generator. For example, this does not explain how 1 is a generator of $\mathbb{Z}$. The formal definition will come later.

5. Properties of groups

We now begin to study groups in more generality, including their basic properties. Throughout, we will generally treat G as an arbitrary group.

When the operation is understood we often will only write the set to denote the group. For example, the operation on $\mathbb{Z}$ is understood to be addition[10]. The most common operation symbols are $+$, $\cdot$, and $\circ$, however $+$ is almost universally reserved for groups with commutative operations (abelian groups).

When the operation is multiplication (or composition), the identity is typically denoted by 1 or e, and the inverse of an element $a \in G$ is denoted a^{-1}. When the operation is addition, the identity is typically denoted by 0, and the inverse of $a \in G$ is denoted $-a$. Because we want to treat G as an arbitrary group, we will use multiplicative notation so that there is no assumption on commutativity.

Proposition 4: Properties of groups

Let G be a group.

(1) The identity element of G is unique.

(2) For all $g \in G$, the inverse element $g^{-1} \in G$ is unique.

(3) For $g, h \in G$, $(gh)^{-1} = h^{-1}g^{-1}$.

(4) Let $a, b, c \in G$.

- (Right cancellation property) If $ba = ca$, then $b = c$.
- (Left cancellation property) If $ab = ac$, then $b = c$.

Proof. (1) Let $e, e' \in G$ be identity elements. Because e is an identity element, then $e = ee'$. Because e' is an identity element, $ee' = e$. Thus, $e = ee' = e'$.

(2) Fix $g \in G$ and let $a, b \in G$ be inverses of G. Then by associativity,

$$a = ea = (bg)a = b(ga) = be = b.$$

(3) First note that, by associativity,

$$(gh)(h^{-1}g^{-1}) = (g(hh^{-1}))g^{-1} = (ge)g^{-1} = gg^{-1} = e,$$
$$(h^{-1}g^{-1})(gh) = h^{-1}((g^{-1}g)h) = h^{-1}(eh) = h^{-1}h = e.$$

That is, $h^{-1}g^{-1}$ satisfies the condition to be *an* inverse of gh. By (2), inverses are unique, so $h^{-1}g^{-1}$ is *the* inverse of gh. Hence, $h^{-1}g^{-1} = (gh)^{-1}$.

(4) Multiply on the left or right by a^{-1} and use the associativity axiom. $\square$

[10] Recall that $\mathbb{Z}$ is *not* a group under multiplication because elements are not invertible in general.

Property (4) above, the cancellation property, implies that we cannot have repetitions in a row or column of the Cayley table[11]. To see this, just note that if $ab = ac$ in the "a row", then (left) cancellation implies that $b = c$.

The following exercise is inspired by (3) above.

Exercise. Let G be a group and $g \in G$. If $g' \in G$ satisfies $gg' = e$ or $g'g = e$, then $g' = g^{-1}$. (A left/right inverse element in a group is a two-sided inverse).

Exercise. Let G be a group and $a, b \in G$. Show that the equations $ax = b$ and $xa = b$ have unique solutions in G.

In multiplicative notation we use exponentials for short hand. Let $g \in G$ with G a group, then

$$g^0 = e, \quad g^1 = g, \quad g^n = g \cdot g \cdots g \ (n \text{ times}), \quad g^{-n} = g^{-1} \cdot g^{-1} \cdots g^{-1} \ (n \text{ times}).$$

For additive notation we use coefficients. Let $a \in A$ with A an abelian group, then

$$0a = 0, \quad 1a = a, \quad na = a + a + \cdots + a \ (n \text{ times}), \quad (-n)a = (-a) + (-a) + \cdots + (-a) \ (n \text{ times}).$$

Theorem 5: Exponential rules for groups

Let G be a group, $g, h \in G$, and $m, n \in \mathbb{Z}$. Then the following hold:

(Multiplicative notation)	(Additive notation)
(1) $g^m g^n = g^{m+n}$;	(1) $mg + ng = (m + n)g$;
(2) $(g^m)^n = g^{mn}$;	(2) $m(ng) = (mn)g$;
(3) $(gh)^n = (h^{-1}g^{-1})^{-n}$.	(3) $m(g + h) = mg + mh$.

[11] Courtesy of my Fall 2017 abstract algebra class, this is known as the *Sudoku rule*.

Definition: Order of a group

The *order of a group* $(G, \cdot)$ is the number of elements in G.

We denote the order of G by $|G|$. If $|G| < \infty$, then G is said to be *finite*. Otherwise, G is *infinite*.

Example. The group $\mathbb{Z}_n$ has order n. (For the groups $\mathbb{Z}_n$, the operation is understood to be addition mod n.) On the other hand, the group $M_2(\mathbb{R})$ has infinite order. (Here the operation is understood to be matrix addition.)

Exercise. Determine all possible Cayley Tables for a group of order 4. Use this to show that every group of order 4 is abelian.

Example (The group $U(n)$)**.** For any positive integer n, let $U(n)$ denote the set of invertible elements (units) in $\mathbb{Z}_n$ under multiplication.

Claim: The elements of $U(n)$ are exactly those integers m, $1 \le m < n$, such that $\gcd(m, n) = 1$.

Proof of claim. Suppose $\gcd(m, n) = 1$. Then there exist[12] integers r, s such that $rm + sn = 1$. But then

$$rm = 1 - sm \equiv 1 \mod n.$$

That is, $r \bmod n$ is the multiplicative inverse of m, so m is a unit in $\mathbb{Z}_n$.

Conversely, suppose m is a unit in $\mathbb{Z}_n$ and let $d = \gcd(m, n)$. Then $xm = 1 \mod n$ for some $x \in \mathbb{Z}_n$. That is $xm = 1 + yn$ for some integer y. But then $xm - yn = 1$. Since $d \mid m$ and $d \mid n$, then d divides $xm - yn$. But then d divides 1, so $d = 1$. □

Now if $m, m' \in U(n)$, then $gcd(m, n) = \gcd(m', n) = 1$ and certainly $\gcd(mm', n) = 1$. Hence, $U(n)$ is closed under multiplication mod n. Multiplication is associative (this follows from associativity of multiplication in $\mathbb{Z}_n$).

The identity element is 1. We have already shown that each element in $U(n)$ is invertible. It remains only to show that the inverse is an element of $U(n)$. Suppose $a \in U(n)$ and $b = a^{-1} \notin U(n)$. That is, $\gcd(b, n) \neq 1$. But then $\gcd(ab, n) \neq 1$ so $ab \notin U(n)$. But $ab = 1 \in U(n)$, a contradiction, so $b \in U(n)$. Thus, $U(n)$ is a group and it is easy to see that it is abelian.

The Cayley table for the group $(U(8), \cdot)$:

	1	3	5	7
1	1	3	5	7
3	3	1	7	5
5	5	7	1	3
7	7	5	3	1

[12] See preliminary notes

We conclude with a return to subgroups. In addition to giving a few more examples, we want to illustrate a mechanism to check more easily whether a given subset of a group is a subgroup.

Since we will use the group $\mathbb{Z}_n$ so frequently, this is a good time to declare that we will drop the equivalence class notation $[x]$ and simply denote this element by x. The understanding is that we are always working with equivalence classes and will denote elements by the preferred representatives $0, 1, \ldots, n-1$.

Example. Determine the subgroups of $\mathbb{Z}_4 = \{0, 1, 2, 3\}$.

Clearly $\{0\}$, $\{0, 2\}$, and $\mathbb{Z}_4$ itself are subgroups. It is left to show that there are no other subgroups. Suppose H is a subgroup containing 1. Then by closure, H also contains 2, 3, and 0. So $H = \mathbb{Z}_4$. Similarly, if K is a subgroup containing 3 then $K = \mathbb{Z}_4$. Hence, the only proper non-trivial subgroup is $\{0, 2\}$.

Example. Consider $\mathbb{Z}_2 \times \mathbb{Z}_2 = \{(0,0), (1,0), (0,1), (1,1)\}$ with addition (mod 2) in each component,

$$(a, b) + (c, d) = (a + b \mod 2, c + d \mod 2).$$

We work out the Cayley Table for this group below.

	(0,0)	(1,0)	(0,1)	(1,1)
(0,0)	(0,0)	(1,0)	(0,1)	(1,1)
(1,0)	(1,0)	(0,0)	(1,1)	(0,1)
(0,1)	(0,1)	(1,1)	(0,0)	(1,0)
(1,1)	(1,1)	(0,1)	(1,0)	(0,0)

Note the similarity between this table and that of $U(8)$. They are the "same" group in a sense that we will make more explicit later.

The subgroups of $\mathbb{Z}_2 \times \mathbb{Z}_2$ are

$$\{(0,0)\}, \quad \{(0,0), (1,0)\}, \quad \{(0,0), (1,0)\}, \quad \{(0,0), (1,1)\}, \quad \mathbb{Z}_2 \times \mathbb{Z}_2.$$

Note that this group has more subgroups than $\mathbb{Z}_4$. Thus, even though they have the same order, $\mathbb{Z}_4$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$ are different groups.

In general, to check that a subgroup is a group we need to verify first that it is a subset and then check that it is a group. The next proposition simplifies that process.

Proposition 6: The Subgroup Test

Let H be a nonempty subset of G. Then H is a subgroup of G if and only if whenever $a, b \in H$, $ab^{-1} \in H$.

Proof. ($\Rightarrow$) Suppose H is a subgroup of G. Then H is a group and so it contains an identity. Thus, $H \neq$. Now suppose $a, b \in H$. Then $b^{-1} \in H$ by closure under inverses and thus $ab^{-1} \in H$ by closure under the operation.

($\Leftarrow$) Suppose H is a nonempty subset of G satisfying the given condition. We claim H is a subgroup. Note that it is not necessary to prove that the operation is associative since the operation is inherited from the group G, and hence is associative.

Since H is nonempty, then there exists some $a \in H$. Hence $e = aa^{-1} \in H$, so H contains an identity element. (In fact, it must contain the identity element of G.) Now let $b \in H$ be any element, then by the given condition $b^{-1} = eb^{-1} \in H$, so H is closed under inverses. Finally, let $x, y \in H$. Then $y^{-1} \in H$ by the above. Hence, $xy = x(y^{-1})^{-1} \in H$, so H is closed under the operation. It follows that H is a group (and hence a subgroup). □

Example (The special linear group $\mathrm{SL}_2(\mathbb{R})$)**.** Let $\mathrm{SL}_2(\mathbb{R})$ denote the subset of determinant one matrices in $\mathrm{GL}_2(\mathbb{R})$. We show that $\mathrm{SL}_2(\mathbb{R})$ is a subgroup of $\mathrm{GL}_2(\mathbb{R})$.

Since $\det(I_2) = 1$, then $I_2 \in \mathrm{SL}_2(\mathbb{R})$ and so $\mathrm{SL}_2(\mathbb{R}) \neq \emptyset$. Now suppose $A, B \in \mathrm{SL}_2(\mathbb{R})$. Then

$$\det(AB^{-1}) = \det(A)\det(B)^{-1} = 1,$$

so $AB^{-1} \in \mathrm{SL}_2(\mathbb{R})$. Thus, $\mathrm{SL}_2(\mathbb{R})$ is a subgroup of $\mathrm{GL}_2(\mathbb{R})$.

Cyclic and Symmetric Groups

1. Cyclic groups

Recall that R_n is the subgroup of rotations in D_n. The element $r \in D_n$ *generates* R_n in the sense that every element of R_n is a power of r. Similarly, every element in C_n (the group of primitive nth roots of unity) is generated by $e^{2pi/n}$. As another example, the group $\mathbb{Z}_n$ can be generated by 1 in the sense that every element of $\mathbb{Z}_n$ is a multiple of 1 (the difference here is that $\mathbb{Z}_n$ is *additive* while R_n and C_n are *multiplicative*). These are all examples of *cyclic groups*, which we study in this chapter. This also gives further insight into the *order* of elements in finite groups.

To start, we define some notation. If G is a group and $a \in G$, then we set

$$\langle a \rangle = \{a^k : k \in \mathbb{Z}\} \quad \text{(multiplicative notation)}$$
$$\langle a \rangle = \{ka : k \in \mathbb{Z}\} \quad \text{(additive notation).}$$

That is, these are the sets of powers (resp. multiples) of a given elements, including powers (resp. multiples) of the inverse. These sets (which turn out to be subgroups) are crucial to understanding the structure of a group.

Theorem 1: The group generated by an element

Let G be a group and $a \in G$. Then $\langle a \rangle$ is the smallest subgroup of G containing a.

Proof. We will prove this theorem using multiplicative notation. The proof for additive notation is similar and left as an exercise.

First we prove that $\langle a \rangle$ is a subgroup. Since $a \in \langle a \rangle$, then $\langle a \rangle \neq \emptyset$. Let $g, h \in \langle a \rangle$, then $g = a^k$ and $h = a^\ell$ for some $k, \ell \in \mathbb{Z}$. Then $gh^{-1} = (a^k)(a^{-\ell}) = a^{k-\ell} \in \langle a \rangle$ and so $\langle a \rangle$ is a subgroup.

Now let H be another subgroup of G containing a. Because H is a group, every power of a lives in H. Thus, $\langle a \rangle \subset H$, proving that $\langle a \rangle$ is the smallest subgroup of G containing a. □

Definition: Cyclic subgroup, cyclic group, generator

Let G be a group. For $a \in G$, $\langle a \rangle$ is the *cyclic subgroup* of G generated by a. If there exists $a \in G$ such that $\langle a \rangle = G$, then G is a *cyclic group* and a a *generator* of G.

Example. The following are examples of cyclic groups.

- The integers $\mathbb{Z}$ (under addition) is a cyclic group generated by either 1 or -1.
- The inters $\mathbb{Z}_n$ (under addition mod n) is a cyclic group. The elements 1 and $n-1$ are always generators. In general, any integer relatively prime to n is a generator. For example, the generators of $\mathbb{Z}_6$ are 1 or 5. (On the other hand, $\langle 2 \rangle = \langle 2, 4, 0 \rangle$).
- The subgroup of rotations R_n in D_n is cyclic. As with $\mathbb{Z}_n$, the elements r and r^{n-1} are generators, as is r^m with $\gcd(m, n) = 1$. However, D_n is not cyclic for $n \geq 3$.
- The group of nth roots of unity C_n is cyclic. A generator is any *primitive* nth root of unity.
- The group of units $U(9)$ is a cyclic group generated by 2. However, $U(n)$ is not cyclic in general. For example, $U(8)$ is not cyclic.

Definition: Order of an element

The *order* of an element a in a group G is the smallest positive integer, if it exists, such that $a^n = e$. When no such n exists then the order is infinite

In additive notation, the condition $a^n = e$ is replaced by $na = 0$.

We write $|a| = n$ to denote the order of the element a (or $|a| = \infty$ if the order is infinite).

Example. Find the order of every element of D_3.

The reflections all have order 2. The rotations r and r^2 have order 3. The identity has order 1.

In general for D_n, reflections have order 2 and the identity has order 1. The order of a rotation is, in general, a little more complicated. We will return to this fact in a little while. For D_4,

Proposition 2: Alternate definition of order of an element

Let G be a group and $a \in G$. Then $|a| = |\langle a \rangle|$.

Proof. Suppose $|a| = n < \infty$. A similar argument applies for $|a| = \infty$. Then

$$\langle a \rangle = \{a, a^2, \dots, a^{n-1}, e\}.$$

Hence $\langle a \rangle \leq n$. Suppose $a^i = a^j$ for some $1 \leq i < j \leq n$. Then $a^{j-i} = e$. But $j - i < n$, contradicting the definition of order. Thus, the a^k with $1 \leq k \neq n$ are all distinct. Thus, $\langle a \rangle = n$. □

Proposition 3: Order of the inverse

Let G be a group and $a \in G$. Then $|a| = |a^{-1}|$.

Proof. First, suppose a has finite order n. Then $a^n = e$ so $(a^{-1})^n = (a^n)^{-1} = e^{-1} = e$. Hence, $|a^{-1}| \leq n = |a|$. This proves that if a has finite order, then so does a^{-1}. Conversely, suppose a^{-1} has finite order m. Then $(a^{-1})^m = e$ so $a^m = ((a^{-1})^{-1})^m = ((a^{-1})^m)^{-1} = e^{-1} = e$. Hence, $|a| \leq m = |a^{-1}|$. This proves that if a^{-1} has finite order, then so does a. Thus, either both have finite order or both have infinite order. If both have finite order, then $|a| \leq |a^{-1}| \leq |a|$, so $|a^{-1}| = |a|$.

Here is an alternate proof which uses the previous proposition. Let $x \in \langle a \rangle$, then $x = a^n$ for some $n \in \mathbb{Z}$. But then $x = (a^{-1})^{-n} \in \langle a^{-1} \rangle$. Thus, $\langle a \rangle \subset \langle a^{-1} \rangle$. Similarly, if $y \in \langle a^{-1} \rangle$, then $y = (a^{-1})^m$ for some $m \in \mathbb{Z}$. Then $y = a^{-m} \in \langle a \rangle$, so $\langle a^{-1} \rangle = \langle a \rangle$. Hence, $\langle a \rangle = \langle a^{-1} \rangle$ so $|a| = |\langle a \rangle| = |\langle a^{-1} \rangle| = |a^{-1}|$. □

It is easy to prove that a cyclic group is abelian. Let $a^k, a^\ell \in \langle a \rangle$. Then

$$a^k \cdot a^\ell = a^{k+\ell} = a^{\ell+k} = a^\ell \cdot a^k.$$

A more interesting result is below. Before we prove it we will recall a preliminary theorem.

The Division Algorithm

Let $a, b \in \mathbb{Z}$ with $b > 0$. Then there exist unique integers q and r such that $a = bq + r$ with $0 \leq r < b$.

Proof. (Existence) Let $S = \{a - bk : k \in \mathbb{Z} \text{ and } a - bk \geq 0\} \subset \mathbb{N}$. If $0 \in S$, then b divides a so choose $q = a/b$ and $r = 0$. Assume $0 \notin S$. If $a \geq 0$, then $a = a - b \cdot 0 \in S$. If $a < 0$, then $a - b(2a) = a(1 - 2b) \in S$. In either case, $S \neq \emptyset$ and so we may apply the Well-Ordering Principle to find a least member of S, say r. By definition of S, there exists an integer q such that $r = a - bq$. That is, $a = bq + r$ and $r \geq 0$ by definition of S. We claim $r < b$. Suppose otherwise. Then

$$a - b(q + 1) = a - bq - b = r - b > 0.$$

But then $a - b(q + 1) \in S$ and $a - b(q + 1) < r$, contradicting the choice of r.

(Uniqueness) Suppose there exist r, r', q, q' such that $a = bq + r$ and $a = bq' + r'$ with $0 \leq r, r' < b$. Then $bq+r = bq'+r'$. Assume $r' \geq r$. Then $b(q-q') = r-r'$ so b divides $r'-r$ and $0 \leq r'-r \leq r' < b$. Thus, $r' - r = 0$ so $r = r'$ and $q = q'$. □

Theorem 4: Every subgroup of a cyclic group is cyclic

If H is a subgroup of a cyclic group G. Then H is cyclic.

Proof. Let $G = \langle a \rangle$ be a cyclic group and H a subgroup of G. If $H = \langle e \rangle$, then H is cyclic, so assume H is non-trivial. Let m be the smallest positive integer such that $a^m \in H$. Such an m exists because H is non-trivial and by the Well-Ordering Principle. We claim $h = a^m$ is a generator for H. Let $h' \in H$. Then $h' = a^k$ for some integer $k > 0$. By the Division Algorithm, there exists some integers q, r such that $k = mq + r$ with $0 \leq r < m$. Now

$$a^k = a^{mq+r} = (a^m)^q a^r = h^q a^r.$$

Then $a^r = a^k h^{-q} = a^{k-mq} \in H$, contradicting the minimality of m. Thus, $r = 0$ and $a^k = h^q \in \langle h \rangle$, so h is a generator of H. □

Recall that $\mathbb{Z}$ is a cyclic group (with generators 1 and -1). By the previous theorem, every subgroup of $\mathbb{Z}$ is cyclic and so has a generator n. This is the set $\langle n \rangle = \{kn : k \in \mathbb{Z}\} = n\mathbb{Z}$. This fully classifies all subgroups of $\mathbb{Z}$.

As mentioned in the introduction, the study of cyclic groups extends beyond cyclic groups themselves. It also gives a way to understand the order of elements in groups. For example, a powerful theorem, which we will prove later, states that the order of an element in a finite group divides the order of the group. These next results will be important pieces of that theorem.

Theorem 5: Order of an element in a cyclic group

Let $G = \langle a \rangle$ be a cyclic group of order n.

(1) Then $a^k = e$ if and only if $n \mid k$.

(2) If $b = a^k$, then $|b| = n/d$ where $d = \gcd(k, n)$.

Proof. (1) If $n \mid k$, then $k = n\ell$ for some positive integer ℓ and

$$a^k = a^{n\ell} = (a^n)^\ell = e^\ell = e.$$

Now suppose $a^k = e$. By the Division Algorithm there exists integers q, r such that $k = nq + r$ with $0 \leq r < n$. Hence,

$$e = a^k = a^{nq+r} = (a^n)^q a^r = ea^r = a^r.$$

By definition of order, n is the smallest positive integer such that $a^n = e$. Thus, $r = 0$ and so $n \mid k$.

(2) Set $m = |b|$. This is the smallest integer such that $e = b^m = a^{km}$. By part (1), this implies that n divides km. Equivalently, n/d divides $m(k/d)$. Since $d = \gcd(k, n)$, then n/d and k/d are relatively prime. Hence, n/d divides m.

On the other hand, k/d is an integer and so n divides $(k/d)n$. Hence, $b^{n/d} = a^{kn/d} = a^{(k/d)n} = e$. But m is the order of b so by part (1), m divides n/d. □

Recall that $\mathbb{Z}_n$ is cyclic (of order n). Of course 1 is a generator of $\mathbb{Z}_n$, but so is any element of order n. To see this, note that if a has order n then so does $\langle a \rangle \subset \mathbb{Z}_n$. By the Pigeonhole Principle, this implies that $\langle a \rangle = \mathbb{Z}_n$. Hence, by the theorem, the generators of $\mathbb{Z}_n$ are those integers r such that $1 \leq r < n$ and $\gcd(r, n) = 1$.

2. Permutation groups

Recall our example of D_3, the symmetries of a triangle with vertices A, B, C. Any symmetry may be regarded as a rearrangement of the vertices and so every symmetry is a bijective function from the set $\{A, B, C\}$ to itself. In this way we may regard D_3 as a *permutation group*. In fact, every dihedral group (group of symmetries) is a permutation group on some set. However, while D_3 captures *every* rearrangement of the vertices, D_4 does not.

Definition: Permutation

A *permutation* is a bijective function on the set X (from X to itself). The set of permutations on X is denoted $\mathcal{S}_X$.

Theorem 6: Group of permutations on a set

For any nonempty set X, $\mathcal{S}_X$ is a group under composition.

Proof. Most of this is contained in preliminary notes. In particular, the composition of two bijective functions is again a bijective function, and the operation of function composition is associative. The identity function is given by $\mathrm{id}(x) = x$ for all $x \in X$, and this is clearly an element of $\mathcal{S}_X$. Finally, a bijective function is invertible and that inverse is again bijective (and hence an element of $\mathcal{S}_X$). □

Definition: Symmetric group, permutation group

Let X be a set. The *symmetric group* on X is the set $\mathcal{S}_X$ under the operation of (function) composition. When $X = \{1, \ldots, n\}$, then $\mathcal{S}_X$ is denoted by $\mathcal{S}_n$ and is called the *symmetric group on n letters*. A subgroup of $\mathcal{S}_n$ is a *permutation group*.

The set X need not be finite, but we will focus almost exclusively on the case of $\mathcal{S}_n$. It should be relatively straightforward for you to convince yourself that the order of $\mathcal{S}_n$ is $n!$.

There are two standard types of notation to represent elements of $\mathcal{S}_n$: two-line and cycle. In two-line notation we write the elements of $\mathcal{S}_n$ as $2 \times n$ matrices. For a given element $\sigma \in \mathcal{S}_n$ we write in the first row $1, \ldots, n$ and in the second the image of each value under σ:

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

Warning. The elements of S_n are functions and therefore we compose right-to-left.

Example. In general, the elements of $\mathcal{S}_n$ do not commute. Consider the following elements of S_3:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \quad \text{and} \quad \tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Compute $\sigma\tau$ and $\tau\sigma$ using two-line notation.

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \quad \text{and} \quad \tau\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Example. Consider the following elements of $\mathcal{S}_4$:

$$\text{id} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \quad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \quad \mu = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}.$$

These elements form a subgroup of $\mathcal{S}_4$ with Cayley table:

	id	σ	τ	μ
id	id	σ	τ	μ
σ	σ	τ	μ	id
τ	τ	μ	id	σ
μ	μ	id	σ	τ

This Cayley table is equivalent to one we've seen before. Where?

A more compact way of representing elements of $\mathcal{S}_n$ is with *cycles*.

Definition: Cycle, cycle length

A permutation $\sigma \in \mathcal{S}_n$ is a *cycle of length k* if there exists $a_1, \ldots, a_k \in \{1, \ldots, n\}$ such that

$$\sigma(a_1) = a_2, \quad \sigma(a_2) = a_3, \quad \ldots \quad \sigma(a_k) = a_1$$

and $\sigma(i) = i$ for $i \notin \{a_1, \ldots, a_k\}$. We denote the cycle by $(a_1 \; a_2 \; \cdots \; a_k)$.

To compose cycles, one *could* translate back to two-line notation but I strongly advise against that. Instead, we compose (from right-to-left) by tracking the image of each element through successive cycles, remembering to close cycles when we get back to where we started.

Example. In the previous example, the elements would be written in cycle notation by

$$\mathrm{id} = (1), \quad \sigma = (1\;4\;3\;2), \quad \tau = (1\;3)(2\;4), \quad \mu = (1\;2\;3\;4).$$

Then we can easily compute the same products we did above. For example,

$$\begin{aligned}
\sigma\tau &= (1\;4\;3\;2)(1\;3)(2\;4) = (2\;3\;4\;1) = (1\;2\;3\;4) = \mu \\
\tau\sigma &= (1\;3)(2\;4)(1\;4\;3\;2) = (1\;2\;3\;4) = \mu \\
\sigma\mu &= (1\;4\;3\;2)(1\;2\;3\;4) = (1)(2)(3)(4) = (1) = \mathrm{id}\,.
\end{aligned}$$

Note in general the permutations *will not* commute.

Example (Inverse of a cycle)**.** The inverse of a cycle is obtained by cycling in the "opposite direction". So, the inverse of $(1\;4\;6\;3\;2\;7)$ in $\mathcal{S}_7$ is $(1\;7\;2\;3\;6\;4)$.

Example. In cyclic notation, the symmetric group on three letters is

$$\mathcal{S}_3 = \{(1), (1\;2\;3), (1\;3\;2), (1\;2), (1\;3), (2\;3)\}.$$

The Cayley Table is

	(1)	(1 2)	(1 3)	(2 3)	(1 2 3)	(1 3 2)
(1)	(1)	(1 2)	(1 3)	(2 3)	(1 2 3)	(1 3 2)
(1 2)	(1 2)	(1)	(1 3 2)	(1 2 3)	(2 3)	(1 3)
(1 3)	(1 3)	(1 2 3)	(1)	(1 3 2)	(1 2)	(2 3)
(2 3)	(2 3)	(1 3 2)	(1 2 3)	(1)	(1 3)	(1 2)
(1 2 3)	(1 2 3)	(1 3)	(2 3)	(1 2)	(1 3 2)	(1)
(1 3 2)	(1 3 2)	(2 3)	(1 2)	(1 3)	(1)	(1 2 3)

Note the similarities between this table and that of D_3. Can you match up the permutations and the symmetries?

Definition: Disjoint cycles

Two cycles $\sigma = (a_1\ a_2\ \cdots\ a_k)$ and $\tau = (b_1\ b_2\ \cdots\ b_\ell)$ are *disjoint* if $a_i \neq b_j$ for all i, j.

Example. The cycles (1 3 5) and (2 7) are disjoint but the cycles (1 3 5) and (3 4 7) are not. Note that (1 3 5)(3 4 7) = (1 3 4 7 5).

Example. Write $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 1 & 5 & 6 & 4 \end{pmatrix}$ as a product of disjoint cycles.

Starting with 1 we have $\sigma_1 = (1\ 3)$. We choose a value, say 2, that is not yet accounted for and continue, $\sigma_2 = (2)$. Next we choose 4 and continue, $\sigma_3 = (4\ 5\ 6)$. This exhausts $\{1, \ldots, 6\}$ and so $\sigma = \sigma_1\sigma_2\sigma_3 = (1\ 3)(2)(4\ 5\ 6)$. Note that (2) is equivalent to the identity so it is appropriate to omit it and write $\sigma = (1\ 3)(4\ 5\ 6)$.

Proposition 7: Disjoint cycles in $\mathcal{S}_n$ commute

If σ and τ are disjoint cycles in $\mathcal{S}_n$, then $\sigma\tau = \tau\sigma$.

Proof. Let $\sigma, \tau \in \mathcal{S}_n$ be disjoint. Write $\sigma = (a_1\ a_2\ \cdots\ a_k)$ and $\tau = (b_1\ b_2\ \cdots\ b_\ell)$. We claim $(\sigma\tau)(x) = (\tau\sigma)(x)$ for all $x \in \{1, \ldots, n\}$.

Suppose $x \notin \{a_1, \ldots, a_k\}$ and $x \notin \{b_1 \ldots, b_\ell\}$. By definition of a cycle, $(\sigma\tau)(x) = \sigma(\tau(x)) = \sigma(x) = x$ and similarly $(\tau\sigma)(x) = \tau(\sigma(x)) = \tau(x) = x$.

Now suppose $x \in \{a_1, \ldots, a_k\}$ (so $x \notin \{b_1, \ldots, b_\ell\}$). Then $\sigma(x) \in \{a_1, \ldots, a_k\}$ and so $\sigma(x) \notin \{b_1, \ldots, b_\ell\}$. Thus $(\sigma\tau)(x) = \sigma(\tau(x)) = \sigma(x)$ and $(\tau\sigma)(x) = \tau(\sigma(x)) = \sigma(x)$.

The proof for $x \in \{b_1, \ldots, b_\ell\}$ is similar. □

Theorem 8: Cycle decomposition

Let $\sigma \in \mathcal{S}_n$. Then σ is a product of disjoint cycles in $\mathcal{S}_n$.

Proof. Set $X = \{1, \ldots, n\}$. First we will decompose X into disjoint pieces and use these to define the cycles. Choose $\sigma \in \mathcal{S}_n$ and define $X_1 = \{1, \sigma(1), \sigma^2(1), \ldots\}$. Then X_1 is finite because σ has finite order. Choose $k \in X \backslash X_1$ and define $X_2 = \{k, \sigma(k), \sigma^2(k), \ldots\}$. Continue in this way. Note that the process *must* end because X is finite. Write $X = X_1 \cup X_2 \cup \cdots \cup X_r$.

Define a cycle σ_i by

$$\sigma_i(x) = \begin{cases} \sigma(x) & x \in X_i \\ x & x \notin X_i. \end{cases}$$

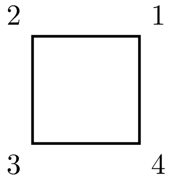
Then $\sigma = \sigma_1\sigma_2 \cdots \sigma_r$. Note that the σ_i are disjoint because the X_i are. □

3. The Dihedral Group

In the remaining sections, we study two examples of permutation groups. The first is familiar.

Previously, we presented the dihedral group D_n ($n \geq 3$ as generated by a rotation r and a reflection s. We observed that D_3 and $\mathcal{S}_3$ have the same Cayley table. This is because each permutation of the set $\{1, 2, 3\}$ corresponds to a symmetry of an equilateral triangle with vertices labeled $\{1, 2, 3\}$.

Of course, it cannot be that D_4 and $\mathcal{S}_4$ have the same Cayley table since $|D_4| = 8$ and $|\mathcal{S}_4| = 24$. However, we can identify D_4 with a *subgroup* of $\mathcal{S}_4$. Consider a square with vertices labeled $\{1, 2, 3, 4\}$.



The rotation r (90° counterclockwise) can be described as the cycle $(1\ 2\ 3\ 4)$ in $\mathcal{S}_4$. We map rotations to permutations in $\mathcal{S}_4$ by

$$1 \mapsto (1) \qquad r \mapsto (1\,2\,3\,4) \qquad r^2 \mapsto (1\,3)(2\,4) \qquad r^3 \mapsto (1\,4\,3\,2).$$

Now recall that for D_4 we can choose s to be any reflection through a vertex. Suppose s reflects on the line through vertices 1 and 3. That is, 1 and 3 are fixed while s swaps vertices 2 and 4. Thus, $s \mapsto (2\ 4)$. Now we can describe all of the reflections as permutations. For example, the reflection rs should correspond to the permutation $(1\ 2\ 3\ 4)(2\ 4)$.

$$s \mapsto (2\,4) \qquad rs \mapsto (1\,2)(3\,4) \qquad r^2 s \mapsto (1\,3) \qquad r^3 s \mapsto (1\,4)(2\,3).$$

Thus, we conclude that D_4 can be identified as the permutation group

$$\{(1), (1\,2\,3\,4), (1\,3)(2\,4), (1\,4\,3\,2)(2\,4), (1\,2)(3\,4), (1\,3), (1\,4)(2\,3)\}$$

4. The Alternating Group

In this section we will define an important subgroup of the symmetric group.

Definition: Transposition

A *transposition* is a cycle of length 2.

Example (Decomposing a cycle as a product of transpositions)**.** Any cycle can be written as the product of transpositions. There are many ways to do this. Here is one. Note that the transpositions are not disjoint in this decomposition.

$$(a_1\ a_2\ a_3\ \cdots\ n) = (a_1\ a_n)(a_1\ a_{n-1})\cdots(a_1\ a_3)(a_1\ a_2).$$

Of course, the above example extends to a permutation by decomposing each cycle. The next theorem is stated here, but will actually be proved later at the end of this section.

Theorem 9: Invariance of parity in cycles

Any decomposition of a cycle contains either an even number or odd number of transpositions.

Assuming the proposition for the time being allows us to define the next terms.

Definition: Even/odd permutation

A permutation is *even* (resp. *odd*) if it can expressed as the product of an even (resp. odd) number of transpositions.

Theorem 10: Subgroup of even permuations

The set of all even permutations in $\mathcal{S}_n$ is a subgroup of $\mathcal{S}_n$.

Proof. The identity is an even permutation so this set is nonempty. Let $\sigma, \tau \in \mathcal{S}_n$ be even permutations. Write, $\sigma = \sigma_1\sigma_2\cdots\sigma_k$ and $\tau = \tau_1\tau_2\cdots\tau_\ell$ with k, ℓ even. Then

$$\sigma\tau^{-1} = (\sigma_1\sigma_2\cdots\sigma_k)(\tau_\ell\tau_{\ell-1}\cdots\tau_1),$$

which is an even permutation. □

Definition: Alternating group

The *alternating group on n letters* is the subgroup A_n of $\mathcal{S}_n$ consisting of all even permutations.

Proposition 11: Order of A_n

The order of A_n is $n!/2$.

Proof. Let A_n and B_n denote the sets of even and odd permutations in $\mathcal{S}_n$, respectively. We will define a bijection between them, implying $|A_n| = |B_n|$. Fix a transposition $\sigma \in \mathcal{S}_n$ and define $\lambda_\sigma : A_n \to B_n$ by $\lambda_\sigma = \sigma\tau$ for all $\tau \in A_n$. Clearly this is well-defined.

If $\lambda_\sigma(\tau) = \lambda_\sigma(\mu)$, then $\sigma\tau = \sigma\mu$ so $\tau = \mu$. Thus λ_σ is 1-1. If $\eta \in B_n$, then $\sigma^{-1}\eta \in A_n$ and $\lambda_\sigma(\sigma^{-1}\eta) = \eta$,so λ_σ is surjective. □

We now return to the proof of Theorem 9, which says that the notion of the *sign* of a permutation is well-defined. The individual pieces of the proof are not difficult, but there are many pieces. We prove this as a series of small lemmas which combine to form the big theorem[1].

Lemma 12: Identities in $\mathcal{S}_n$

Let a, b, c, d be distinct elements of $\{1, \ldots, n\}$. In $\mathcal{S}_n$ we have the following identities:

$$(c\ d)(a\ b) = (a\ b)(c\ d) \quad \text{and} \quad (b\ c)(a\ b) = (a\ c)(b\ c).$$

Proof. The first statement follows from the fact that disjoint cycles commute. For the second, we again apply each the function on each side to an element of $\{1, \ldots, n\}$ and verify that the results are the same, proving that they are in fact equal as functions (and hence as elements of $\mathcal{S}_n$).

First we apply both functions to a

$$(b\ c)(a\ b)(a) = (b\ c)(b) = c \quad \text{and} \quad (a\ c)(b\ c)(a) = (a\ c)(a) = c.$$

The cases of applying to b or c are similar and left as an exercise. If $x \neq a, b, c$, then clearly both sides applied to x return x.

Applying c,

$$(b\ c)(a\ b)(c) = (b\ c)(c) = b \quad \text{and} \quad (a\ c)(b\ c)(c) = (a\ c)(b) = b.$$

Now applying b,

$$(b\ c)(a\ b)(b) = (b\ c)(a) = a \quad \text{and} \quad (a\ c)(b\ c)(b) = (a\ c)(c) = a.$$

□

[1] Like Voltron!

Lemma 13: Decomposing the identity

If the identity $(1) \in \mathcal{S}_n$ is written as a product of transpositions $(1) = \tau_1\tau_2 \dots \tau_k$, then k is even.

Proof. The identity itself is not a transposition, so $k \neq 1$. If $k = 2$ then we are done[2], suppose inductively that for some $k \geq 3$ and for every decomposition of the identity into transpositions of fewer than k transpositions, the number of transpositions is even.

Write

$$(\star) \qquad (1) = (a_1\ b_1)(a_2\ b_2) \cdots (a_k\ b_k).$$

In order for this to be valid, one of the $(a_i\ b_i)$, $i = 2, 3, \dots, k$, must move a_1. Since $(a_i\ b_i) = (b_i\ a_i)$, then without loss of generality we may assume $a_1 = a_i$ for some $i > 1$. Now using our identities from Lemma 12, we can move that a_i so that it appears in the cycle following $(a_1\ b_1)$. That is, we may assume $a_2 = a_1$ so the first two terms in $(\star)$ are $(a_1\ b_1)$ and $(a_1\ b_2)$.

If $b_1 = b_2$, then $(a_1\ b_1)(a_1\ b_2) = (1)$ and so we can remove it from our decomposition. Hence, $(\star)$ can be expressed in terms of $k-2$ transpositions and by our inductive hypothesis, $k-2$ is even, whence k is even.

Now assume $b_1 \neq b_2$. Note that $(a_1\ b_1)(a_1\ b_2) = (a_1\ b_2)(b_1\ b_2)$ with $a_1 \neq b_1, b_2$. That is, $(b_1\ b_2)$ does not move a_1. Again, there must be some cycle $(a_i\ b_i)$ with $i = 3, 4, \dots, k$ that moves a_1. Repeating our argument above, this cycle either allows us to eliminate a pair of cycles, or else we can write the first three terms in such a way that only the first moves a_1.

This argument continues but ultimately must end because k is finite. However, if in this process we ultimately write (1) as a decomposition where only the first moves a_1 then we arrive at a contradiction. Thus, we conclude that at some step we are able to reduce the number of cycles and apply our inductive hypothesis. □

Proof of Theorem 9. Let $\sigma \in \mathcal{S}_n$. Suppose we have two decompositions,

$$\sigma = \tau_1\tau_2 \cdots \tau_r = \tau_1'\tau_2' \cdots \tau_s'$$

where each τ_i, τ_i' is a transposition. Then

$$(1) = \sigma\sigma^{-1} = \tau_1\tau_2 \cdots \tau_r\tau_s'\tau_{s-1}' \cdots \tau_2'\tau_1'.$$

Thus, (1) is the product of $r+s$ transpositions, so $r+s$ is even by Lemma 13. Hence, r and s have the same parity. (Said another way, $r \equiv s \mod 2$).

[2] Some people (not me) like to write this sort of phrase as "then we win".

Group Homomorphism

We began this course by studying a fundamental object in algebra (groups) and the corresponding sub-objects (subgroups). Now we study maps between those objects and this leads naturally to a discussion on quotient objects. This framework appears frequently in mathematics. For example, in 331 you learned about sets, subsets, and functions (between sets). In Topology, for instance, one studies topological spaces, their subspaces, and continuous functions between topological spaces.

1. Homomorphisms

A homomorphism as a *structure preserving map* between groups that respects the two operations.

Definition: Homomorphism

A *homomorphism* is a function $\phi : (G, \cdot) \to (H, \circ)$ between groups such that

$$\phi(g_1 \cdot g_2) = \phi(g_1) \circ \phi(g_2) \quad \text{for all } g_1, g_2 \in G.$$

Example. The following are examples of homomorphisms.

(1) Define $\phi : \mathrm{GL}_2(\mathbb{R}) \to \mathbb{R}^\times$ by $\phi(M) = \det(M)$. For $A, B \in \mathrm{GL}_2(\mathbb{R})$,

$$\phi(AB) = \det(AB) = \det(A)\det(B) = \phi(A)\phi(B).$$

Hence, ϕ is a homomorphism. (Note however that the determinant map $M_2(\mathbb{R}) \to \mathbb{R}$ is not a homomorphism because $\det(M + N) \neq \det(M) + \det(N)$ in general).

(2) Define $\rho : \mathbb{Z} \to \mathbb{Z}$ by $\rho(n) = 2n$. For $x, y \in \mathbb{Z}$,

$$\rho(x + y) = 2(x + y) = 2x + 2y = \rho(x) + \rho(y).$$

Thus, ρ is a homomorphism.

(3) Define a map $\psi : \mathbb{Z} \to \mathbb{R}^\times$ by $\psi(n) = 2^n$. For $m, n \in \mathbb{Z}$,

$$\psi(m + n) = 2^{m+n} = 2^m 2^n = \psi(m)\psi(n).$$

Hence, ψ is a homomorphism.

(4) Define $\mu : \mathbb{Z}_2 \mapsto \mathbb{Z}_4$ by $\mu(0) = 0$ and $\mu(1) = 2$. To check that this is a homomorphism we must verify all of the possible compositions in $\mathbb{Z}_2$:

$$\mu(0 + 0) = \mu(0) = 0 = 0 + 0 = \mu(0) + \mu(0)$$
$$\mu(0 + 1) = \mu(1) = 2 = 0 + 2 = \mu(0) + \mu(1)$$
$$\mu(1 + 1) = \mu(0) = 0 = 2 + 2 = \mu(1) + \mu(1).$$

Thus, μ is a homomorphism.

These properties of homomorphisms we can prove immediately. Others we will prove later.

Proposition 1: The identity and inverses are preserved

Let $\phi : G \to H$ be a homomorphism of groups.

(1) The identity of G is mapped to the identity of H under ϕ. That is, $\phi(e_G) = e_H$.

(2) For any $g \in G$, $\phi(g^{-1}) = \phi(g)^{-1}$.

Proof. (1) We have

$$\phi(e_G)e_H = \phi(e_G) = \phi(e_G e_G) = \phi(e_G)\phi(e_G).$$

By left cancellation, $\phi(e_G) = e_H$.

(2) Let $g \in G$, then by (1),

$$e_H = \phi(gg^{-1}) = \phi(g)\phi(g^{-1}).$$

By uniqueness of the inverse, $\phi(g^{-1}) = \phi(g)^{-1}$. □

There are two important subgroups relative to every homomorphism: the *image* and the *kernel*. These appear in linear algebra in the context of the column and null space of a matrix.

Definition: Image and kernel (of a homomorphism)

Let $\phi : G \to H$ be a group homomorphism. The *image* of ϕ is the set $\operatorname{Im}\phi = \{\phi(g) : g \in G\}$. The *kernel* of $\phi : G \to H$ is the set $\mathrm{Ker}\phi = \{g \in G : \phi(g) = e\}$.

Example. We compute the image and kernel in some of our examples above.

(1) Consider $\phi : \mathrm{GL}_2(\mathbb{R}) \to \mathbb{R}^\times$ given by $\phi(M) = \det(M)$. Then $\operatorname{Im}\phi = \mathbb{R}^\times$ and $\mathrm{Ker}\phi = \mathrm{SL}_2(\mathbb{R})$. Note that ϕ is surjective, but it is not injective.

(2) Consider $\rho : \mathbb{Z} \to \mathbb{Z}$ by $\rho(n) = 2n$. Then $\operatorname{Im}\rho = 2\mathbb{Z}$ and $\mathrm{Ker}\rho = \{0\}$. This map is injective but not surjective.

(3) Consider $\psi : \mathbb{Z} \to \mathbb{R}^\times$ by $\psi(n) = 2^n$. Then $\operatorname{Im}\psi = \{2^n : n \in \mathbb{Z}\}$ and $\mathrm{Ker}\psi = \{0\}$. This map is also injective but not surjective.

(4) Consider $\mu : \mathbb{Z}_2 \mapsto \mathbb{Z}_4$ by $\mu(0) = 0$ and $\mu(1) = 2$. Then $\operatorname{Im}\mu = \{0, 2\}$ and $\mathrm{Ker}\mu = \{0, 2\}$. This map is neither injective nor surjective.

Recall that, generically, we use the notation ϕ^{-1} to refer to the *preimage* of a set. In general, the inverse of a homomorphism is not itself a homomorphism (because it may not be a function).

Proposition 2: Subgroups are mapped to subgroups (and back again)

Let $\phi : G \to H$ be a homomorphism of groups.

(1) If K is a subgroup of G, then $\phi(K) = \{\phi(k) : k \in K\}$ is a subgroup of H.

(2) If L is a subgroup of H, then $\phi^{-1}(L) = \{g \in G : \phi(g) \in L\}$ is a subgroup of G.

Proof. (1) Let K be a subgroup of G. Since $e_G \in K$, then $e_H = \phi(e_G) \in \phi(K)$, so $\phi(K) \neq \emptyset$. Let $a, b \in \phi(K)$, so there exists $x, y \in K$ such that $\phi(x) = a$ and $\phi(y) = b$. Then

$$ab^{-1} = \phi(x)\phi(y)^{-1} = \phi(xy^{-1}).$$

Since K is a subgroup, $xy^{-1} \in K$, so $ab^{-1} = \phi(xy^{-1}) \in \phi(K)$. Thus, $\phi(K)$ is a subgroup of H by the subgroup test.

(2) Let L be a subgroup of H. Since $e_H \in L$ and $\phi(e_G) = e_H$, then $e_G \in \phi^{-1}(L)$. Thus, $\phi^{-1}(L) \neq \emptyset$. Choose $c, d \in \phi^{-1}(L)$, so $\phi(c), \phi(d) \in L$. Then

$$\phi(cd^{-1}) = \phi(c)\phi(d^{-1}) = \phi(c)\phi(d)^{-1} \in L$$

because L is a subgroup of H. Thus, $cd^{-1} \in \phi^{-1}(L)$ and so $\phi^{-1}(L)$ is a subgroup of G. □

As we observed earlier, the kernel turns out to be a subgroup of the domain. In fact it is a *normal* subgroup, a fact we will return to later (once we have defined the term normal).

Theorem 3: Kernels are subgroups

Let $\phi : G \to H$ be a group homomorphism. Then $\text{Ker}\phi$ is a subgroup of G.

Proof. Proof #1: Because $\phi(e_G) = e_H$, then $e_G \in \text{Ker}\phi$, so $\text{Ker}\phi \neq \emptyset$. Let $g_1, g_2 \in \text{Ker}\phi$. Then

$$\phi(g_1 g_2^{-1}) = \phi(g_1)\phi(g_2)^{-1} = e_H e_H^{-1} = e_H.$$

Thus, $\text{Ker}\phi$ is a subgroup by the subgroup test.

Proof #2: By definition, $\text{Ker}\phi = \phi^{-1}(\{e_H\})$. That is, the kernel is the preimage of the trivial subgroup of H, and hence a subgroup by the previous proposition. □

Kernels are important in detecting *injectivity* of a group homomorphism.

Lemma 4: Injective is equivalent to trivial kernel

Let $\phi : G \to H$ be a group homomorphism. Then ϕ is injective if and only if $\text{Ker}\phi = \{e_G\}$.

Proof. Assume ϕ is injective and let $g \in \text{Ker}\phi$. Then $\phi(g) = e_H = \phi(e_G)$. By injectivity, $g = e_G$. Hence, $\text{Ker}\phi = \{e_G\}$.

Now assume $\text{Ker}\phi = \{e_G\}$. We claim ϕ is injective. If $\phi(g_1) = \phi(g_2)$ for some $g_1, g_2 \in G$, then

$$e_H = \phi(g_1)\phi(g_2)^{-1} = \phi(g_1 g_2^{-1}).$$

Thus, $g_1 g_2^{-1} \in \text{Ker}\phi$. By our assumption, $g_1 g_2^{-1} = e_G$, so $g_1 = g_2$. That is, ϕ is injective. □

We have already seen isomorphisms in practice. The groups $\mathbb{Z}_n$ (integers mod n), R_n (rotations in D_n), and C_n (nth roots of unity) all appear to be the same group in that their Cayley tables are the same up to changing notation. Isomorphisms formalize this intuition.

Definition: Isomorphic, isomorphism

Two groups $(G, \cdot)$ and $(H, \circ)$ are said to be *isomorphic* if there exists a bijective homomorphism $\phi : G \to H$. The map ϕ in this case is called an *isomorphism*.

To check that a given map ϕ is an isomorphism, we must show that it is a homomorphism, that it is injective, and that it is surjective. By Lemma 4, we can verify injectivity by showing that the kernel is trivial.

Example. The following are examples of isomorphisms.

(1) Let $G = \langle a \rangle$ be a cyclic group of infinite order. Define a map $\phi : \mathbb{Z} \to G$ by $\phi(n) = a^n$. We first show that ϕ is a homomorphism. Let $m, n \in \mathbb{Z}$. Then

$$\phi(m + n) = a^{m+n} = a^m a^n = \phi(m)\phi(n).$$

Thus, ϕ is indeed a homomorphism. It is clear that ϕ is surjective since the preimage of $a^k \in G$ is k. We show that ϕ is injective using kernels. Let $x \in \text{Ker}\phi$, then $e_G = \phi(x) = a^x$. Because G has infinite order, this implies that $x = 0$, so $\text{Ker}\phi = \{0\}$. Thus, ϕ is injective and thus an isomorphism.

(2) Let $G = \langle a \rangle$ be a cyclic group of order $n < \infty$. In a similar way to the previous exercise, we can define a map $\phi : \mathbb{Z}_n \to G$ and prove that ϕ is an isomorphism. This is a bit harder, however, because $\mathbb{Z}_n$ is a set of equivalence classes and one must verify ϕ is well-defined. There is a *better* way to do this one which we will discuss later.

(3) The groups D_3 and $\mathcal{S}_3$ are isomorphic. This is not difficult to see by comparing Cayley Tables. One map (of many) between the groups is determined by $\phi(r) = (1\ 2\ 3)$ and $\phi(s) = (1\ 2)$.

Theorem 5: Properties of isomorphisms

Let $\phi : G \to H$ be an isomorphism of groups.

(1) $\phi^{-1} : H \to G$ is an isomorphism.

(2) $|G| = |H|$.

(3) If G abelian, then H is abelian.

(4) If G is cyclic, then H is cyclic.

(5) If G has a subgroup of order n, then H has a subgroup of order n.

(6) If $g \in G$ has order n, then $|\phi(g)| = n$.

Proof. (1) By standard results on functions, ϕ^{-1} exists and is bijective (see Preliminary notes). We need only verify that it is a homomorphism.

Let $x, y \in H$. Because ϕ is surjective, there exist (unique) elements $a, b \in G$ such that $\phi(a) = x$ and $\phi(b) = y$. Since ϕ is a homomorphism and $\phi^{-1}\phi = \mathrm{id}_G$, then

$$\phi^{-1}(x)\phi^{-1}(y) = ab = \phi^{-1}(\phi(ab)) = \phi^{-1}(\phi(a)\phi(b)) = \phi^{-1}(xy).$$

Hence, ϕ^{-1} is a homomorphism.

(2) This follows because ϕ is bijective.

(3) Assume G is abelian and let $h_1, h_2 \in H$. Then there exist (unique) elements $g_1, g_2 \in G$ such that $\phi(g_1) = h_1$ and $\phi(g_2) = h_2$. Then

$$h_1h_2 = \phi(g_1)\phi(g_2) = \phi(g_1g_2) = \phi(g_2g_1) = \phi(g_2)\phi(g_1) = h_2h_1.$$

Thus, H is abelian.

(4) Let a be a generator for G and let $b = \phi(a) \in H$. If $h \in H$, then by surjectivity, $\phi(g) = h$ for some $g \in G$. But G is cyclic so $g = a^k$ for some $k \in \mathbb{Z}$. Hence,

$$h = \phi(g) = \phi(a^k) = \phi(a)^k = b^k.$$

That is, b is a generator for H, so h is cyclic.

(5) Let K be a subgroup of G of order n. Then $\phi(K)$ is a subgroup of H because ϕ is a homomorphism. Moreover, $|\phi(K)| = n$ because ϕ is bijective.

(6) Suppose $g \in G$ has order n. (Note n here may be infinity.) Then $|\langle g \rangle| = n$. The proof of part (4) shows that $\phi(\langle g \rangle) = \langle \phi(g) \rangle$. Hence, by part (5), $|\langle \phi(g) \rangle| = n$. Hence, $|\phi(g)| = n$. □

2. Cosets

Cosets are arguably one of the strangest structures that students encounter in abstract algebra, along with factor groups, which are strongly related. Here's a motivating question for this section: if H is a subgroup of a group G, then how are $|H|$ and $|G|$ related? A partial answer to this is contained in Lagrange's Theorem.

Definition: Left and right cosets

Let H be a subgroup of a group G. A *left coset* of H with representative in $g \in G$ is the set

$$gH = \{gh : h \in H\}.$$

A *right coset* of H with representative in $g \in G$ is the set

$$Hg = \{hg : h \in H\}.$$

Warning. Cosets are **NOT** subgroups in general!

Example. Let $K = \{(1), (1\ 2)\}$ in $\mathcal{S}_3$.

The left cosets are

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

The right cosets are

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

Note that in general the left and right cosets are different.

Example. Let $L = \{(1), (1\ 2\ 3), (1\ 3\ 2)\}$ in $\mathcal{S}_3$.

The left cosets are

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

The right cosets are

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

In additive notation, we write

$$g + H = \{g + h : h \in H\}.$$

Additive groups are abelian by definition, so $g + H = H + g$, that is, the left and right cosets are equal. Hence, in this setting we will speak only of the *cosets* with no ambiguity.

Example. We compute the cosets of $H = \langle 3 \rangle = \{0, 3\}$ in $\mathbb{Z}_6$. The cosets are

$$0 + H = \{0, 3\} = 3 + H \quad 1 + H = \{1, 4\} = 4 + H \quad 2 + H = \{2, 5\} = 5 + H.$$

In this case, the left and right cosets are the same.

Lemma 6: Properties of cosets

Let H be a subgroup of G and suppose $g_1, g_2 \in G$. The following are equivalent.

(1) $g_1 H \subset g_2 H$

(2) $g_1 H = g_2 H$

(3) $H g_1^{-1} = H g_2^{-1}$

(4) $g_2 \in g_1 H$

(5) $g_1^{-1} g_2 \in H$.

Proof. (1) ⇒ (2) Assume $g_1 H \subset g_2 H$. We claim the opposite inclusion holds. Let $g_2 h \in g_2 H$ for some $h \in H$. Since $g_1 \in g_1 H \subset g_2 H$, then $g_1 = g_2 h'$ for some h'. But then $g_2 = g_1 (h')^{-1}$ and so

$$g_2 h = g_1 (h')^{-1} h = g_1((h')^{-1} h) \in g_1 H.$$

(2) ⇒ (3) Assume $g_1 H = g_2 H$. Let $h g_1^{-1} \in H g_1^{-1}$ with $h \in H$. Note that $g_1 \in g_1 H = g_2 H$, so $g_1 = g_2 h'$ for some $h' \in H$. Then $g_1^{-1} = (h')^{-1} g_2^{-1}$ and so

$$h g_1^{-1} = h((h')^{-1} g_2^{-1}) = (h(h')^{-1}) g_2^{-1} \in H g_2^{-1}.$$

Thus, $H g_1^{-1} \subset H g_2^{-1}$. A similar proof shows the reverse inclusion and the result follows.

(3) ⇒ (4) Assume $H g_1^{-1} = H g_2^{-1}$. Then $g_2^{-1} \in H g_2^{-1} = H g_1^{-1}$ so $g_2^{-1} = h g_1^{-1}$ for some $h \in H$. Thus, $g_2 = g_1 h^{-1} \in g_1 H$.

(4) ⇒ (5) Assume $g_2 \in g_1 H$. Then $g_2 = g_1 h$ for some $h \in H$. Thus, $g_1^{-1} g_2 = h \in H$.

(5) ⇒ (1) Assume $g_1^{-1} g_2 \in H$. Let $g_1 h \in g_1 H$ for some $h \in H$. By our assumption, $g_1^{-1} g_2 = h'$ for some $h' \in H$. Thus, $g_1 = g_2 (h')^{-1}$ and $g_1 h = g_2((h')^{-1} h) \in g_2 H$. Therefore $g_1 H \subset g_2 H$. □

The next result says that cosets of a subgroup *partition* the group. Though stated for left cosets, the result is clearly true also for right cosets. A separate argument is presented in your homework.

Proposition 7: Cosets partition the group

Let H be a subgroup of a group G. If $g_1, g_2 \in G$, then the left cosets $g_1 H$ and $g_2 H$ are either equal ($g_1 H = g_2 H$) or disjoint ($g_1 H \cap g_2 H = \emptyset$). Hence, every element belongs to exactly one left cosets.

Proof. Suppose $g_1 H \cap g_2 H \neq \emptyset$ and let $x \in g_1 H \cap g_2 H$. We claim $g_1 H = g_2 H$. Then $g_1 h = x = g_2 h'$ for some $h, h' \in H$. But then $g_2 = g_1 h (h')^{-1} \in g_1 H$. Hence, by the lemma, $g_1 H = g_2 H$.

Now for any $g \in G$, $g \in gH$, so g belongs to *at least* one left coset. By the above argument, if g belongs to two cosets then those are equal. □

3. Lagrange's Theorem

Lagrange's Theorem is an important step in understanding the structure of (finite) groups. Stated simply, it says that the order of a subgroup of a finite group divides the order of the group. The converse of Lagrange's Theorem is false in general. If $n \mid |G|$, this *does not* imply that there exists a subgroup H of G with $|H| = n$. For example, A_4 has no subgroup of order 6.

Definition: Index

The *index* of a subgroup H in a group G, denoted $[G : H]$, is the number of left cosets in G.

Example. In the previous examples, we have $[\mathbb{Z}_6 : H] = 3$, $[\mathcal{S}_3 : K] = 3$, and $[\mathcal{S}_3 : L] = 2$.

Theorem 8: Number of left cosets equals number of right cosets

The number of left cosets of a subgroup H in a group G equals the number of right cosets.

Proof. Denote by $\mathcal{L}_H$ the set of left cosets of H in G and by $\mathcal{R}_H$ the set of right cosets of H in G. We will establish a bijection between these sets, which shows that $|\mathcal{L}_H| = |\mathcal{R}_H|$.

Define a map

$$\begin{aligned} \phi : \mathcal{L}_H &\to \mathcal{R}_H \\ gH &\mapsto Hg^{-1}. \end{aligned}$$

First we need to show that this map is well-defined. Suppose $g_1H = g_2H$. By the lemma, $\phi(g_1H) = Hg_1^{-1} = Hg_2^{-1} = \phi(g_2H)$. Thus, ϕ is well-defined.

Suppose $\phi(g_1H) = \phi(g_2H)$, then $Hg_1^{-1} = Hg_2^{-1}$. Again, the lemma implies $g_1H = g_2H$, so ϕ is injective. Let $Hg \in \mathcal{R}_H$. Then $\phi(g^{-1}H) = H(g^{-1})^{-1} = Hg$, so ϕ is surjective.

Thus, ϕ is bijective and so the result holds. □

Lemma 9: All cosets have the same order

Let H be a subgroup of G. For all $g \in G$, $|H| = |gH|$.

Proof. Fix $g \in G$ and define a map $\phi : H \to gH$ by $h \mapsto gh$. It is clear that ϕ is well-defined. We will show that it is bijective. It will then follow immediately that $|H| = |gH|$.

Suppose $\phi(h) = \phi(h')$ for $h, h' \in H$. Then $gh = gh'$ and so by left cancellation, $h = h'$. Thus ϕ is injective. Next let $gh \in gH$ for some $h \in H$, then $\phi(h) = gh$ so the map is surjective. □

The proof of Lagrange's Theorem is now simple because we've done the legwork already.

Theorem 10: Lagrange's Theorem

Let G be a finite group and H a subgroup of G. Then $|G|/|H| = [G : H]$.

Proof. The group G is partitioned into $[G : H]$ distinct left cosets. Each coset has exactly $|H|$ elements by the previous lemma. Hence, $|G| = |H|[G : H]$. □

Warning. Lagrange's theorem does not work (or even make sense) for infinite groups. There are variations for special types of infinite groups but we will not discuss these.

One immediate consequence of Lagrange's Theorem is that the order of a subgroup divides the order of a group (when the group has finite order). This holds for order of elements as well.

Corollary 11: Order of an element divides the order of the group

Suppose G is a finite group and $g \in G$. Then the order of g divides $|G|$.

Proof. Note that $|g| = \langle g \rangle$. Thus, for the first statement we simply apply Lagrange's Theorem with $H = \langle g \rangle$. Now set $d = |G|$. Then $|g| = k \mid d$ by the above. Thus, $d = k\ell$ for some integer ℓ. Then $g^d = g^{k\ell} = (g^k)^\ell = e^\ell = e$. □

Now suppose that G is a group of prime order p. If $g \in G$ is not the identity, then $|g| > 1$ and $|g|$ divides p, so $|g| = p$. That is, g is a generator for G. Hence, *all groups of prime order are cyclic*. It follows that $G \cong \mathbb{Z}_p$ (though we have no quite proved this yet).

The last result of this section shows that index plays nicely with *towers* of subgroups.

Corollary 12: Index is multiplicative

Let H, K be subgroups of a finite group G such that $K \subset H \subset G$. Then

$$[G : K] = [G : H][H : K].$$

Proof. By Lagrange's Theorem,

$$[G : K] = \frac{|G|}{|K|} = \frac{|G|}{|H|} \cdot \frac{|H|}{|K|} = [G : H][H : K].$$ □

4. Normal subgroups and factor groups

Normal subgroups lead to factor groups, which reveal "hidden" structure in groups not revealed. Factor groups are a group structure on the set of cosets of a group by certain subgroups. This is also setup for several fundamental results called "isomorphism theorems". We will only prove one of these here, but this will prove immensely useful.

Definition: Normal subgroup

A subgroup N of a group G is *normal* if $gN = Ng$ for all $g \in G$.

Example. The following are examples of normal subgroups.

(1) The alternating group A_3 (which we called L above) is normal in $\mathcal{S}_3$. But the subgroup $K = \{(1), (1\ 2)\}$ of $\mathcal{S}_3$ we considered is *not* normal.

(2) The subgroup R_n of D_n is normal. There are two left cosets: one is R_n itself and the other consists of all of the reflections. The same is true of the right cosets.

(3) If G is abelian, then *every* subgroup is normal.

(4) The trivial subgroup of a group G, as well as the group itself, are normal subgroups.

Theorem 13: Equivalent definitions of normality

Let G be a group and N a subgroup. The following are equivalent.

(1) The subgroup N is normal ($gN = Ng$ for all $g \in G$)

(2) For all $g \in G$, $gNg^{-1} \subset N$.

(3) For all $g \in G$, $gNg^{-1} = N$.

Proof. (1) ⇒ (2) Assume N is normal and fix $g \in G$. Let $n \in N$, then by the definition of normal, $gN = Ng$ so $gn = n'g$ for some $n' \in N$. Thus, $gng^{-1} = n' \in N$, so $gNg^{-1} \subset N$.

(2) ⇒ (3) Assume $gNg^{-1} \subset N$ for all $g \in G$. We claim the opposite inclusion holds (for all $g \in G$). Let $n \in N$ and fix $g \in G$. By our assumption, $g^{-1}ng \subset N$ (since it holds for all $g \in G$). Thus, $g^{-1}ng = n'$ for some $n' \in N$, so $n = gn'g^{-1} \in gNg^{-1}$. Therefore, $N \subset gNg^{-1}$.

(3) ⇒ (1) Assume $gNg^{-1} = N$ for all $g \in G$. We will show $gN \subset Ng$. The proof that $Ng \subset gN$ is similar. Let $g \in G$ and $n \in N$. Then $gng^{-1} = n'$ for some $n' \in N$. Thus, $gn = n'g \in Ng$. □

Example (Kernels are normal)**.** Let $\phi : G \to H$ be a group homomorphism.We have already shown that Kerϕ is a subgroup. Now for any $x \in \text{Ker}\phi$ and any $g \in G$ we have

$$\phi(gxg^{-1}) = \phi(g)\phi(x)\phi(g)^{-1} = \phi(g)\phi(g)^{-1} = e.$$

Thus, Kerϕ is a normal subgroup of G.

Let N be a normal subgroup of a group G. We denote by G/N the set of cosets. In this setup (because N is normal), there is no need to differentiate between left and right cosets. However, we will typically work with left cosets.

Lemma 14: Binary operation on G/N

Let N be a normal subgroup of a group G. There is a binary operation on G/N given by

$$(aN)(bN) = (ab)N$$

for all $aN, bN \in G/N$.

Proof. We must verify that this operation is well-defined. Said another way, we must verify that the product is independent of choice of coset representative.

Let $aN = cN$ and $bN = dN$. We claim $(aN)(bN) = (cN)(dN)$. Since $c \in aN$ and $d \in bN$, then there exists $n_1, n_2 \in N$ such that $c = an_1$ and $d = bn_2$. Note that $n_iN = N$ because N is a subgroup and $n_i \in N$, and that $bN = Nb$ by normality. Thus, we have the following equalities of sets:

$$(cd)N = (an_1)(bn_2)N = (an_1)b(n_2N) = (an_1)(bN) = (an_1)(Nb) = a(n_1N)b = aNb = (ab)N. \qquad \square$$

Theorem 15: Group structure on G/N

Let N be a normal subgroup of G. The cosets of N in G form a group G/N (with the operation above) of order $[G : N]$.

Proof. We have already shown that the operation is binary. We verify the remaining group axioms. Throughout, let $aN, bN, cN \in G/N$. The operation is associative since

$$(aN)[(bN)(cN)] = (aN)(bcN) = (a(bc))N = ((ab)c)N = (abN)(cN) = [(aN)(bN)](cN).$$

The identity element in G/N is just the coset of the identity in G,

$$(eN)(aN) = (ea)N = aN.$$

Finally, the inverse of aN is the coset $a^{-1}N$,

$$(aN)(a^{-1}N) = (aa^{-1})N = eN. \qquad \square$$

Definition: Factor group

Let G be a group and N a normal subgroup. The set G/N along with the operation $(aN)(bN) = (ab)N$ is the *factor group* of G by N.

Example (The factor group D_n/R_n)**.** We have already shown that R_n is a normal subgroup of D_n. There are two cosets: R_n, sR_n. The Cayley table for these is

	R_n	sR_n
R_n	R_n	sR_n
sR_n	sR_n	R_n

It is clear that this factor group is isomorphic to $\mathbb{Z}_2$.

For an abelian group, where cosets are denoted $a + N$, we denote the above binary operation by

$$(a + N) + (b + N) = (a + b) + N.$$

Example (The factor group $\mathbb{Z}/3\mathbb{Z}$)**.** The group $\mathbb{Z}$ is abelian so the subgroup $H = 3\mathbb{Z}$ is normal. There are three cosets: $H, 1 + H, 2 + H$. This group has the following Cayley table:

	H	$1+H$	$2+H$
H	H	$1+H$	$2+H$
$1+H$	$1+H$	$2+H$	H
$2+H$	$2+H$	H	$1+H$

It is clear that this factor group is (again) isomorphic to $\mathbb{Z}_3$.

We have previously seen that the kernel of a homomorphism is a normal subgroup of the domain, so *Kernels are Normals*. The next result completes this idiom.

Proposition 16: Normals are Kernels

Let G be a group and N a normal subgroup. The map $\pi : G \to G/N$ given by $\pi(g) = gN$ is a group homomorphism and $\text{Ker}\pi = N$.

Proof. Let $g, h \in G$. Then $\pi(g)\pi(h) = (gN)(hN) = (gh)N = \pi(gh)$. Hence, π is a homomorphism. Now let $x \in \text{Ker}\pi$. Then $N = \pi(x) = xN$, so $x \in N$. Thus, $\text{Ker}\pi \subset N$. Conversely, if $y \in N$, then $\pi(y) = yN = N$, so $N \subset \text{Ker}\pi$. □

The map π in the above proposition is called the *quotient map*.

The isomorphism theorems give a more direct way of viewing the isomorphisms we just considered. There are three of them, but we will only consider the first one right now.

Our next result actually generalizes an earlier result that a homomorphism is an isomorphism if and only if its kernel is trivial. The first isomorphism theorem says that the factor group of a group by the kernel of an homomorphism is isomorphic to the image of the homomorphism. It is a powerful tool in proving that two groups are isomorphic.

Theorem 17: First Isomorphism Theorem

Let $\phi : G \to H$ be a homomorphism. Then $G/\text{Ker}\phi \cong \phi(G)$.

Proof. Let $K = \text{Ker}\phi$. By our work above, K is a normal subgroup and thus G/K is well-defined. Define a map $\psi : G/K \to \phi(G)$ by $\psi(gK) = \phi(g)$. We claim that ψ is a (well-defined) group isomorphism.

Suppose $g_1K = g_2K$ for some $g_1K, g_2K \in G/K$. We must show that $\psi(g_1K) = \psi(g_2K)$. Because K is normal, our hypothesis implies that $g_2^{-1}g_1 \in K$. Therefore, $\psi(g_2^{-1}g_1) = e$, so $\psi(g_2) = \phi(g_1)$. But then $\psi(g_1K) = \phi(g_2K)$. Thus, ψ is well-defined.

We now check that ψ is a homomorphism. Let $g_1K, g_2K \in G/K$. Then

$$\psi((g_1K)(g_2K)) = \psi((g_1g_2)K) = \phi(g_1g_2) = \phi(g_1)\phi(g_2) = \psi(g_1K)\psi(g_2K).$$

It is left to show that ψ is bijective. Clearly ψ is surjective since for all $\phi(g) \in \phi(G)$, $\psi(gK) = \phi(g)$. To show injectivity we reverse the argument above for well-definedness. Suppose $\psi(g_1K) = \psi(g_2K)$ for some $g_1K, g_2K \in G/K$. Then $\phi(g_1) = \phi(g_2)$ so $g_2^{-1}g_1 \in K$. Because K is normal, $g_1K = g_2K$, so ψ is injective. □

The first isomorphism tells us that in fact we should never try to define a homomorphism whose domain is a factor group. Instead, we should show the corresponding homomorphism is the kernel of some (surjective) homomorphism.

Example ($\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$)**.** Consider the map $\phi : \mathbb{Z} \mapsto \mathbb{Z}_n$ given by $\phi(k) = k \mod n$. This map is a surjective homomorphism (check!) and $\text{Ker}\phi = n\mathbb{Z}$. By the First Isomorphism Theorem, $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$.

Example ($\mathrm{GL}_2(\mathbb{R})/\,\mathrm{SL}_2(\mathbb{R}) \cong \mathbb{R}^\times$)**.** Let $\phi : \mathrm{GL}_2(\mathbb{R}) \to \mathbb{R}^\times$ be the map given by $\phi(A) = \det(A)$. We have already shown that this is a homomorphism. For $r \in \mathbb{R}^\times$, let

$$B_r = \begin{pmatrix} r & 0 \\ 0 & 1 \end{pmatrix}.$$

Then $B_r \in \mathrm{GL}_2(\mathbb{R})$ and $\phi(B_r) = r$, so ϕ is surjective. Now

$$\text{Ker}\phi = \{M \in \mathrm{GL}_2(\mathbb{R}) : \det(M) = 1\} = \mathrm{SL}_2(\mathbb{R}).$$

Thus, by the First Isomorphism Theorem, $\mathrm{GL}_2(\mathbb{R})/\,\mathrm{SL}_2(\mathbb{R})$ is isomorphic to $\mathbb{R}^\times$.

Finite Abelian Groups

This set of notes is dedicated to an important classification theorem in group theory. While we will focus on the theorem statement, the proof illustrates the power in using factor groups.

1. Direct products

We have seen some examples of direct products previously. For example, the *Klein-4* group $\mathbb{Z}_2 \times \mathbb{Z}_2$ where the operation is addition mod 2 in each coordinate. Recall that for any two sets A and B, the cartesian product of two sets A and B is the set $A \times B = \{(a,b) : a \in A, b \in B\}$. Recall that the order of this set is $|A \times B| = |A| \cdot |B|$.

Proposition 1: External direct product of groups

Let $(G,\cdot)$ and $(H,\circ)$ be groups. Then $G \times H$ is a group under the operation

$$(g_1,h_1) \star (g_2,h_2) = (g_1 \cdot g_2, h_1 \circ h_2) \quad \text{for all } (g_1,h_1),(g_2,h_2) \in G \times H.$$

Proof. Let $(g_1,h_1),(g_2,h_2) \in G \times H$. Then $g_1 \cdot g_2 \in G$ because G is a group and so $\cdot$ is binary. Similarly, $h_1 \circ h_2 \in H$. Hence, $(g_1,h_1) \star (g_2,h_2) = (g_1 \cdot g_2, h_1 \circ h_2) \in G \times H$. Let $(g,h) \in G \times H$. Then

$$(e_G,e_H) \star (g,h) = (e_G \cdot g, e_H \circ h) = (g,h) = (g \cdot e_G, h \circ e_H) = (g,h) \star (e_G,e_H),$$

so (e_G,e_H) is an identity element in $G \times H$. Similarly, since G and H are closed under inverses then $(g^{-1},h^{-1}) \in G \times H$ so

$$(g^{-1},h^{-1}) \star (g,h) = (g^{-1} \cdot g, h^{-1} \circ h) = (e_G,e_H) = (g \cdot g^{-1}, h \circ h^{-1}) = (g,h) \star (g^{-1},h^{-1}),$$

so $G \times H$ is closed under inverses. Finally we check associativity. Let $(g_1,h_1),(g_2,h_2),(g_3,h_3) \in G \times H$. Then

$$\begin{aligned}(g_1,h_1) \star ((g_2,h_2),(g_3,h_3)) = (g_1,h_1) \star (g_2 \cdot g_3, h_2 \circ h_3) &= (g_1 \cdot (g_2 \cdot g_3), h_1 \circ (h_2 \circ h_3)) \\ &= ((g_1 \cdot g_2) \cdot g_3, (h_1 \circ h_2) \circ h_3) \quad \text{(by associativity of } \cdot \text{ and } \circ) \\ &= (g_1 \cdot g_2, h_1 \circ h_2) \star (g_3,h_3) = ((g_1,h_1) \star (g_2,h_2)) \star (g_3,h_3).\end{aligned}$$

We conclude that $G \times H$ is a group under $\star$. □

Definition: External direct product

The *external direct prduct* of groups $(G,\cdot)$ and $(H,\circ)$ is the set $G \times H$ with operation

$$(g_1,h_1) \star (g_2,h_2) = (g_1 \cdot g_2, h_1 \circ h_2) \quad \text{for all } (g_1,h_1),(g_2,h_2) \in G \times H.$$

The proof of the next result is left as a homework exercise.

Lemma 2: Order of elements in a direct product

Let G and H be groups and $(a,b) \in G \times H$. Then $|(a,b)| = \operatorname{lcm}(|a|,|b|)$.

Recall that if $|G| = p$, p prime, then $G \cong \mathbb{Z}_p$. Here is a related result.

Proposition 3: Direct product of cyclic groups

Let m, n be positive integers, then $\mathbb{Z}_m \times \mathbb{Z}_n \cong \mathbb{Z}_{mn}$ if and only if $\gcd(m,n) = 1$.

Proof. As a set $|\mathbb{Z}_m \times \mathbb{Z}_n| = mn$. Since $\mathbb{Z}_{mn}$ is the unique (up to isomorphism) cyclic group of order mn, it suffices to prove that $\mathbb{Z}_m \times \mathbb{Z}_n$ is cyclic if and only if $\gcd(m,n) = 1$.

Let a and b be generators of $\mathbb{Z}_m$ and $\mathbb{Z}_n$, respectively. Suppose $\gcd(m,n) = 1$, then by the lemma,

$$|(a,b)| = \operatorname{lcm}(|a|,|b|) = \frac{mn}{\gcd(m,n)} = mn.$$

Suppose $\gcd(m,n) > 1$. If $(x,y) \in \mathbb{Z}_m \times \mathbb{Z}_n$, then $|x| \leq m$ and $|y| \leq n$ and so again by the lemma,

$$|(x,y)| \leq \frac{mn}{\gcd(m,n)} < mn,$$

so $\mathbb{Z}_m \times \mathbb{Z}_n$ is not cyclic. □

Example (Decomposing $\mathbb{Z}_6$)**.** We will show that $\mathbb{Z}_6$ is the external direct product of its subgroups. The cyclic group $\mathbb{Z}_6$ has subgroups $A = \langle 2 \rangle$ and $B = \langle 3 \rangle$. Note that $\langle 2 \rangle \cong \mathbb{Z}_3$ and $\langle 3 \rangle \cong \mathbb{Z}_2$. By the above proposition, $\mathbb{Z}_6 \cong \mathbb{Z}_3 \times \mathbb{Z}_2$. Thus, $\mathbb{Z}_6$ is the external direct product of A and B.

Example (Two non-isomorphic groups of order 4)**.** The gcd requirement in the proposition is important. The groups $\mathbb{Z}_4$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$ are *not* isomorphic. We can see this by noting that $\mathbb{Z}_2 \times \mathbb{Z}_2$ (the Klein-4 group) has no elements of order 4, whereas $\mathbb{Z}_4$ has two.

For a group G with subgroups H, N, let $HN = \{hn : h \in H, n \in N\}$. In general this is *not* a subgroup of G.

Lemma 4: H, N commute implies HN is a subgroup

Let G be a group with subgroups H, N. If $hn = nh$ for all $h \in H$, $n \in N$, then HN is a subgroup of G.

Proof. Since H and N are subgroups then $e \in H$ and $e \in N$. Hence, $e = e \cdot e \in HN$. Let $h_1n_1, h_2n_2 \in HN$. Then

$$(h_1n_1)(h_2n_2)^{-1} = (h_1n_1)(n_2^{-1}h_2^{-1}) = (h_1h_2^{-1})(n_1n_2^{-1}).$$

Since H is a subgroup, $h_1h_2 \in H$. Since N is a subgroup, then $n_1n_2^{-1} \in N$. Thus, $(h_1n_1)(h_2n_2)^{-1} \in HN$, so HN is a subgroup. □

In additive notation, HN is $H + N = \{h + n : h \in H, n \in N\}$. It is clear from the lemma that $H + N$ is always a subgroup of G.

The next result is a homework exercise.

Proposition 5: N normal implies HN is a subgroup

Let G be a group with subgroups H, N. If N is normal, then HN is a subgroup of G.

We are now ready to discuss another form of combining two (sub)groups.

Definition: Internal direct product

A group G is the *interal direct product* of subgroups H and K provided

(1) $G = HK$ (as sets),

(2) $H \cap K = \{e\}$, and

(3) $hk = kh$ for all $h \in H$, $k \in K$.

Example (Decomposing $U(8)$)**.** We will show that $U(8)$ is an internal direct product of two subgroups. Let $G = U(8)$, $H = \{1, 3\}$, $K = \{1, 5\}$. Then $G = HK$ (because $3 \cdot 5 \equiv 7 \mod 8$). Clearly $H \cap K = \{1\}$, and because G is abelian we have $hk = kh$ for all $h \in H$, $k \in K$. Thus, G is the internal direct product of H and K.

Example ($\mathcal{S}_3$ is not an IDP)**.** We will show that $\mathcal{S}_3$ is *not* an internal direct product of its subgroups. $\mathcal{S}_3$ is *not* an internal direct product of its subgroups. Since $|\mathcal{S}_3| = 6$ then without loss of generality we need subgroups H, K with $|H| = 3$ and $|K| = 2$. But then $H \cong \mathbb{Z}_3$ and $K \cong \mathbb{Z}_2$. The condition $hk = kh$ for all $h \in H$, $k \in K$ now implies that $\mathcal{S}_3$ is abelian, a contradiction.

Lemma 6

Let G be a group with subgroup H and K satisfying $hk = kh$ for all $h \in H$ and $k \in K$. Then the map $\phi : H \times K \to G$ given by $\phi(h, k) = hk$ is a homomorphism.

Proof. Let $(h_1, k_1), (h_2, k_2) \in H \times K$. Then

$$\phi((h_1, k_1)(h_2, k_2)) = \phi(h_1h_2, k_1k_2) = (h_1h_2)(k_1k_2) = (h_1k_1)(h_2k_2) = \phi(h_1, k_1)\phi(h_2, k_2).$$

Thus, ϕ is a homomorphism. □

The next theorem says that if G is the internal direct product of subgroups H and K, then it is also the external direct product of those subgroups.

Theorem 7: Internal direct product is an external direct product

If G is the internal direct product of subgroups H and K, then $G \cong H \times K$.

Proof. Let $\phi : H \times K \to G$ be as in the lemma. By IDP(3), ϕ is a homomorphism. Let $g \in G$. By IDP(1), $G = HK$, so $g = hk$ for some $h \in H$ and $k \in K$. Hence, $(h, k) \in H \times K$ and $\phi(h, k) = hk = g$, so ϕ is surjective.

It remains only to show that ϕ is injective. Let $(x, y) \in \text{Ker}\phi$. Then $e = \phi(x, y) = xy$. But then $y^{-1} = x \in H \cap K$, so $x = e$ and $y = e$ by IDP(2). Thus, Kerϕ is trivial so ϕ is injective. It follows that ϕ is an isomorphism. □

Example (The group $U(8)$ again)**.** By a previous example, $U(8) \cong \mathbb{Z}_2 \times \mathbb{Z}_2$.

Another nice corollary of this lemma provides an alternative way to check property (1) in the definition of an IDP for finite groups. Suppose $|G| = n < \infty$ with subgroups H and K. Let $\phi : H \times K \to G$ be the map in the lemma. Then clearly $\Im\phi = HK$ and it is not difficult to see that Ker$\phi \cong H \cap K$. Thus, by the First Isomorphism Theorem, $(H \times K)/(H \cap K) \cong HK$. Thus,

$$|HK| = \frac{|H||K|}{|H \cap K|}.$$

Hence, if $|H \times K| = n$ and $|H \cap K| = 1$ (this is equivalent to condition (2) in the definition), then we conclude that $|HK| = n$. By pigeonhole, $HK = G$.

All of our work thus far generalizes to more than two subgroups. The proof of the proposition below is left as an exercise.

Definition: Internal direct product (general)

A group G is the *internal direct product* of subgroups $H_1, H_2, \dots, H_n$ provided

(1) $G = H_1H_2\cdots H_n = \{h_1h_2\cdots h_n : h_i \in H_i\}$ (as sets),

(2) $H_i \cap \langle \bigcup_{j\neq i} H_j\rangle = \{e\}$, and

(3) $h_ih_j = h_jh_i$ for all $h_i \in H_i$, $h_j \in H_j$, $i \neq j$.

Proposition 8: Internal direct product is an external direct product (general)

If a group G is the internal direct product of subgroups $H_1, H_2, \dots, H_n$, then

$$G \cong H_1 \times H_2 \times \cdots \times H_n.$$

We are now ready to state our main theorem of this set. Ultimately, this is a generalization of the fact that $\mathbb{Z}_{mn} \cong \mathbb{Z}_m \times \mathbb{Z}_n$ if and only if $\gcd(m, n) = 1$. For the proof of this theorem, we will need more technology.

Theorem 9: The Fundamental Theorem of Finite Abelian Groups

Every finite abelian group is isomorphic to a direct product of cyclic groups of prime power order.

The Fundamental Theorem implies that every finite abelian group can be written (up to isomorphism) in the form

$$\mathbb{Z}_{p_1^{\alpha_1}} \times \mathbb{Z}_{p_2^{\alpha_2}} \times \cdots \times \mathbb{Z}_{p_n^{\alpha_n}},$$

with p_i prime (not necessarily distinct) and $\alpha_i \in \mathbb{N}$.

Example. Every finite abelian group of order $540 = 2^2 \cdot 3^3 \cdot 5$ is isomorphic to exactly one of the following:

(1) $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_5$

(2) $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_9 \times \mathbb{Z}_5$

(3) $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_{27} \times \mathbb{Z}_5$

(4) $\mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_5$

(5) $\mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_9 \times \mathbb{Z}_5$

(6) $\mathbb{Z}_4 \times \mathbb{Z}_{27} \times \mathbb{Z}_5$

2. Proof of the Fundamental Theorem (Part I)

The first part of the proof involves breaking down abelian groups into direct products of subgroups where the orders of the subgroups are relatively prime. This is a generalization of the idea used above to show that $\mathbb{Z}_{mn} \cong \mathbb{Z}_m \times \mathbb{Z}_n$ when $\gcd(m, n) = 1$.

Lemma 10: Decomposing abelian groups

Let G be an abelian group with $|G| = mn$ and $\gcd(m, n) = 1$. Let H be the set of elements in G whose elements have an order dividing m and define K similarly for n. Then $G \cong H \times K$.

Proof. Since $|e| = 1$, then $e \in H$. Let $h_1, h_2 \in H$, then the orders of h_1 and h_2 divide m. Because G is abelian, $(h_1h_2^{-1})^m = h_1^m(h_2^m)^{-1} = e$. Thus, $|h_1h_2^{-1}|$ divides m and this implies $h_1h_2^{-1} \in H$, so H is a subgroup. The proof that K is a subgroup is similar.

Now we check the properties of HK to be an IDP for G. Clearly $hk = kh$ for all $h \in H$ and $k \in K$ because G is abelian. If $g \in H \cap K$, then $|g|$ divides m and $|g|$ divides n. Since $\gcd(m, n) = 1$, then $|g| = 1$, so $g = e$. Thus, $H \cap K = \{e\}$.

For the last condition, let $g \in G$ and write $|g| = rs$ with $r \mid m$ and $s \mid n$. Because m and n are relatively prime, so are r and s. Thus, there exists integers a, b such that $1 = ar + bs$. Set $h = g^{bs}$ and $k = g^{ar}$. Since $r \mid m$, then $rs \mid bsm$ and so

$$h^m = g^{bsm} = 1.$$

That is, the order of h divides m, so $h \in H$. A similar argument shows $k \in K$. Now,

$$g = g^1 = g^{ar+bs} = g^{ar}g^{bs} = kh = hk \in HK.$$

Thus, $G = HK$. It follows that G is the IDP of H and K, so $G \cong H \times K$. □

Now suppose $|G| = p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_n^{\alpha_n}$ with the p_i distinct primes. Set $G_i = \{g \in G : |g| = p_i^k, k \in \mathbb{Z}\}$. The proof above shows that each G_i is a subgroup (because the only factors of a prime power are other powers of that prime). Now as a consequence of the theorem (by induction) we have

$$G \cong G_1 \times G_2 \times \cdot \times G_n.$$

It suffices to prove the fundamental theorem for the G_i.

Definition: p-group

Let p be a prime. A group G is a *p-group* if every element in G has order a power of p.

Example. The following are examples of p-groups:

- The groups $\mathbb{Z}_2$, $\mathbb{Z}_4$, and $\mathbb{Z}_2 \times \mathbb{Z}_2$ are all 2-groups, as is D_4.
- The group $\mathbb{Z}_{27}$ is a 3-group.
- The groups G_i defined above are p_i-groups.

By Lagrange's Theorem, every group of order p^n, p a prime, is automatically a p-group since the order of every element must divide p^n. We will prove a converse to this for finite abelian groups.

First we make a couple brief remarks. Suppose G is a group and N a normal subgroup of G. If $a \in G$ has order m, then $(aN)^m = a^m N = N$. Hence, the order of aN divides m, the order of a. We will use this fact several times in what follows.

Also recall that if $g \in G$ has order n, then by a result from cyclic groups,

$$|g^k| = \frac{|g|}{\gcd(|g|, k)} = \frac{n}{\gcd(n, k)}.$$

Thus, if $G = \langle g \rangle$ is a cyclic group of order n and $p \mid n$, then $g^{n/p}$ has order p.

Lemma 11: Elements of prime order

Let G be a finite abelian group of order n. If p is a prime dividing n, then G contains an element of order p

Proof. Assume the statement is false. Then there exists a group G of order n whose order is divisible by p but G does not contain an element of order p. Also, we may choose n minimal amongst all such G with this property.

Let $g \in G$ be a non-identity element and let $H = \langle g \rangle$. Since no element of G has order p, then $\gcd(|g|, p) = 1$ by the discussion above. But $|G| = |g| \cdot |G/H|$. Thus, $p \mid |G/H|$. Since $|G/H| < |G|$, then by our choice of G we conclude that G/H has an element of order p, say aH. But the order of aH divides the order of $a \in G$, so p divides the order of a. Then again by our discussion above, this implies that some power of a has order p. Thus, we conclude that G contains an element of order p, a contradiction. □

An immediate consequence of the previous lemma is that a finite p-group must have order a power of p. This follows from that fact that of G is a finite p-group and q is another prime dividing the order of G, then G has an element of order q. But this contradicts the definition of a p-group.

3. Proof of the Fundamental Theorem (Part II)

In this section, we prove the Fundamental Theorem for finite p-groups. This will conclude the proof since we have already shown that we can decompose any finite abelian group into a product of p-groups. We begin with a technical result that will help in the proof of the first proposition.

Lemma 12

Let G be a finite abelian p-group that is not cyclic. Suppose that $g \in G$ has maximal order. If $h \in G \backslash \langle g \rangle$ has smallest possible order, then $|h| = p$.

Proof. Let $g \in G$ be of maximal order in G, say $|g| = p^m$ for some $m \leq n$. Let $x \in G$ with $|x| = p^j$, $j \leq m$. Then $x^{p^m} = (x^{p^j})^{p^{m-j}} = e^{p^{m-j}} = e$. Since G is not cyclic, $G \neq \langle g \rangle$. Choose $h \in G \backslash \langle g \rangle$ where h has smallest possible order, say $|h| = p^\ell$. Since $e \in \langle g \rangle$, then $h \neq e$ and so $\ell > 0$. But $|h^p| = p^{\ell - l}$ and so $|h^p| < |h|$, whence $h^p \in \langle g \rangle$. That is, $h^p = g^r$ for some r.

By the above,

$$(g^r)^{p^{\ell-1}} = (h^p)^{p^{\ell-1}} = h^{p^\ell} = e.$$

Then $|g^r| \leq p^{\ell-1} < p^m$ and so g^r is not a generator for $\langle g \rangle$. Consequently, $r = ps$ for some integer s $(\star)$. Define $a = g^{-s}h$, so $a \notin \langle g \rangle$. Then

$$a^p = g^{-sp}h^p = g^{-r}h^p = h^{-p}h^p = e.$$

Since a has minimal non-trivial order and $a \notin \langle g \rangle$, then $|h| = p$. □

The point $(\star)$ above is subtle. Recall $|g| = p^m$, so We have

$$|g^r| = \frac{|g|}{\gcd(|g|, r)} = \frac{p^m}{\gcd(p^m, r)}.$$

If p does not divide r, then $\gcd(p^m, r) = 1$ so $|g^r| = p^m$ and thus g^r is a generator of $\langle g \rangle$, a contradiction.

Lemma 13

Let G be a finite group, N a normal subgroup of G, and $g \in G$ an element of maximal order in G. If $\langle g \rangle \cap N = \{e\}$, then $|gN| = |g|$ and so gN is an element of maximal order in G/N.

Proof. If $g \in G$ has maximal order n in G, then $|gN| \leq |g|$ by the above. On the other hand, if $|gN| = k$, then $N = (gN)^k = g^kN$ and so $g^k \in N$. But $\langle g \rangle \cap N = \{e\}$ by hypothesis, so $g^k = e$. That is, $|g|$ divides k and so $|gN| = |g|$. By the above argument, no element of G/N has order greater than the maximal order in G and so the conclusion follows. □

Proposition 14: Decomposing a finite abelian p-group

Let G be a finite abelian p-group and suppose that $g \in G$ has maximal order. Then G is the internal direct product of $\langle g \rangle$ and some subgroup K. Hence, $G \cong \langle g \rangle \times K$.

Proof. Because G is a finite abelian p-group, $|G| = p^n$ for some $n \in \mathbb{N}$. The case $n = 1$ (or $n = 0$) implies G is cyclic in which case this result is trivial. Assume now that $n > 1$ and that G is not cyclic. We inductively assume that the lemma holds for all k such that $1 \leq k < n$.

Let $g \in G$ be of maximal order in G, say $|g| = p^m$ for some $m \leq n$. Choose $h \in G \backslash \langle g \rangle$ where h has smallest possible order and set $H = \langle h \rangle$. By a lemma above, $|H| = |h| = p$. Thus, $|G/H| = |G|/|H| = p^n/p = p^{n-1}$ and so we can apply the inductive hypothesis to G/H.

Since H has no nontrivial subgroups, then $\langle g \rangle \cap H = \{e\}$. Thus, gH is an element of G/H of maximal order. By the inductive hypothesis, G/H is the internal direct product of $\langle gH \rangle$ and some subgroup K' of G/H. Then $K' = K/H$ for some subgroup[1] K of G containing H. We claim that G is the internal direct product of $\langle g \rangle$ and K.

Let $b \in G$. Because G/H is the IDP of $\langle gH \rangle$ and K/H, then $bH = (gH)^t(yH)$ with $t \in \mathbb{N}$ and $y \in K$. So, $bH = g^t yH$ and so there exists $h, h' \in H$ such that $bh = g^t yh'$. Thus, $b = g^t(yh'h^{-1})$ and since $H \subset K$, then $yh'h^{-1} \in K$. Thus, $b \in \langle g \rangle K$.

The group G is abelian so it is left only to prove that $\langle g \rangle \cap K = \emptyset$. Suppose there exists $b \in \langle g \rangle \cap K$. Then $bH \in \langle gH \rangle \cap K/H$. Recall that G/H is the IDP of $\langle gH \rangle$ and K/H, so $\langle gH \rangle \cap K/H = H$ (the identity in G/H). Hence, $bH = H$ so $b \in H$. But this implies that $b \in \langle g \rangle \cap H = \{e\}$. It follows that G is the internal direct product of $\langle g \rangle$ and K. □

By Lagrange's Theorem, the subgroup K appearing in the proposition must also be a finite abelian p-group whose order is less than that of G. That is, unless $K = \{e\}$. But in that case G is cyclic and so we conclude that $G \cong \mathbb{Z}_{p^k}$ for some k. Thus, either we are done or we proceed by induction. This concludes the proof of the fundamental theorem of finite abelian groups.

[1] This statement is due to the Correspondence Theorem, which is a consequence of the First Isomorphism Theorem.

Introduction to Rings

Calling $\mathbb{Z}$ a group (under addition) obscures the fact that there are actually two well-defined (binary) operations on $\mathbb{Z}$: addition and multiplication. Moreover, these two operations play nicely together (via the distributive law). Hence, $\mathbb{Z}$ is a *ring.* We will study rings and a few families of interesting examples, drawing many parallels with our knowledge of groups.

1. Rings

Definition: Ring

A *ring* is a set R along with two binary operations (typically $+$ and $\cdot$) satisfying:

(1) $(R, +)$ is an additive abelian group;

(2) $\cdot$ is associative: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in R$;

(3) the left and right distributive properties hold: for all $a, b, c \in R$,

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \quad \text{and} \quad (a + b) \cdot c = (a \cdot c) + (b \cdot c).$$

Remark. Because $(R, +)$ is assumed to be an (additive) abelian group, we denote the additive inverse of an element $a \in R$ by $-a$. If an element $a \in R$ has a multiplicative inverse we denote it by a^{-1}. Note that R need not be closed under multiplicative inverses.

Example. It is now easy to verify that $\mathbb{Z} = (\mathbb{Z}, +, \cdot)$ is a ring formally. Of course, we have already discussed that $(\mathbb{Z}, +)$ is an (additive) abelian group. You learned, quite some time ago, that multiplication of the integers is associative. Moreover, the left and right distributive properties hold. In fact, because multiplication on $\mathbb{Z}$ is commutative, then there is no distinction between these two distributive properties.

Example. In addition to $\mathbb{Z}$, many of the groups we studied earlier in the semester are actually rings under the correct operations. In particular, $\mathbb{R}, \mathbb{Q}, \mathbb{C}$ are all rings under addition and multiplication. Similarly, $\mathbb{Z}_n$ is a ring under addition and multiplication mod n. Finally, $M_n(\mathbb{R})$ is a ring under matrix addition and matrix multiplication.

The next example checks the ring axioms in more detail.

Example (The ring $\mathcal{F}$)**.** Let $\mathcal{F}$ be the set of functions $f : \mathbb{R} \to \mathbb{R}$. We define two operations of $\mathcal{F}$. For any $f, g \in \mathcal{F}$ and $x \in \mathbb{R}$, set

$$(f+g)(x) = f(x) + g(x) \quad \text{and} \quad (fg)(x) = f(x)g(x).$$

These operations are called *pointwise addition* and *pointwise multiplication*, respectively. The sum or product of two real numbers is a real number so both of these operations are binary on $\mathcal{F}$.

We previously showed that $(\mathcal{F}, +)$ is an abelian group under pointwise addition. Recall that the zero function z defined by $z(x) = 0$ for all $x \in \mathbb{R}$ is the additive identity and that the inverse of a function $f \in \mathcal{C}$ is $(-f)$ defined by $(-f)(x) = -(f(x))$ for all $x \in RR$.

Next we check that pointwise multiplication is associative. Let $f, g, h \in \mathcal{F}$ and $x \in \mathbb{R}$. Then, by associativity of real numbers,

$$((fg)h)(x) = (f(x)g(x))h(x) = f(x)(g(x)h(x)) = (f(gh))(x).$$

Thus, $(fg)h = f(gh)$ and so pointwise multiplication is associative.

It remains to check the distributive properties. Keep the notation as above, then by the distributive property of real numbers,

$$(f(g+h))(x) = f(x)(g+h)(x) = f(x)(g(x)+h(x)) = f(x)g(x) + f(x)h(x) = (fg+fh)(x).$$

Thus, $f(g+h) = fg + fh$ so the left distributive property holds. Similarly, the right distributive property holds. It follows that $\mathcal{F}$ is a ring under these operations.

We'll now discuss a variety of properties that a ring may or may not possess.

Definition: Ring with unity, commutative ring, left/right zero divisor, domain, integral domain, unit, division ring, field

Let R be a ring.

(1) A multiplicative identity in R is called a *unity.*

(2) If $ab = ba$ for all $a, b \in R$, then R is said to be *commutative.*

(3) If $ab = 0$ implies $a = 0$ or $b = 0$ for any $a, b \in R$, then R is said to be a *domain.* A commutative domain with unity is an *integral domain.*

(4) An element $u \in R$ is a *unit* if $u^{-1} \in R$. A ring with unity in which every nonzero element is a unit is a *division ring.* A commutative division ring is a *field.*

If R is a division ring, then $(R\backslash\{0\}, \cdot)$ is a group. If R is a field, then $(R\backslash\{0\}, \cdot)$ is an *abelian* group.

Example. We consider these properties for the examples discussed above.

(1) The rings $\mathbb{R}, \mathbb{Q}, \mathbb{C}$ are fields, while $\mathbb{Z}$ is an integral domain but not a field.
(2) The identity matrix is a unity for $M_n(\mathbb{R})$, which is not a domain.
(3) An example of a commutative ring without unity is $2\mathbb{Z}$.
(4) The ring $\mathbb{Z}_n$ is an integral domain if and only if n is prime. In particular, for a prime p, $\mathbb{Z}_p$ is a field.
(5) The ring $\mathcal{F}$ is commutative since for all $f, g \in \mathcal{F}$ and $x \in \mathbb{R}$, the commutative property for $\mathbb{R}$ implies

$$(fg)(x) = f(x)g(x) = g(x)f(x) = (gf)(x),$$

so $fg = gf$. The ring does contain a unity, the function e defined by $e(x) = 1$ for all $x \in \mathbb{R}$. However, $\mathcal{F}$ is not a domain. Consider the piecewise functions:

$$f_1 = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 1, \end{cases} \qquad f_2 = \begin{cases} 1 & \text{if } x < 0 \\ 0 & \text{if } x \geq 1. \end{cases}$$

Then $f_1, f_2 \neq 0$, but $(f_1 f_2)(x) = 0$ for all x, so $f_1 f_2 = z$.

Example (The quaternions)**.** Recall the quaternion group is $Q = \{\pm 1, \pm i, \pm j, \pm k\}$ under multiplication with identity element 1 satisfying $(-1)^2 = 1$, $i^2 = j^2 = k^2 = -1$, and

$$ij = k, \quad ji = -k, \quad jk = i, \quad kj = -i, \quad ki = j, \quad ik = -j.$$

Let $\mathbb{H} = \{a + bi + cj + dk : a, b, c, d \in \mathbb{R}\}$, so $\mathbb{H}$ is a real vector space with basis $\{1, i, j, k\}$ Define addition and multiplication on $\mathbb{H}$ as follows. Let $a_1 + b_1 i + c_1 j + d_1 k, a_2 + b_2 i + c_2 j + d_2 k \in \mathbb{H}$, then

$$(a_1 + b_1 i + c_1 j + d_1 k) + (a_2 + b_2 i + c_2 j + d_2 k) = (a_1 + a_2) + (b_1 + b_2)i + (c_1 + c_2)j + (d_1 + d_2)k$$
$$(a_1 + b_1 i + c_1 j + d_1 k)(a_2 + b_2 i + c_2 j + d_2 k) = \alpha + \beta i + \gamma j + \delta k$$

where

$$\alpha = a_1 a_2 - b_1 b_2 - c_1 c_2 - d_1 d_2 \qquad \beta = a_1 b_2 + a_2 b_1 + c_1 d_2 - d_1 c_2$$
$$\gamma = a_1 c_2 - b_1 d_2 + c_1 a_2 - d_1 b_2 \qquad \delta = a_1 d_2 + b_1 c_2 - c_1 b_2 - d_1 a_2.$$

Thus, $\mathbb{H}$ is a ring with identity. The multiplication operation is noncommutative (since $ij \neq ji$). One can now check that for $a + bi + cj + dk \neq 0$,

$$(a + bi + cj + dk) \left(\frac{a - bi - cj - dk}{a^2 + b^2 + c^2 + d^2} \right) = 1.$$

Thus, $\mathbb{H}$ is a (noncommutative) division ring.

Proposition 1: Basic properties of rings

Let R be a ring with $a, b \in R$. Then

(1) $a0 = 0a = 0$.

(2) $a(-b) = (-a)b = -(ab)$.

(3) $(-a)(-b) = ab$.

Proof. (1) We have, $a0 = a(0+0) = a0 + a0$, so $a0 = 0$.

(2) By (1) and the (left) distributive property, $ab + a(-b) = a(b-b) = a0 = 0$. Thus, $a(-b) = -(ab)$. Similarly $(-a)b = -(ab)$.

(3) From (2), it follows that $(-a)(-b) = -(a(-b)) = -(-(ab)) = ab$, proving (3). □

Proposition 2: Basic properties of rings with unity

Let R be a ring with multiplicative identity 1.

(1) The multiplicative identity is unique.

(2) If $a \in R$ is a unit, then $ab = 0$ and $ba = 0$ imply $b = 0$.

(3) If $a \in R$ is a unit, then its multiplicative inverse is unique.

Proof. (1) We use the same proof as for groups. If $1, 1'$ are identity elements, then $1 = 11' = 1'$, so $1 = 1'$.

(2) For (2), let a^{-1} be an inverse of a and suppose $ab = 0$. $0 = a^{-1}0 = a^{-1}(ab) = (a^{-1}a)b = 1b = b$. Similarly, $ba = 0$ implies $b = 0$.

(3) Suppose that b, c are multiplicative inverses of a. Then $ba = 1 = ca$, so $(b-c)a = 0$. Thus, by (2), $b - c = 0$, or $b = c$. □

Definition: Subring

A subset S of a ring R is a *subring* if S is a ring under the inherited operations from R.

Example. $\mathbb{Z}_n$ is *not* a subring of $\mathbb{Z}$. However, $\mathbb{Z}$ is a subring of $\mathbb{R}$.

Proposition 3: Subring Test

Let R be a ring and S a nonempty subset of R. Then S is a subring of R if and only if for all $s_1, s_2 \in S$, $s_1 s_2 \in S$ and $s_1 - s_2 \in S$.

Proof. Let S be a nonempty subset of R and let $s_1, s_2 \in S$. By the subgroup test, $(S, +)$ is an abelian group if and only if $s_1 - s_2 \in S$. If S is a subring, then $s_1 s_2 \in S$ by closure. Conversely, if $s_1 s_2 \in S$, then multiplication is a binary operation with inherited associativity. □

Example (The subring $2\mathbb{Z}$ of $\mathbb{Z}$)**.** Clearly $2 \in 2\mathbb{Z}$ so $2\mathbb{Z} \neq \emptyset$. Let $2a, 2b \in 2\mathbb{Z}$. Then we have $2a - 2b = 2(a-b) \in 2\mathbb{Z}$ and $(2a)(2b) = 2(2ab) \in 2\mathbb{Z}$. Thus, $2\mathbb{Z}$ is a subring of $\mathbb{Z}$ by the subring test.

Note that while $\mathbb{Z}$ has unity, $2\mathbb{Z}$ does not. It is not necessary for a ring to inherit this.

Example (Upper triangular matrices)**.** Let T be the set of 2×2 real upper-triangular matrices. We will prove that T is a subring of $M_2(\mathbb{R})$.

Clearly, $I_2 \in T$ so $T \neq \emptyset$. Let $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}, \begin{pmatrix} a' & b' \\ 0 & c' \end{pmatrix} \in T$. Then

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} - \begin{pmatrix} a' & b' \\ 0 & c' \end{pmatrix} = \begin{pmatrix} a-a' & b-b' \\ 0 & c-c' \end{pmatrix} \in T$$

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} a' & b' \\ 0 & c' \end{pmatrix} = \begin{pmatrix} aa' & ab'+bc' \\ 0 & cc' \end{pmatrix} \in T.$$

Thus, T is a subring of $M_2(\mathbb{R})$.

Example (Gaussian Integers)**.** The Gaussian integers are defined as $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\} \subset \mathbb{C}$.

Clearly $0 \in \mathbb{Z}[i]$ so $\mathbb{Z}[i] \neq \emptyset$. Let $a + bi, c + di \in \mathbb{Z}[i]$. Then

$$(a+bi) + (c+di) = (a+c) + (b+d)i \in \mathbb{Z}[i]$$

$$(a+bi)(c+di) = (ac-bd) + (ad+bc)i \in \mathbb{Z}[i].$$

Thus, $\mathbb{Z}[i]$ is a subring of $\mathbb{C}$.

Ideals take the place of normal subgroups in ring theory in the sense that they are the right structure to allow us to define factor rings.

Definition: Ideal

An *ideal* in a ring R is a subring I of R such that if $x \in I$ and $r \in R$, then $xr \in I$ and $rx \in I$.

Example (Examples of ideals)**.** (1) Every ring R has two ideals: $\{0\}$ (called the *trivial ideal*) and R itself. An ideal I of R that is not one of these is called a *proper* ideal.

(2) We showed previously that $2\mathbb{Z}$ is a subring of $\mathbb{Z}$. If $r \in \mathbb{Z}$ and $x \in 2\mathbb{Z}$, then $rx = xr \in 2\mathbb{Z}$ because it is even, and so $2\mathbb{Z}$ is an ideal of $\mathbb{Z}$.

For a commutative ring, the conditions $xr \in I$ and $rx \in I$ are the same. For a noncommutative ring R, the story of ideals is a little different. A *left ideal* I is a subring satisfying $rx \in I$ for every $r \in R$, $x \in I$. A *right ideal* I is a subring satisfying $xr \in I$ for every $r \in R$, $x \in I$. A *two-sided ideal* (or just *ideal*) is both a left and right ideal.

The next proposition is a modified version of the Subring Test for Ideals. Note that we do not need to prove, separately, that I is closed under multiplication because we prove that it is closed under multiplication by *any* element of R.

Proposition 4: Ideal Test

Let R be a ring and I a nonempty subset of R. Then I is an ideal of R if and only if for all $a, b \in I$ and all $r \in R$, $a - b \in I$ and $ra, ar \in I$.

Proof. If I is an ideal, then clearly the given conditions hold. Now suppose the conditions holds. Since $I \subset R$, then the multiplicative condition applies to elements only in I. Hence, along with the first condition, I is a subring of R. The last condition now implies that I is an ideal. □

Example (An ideal of $\mathcal{F}$)**.** Define $I = \{f \in \mathcal{F} : f(0) = 0\}$. We claim I is an ideal in $\mathcal{F}$.

Clearly the zero function satisfies $z(0) = 0$, so $z \in I$. Let $f, g \in I$ and $h \in \mathcal{F}$. Then

$$\begin{aligned}(f - g)(0) &= f(0) - g(0) = 0 - 0 = 0 \\ (fh)(0) &= f(0)h(0) = 0h(0) = 0 \\ (hf)(0) &= h(0)f(0) = h(0)0 = 0.\end{aligned}$$

Thus, by the Ideal Test, I is an ideal of $\mathcal{F}$.

Proposition 5: Ideal generated by an element

Let R be a commutative ring and $a \in R$. The set $\langle a \rangle = \{ar : r \in R\}$ is an ideal in R.

Proof. Clearly $a \in \langle a \rangle$ so $\langle a \rangle \neq \emptyset$. If $x, y \in \langle a \rangle$, then $x = ar$ and $y = ar'$ for some $r, r' \in R$. Thus, $x - y = ar - ar' = a(r - r') \in \langle a \rangle$. Now if $s \in R$, then $xs = (ar)s = a(rs) \in \langle a \rangle$ and by commutativity, $sx = s(ar) = a(sr) \in \langle a \rangle$. Thus, $\langle a \rangle$ is an ideal by the ideal test. □

Definition: Principal ideal, PID

The set $\langle a \rangle$ in the previous proposition is called the *principal ideal generated by a*. A integral domain R in which every ideal is principal is called a *principal ideal domain* (PID).

Theorem 6: $\mathbb{Z}$ is a PID

Let I be an ideal in $\mathbb{Z}$. Then I is principal.

Proof. Since the trivial ideal is principal, assume I is a nontrivial ideal in $\mathbb{Z}$. The set $I \cap \mathbb{N} \subset \mathbb{N}$ is nonempty[1] and so has a least element n by the Well-Ordering Principle. We claim $I = \langle n \rangle = n\mathbb{Z}$. Clearly $n\mathbb{Z} \subset I$ so we need only show the opposite inclusion.

Let $a \in I$ be positive. By the division algorithm we have $a = nq + r$ for some $q, r \in \mathbb{Z}$ with $0 \leq r < n$. Then $r = a - nq \in I$, a contradiction unless $r = 0$, so $a \in \langle n \rangle$. If $a \in I$ is negative, then $-a \in \langle n \rangle$ by the above argument so $a = (-1)(-a) \in \langle n \rangle$. □

[1] If $x \in I$ is negative, then $-x = (-1)x \in I \cap \mathbb{N}$ because I is a subring and hence closed under taking additive inverses.

3. Homomorphisms

We now extend notions of homomorphisms, cosets, and factor groups to rings.

Definition: Ring homomorphism, image, kernel, isomorphism

A map $\phi : R \to S$ of rings is a *(ring) homomorphism* if

$$\phi(a+b) = \phi(a) + \phi(b) \quad \text{and} \quad \phi(ab) = \phi(a)\phi(b).$$

An *isomorphism* (of rings) is a bijective homomorphism. The *kernel* and *image* of ϕ are defined, respectively, as the sets

$$\text{Ker}\phi = \{x \in R : \phi(x) = 0_s\} \quad \text{and} \quad \operatorname{Im}\phi = \{\phi(a) : a \in R\}.$$

Example (A map $\mathbb{Z} \to \mathbb{Z}_n$)**.** Fix a positive integer n. Define a map $\phi : \mathbb{Z} \to \mathbb{Z}_n$ by $\phi(a) = a \mod n$. Let $a, b \in \mathbb{Z}$, then

$$\phi(a+b) = (a+b) \mod n = (a \mod n) + (b \mod n) = \phi(a) + \phi(b)$$
$$\phi(ab) = (ab) \mod n = (a \mod n)(b \mod n) = \phi(a)\phi(b).$$

Hence, ϕ is a homomorphism. The image is all of $\mathbb{Z}_n$ and the kernel is $n\mathbb{Z}$.

Example (Evaluation map in $\mathcal{F}$)**.** Fix $\alpha \in \mathbb{R}$. The *evaluation map* $\phi_\alpha : \mathcal{F} \to \mathbb{R}$ is defined by $\phi_\alpha(f) = f(\alpha)$. Let $f, g \in \mathcal{F}$.

$$\phi_\alpha(f+g) = (f+g)(\alpha) = f(\alpha) + g(\alpha) = \phi_\alpha(f) + \phi_\alpha(g),$$
$$\phi_\alpha(fg) = (fg)(\alpha) = f(\alpha)g(\alpha) = \phi_\alpha(f)\phi_\alpha(g).$$

Hence, the evaluation map is a homomorphism.

For $r \in \mathbb{R}$, let c_r be the constant function at r. Then $\phi_\alpha(c_r) = c_r(\alpha) = r$, so ϕ_α is surjective ($\operatorname{Im}\phi = \mathbb{R}$). The kernel of ϕ consists of all functions which have a root at α.

Proposition 7: Kernels are ideals

Let $\phi : R \to S$ be a ring homomorphism. Then $\text{Ker}\phi$ is an ideal of R.

Proof. Since ϕ is a homomorphism $(R, +) \to (S, +)$, then $\phi(0_R) = 0_S$, then $0_R \in \text{Ker}\phi$. Let $x, y \in \text{Ker}\phi$ and $r \in R$. Then

$$\phi(x-y) = \phi(x) - \phi(y) = 0 - 0 = 0$$
$$\phi(rx) = \phi(r)\phi(x) = \phi(r)0 = 0$$
$$\phi(xr) = \phi(x)\phi(r) = 0\phi(r) = 0.$$

Thus, $\text{Ker}\phi$ is an ideal by the ideal test. □

We can define "quotient objects" in rings as well. Here, again, ideals take the place of normal subgroups.

Theorem 8: Multiplication on cosets

Let I be an ideal of a ring R. The factor group R/I is a ring with multiplication defined by

$$(r+I)(s+I) = rs + I.$$

Proof. Because $(R,+)$ is an abelian group, I is a normal subgroup under addition. Thus,R/I is an abelian group under addition. It is only left to show that the multiplication operation is well-defined, that it is associative, and that the distributive properties hold. Let $r+I, s+I \in R/I$. Suppose $r+I = r'+I$ and $s+I = s'+I$ for some $r', s' \in R$. We will show that $r's'+I = rs+I$.

Our hypothesis implies that $r' \in r+I$ so $r' = r+a$ for some $a \in I$. Similarly, $s' = s+b$ for some $b \in I$. Then

$$r's' = (r+a)(s+b) = rs + rb + as + ab.$$

Since I is an ideal, $rb + as + ab \in I$, whence $r's' \in rs + I$. Thus, because cosets are either equal or disjoint, $r's' + I = rs + I$.

Checking associativity and the distributive properties are left as an exercise. □

Definition: Factor ring

Let I be an ideal of a ring R. The set R/I with addition and multiplication operations defined by

$$(a+I)+(b+I) = (a+b)+I \quad \text{and} \quad (a+I)(b+I) = ab+I,$$

respectively, for all $a+I, b+I$, is called the *factor ring* of R by I.

Theorem 9: Ideals are kernels

Let I be an ideal of a ring R. The map $\phi : R \to R/I$ given by $r \mapsto r + I$ is a surjective ring homomorphism with kernel I.

Proof. It is clear that ϕ is a surjective group homomorphism. We need only show that it respects multiplication. Let $r, s \in R$. Then

$$\phi(r)\phi(s) = (r + I)(s + I) = rs + I = \phi(rs). \qquad \square$$

Theorem 10: First Isomorphism Theorem for Rings

Let $\phi : R \to S$ be a ring homomorphism. Then $R/\mathrm{Ker}\phi \cong \phi(R)$.

Proof. Let $K = \mathrm{Ker}\phi$. Define $\psi : R/K \to \phi(R)$ by $\psi(r+K) = \phi(r)$. From the corresponding result in group theory, we have that ϕ is a well-defined (group) homomorphism with respect to addition. It remains only to check that ψ respects multiplication. Let $r + K, s + K \in R/K$. Then

$$\psi((r + K)(s + K)) = \psi(rs + K) = \phi(rs) = \phi(r)\phi(s) = \psi(r + K)\psi(s + K).$$

This proves the claim. □

4. Polynomials

Definition: Polynomial, coefficients, degree, leading coefficient, monic

A *polynomial over R in indeterminate x* is an expression of the form

$$f(x) = \sum_{i=0}^{n} a_i x^i$$

where $a_i \in R$. The elements a_i are the *coefficients* of f. The *degree* of f is the largest m such that $0 \neq a_m$ if such an m exists. We write $\deg(f) = m$ and say a_m is the *leading coefficient.* Otherwise $f = 0$ and we set $\deg(f) = -\infty$. A nonzero polynomial with leading coefficient 1 is called *monic.*

We denote the set of polynomials over R by $R[x]$.

Let $p(x), q(x) \in R[x]$ be nonzero polynomials over R with degrees n and m, respectively. Write

$$\begin{aligned} p(x) &= a_0 + a_1 x + \cdots + a_n x^n \\ q(x) &= b_0 + b_1 x + \cdots + b_m x^m. \end{aligned}$$

The polynomials $p(x)$ and $q(x)$ are equal ($p(x) = q(x)$) if and only if $n = m$ and $a_i = b_i$ for all i. We can define two binary operations, addition and multiplication, on $R[x]$.

(Addition) Suppose $n \geq m$ and set $b_i = 0$ for $i > m$. (This is similar if $m > n$.) Then

$$p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n \in R[x].$$

Since R is a ring, then $a_i + b_i \in R$ for all i. Hence, $p(x) + q(x) \in R[x]$, so addition is a binary operation on $R[x]$.

(Multiplication) Set

$$p(x)q(x) = c_0 + c_1 x + \cdots + c_{m+n} x^{m+n},$$

where

$$c_i = \sum_{k=0}^{i} a_0 b_i + a_1 b_{i-1} + \cdots + a_{i-1} b_1 + a_i b_0.$$

Since R is a ring, then $a_i b_j \in R$ for all i. Thus $p(x)q(x) \in R[x]$ for all i. So, multiplication is a binary operation on $R[x]$.

Example. Suppose $p(x) = 3 + 2x^3$ and $q(x) = 2 - x^2 + 4x^4$ are polynomials in $\mathbb{Z}[x]$. Note that $\deg(p(x)) = 3$ and $\deg(q(x)) = 4$. Compute $p(x) + q(x)$ and $p(x)q(x)$.

$$p(x) + q(x) = (3 + 2) + (0 + 0)x + (0 - 1)x^2 + (2 + 0)x^3 + (0 + 4)x^4 = 5 - x^2 + 2x^3 + 4x^4$$

and

$$p(x)q(x) = (3 + 2x^3)(2 - x^2 + 4x^4) = 6 - 3x^2 + 4x^3 + 12x^4 - 2x^5 + 8x^7.$$

Theorem 11: Polynomial ring over R

Let R be a ring. The set $R[x]$ is a ring under the operations of (polynomial) addition and (polynomial) multiplication.

Proof. Above we showed that addition and multiplication are binary operations. It is easy to check that $(R[x], +)$ is an abelian group where the zero polynomial is the (additive) identity. Associativity of multiplication and the distributive property are easy (albeit annoying) proofs. The details are left as an exercise. □

Definition: Polynomial ring over R

Let R be a ring. The set $R[x]$ with the operations of (polynomial) addition and (polynomial) multiplication is called the *polynomial ring over R*.

If y is another indeterminate, then it makes sense to define $(R[x])[y]$. Note that $(R[x])[y] \cong (R[y])[x]$. Both of these rings will be identified with the ring $R[x, y]$ and call this the *ring of polynomials in two indeterminates x and y with coefficients in R*. Similarly (or inductively), one can then define the *ring of polynomials in n indeterminates with coefficients in R*, denoted $R[x_1, \dots, x_n]$.

Proposition 12: Properties of polynomial rings

Let R be a ring.

(1) If R is commutative, then so is $R[x]$.

(2) If R has (multiplicative) identity, then so does $R[x]$.

(3) If R is an integral domain, then so is $R[x]$.

Proof. The first two are left as easy exercises.

For (3), let $p(x), q(x) \in R[x]$ be nonzero polynomials with degrees n and m, respectively. Write

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$
$$q(x) = b_0 + b_1 x + \cdots + b_m x^m.$$

Then the leading term of $p(x)q(x)$ is $a_n b_m x^{n+m}$. By hypothesis, $a_n, b_m \neq 0$ and because R is an integral domain, $a_n b_m \neq 0$ so $p(x)q(x) \neq 0$. □

Remark. Let R be an integral domain. What we proved in part (3) of the proposition is

$$\deg(p(x)q(x)) = \deg(p(x)) + \deg(q(x)),$$

for *any* polynomials $p(x), q(x) \in R[x]$. This justifies why we set $\deg(0) = -\infty$.

Let S be a commutative ring with identity and R a subring of S containing 1. Let $\alpha \in S$. For $p(x) = a_0 + a_1 x + \cdots + a_n x^n$, we set

$$p(\alpha) = a_0 + a_1\alpha + \cdots + a_n\alpha^n \in S.$$

We can define a map $\phi_\alpha : R[x] \to S$ by $\phi_\alpha(p(x)) = p(\alpha)$. This is the *evaluation homomorphism* at α discussed above, but specialized to polynomial rings.

Definition: Root of a polynomial

We say $\alpha \in R$ is a *root* (or *zero*) of $p(x) \in R[x]$ if $\phi_\alpha(p(x)) = 0$.

Example (The polynomial x^2+1 in $\mathbb{R}[x]$)**.** The ring $\mathbb{R}$ is a subring of $\mathbb{C}$. Consider the evaluation map $\phi_i : \mathbb{R}[x] \to \mathbb{C}$. Then

$$\phi_i(x^2+1) = i^2 + 1 = 0,$$

so i is a root of x^2+1. Similarly, $\phi_{-i}(x^2+1) = 0$.

The codomain of the evaluation homomorphism is important. Indeed, if we consider only maps $\phi : \mathbb{R}[x] \to \mathbb{R}$, then x^2+1 has no roots.

Example (The polynomial x^2+x in $\mathbb{Z}_6[x]$)**.** Consider the polynomial x^2+x in $\mathbb{Z}_6[x]$. Then

$$\phi_2(x^2+x) = 2^2 + 2 = 6 \equiv 0 \mod 6.$$

Thus, 2 is a root of x^2+x in $\mathbb{Z}_6$. Similarly,

$$\phi_3(x^2+x) = 3^2 + 3 = 12 \equiv 0 \mod 6.$$

Thus, 3 is also a root of x^2+x. One can check that

$$(x-2)(x-3) = x^2 - 2x - 3x + 6 = x^2 - 5x + 6 \equiv x^2 + x \mod 6.$$

We will now prove a version of the division algorithm for polynomials. This will be applied to determine when polynomials are irreducible over certain rings.

Theorem 13: Division algorithm for polynomials

Let F be a field and $f(x), g(x) \in F[x]$ with $g(x) \neq 0$. Then there exist unique polynomials $q(x), r(x) \in F[x]$ such that

$$f(x) = g(x)q(x) + r(x),$$

where $\deg r(x) < \deg g(x)$.

Proof. For simplicity throughout we will write $f = f(x), g = g(x), etc.$ Set $\deg f = n$ and $\deg g = m$. If $n < m$, then let $q = 0$ and $r = f$. Note that this may happen if $f = 0$.

Now suppose $m \leq n$. Write

$$f = a_0 + a_1 x + \cdots + a_n x^n$$

$$g = b_0 + b_1 x + \cdots + b_m x^m.$$

Because $b_m \neq 0$ and F is a field, b_m^{-1} exists.

If $n = 0$, then $m = 0$ so $f = a_0$ and $g = b_0$. Set $q = a_0 b_0^{-1}$ and $r = 0$.

We proceed inductively. That is, assume the division algorithm holds when $\deg f < n$. Note that $f - a_n b_m^{-1} x^{n-m} g$ has degree strictly less than n. Hence, there exists q_0, r such that $f - a_n b_m^{-1} x^{n-m} g = q_0 g + r$ with $\deg r < \deg g$. This implies that

$$f = (q_0 + a_n b_m^{-1} x^{n-m})g + r.$$

Setting $q = q_0 + a_n b_m^{-1} x^{n-m}$ completes the existence part of the proof.

For uniqueness, assume there exists $q, q', r, r' \in F[x]$ such that $f = qg + r = q'g + r'$ with $\deg r, \deg r' < \deg g$. Then $g(q - q') = r' - r$. If $q \neq q'$, then because $g \neq 0$ we have

$$\deg(r' - r) = \deg(g(q - q')) \geq \deg g.$$

This is a contradiction since both r' and r have degrees less than $\deg g$. Thus, $q = q'$ and so $r = r'$. □

Now we consider some consequences of the Division Algorithm.

Definition: Factor

Let F be a field. We say $q(x)$ is a *factor* of $p(x)$ if $q(x)$ divides $p(x)$.

Corollary 14: Factors correspond to roots

Let F be a field. An element $\alpha \in F$ is a root of $p(x) \in F[x]$ if and only if $(x-\alpha)$ divides $p(x)$.

Proof. By the division algorithm, $p(x) = (x-\alpha)q(x) + r(x)$ for some $q(x), r(x) \in F[x]$ with $\deg r(x) < \deg(x-\alpha) = 1$. Applying the evaluation homomorphism we get $p(\alpha) = (\alpha-\alpha)q(\alpha) + r(\alpha) = r(\alpha)$. If $x-\alpha$ divides $p(x)$, then $r(x) = 0$ so $p(\alpha) = 0$ and α is a root of $p(x)$. Conversely, if α is a root, then $r(\alpha) = 0$. But $\deg r < 1$ and since $r(x)$ cannot be a nonzero constant, then $r(x) = 0$. □

Corollary 15: Degree is less than or equal to number of roots

Let F be a field. A nonzero polynomial $p(x) \in F[x]$ of degree n can have at most n distinct roots in F.

Proof. First note that a polynomial of degree 0 has no roots. Let $p(x) \in F[x]$ have degree n. If $n = 1$, then $p(x)$ has 1 root. We proceed by induction on $\deg p(x)$. Suppose all polynomials of degree m, $m \geq 1$, have at most m distinct roots. Let $\deg p(x) = m+1$ and let α be a root of $p(x)$. By the previous corollary, $p(x) = (x-\alpha)q(x)$ for some $q(x)$ with $\deg q(x) = m$. Thus, m has at most m distinct roots and so $p(x)$ has at most $m+1$ distinct roots. □

The next proof is very similar to the corresponding result for $\mathbb{Z}$.

Corollary 16: Polynomial ring over a field is a PID

Let F be a field and I an ideal in $F[x]$. Then I is principal.

Proof. Let I be a nonzero ideal in $F[x]$. Set $S = \{\deg p(x) : p(x) \in I, p(x) \geq 0\}$. Since $I \neq 0$, then $S \neq \emptyset$ and $S \subset \mathbb{N}$. Thus, by the Well-Ordering Principal, S has a least element, d. Let $g(x) \in I$ be a polynomial of degree d. We claim $I = \langle g(x) \rangle$. It is clear that $\langle g(x) \rangle \subset I$ and so it is left only to prove the reverse inclusion.

Let $f(x) \in I$. If $\deg f(x) = d$, then either $f(x) = ag(x)$ for some $a \in F$ or else $\deg(f(x) - g(x)) < d$, a contradiction since $f(x) - g(x) \in I$. Now assume $\deg f(x) > d$. By the division algorithm, $f(x) = g(x)q(x) + r(x)$ for some $q(x), r(x) \in I$ with $\deg r(x) < \deg g(x)$. But then $r(x) = f(x) - g(x)q(x) \in I$. If $\deg r(x) \geq 0$, then this contradicts the minimality of d. Thus, $r(x) = 0$ and $f(x) \in \langle g(x) \rangle$ □

Warning. The above result *does not* hold for $F[x, y]$. In particular, the ideal $\langle x, y \rangle$ is not principal.