Abstract Algebra

PDF generated using the open source mwlib toolkit. See http://code.pediapress.com/ for more information.
PDF generated at: Fri, 13 Dec 2013 04:15:53 UTC

Contents

Articles

Algebras

Further Abstract Algebra

References

Article Licenses

Introduction

Sets

Sets

In the so-called naive set theory, which is sufficient for the purpose of studying abstract algebra, the notion of a set is not rigorously defined. We describe a set as a well-defined aggregation of objects, which are referred to as members or elements of the set. If a certain object is an element of a set, it is said to be contained in that set. The elements of a set can be anything at all, but in the study of abstract algebra, elements are most frequently numbers or mathematical structures. The elements of a set completely determine the set, and so sets with the same elements as each other are equal. Conversely, sets which are equal contain the same elements.

For an element x and a set A , we can say either $x \in A$, that is, x is contained in A , or $x \notin A$, that is, x is not contained in A . To state that multiple elements $a, b, c, \ldots, n$ are contained in A , we write $a, b, c, \ldots, n \in A$.

The axiom of extensionality

Using this notation and the symbol $\longrightarrow$, which represents logical implication, we can restate the definition of equality for two sets A and B as follows:

$A = B$ if and only if $x \in A \rightarrow x \in B$ and $x \in B \rightarrow x \in A$.

This is known as the axiom of extensionality.

Comprehensive notation

If it is not possible to list the elements of a set, it can be defined by giving a property that its elements are sole to possess. The set of all objects x with some property $Q(x)$ can be denoted by $\{x : Q(x)\}$. Similarly, the set of all elements x of a set A with some property $Q(x)$ can be denoted by $\{x \in A : Q(x)\}$. The colon : here is read as "such that". The vertical bar | is synonymous with the colon in similar contexts. This notation will appear quite often in the rest of this book, so it is important for the reader to familiarize himself with this now.

As an example of this notation, the set of integers can be written as $\mathbb{Z} = \{x : x \text{ is an integer}\}$, and the set of even integers can be written as $\{x \in \mathbb{Z} : x \text{ is even}\}$.

Set inclusion

For two sets A and B , we define set inclusion as follows: A is included in, or a subset of, B , if and only if every member of A is a member of B . In other words,

$A \subseteq B \Leftrightarrow x \in A \rightarrow x \in B$,

where the symbol $\subseteq$ denotes "is a subset of".

By the aforementioned axiom of extensionality, one finds that $A = B \Leftrightarrow A \subseteq B \text{ and } B \subseteq A$.

The empty set

One can define an empty set, written $\emptyset$, such that $\forall x(x \notin \emptyset)$, where $\forall$ denotes universal quantification (read as "for all" or "for every"). In other words, the empty set is defined as the set which contains no elements. The empty set can be shown to be unique.

Since the empty set contains no elements, it can be shown to be a subset of every set. Similarly, no set but the empty set is a subset of the empty set.

Proper set inclusion

For two sets A and B , we can define proper set inclusion as follows: A is a proper subset of B if and only if A is a subset of B , and A does not equal B . In other words, there is at least one member in B not contained in A

$$A \subset B \Leftrightarrow (A \subseteq B \wedge A \neq B),$$

where the symbol $\subset$ denotes "is a proper subset of" and the symbol $\wedge$ denotes logical *and*.

Cardinality of sets

The cardinality of a set A , denoted by $|A|$, can be said informally to be a measure of the number of elements in A . However, this description is only rigorously accurate for finite sets. To find the cardinality of infinite sets, more sophisticated tools are needed.

Set intersection

For sets A and B , we define the intersection of A and B by the set $A \cap B$ which contains all elements which are common to both A and B . Symbolically, this can be stated as follows:

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\} .$$

Because every element of $A \cap B$ is an element of A and an element of B , $A \cap B$ is, by the definition of set inclusion, a subset of A and B .

If the sets A and B have no elements in common, they are said to be disjoint sets. This is equivalent to the statement that A and B are disjoint if $A \cap B = \emptyset$.

Set intersection is an associative and commutative operation; that is, for any sets A , B , and C , $(A \cap B) \cap C = A \cap (B \cap C)$ and $A \cap B = B \cap A$.

By the definition of intersection, one can find that $A \cap \emptyset = \emptyset$ and $A \cap A = A$. Furthermore, $A \subseteq B \Leftrightarrow A \cap B = A$.

One can take the intersection of more than two sets at once; since set intersection is associative and commutative, the order in which these intersections are evaluated is irrelevant. If A_i are sets for every $i \in I$, we can denote the intersection of all A_i by

$$\bigcap_{i \in I} A_i = \{x \mid (\forall i \in I) x \in A_i\}$$

In cases like this, I is called an index set, and A_i are said to be indexed by I .

In the case of $I = \{1, 2, ..., n\}$ one can either write $A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n$ or

$$\bigcap_{i=1}^{n} A_i .$$

Set union

For sets A and B, we define the union of A and B by the set $A \cup B$ which contains all elements which are in either A or B or both. Symbolically,

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\} .$$

Since the union $A \cup B$ of sets A and B contains the elements of A and B, $A \subseteq A \cup B$ and $B \subseteq A \cup B$.

Like set intersection, set union is an associative and commutative operation; for any sets A, B, and C, $(A \cup B) \cup C = A \cup (B \cup C)$ and $A \cup B = B \cup A$.

By the definition of union, one can find that $A \cup \emptyset = A \cup A = A$. Furthermore, $A \subseteq B \Leftrightarrow A \cup B = B$.

Just as with set intersection, one can take the union of more than two sets at once; since set union is associative and commutative, the order in which these unions are evaluated is irrelevant. Let A_i be sets for all $i \in I$. Then the union of all the A_i is denoted by

$$\bigcup_{i \in I} A_i = \{x \mid (\exists i \in I) x \in A_i\}$$

For the union of a finite number of sets $A_1, A_2, A_3, \ldots, A_n$, that is, $I = \{1, 2, ..., n\}$ one can either write $A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n$ or abbreviate this as

$$\bigcup_{i=1}^{n} A_i .$$

Distributive laws

Set union and set intersection are distributive with respect to each other. That is,

$$A \cap (B \cup C) = (B \cup C) \cap A = (A \cap B) \cup (A \cap C) \text{ and}$$
$$A \cup (B \cap C) = (B \cap C) \cup A = (A \cup B) \cap (A \cup C).$$

Cartesian product

The Cartesian product of sets A and B, denoted by $A \times B$, is the set of all ordered pairs which can be formed with the first object in the ordered pair being an element of A and the second being an element of B. This can be expressed symbolically as

$$A \times B = \{(x, y) \mid x \in A \text{ and } y \in B\} .$$

Since different ordered pairs result when one exchanges the objects in the pair, the Cartesian product is not commutative. The Cartesian product is also not associative. The following identities hold for the Cartesian product for any sets A, B, C, D:

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

The Cartesian product of any set and the empty set yields the empty set; symbolically, for any set A, $A \times \emptyset = \emptyset \times A = \emptyset$.

The Cartesian product can easily be generalized to the *n*-ary Cartesian product, which is also denoted by $\times$. The *n*-ary Cartesian product forms ordered *n*-tuples from the elements of n sets. Specifically, for sets $A_1, A_2, A_3, \ldots, A_n$,

$$A_1 \times A_2 \times A_3 \times \cdots \times A_n = \{(x_1, \ldots, x_n) : x_i \in A_i\} .$$

This can be abbreviated as

$$\prod_{i=1}^{n} A_i \,.$$

In the *n*-ary Cartesian product, each x_i is referred to as the i -th coordinate of $(x_1, \ldots, x_n)$.

In the special case where all the factors are the same set A , we can generalize even further. Let A^Ωbe the set of all functions $f\ :\ \Omega \longrightarrow A$. Then, in analogy with the above, $A^\Omega$is effectively the set of " Ω-tuples" of elements in A , and for each such function $f$and each $i \in \Omega$, we call $f(i)$the i -th coordinate of f. As one might expect, in the simple case when $\Omega = \{1, 2, ..., n\}$ for an integer n , this construction is equivalent to $\prod_{i=1}^{n} A$, which we can abbreviate further as A^n . We also have the imporant case of $\Omega = \mathbb{N}$, giving rise to the set of all infinite sequences of elements of A , which we can denote by A^∞. We will need this contruction later, in paticular when dealing with polynomial rings.

Disjoint union

Let A and B be any two sets. We then define their disjoint union, denoted $A \coprod B$ to be the following: First create copies A' and B' of A and B such that $A' \cap B' = \emptyset$. Then define $A \coprod B = A' \cup B'$. Notice that this definition is not explicit, like the other operations defined so far. The definition does not output a single set, but rather a family of sets. However, these are all "the same" in a sense which will be defined soon. In other words, there exists bijective functions between them.

Luckily, if a disjoint union is needed for explicit computation, one can easily be constructed, for example $A \coprod B = (\{0\} \times A) \cup (\{1\} \times B)$.

Set difference

The set difference, or relative set complement, of sets A and B, denoted by $A \setminus B$, is the set of elements contained in A which are not contained in B . Symbolically,

$$A \setminus B = \{x \mid x \in A \text{ and } x \notin B\}\,.$$

By the definition of set difference, $A \setminus B \subseteq A$.

The following identities hold for any sets A, B, C :

$$A \setminus B = A \Leftrightarrow A \cap B = \emptyset\,,$$

$$A \setminus B = \emptyset \Leftrightarrow A \subseteq B\,,$$

$$A \setminus B = A \setminus C \Leftrightarrow A \cap B = A \cap C\,,$$

$$A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C),$$

$$A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C),$$

$$A \setminus (B \setminus C) = (C \cap A) \cup (A \setminus B),$$

$$A \setminus \emptyset = A\,,$$

$$A \setminus A = \emptyset\,,$$

$$\emptyset \setminus A = \emptyset\,.$$

The set difference of two Cartesian products can be found as $(A \times B) \setminus (C \times D) = [A \times (B \setminus D)] \cup [(A \setminus C) \times B]$.

The universal set and set complements

We define some arbitrary set U for which every set under consideration is a subset of U as the universal set, or universe. The complement of any set is then defined to be the set difference of the universal set and that set. That is, for any set $A \subseteq U$, the complement of A is given by $A^C = U \setminus A$. The following identities involving set complements hold true for any sets A and B:

De Morgan's laws for sets:

$$(A \cup B)^C = A^C \cap B^C,$$
$$(A \cap B)^C = A^C \cup B^C,$$

Double complement law:

$$A^{CC} = A,$$

Complement properties:

$$A \cup A^C = U,$$
$$A \cap A^C = \emptyset,$$
$$\emptyset^C = U,$$
$$U^C = \emptyset,$$
$$A \subseteq B \rightarrow B^C \subseteq A^C.$$

The set complement can be related to the set difference with the identities $A \setminus B = A \cap B^C$ and $(A \setminus B)^C = A^C \cup B$.

Symmetric difference

For sets A and B, the symmetric set difference of A and B, denoted by $A \bigtriangleup B$ or by $A \ominus B$, is the set of elements which are contained either in A or in B but not in both of them. Symbolically, it can be defined as

$$A \bigtriangleup B = \{x : (x \in A) \oplus (x \in B)\}.$$

More commonly, it is represented as

$$A \bigtriangleup B = (A \setminus B) \cup (B \setminus A)$$ or as

$$A \bigtriangleup B = (A \cup B) \setminus (A \cap B).$$

The symmetric difference is commutative and associative so that $A \bigtriangleup B = B \bigtriangleup A$ and $(A \bigtriangleup B) \bigtriangleup C = A \bigtriangleup (B \bigtriangleup C)$. Every set is its own symmetric-difference inverse, and the empty set functions as an identity element for the symmetric difference, that is, $A \bigtriangleup A = \emptyset$ and $A \bigtriangleup \emptyset = A$. Furthermore, $A \bigtriangleup B = \emptyset$ if and only if $A = B$.

Set intersection is distributive over the symmetric difference operation. In other words, $A \cap (B \bigtriangleup C) = (A \cap B) \bigtriangleup (A \cap C)$.

The symmetric difference of two set complements is the same as the symmetric difference of the two sets: $A^C \bigtriangleup B^C = A \bigtriangleup B$.

Notation for specific sets

Commonly-used sets of numbers in mathematics are often denoted by special symbols. The conventional notations used in this book are given below.

Natural numbers with 0: $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ or $\mathbb{N} = \{0, 1, 2, \ldots\}$
Natural numbers without 0: $\mathbb{N}^* = \{1, 2, 3, \ldots\}$
Integers: $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$
Rational numbers: $\mathbb{Q} = \left\{ \frac{p}{q} : p \in \mathbb{Z} \text{ and } q \in \mathbb{N}^* \right\}$
Real numbers: $\mathbb{R}$
Complex numbers: $\mathbb{C}$

Equivalence relations and congruence classes

We often wish to describe how two mathematical entities within a set are related. For example, if we were to look at the set of all people on Earth, we could define "is a child of" as a relationship. Similarly, the $\geq$ operator defines a relation on the set of integers. A binary relation, hereafter referred to simply as a relation, is a binary proposition defined on any selection of the elements of two sets.

Formally, a relation is any arbitrary subset of the Cartesian product between two sets X and Y so that, for a relation R , $R \subseteq X \times Y$. In this case, X is referred to as the domain of the relation and Y is referred to as its codomain. If an ordered pair (x, y) is an element of R (where, by the definition of R , $x \in X$ and $y \in Y$), then we say that x is related to y by R . We will use $R(x)$ to denote the set

$$\{y \in Y : (x, y) \in R\} .$$

In other words, $R(x)$ is used to denote the set of all elements in the codomain of R to which some x in the domain in related.

Equivalence relations

To denote that two elements x and y are related for a relation R which is a subset of some Cartesian product $X \times X$, we will use an infix operator. We write $x \sim y$ for some $x, y \in X$ and $(x, y) \in R$.

There are very many types of relations. Indeed, further inspection of our earlier examples reveals that the two relations are quite different. In the case of the "is a child of" relationship, we observe that there are some people A,B where neither A is a child of B, nor B is a child of A. In the case of the $\geq$ operator, we know that for any two integers $m, n \in Z$ exactly one of $m \geq n$ or $n > m$ is true. In order to learn about relations, we must look at a smaller class of relations.

In particular, we care about the following properties of relations:

- Reflexivity: A relation $R \subseteq X \times X$ is reflexive if $a \sim a$ for all $a \in X$.
- Symmetry: A relation $R \subseteq X \times X$ is symmetric if $a \sim b \implies b \sim a$ for all $a, b \in X$.
- Transitivity: A relation $R \subseteq X \times X$ is transitive if $a \sim b \wedge b \sim c \implies a \sim c$ for all $a, b, c \in X$.

One should note that in all three of these properties, we quantify across *all* elements of the set X .

Any relation $R \subseteq X \times X$ which exhibits the properties of reflexivity, symmetry and transitivity is called an **equivalence relation** on X . Two elements related by an equivalence relation are called equivalent under the equivalence relation. We write $a \sim_R b$ to denote that a and b are equivalent under R . If only one equivalence relation is under consideration, we can instead write simply $a \sim b$. As a notational convenience, we can simply say that $\sim$ is an equivalence relation on a set X and let the rest be implied.

Example: For a fixed integer p, we define a relation $\sim_p$ on the set of integers such that $a \sim_p b$ if and only if $a - b = kp$ for some $k \in Z$. Prove that this defines an equivalence relation on the set of integers.
Proof:

- Reflexivity: For any $a \in X$, it follows immediately that $a - a = 0 = 0p$, and thus $a \sim_p a$ for all $a \in G$.
- Symmetry: For any $a, b \in X$, assume that $a \sim_p b$. It must then be the case that $a - b = kp$ for some integer k, and $b - a = (-k)p$. Since k is an integer, $-k$ must also be an integer. Thus, $a \sim_p b \implies b \sim_p a$ for all $a, b \in G$.
- Transitivity: For any $a, b, c \in X$, assume that $a \sim_p b$ and $b \sim_p c$. Then $a - b = k_1p$ and $b - c = k_2p$ for some integers k_1, k_2. By adding these two equalities together, we get $(a - b) + (b - c) = (k_1p) + (k_2p) \Leftrightarrow a - c = (k_1 + k_2)p$, and thus $a \sim_p c$.

Q.E.D.

Remark. In elementary number theory we denote this relation $a \equiv b (\text{mod } p)$ and say a *is equivalent to* b *modulo* p.

Equivalence classes

Let $\sim$ be an equivalence relation on X. Then, for any element $a \in X$ we define the equivalence class of a as the subset $[a] \subseteq X$ given by

$$[a] = \{b \in X | a \sim b\}$$

Theorem: $b \in [a] \implies [b] = [a]$

Proof: Assume $b \in [a]$. Then by definition, $a \sim b$.

- We first prove that $[b] \subseteq [a]$. Let p be an arbitrary element of $[b]$. Then $p \sim b$ by definition of the equivalence class, and $p \sim a$ by transitivity of equivalence relations. Thus, $p \in [b] \implies p \in [a]$ and $[b] \subseteq [a]$.
- We now prove that $[a] \subseteq [b]$ Let q be an arbitrary element of $[a]$. Then, by definition $q \sim a$. By transitivity, $q \sim b$, so $q \in [b]$. Thus, $q \in [a] \implies q \in [b]$ and $[a] \subseteq [b]$.

As $[a] \subseteq [b]$ and as $[b] \subseteq [a]$, we have $[b] = [a]$.
Q.E.D.

Partitions of a set

A partition of a set X is a disjoint family of sets X_i, $i \in I$, such that $\bigcup_{i \in I} X_i = X$.

Theorem: An equivalence relation $\sim$ on X induces a unique partition of X, and likewise, a partition induces a unique equivalence relation on X, such that these are equivalent.

Proof: (Equivalence relation induces Partition): Let P be the set of equivalence classes of $\sim$. Then, since $a \in [a]$ for each $a \in X$, $\cup P = X$. Furthermore, by the above theorem, this union is disjoint. Thus the set of equivalence relations of $\sim$ is a partition of X.
(Partition induces Equivalence relation): Let X_i, $i \in I$ be a partition of X. Then, define $\sim$ on X such that $a \sim b$ if and only if both a and b are elements of the same X_i for some $i \in I$. Reflexivity and symmetry of $\sim$ is immediate. For transitivity, if $a, b \in X_i$ and $b, c \in X_i$ for the same $i \in I$, we neccesarily have $a, c \in X_i$, and transitivity follows. Thus, $\sim$ is an equivalence relation with X_i, $i \in I$ as the equivalence classes.
Lastly obtaining a partition P from $\sim$ on X and then obtaining an equivalence equation from P obviously returns $\sim$ again, so $\sim$ and P are equivalent structures.

Q.E.D.

Quotients

Let $\sim$ be an equivalence relation on a set X . Then, define the set $X/\sim$ as the set of all equivalence classes of X . In order to say anything interesting about this construction we need more theory yet to be developed. However, this is one of the most important constructions we have, and one that will be given the much attention throughout the book.

Functions

Definition

A function f is a triplet (A, B, G) such that:

- A is a set, called the domain of f
- B is a set, called the codomain of f
- G is a subset of $A \times B$, called the graph of f

In addition the following two properties hold:

1. $\forall x \in A, \exists y \in B \mid (x, y) \in G$.
2. $\forall x \in A, y \in B, y' \in B, \text{ then } (x, y) \in G \text{ and } (x, y') \in G \Rightarrow y = y'$.

$\forall x \in A$ we write $\mathrm{f}(x)$ for the unique $y \in B$ such that $(x, y) \in G$.

We say that f is a function from A to B , which we write:

$$\mathrm{f} : A \to B$$

Example

Let's consider the function from the reals to the reals which squares its argument. We could define it like this:

$$\mathrm{f} : \mathbb{R} \to \mathbb{R}$$

$$\mathrm{f} : x \mapsto x^2$$

Remark

As you see in the definition of a function above, the domain and codomain are an integral part of the definition. In other words, even if the values of $\mathrm{f}(x)$ don't change, changing the domain or codomain changes the function.

Let's look at the following four functions.

The function:

$$\mathrm{f}_1 : \mathbb{R} \to \mathbb{R}$$

$$\mathrm{f}_1 : x \mapsto x^2$$

is neither injective nor surjective (these terms will be defined later).

The function:

$$\mathrm{f}_2 : \mathbb{R} \to \mathbb{R}_{\geq 0}$$

$$\mathrm{f}_2 : x \mapsto x^2$$

is not injective but surjective.

The function:

$$\mathrm{f}_3 : \mathbb{R}_{\geq 0} \to \mathbb{R}$$

$$\mathrm{f}_3 : x \mapsto x^2$$

is injective but not surjective.

The function:

$$\mathrm{f}_4 : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$$

$$\mathrm{f}_4 : x \mapsto x^2$$

is injective and surjective

As you see, all four functions have the same mapping but all four are different. That's why just giving the mapping is insufficient; a function is only defined if its domain and codomain are known.

Image and preimage

For a set E, we write $\mathcal{P}(E)$ for the set of subsets of E.

Let $\mathrm{f} : A \to B$. We will now define two related functions.

The **image** function:

$$\mathrm{f} : \mathcal{P}(A) \to \mathcal{P}(B), S \subseteq A \mapsto \{\mathrm{f}(x) \mid x \in S\}$$

The **preimage** function:

$$\mathrm{f}^{-1} : \mathcal{P}(B) \to \mathcal{P}(A), T \subseteq B \mapsto \{x \in A \mid \mathrm{f}(x) \in T\}$$

Note that the image and preimage are written respectively like $\mathrm{f}$and its inverse (if it exists). There is however no ambiguity because the domains are different. Note also that the image and preimage are not necessarily inverse of one another. (See the section on bijective functions below).

We define $\mathrm{Im}_\mathrm{f} := \mathrm{f}(A)$, which we call the **image** of f.

For any $y \in B$, we call $\mathrm{f}^{-1}(\{y\})$ the **support** of y.

Example

Let's take again the function:

$$\mathrm{f} : \mathbb{R} \to \mathbb{R}$$

$$\mathrm{f} : x \mapsto x^2$$

Let's consider the following examples:

$$\mathrm{f}^{-1}(\{4\}) = \{-2, 2\}$$

$$\mathrm{f}^{-1}(\mathbb{R}_{<0}) = \emptyset$$

$$\mathrm{f}(\mathbb{R}_{\geq 0}) = \mathbb{R}_{\geq 0}$$

Further definitions

Let $\mathrm{f} : B \to C$ and $\mathrm{g} : A \to B$. We define $\mathrm{f} \circ \mathrm{g} : A \to C$ by $(\mathrm{f} \circ \mathrm{g})(x) := \mathrm{f}(\mathrm{g}(x))$, which we call the **composition** of $\mathrm{f}$and g.

Let A be a set. We define the **identity function** on A as

$$\mathrm{id}_\mathrm{A} : A \to A, x \mapsto x$$

Properties

Definition: A function $\mathrm{f} : A \to B$ is **injective** if

$$\forall x \in A, x' \in A, \mathrm{f}(x) = \mathrm{f}(x') \Rightarrow x = x'$$

Lemma: Consider a function $\mathrm{f} : A \to B$ and suppose $A \neq \emptyset$. Then $\mathrm{f}$is injective if and only if there exists a function $\mathrm{g} : B \to A$ with $\mathrm{g} \circ \mathrm{f} = \mathrm{id}_\mathrm{A}$.

Proof:

$' \Rightarrow '$:

Suppose $\mathrm{f}$is injective. As $A \neq \emptyset$ let's define m as an arbitrary element of A . We can then define a suitable function $\mathrm{g} : B \longrightarrow A$

$$\mathrm{g}(y) := \begin{cases} \text{the unique } x \in A \mid \mathrm{f}(x) = y & \text{, if } y \in \mathrm{Im}_\mathrm{f} \\ m & \text{, else} \end{cases}$$

It is now easy to verify that $\mathrm{g} \circ \mathrm{f} = \mathrm{id}_\mathrm{A}$.

$' \Leftarrow '$:

Suppose there is a function $\mathrm{g} : B \longrightarrow A$ with $\mathrm{g} \circ \mathrm{f} = \mathrm{id}_\mathrm{A}$. Then $\forall x, x' \in A, \mathrm{f}(x) = \mathrm{f}(x') \Rightarrow \mathrm{g}(\mathrm{f}(x)) = \mathrm{g}(\mathrm{f}(x')) \Rightarrow x = x'$. $\mathrm{f}$is thus injective.

Q.E.D.

Definition: A function $\mathrm{f} : A \longrightarrow B$ is **surjective** if

$$\forall y \in B, \exists x \in A \mid \mathrm{f}(x) = y$$

Lemma: Consider a function $\mathrm{f} : A \longrightarrow B$. Then $\mathrm{f}$is surjective if and only if there exists a function $\mathrm{g} : B \longrightarrow A$ with $\mathrm{f} \circ \mathrm{g} = \mathrm{id}_\mathrm{B}$.

Proof:

$' \Rightarrow '$:

Suppose $\mathrm{f}$is surjective. We can define a suitable function $\mathrm{g} : B \longrightarrow A$ as follows:

$$\mathrm{g}(y) := \text{an } x \in A \mid \mathrm{f}(x) = y$$

It is now easy to verify that $\mathrm{f} \circ \mathrm{g} = \mathrm{id}_\mathrm{B}$.

$' \Leftarrow '$:

Suppose there is a function $\mathrm{g} : B \longrightarrow A$ with $\mathrm{f} \circ \mathrm{g} = \mathrm{id}_\mathrm{B}$. Then $\forall y \in B, \text{ let } x := \mathrm{g}(y)$. Then $\mathrm{f}(x) = \mathrm{f}(\mathrm{g}(y)) = y$. $\mathrm{f}$is thus surjective.

Q.E.D.

Definition: A function $\mathrm{f} : A \longrightarrow B$ is **bijective** if it is both injective and surjective.

Lemma: A function $\mathrm{f} : A \longrightarrow B$ is bijective if and only if there exists a function $\mathrm{g} : B \longrightarrow A$ with $\mathrm{g} \circ \mathrm{f} = \mathrm{id}_\mathrm{A}$ and $\mathrm{f} \circ \mathrm{g} = \mathrm{id}_\mathrm{B}$. Furthermore it can be shown that such a $\mathrm{g}$is unique. We write it $\mathrm{f}^{-1} : B \longrightarrow A$ and call it the **inverse** of f.

Proof:

Left as an exercise.

Example: If A and B are sets such that $B \subseteq A$, there exists an obviously injective function $i : B \longrightarrow A$, called the inclusion $B \subseteq A$, such that $i(b) = b$ for all $b \in B$.

Example: If $\sim$ is an equivalence relation on a set X , there is an obviously surjective function $\pi : X \longrightarrow X/\sim$, called the canonical projection onto $X/\sim$, such that $\pi(x) = [x]$ for all $x \in X$.

Theorem: Define the equivalence relation $\sim$ on A such that $a \sim b$ if and only if $f(a) = f(b)$. Then, if $f : A \longrightarrow B$ is any function, f decomposes into the composition

$$A \xrightarrow{\pi} A/\sim \xrightarrow{\tilde{f}} \mathrm{im} f \xrightarrow{i} B$$

where π is the canonical projection, i is the inclusion $\mathrm{im}\, f \subseteq B$, and $\tilde{f}$ is the bijection $\tilde{f}([a]) = f(a)$ for all $a \in A$.

Proof: The definition of $\tilde{f}$ immediately implies that $f = i \circ \tilde{f} \circ \pi$, so we only have to prove that $\tilde{f}$ is well defined and a bijection. Let $a, a' \in A$. Then $[a] = [a'] \Rightarrow a \sim a' \Rightarrow f(a) = f(a')$. This shows that the value of $\tilde{f}([a])$ is independent of the representative chosen from $[a]$, and so it is well-defined.

For injectivity, we have $\tilde{f}([a]) = \tilde{f}([a']) \Rightarrow f(a) = f(a') \Rightarrow [a] = [a']$, so $\tilde{f}$ is injective.

For surjectivity, let $b \in \operatorname{im} f$. Then there exists an $a \in A$ such that $f(a) = b$, and so $\tilde{f}([a]) = b$ by definition of $\tilde{f}$. Since b is arbitrary in $\operatorname{im} f$, this proves that $\tilde{f}$ is surjective.
Q.E.D.

Definition: Given a function $f : X \longrightarrow Y$, f is a

(i) *Monomorphism* if given any two functions $g, h : W \longrightarrow X$ such that $f \circ g = f \circ h$, then $g = h$.

(ii) *Epimorphism* if given any two functions $g, h : Y \longrightarrow Z$ such that $g \circ f = h \circ f$, then $g = h$.

Theorem: A function between sets is

(i) a monomorphism if and only if it is injective.

(ii) an epimorphism if and only if it is surjective.

Proof: (i) Let $f : B \longrightarrow C$ be a monomorphism. Then, for any two functions $g, h : A \longrightarrow B$, $f(g(a)) = f(h(a)) \Rightarrow g(a) = h(a)$ for all $a \in A$. This is the definition if injectivity. For the converse, if f is injective, it has a left inverse f'. Thus, if $f(g(a)) = f(h(a))$ for all $a \in A$, compose with f' on the left side to obtain $g(a) = h(a)$, such that f is a monomorphism.
(ii) Let $f : A \longrightarrow B$ be an epimorphism. Then, for any two functions $g, h : B \longrightarrow C$, $g(f(a)) = h(f(a)) \Rightarrow g(b) = h(b)$ for all $a \in A$ and $b \in B$. Assume $\operatorname{im} f \neq B$, that is, that f is not surjective. Then there exists at least one $b \in B$ not in $\operatorname{im} f$. For this b choose two functions g, h which coincide on $\operatorname{im} f$ but disagree on $\{b\}$. However, we still have $g(f(a)) = h(f(a))$ for all $a \in A$. This violates our assumtion that f is an epimorphism. Consequentally, f is surjective. For the converse, assume f is surjective. Then the epimorphism property immediately follows.
Q.E.D.

Remark: The equivalence between monomorphism and injectivity, and between epimorphism and surjectivity is a special property of functions between sets. This not the case in general, and we will see examples of this when discussing structure-preserving functions between groups or rings in later sections.

Example: Given any two sets A and B, we have the canonical projections $\pi_A : A \times B \longrightarrow A$ sending (a, b) to a, and $\pi_B : A \times B \longrightarrow B$ sending (a, b) to b. These maps are obviously surjective.
In addition, we have the natural inclusions $i_A : A \longrightarrow A \coprod B$ and $i_B : B \longrightarrow A \coprod B$ which are obviously injective as stated above.

Universal properties

The projections and inclusions described above are special, in that they satisfy what are called *universal properties*. We will give the theorem below. The proof is left to the reader.

Theorem: Let A, B, C be any sets.

(i) Let $f : C \longrightarrow A$ and $g : C \longrightarrow B$. Then there exists a *unique* function $u : C \longrightarrow A \times B$ such that $f = \pi_A \circ u$ and $g = \pi_B \circ u$ are simultaneously satisfied. u is sometimes denoted $f \times g$.

(ii) Let $f : A \longrightarrow C$ and $g : B \longrightarrow C$. Then there exists a *unique* function $u : A \coprod B \longrightarrow C$ such that $f = u \circ i_A$ and $g = u \circ i_B$ are simultaneously satifsied.
The canonical projections onto quotients also satisfy a universal property.

Theorem: Define the equivalence relation $\sim$ on X and let $f : X \longrightarrow Y$ be any function such that $a \sim b \Rightarrow f(a) = f(b)$ for all $a, b \in X$. Then there exists a *unique* function $\bar{f} : X/\sim \longrightarrow Y$ such that $f = \bar{f} \circ \pi$, where $\pi : X \longrightarrow X/\sim$ is the canonical projection.

Binary Operations

Definition

A binary operation on a set A is a function $* : A \times A \rightarrow A$. For $a, b \in A$, we usually write $*(a, b)$ as $a * b$. The property that $a * b \in A$ for all $a, b \in A$ is called *closure under* $*$.

Example: Addition between two integers produces an integer result. Therefore addition is a binary operation on the integers. Whereas division of integers is an example of an operation that is *not* a binary operation. $1/2$ is not an integer, so the integers are not closed under division.

To indicate that a set A has a binary operation $*$ defined on it, we can compactly write $(A, *)$. Such a pair of a set and a binary operation on that set is collectively called a binary structure. A binary structure may have several interesting properties. The main ones we will be interested in are outlined below.

Definition: A binary operation $*$ on A is *associative* if for all $a, b, c \in A$, $(a * b) * c = a * (b * c)$.

Example: Addition of integers is associative: $(1 + 2) + 3 = 6 = 1 + (2 + 3)$. Notice however, that subtraction is not associative. Indeed, $2 = 1 - (2 - 3) \neq (1 - 2) - 3 = -4$.

Definition: A binary operation $*$ on A is *commutative* is for all $a, b \in A$, $a * b = b * a$.

Example: Multiplication of rational numbers is commutative: $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd} = \frac{ca}{bd} = \frac{c}{d} \cdot \frac{a}{b}$. Notice that division is not commutative: $2 \div 3 = \frac{2}{3}$while $3 \div 2 = \frac{3}{2}$. Notice also that commutativity of multiplication depends on the fact that multiplication of integers is commutative as well.

Exercise

- Of the four arithmetic operations, addition, subtraction, multiplication, and division, which are associative? commutative?

Answer

operation	associative	commutative
Addition	yes	yes
Multiplication	yes	yes
Subtraction	No	No
Division	No	No

Linear Algebra

The reader is expected to have some familiarity with linear algebra. For example, statements such as

> Given vector spaces V and W with bases B and C and dimensions n and m , respectively, a linear map $f : V \rightarrow W$ corresponds to a unique $m \times n$ matrix, dependent on the particular choice of basis.

should be familiar. It is impossible to give a summary of the relevant topics of linear algebra in one section, so the reader is advised to take a look at the linear algebra book.

In any case, the core of linear algebra is the study of linear functions, that is, functions with the property $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$, where greek letters are scalars and roman letters are vectors.

The core of the theory of finitely generated vector spaces is the following:

Every finite-dimensional vector space V is isomorphic to $\mathbb{F}^n$ for some field $\mathbb{F}$ and some $n \in \mathbb{N}$, called the dimension of V. Specifying such an isomorphism is equivalent to choosing a basis for V. Thus, any linear map between vector spaces $f : V \rightarrow W$ with dimensions n and m and given bases ϕ and ψ induces a unique linear map $[f]_{\phi}^{\psi} : \mathbb{R}^n \rightarrow \mathbb{R}^m$. These maps are presicely the $m \times n$ matrices, and the matrix in question is called the *matrix representation* of f relative to the bases ϕ, ψ .

Remark: The idea of identifying a basis of a vector space with an isomorphism to $\mathbb{F}^n$ may be new to the reader, but the basic principle is the same.

Number Theory

As numbers of various number systems form basic units with which one must work when studying abstract algebra, we will now define the natural numbers and the rational integers as well as the basic operations of addition and multiplication. Using these definitions, we will also derive important properties of these number sets and operations. Following this, we will discuss important concepts in number theory; this will lead us to discussion of the properties of the integers modulo *n*.

The Peano postulates and the natural numbers

Definition: Using the undefined notions "1" and "successor" (denoted by $'$), we define the set of natural numbers without zero $\mathbb{N}^*$, hereafter referred to simply as the natural numbers, with the following axioms, which we call the Peano postulates:

> Axiom 1. $\exists 1(1 \in \mathbb{N}^*)$.
>
> Axiom 2. $\forall a(a \in \mathbb{N}^* \implies \exists b(b = a'))$.
>
> Axiom 3. $\neg \exists a(a' = 1)$.
>
> Axiom 4. $\forall a \in \mathbb{N}^*(\forall b \in \mathbb{N}^*(a' = b' \implies a = b))$.
>
> Axiom 5. $\forall A \subseteq \mathbb{N}^*((1 \in A) \land \forall a \in A(a' \in A) \implies A = \mathbb{N}^*)$.

We can prove theorems for natural numbers using mathematical induction as a consequence of the fifth Peano Postulate.

Addition

Definition: We recursively define addition for the natural numbers as a composition using two more axioms; the other properties of addition can subsequently be derived from these axioms. We denote addition with the infix operator +.

Axiom 6. $\forall a \in \mathbb{N}^*(a+1=a')$.

Axiom 7. $\forall a \in \mathbb{N}^*(\forall b \in \mathbb{N}^*(a+b'=(a+b)'))$.

Axiom 6 above relies on the first Peano postulate (for the existence of 1) as well as the second (for the existence of a successor for every number).

Henceforth, we will assume that proven theorems hold for all $a, b, c, \ldots, n$ in $\mathbb{N}^*$.

Multiplication

Definition: We similarly define multiplication for the natural numbers recursively, again using two axioms:

Axiom 8. $a(1)=a$.

Axiom 9. $ab'=ab+a$.

Properties of addition

We start by proving that addition is associative.

Theorem 1: Associativity of addition: $(a+b)+c=a+(b+c)$.

Proof: Base case: By axioms 6 and 7, $a+(b+1)=(a+b)'$.

By axiom 6, $a+(b+1)=(a+b)+1$.

Inductive hypothesis: Suppose that, for $k'>1$, $(a+b)+k=a+(b+k)$.

Inductive step: By axiom 7, $(a+b)+k'=[(a+b)+k]'$.

By the inductive hypothesis, $(a+b)+k'=[a+(b+k)]'$.

By axiom 7, $(a+b)+k'=a+(b+k)'$.

By axiom 7, $(a+b)+k'=a+(b+k')$.

By induction, $(a+b)+c=a+(b+c)$. QED.

Lemma 1: $a+1=1+a$.

Proof: Base case: 1+1=1+1.

Inductive hypothesis: Suppose that, for $k'>1$, $k+1=1+k$.

Inductive step: By axiom 6, $k'+1=(k+1)+1$.

By the inductive hypothesis, $k'+1=(1+k)+1$.

By theorem 1, $k'+1=1+(k+1)$.

By axiom 6, $k'+1=1+k'$.

By induction, $a+1=1+a$. QED.

Theorem 2: Commutativity of addition: $a+b=b+a$.

Proof: Base case: By lemma 1, $a+1=1+a$.

Inductive hypothesis: Suppose that, for $k'>1$, $a+k=k+a$.

By axiom 6, $a+k'=a+(k+1)$.

By theorem 1, $a+k'=(a+k)+1$.

By the inductive hypothesis, $a+k'=(k+a)+1$.

By theorem 1, $a + k' = k + (a + 1)$.

By lemma 1, $a + k' = k + (1 + a)$.

By theorem 1, $a + k' = (k + 1) + a$.

By axiom 6, $a + k' = k' + a$.

By induction, $a + b = b + a$. QED.

Theorem 3: $a + b = a + c \implies b = c$.

Proof: Base case: Suppose $1 + b = 1 + c$.

By theorem 2, $b + 1 = c + 1$.

By axiom 6, $b' = c'$.

By axiom 4, $b = c$.

Inductive hypothesis: Suppose that, for $k' > 1$, $k + b = k + c \implies b = c$.

Inductive step: Suppose $k' + b = k' + c$.

By axiom 6, $(k + 1) + b = (k + 1) + c$.

By theorem 2, $(1 + k) + b = (1 + k) + c$.

By theorem 1, $1 + (k + b) = 1 + (k + c)$.

By the base case, $k + b = k + c$. Thus, $k' + b = k' + c \implies k + b = k + c$.

By the inductive hypothesis, $k' + b = k' + c \implies b = c$.

By induction, $a + b = a + c \implies b = c$. QED.

Properties of multiplication

Theorem 4: Left-distributivity of multiplication over addition: $a(b + c) = ab + ac$.

Proof: Base case: By axioms 6 and 9, $a(b + 1) = ab + a$.

By axiom 8, $a(b + 1) = ab + a(1)$.

Inductive hypothesis: Suppose that, for $k' > 1$, $a(b + k) = ab + ak$.

Inductive step: By axiom 7, $a(b + k') = a(b + k)'$.

By axiom 9, $a(b + k') = a(b + k) + a$.

By the inductive hypothesis, $a(b + k') = (ab + ak) + a$.

By theorem 1, $a(b + k') = ab + (ak + a)$.

By axiom 9, $a(b + k') = ab + ak'$.

By induction, $a(b + c) = ab + ac$. QED.

Theorem 5: $1a = a$.

Proof: Base case: By axiom 8, 1(1)=1.

Inductive hypothesis: Suppose that, for $k' > 1$, $1k = k$.

Inductive step: By axiom 6, $1k' = 1(k + 1)$.

By theorem 4, $1k' = 1k + 1(1)$.

By the base case, $1k' = 1k + 1$.

By the inductive hypothesis, $1k' = k + 1$.

By axiom 6, $1k' = k'$.

By induction, $1a = a$. QED.

Theorem 6: $a'b = ab + b$.

Proof: Base case: By axiom 8, $a'(1) = a'$.

By axiom 6, $a'(1) = a + 1$.

By axiom 8, $a'(1) = a(1) + 1$.

Inductive hypothesis: Suppose that, $k' > 1$, $a'k = ak + k$.

Inductive step: By axiom 9, $a'k' = a'k + a'$.

By the inductive hypothesis, $a'k' = ak + k + a'$.

By axiom 6, $a'k' = ak + k + (a + 1)$.

By theorem 1, $a'k' = ak + (k + a) + 1$.

By theorem 2, $a'k' = ak + (a + k) + 1$.

By theorem 1, $a'k' = ak + a + k + 1$

By axiom 9, $a'k' = ak' + k + 1$.

By theorem 1, $a'k' = ak' + (k + 1)$.

By axiom 6, $a'k' = ak' + k'$.

By induction, $a'b = ab + b$. QED.

Theorem 7: Associativity of multiplication: $(ab)c = a(bc)$.

Proof: Base case: By axiom 8, $ab(1) = ab = a\left[b(1)\right]$.

Inductive hypothesis: Suppose that, for $k' > 1$, $(ab)k = a(bk)$.

Inductive step: By axiom 9, $(ab)k' = (ab)k + ab$.

By the inductive hypothesis, $(ab)k' = a(bk) + ab$.

By theorem 4, $(ab)k' = a(bk + b)$.

By axiom 9, $(ab)k' = a(bk')$.

By induction, $(ab)c = a(bc)$. QED.

Theorem 8: Commutativity of multiplication: $ab = ba$.

Proof: Base case: By axiom 8 and theorem 5, $a(1) = a = 1a$.

Inductive hypothesis: Suppose that, for $k' > 1$, $ak = ka$.

Inductive step: By axiom 9, $ak' = ak + a$.

By the inductive hypothesis, $ak' = ka + a$.

By theorem 6, $ak' = k'a$.

By induction, $ab = ba$. QED.

Theorem 9: Right-distributivity of multiplication over addition: $(a + b)c = ac + bc$.

Proof: By theorems 4 and 7, $(a + b)c = ca + cb$.

By theorem 7, $(a + b)c = ac + bc$. QED.

The integers

The set of rational integers $\mathbb{Z}$ can be constructed from ordered pairs of natural numbers (a, b). We define an equivalence relation on the set of all such ordered pairs such that

$$(a, b) \equiv (c, d) \Leftrightarrow a + d = c + b.$$

Then the set of rational integers is the set of all equivalence classes of such ordered pairs. We will denote the equivalence class of which some pair (a, b) is a member with the notation $[(a, b)]$. Then, for any natural numbers a and b, $[(a, b)]$ represents a rational integer.

Integer addition

Definition: We define addition for the integers as follows:

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

Using this definition and the properties for the natural numbers, one can prove that integer addition is both associative and commutative.

Integer multiplication

Definition: Multiplication for the integers, like addition, can be defined with a single axiom:

$$[(a, b)] [(a, b)] = [(ac + bd, ad + bc)] .$$

Again, using this definition and the previously-proven properties of natural numbers, it can be shown that integer multiplication is commutative and associative, and furthermore that it is both left- and right-distributive with respect to integer addition.

Group Theory

Group

In this section we will begin to make use of the definitions we made in the section about binary operations. In the next few sections, we will study a specific type of binary structure called a *group*. First, however, we need some preliminary work involving a less restrictive type of binary structure.

Monoids

Definition 1: A *monoid* is a binary structure $(M, *)$satisfying the following properties:

(i) $(a * b) * c = a * (b * c)$for all $a, b, c \in M$.

(ii) There exists an element $e \in M$ such that $a * e = a = e * a$ for all $a \in M$.

The element e in (ii) is called an identity element of M.

Now we have our axioms in place, we are faced with a pressing question; what is our first theorem going to be? Since the first few theorems are not dependent on one another, we simply have to make an arbitrary choice. We choose the following:

Theorem 2: The identity element of M is unique.

Proof: Assume e and e' are both identity elements of M. Then both satisfy condition (ii) in the definition above. In particular, $e = e * e' = e'$, proving the theorem. ▯

This theorem will turn out to be of fundamental importance later when we define groups.

Theorem 3: If $a_1, a_2, \ldots, a_n$ are elements of M for some $n \in \mathbb{N}$, then the product $a_1 * a_2 * \cdots * a_n$ is unambiguous.

Proof: We can prove this by induction. The cases for $n = 1$and $n = 2$are trivially true. Then, assume that the statement is true for all $k \leq n - 1$. Then, the product $a_1 * \cdots * a_n$, inserting parentheses, can be "partitioned" into $(a_1 * \cdots * a_i) * (a_{i+1} * \cdots * a_n)$. Then, since $1 \leq i \leq n - 1$, both parts of the product have a number of elements less than or equal to k and are thus unambiguous. We compute these products and are left with $(b_1) * (b_2) = b_1 * b_2$which is unambiguous, proving the theorem. ▯

This is about as far as we are going to take the idea of a monoid. We now proceed to groups.

Groups

Definition 4: A *group* is a monoid $(G, *)$that also satisfies the property

(iii) For each $a \in G$, there exists and element $a' \in G$ such that $a * a' = a' * a = e$.

Such an element a' is called an inverse of a. When the operation on the group is understood, we will conveniently refer to $(G, *)$as G. In addition, we will gradually stop using the symbol $*$ for multiplication when we are dealing with only one group, or when it is understood which operation is meant, instead writing products by juxtaposition, $a * b \equiv ab$.

Remark 5: Notice how this definition depends on Theorem 2 to be well defined. Therefore we could not state this definition before at least proving uniqueness of the identity element. Alternatively, we could have included the existence of a distinguished identity element in the definition. In the end, the two approaches are logically equivalent.

Theorem 6: The inverse of any element is unique.

Proof: Let $g \in G$ and let g' and g'' be inverses of g. Then, $g' = g' * e = g' * g * g'' = e * g'' = g''$. □
Thus, we can speak of *the* inverse of an element, and we will denote this element by a^{-1}. We also observe this nice property:

Corollary 7: $(a^{-1})^{-1} = a$.

Proof: This follows immediately since $a * a^{-1} = a^{-1} * a = e$.

The next couple of theorems may appear obvious, but in the interest of keeping matters fairly rigorous, we allow ourselves to state and prove seemingly trivial statements.

Theorem 8: Let G be a group and $a, b \in G$. Then $(a * b)^{-1} = b^{-1} * a^{-1}$.

Proof: The result follows by direct computation: $(a * b) * (b^{-1} * a^{-1}) = a * b * b^{-1} * a^{-1} = a * e * a^{-1} = a * a^{-1} = e$. □

Theorem 9: Let $a, b, c \in G$. Then, $a * b = a * c$ if and only if $b = c$. Also, $a * c = b * c$ if and only if $a = b$.

Proof: We will prove the first assertion. The second is identical. Assume $a * b = a * c$. Then, multiply on the left and right by a^{-1} to obtain $b = c$. Secondly, assume $b = c$. Then, multiply on the left by a to obtain $a * b = a * c$. □

Theorem 10: The equation $a * x = b$ has a unique solution in G for any $a, b \in G$.

Proof: We must show existence and uniqueness. For existence, observe that $g = a^{-1} * b$ is a solution in G . For uniqueness, multiply both sides of the equation on the left by a^{-1} to show that this is the only solution. □

Notation: Let G be a group and $a \in G$. We will often encounter a situation where we have a product $\underbrace{a * a * \cdots * a}_{\text{n terms}}$. For these situations, we introduce the shorthand notation $a^n = \underbrace{a * a * \cdots * a}_{\text{n terms}}$ if n is positive, and $a^n = \underbrace{a^{-1} * a^{-1} * \cdots * a^{-1}}_{|\text{n}|\text{ terms}}$ if n is negative. Under these rules, it is straightforward to show $g^n * g^m = g^{n+m}$ and $(a^n)^{-1} = a^{-n}$ and $a^0 = e$ for all $a \in G$.

Definition 11: (i) The *order* of a group G , denoted $|G|$ or $o(G)$, is the number of elements of G if G is finite, and ∞ otherwise.

(ii) The order of an element of $g \in G$, similarly denoted $|g|$ or $o(g)$, is defined as the lowest positive integer n such that $g^n = e$ if such an integer exists, and ∞ otherwise.

Theorem 12: Let G be a group and $a, b \in G$. Then $|ab| = |ba|$.

Proof: Let the order of ab be n . Then, $(ab)^n = abab...ab = e$, n being the smallest positive integer such that this is true. Now, multiply by b on the left and a on the right to obtain $(ba)^{n+1} = ba$ implying $(ba)^n = e$. Thus, we have shown that $|ba| \leq |ab|$. A similar argument in the other direction shows that $|ab| \leq |ba|$. Thus, we must have $|ab| = |ba|$, proving the theorem. □

Corollary 13: Let G be a group with $a, b \in G$. Then, $|aba^{-1}| = |b|$.

Proof: By Theorem 7, we have that $|aba^{-1}| = |ba^{-1}a| = |be| = |b|$. □

Theorem 14: An element of a group not equal to the identity has order 2 if and only if it is its own inverse.

Proof: Let g have order 2 in the group G . Then, $g^2 = gg = e$, so by definition, $g^{-1} = g$. Now, assume $g^{-1} = g$ and $g \neq e$. Then $e = gg^{-1} = gg = g^2$. Since $g \neq e$, 2 is the smallest positive integer satisfying this property, so g has order 2. □

Definition 15: Let G be a group such that for all $a, b \in G$, $ab = ba$. Then, G is said to be *commutative* or *abelian*.

When we are dealing with an abelian group, we will sometimes use so-called additive notation, writing $+$for our binary operation and replacing $a^n$with na. In such cases, we only need to keep track of the fact that $n$is an integer while $a$is a group element. We will also talk about the *sum* of elements rather than their product.

Abelian groups are in many ways nicer objects than general groups. They also admit more structure where ordinary groups do not. We will see more about this later when we talk about structure-preserving maps between groups.

Definition 16: Let G be a group. A subset $S \subseteq G$ is called a *generating set* for G if every element in G can be written in terms of elements in S.

Now that we have our definitions in place and have a small arsenal of theorems, let us look at three (really, two and a half) important families of groups.

Multiplication tables

We will now show a convenient way of representing a group structure, or more precisely, the multiplication rule on a set. This notion will not be limited to groups only, but can be used for any structure with any number of operations. As an example, we give the group multiplication table for the *Klein 4-group* K_4. The multiplication table is structured such that $g * h$ is represented by the element in the " (g, h)-position", that is, in the intersection of the g-row and the h-column.

$*$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

This next group is for the group of integers under addition modulo 4, called $\mathbb{Z}_4$. We will learn more about this group later.

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

We can clearly see that $K_4$and $\mathbb{Z}_4$are "different" groups. There is no way to relabel the elements such that the multiplication tables coincide. There is a notion of "equality" of groups that we have not yet made precise. We will get back to this in the section about group homomorphisms.

The reader might have noticed that each row in the group table features each element of the group exactly once. Indeed, assume that an element $k \in G$ appeared twice in some row of the multiplication table for G. Then there would exist $g, h, h' \in G$ such that $gh = gh'$, implying $h = h'$and contradicting the assumption of k appearing twice. We state this as a theorem:

Theorem 17: Let G be a group and $a \in G$. Then $aG = \{a * g \mid g \in G\} = G$.

Using this, the reader can use a multiplication table to find all groups of order 3. He/she will find that there is only one possibility.

Problems

Problem 1: Show that $M_n(\mathbb{R})$, the set of $n \times n$ matrices with real entries, forms a group under the operation of matrix addition.

Problem 2: Let V, W be vector spaces and $\mathrm{Hom}(V, W)$ be the set of linear maps from V to W. Show that $\mathrm{Hom}(V, W)$ forms an abelian group by defining $(f + g)(v) = f(v) + g(v)$.

Problem 3: Let $\mathbb{H}$ be generated by the elements i, j, k, m such that $i^2 = j^2 = k^2 = m$, $m^2 = e$ and $ij = mji = k$. Show that $\mathbb{H}$ forms a group. This group is called the group of *quaternions*, and is a 4-dimensional version of the complex numbers. Are any of the conditions above redundant?

Problem 4: Let S be any nonempty set and consider the set G^S. Show that G^S has a natural group structure.

Answer

G^S is the set of functions $f : S \longrightarrow G$. Let $f_1, f_2 \in G^S$ and define the binary operation $(f_1 * f_2)(x) = f_1(x)f_2(x)$ for all $x \in S$. Then G^S is a group with identity 0 such that $0(x) = e$ for all $x \in X$ and inverses $f^{-1}(x) = f(x)^{-1}$ for all $f \in G^S, x \in X$.□

Problem 5: Let G be a group with two distinct elements a and b, both of order 2. Show that G has a third element of order 2.

Answer

We consider first the case where $ab = ba$. Then $(ab)^2 = abab = aabb = e$ and ab is distinct from e, a and b. If $ab \neq ba$, then $(aba^{-1})^2 = abaa^{-1}ba^{-1} = abba^{-1} = e$ and aba^{-1} is distinct from e, a and b. □

Problem 6: Let G be a group with one and only one element f of order 2. Show that $\prod_{g \in G} g = f$.

Answer

Since the product of two elements generally depends on the order in which we multiply them, the stated product is not neccesarily well defined. However, it works out in this case.

Since every element of G appears once in the product, for every element $g \in G$, the inverse of g must appear somewhere in the product. That, is, unless $|g| = 2$in which case g is its own inverse by Theorem 14. Now, applying Corollary 13 to the product shows that its order is that same as the order of the product of all elements of order 2 in G. But there is only one such element, f, so the order of the product is 2. Since the only element in G having order 2 is f, the equality follows. □

Subgroup

Subgroups

We are about to witness a universal aspect of mathematics. That is, whenever we have any sort of structure, we ask ourselves: does it admit substructures? In the case of groups, the answer is yes, as we will immediately see.

Definition 1: Let G be a group. Then, if $H \subseteq G$ is a subset of G which is a group in its own right under the same operation as G , we call H a *subgroup* of G and write $H \leq G$.

Example 2: Any group G has at least 2 subgroups; G itself and the trivial group $\{e\}$. These are called the *improper* and *trivial* subgroups of G , respectively.

Naturally, we would like to have a method of determining whether a given subset of a group is a subgroup. The following two theorems provide this. Since H naturally inherits the associativity property from G , we only need to check closure.

Theorem 3: A subset H of a group G is a subgroup if and only if

(i) H is closed under the operation on G . That is, if $a, b \in H$, then $ab \in H$,

(ii) $e \in H$,

(iii) H is closed under the taking of inverses. That is, if $a \in H$, then $a^{-1} \in H$.

Proof: The left implication follows directly from the group axioms and the definition of subgroup. For the right implication, we have to verify each group axiom for H . Firstly, since H is closed, it is a binary structure, as required, and as mentioned, H inherits associativity from G. In addition, H has the identity element and inverses, so H is a group, and we are done. □

There is, however, a more effective method. Each of the three criteria listed above can be condensed into a single one.

Theorem 4: Let G be a group. Then a subset $H \subseteq G$ is a subgroup is and only if $a, b \in H \Rightarrow ab^{-1} \in H$.

Proof: Again, the left implication is immediate. For the right implication, we have to verify the (i)-(iii) in the previous theorem. First, assume $a \in H$. Then, letting $b = a$, we obtain $aa^{-1} = e \in H$, taking care of (ii). Now, since $e, a \in H$ we have $ea^{-1} = a^{-1} \in H$ so H is closed under taking of inverses, satisfying (iii). Lastly, assume $a, b \in H$. Then, since $b^{-1} \in H$, we obtain $a(b^{-1})^{-1} = ab \in H$, so H is closed under the operation of G , satisfying (i), and we are done. □

All right, so now we know how to recognize a subgroup when we are presented with one. Let's take a look at how to find subgroups of a given group. The next theorem essentially solves this problem.

Theorem 5: Let G be a group and $g \in G$. Then the subset $\{g^n \mid n \in \mathbb{Z}\}$ is a subgroup of G , denoted $\langle g \rangle$ and called the *subgroup generated by* g . In addition, this is the *smallest* subgroup containing g in the sense that if H is a subgroup and $g \in H$, then $\langle g \rangle \leq H$.

Proof: First we prove that $\langle g \rangle$ *is* a subgroup. To see this, note that if $h, k \in \langle g \rangle$, then there exists integers $n, m \in \mathbb{Z}$ such that $h = g^n$, $k = g^m$. Then, we observe that $hk^{-1} = g^n g^{-m} = g^{n-m} \in \langle g \rangle$ since $n - m \in \mathbb{Z}$, so $\langle g \rangle$ is a subgroup of G , as claimed. To show that it is the smallest subgroup containing g , observe that if H is a subgroup containing g , then by closure under products and inverses, $g^n \in H$ for all $n \in Z$. In other words, $\langle g \rangle \subseteq H$. Then automatically $\langle g \rangle \leq H$ since $\langle g \rangle$ is a subgroup of G . □

Theorem 6: Let H and H' be subgroups of a group G . Then $H \cap H'$ is also a subgroup of G .

Proof: Since both H and H' contain the identity element, their intersection is nonempty. Let $a, b \in H \cap H'$. Then $a, b \in H$ and $a, b \in H'$. Since both H and H' are subgroups, we have $ab^{-1} \in H$ and $ab^{-1} \in H'$. But then $ab^{-1} \in H \cap H'$ (why?). Thus $H \cap H'$ is a subgroup of G . □

Theorem 6 can easily be generalized to apply for any arbitrary intesection $\bigcap_{i \in I} H_i$ where H_i is a subgroup for every i in an arbitrary index set I. The reasoning is identical, and the proof of this generalization is left to the reader to formalize.

Definition 7: Let G be a group and H be a subgroup of G. Then $gH = \{gh \mid h \in H\}$ is called a *left coset* of H. The set of all left cosets of H in G is denoted G/H. Likewise, $Hg = \{hg \mid h \in H\}$ is called a *right coset*, and the set of all right cosets of H in G is denoted $H\backslash G$.

Lemma 8: Let G be a group and H be a subgroup of G. Then every left coset has the same number of elements.

Proof: Let $g \in G$ and define the function $f : H \longrightarrow gH$ by $h \mapsto gh$. We show that f is a bijection. Firstly, $gh = gh' \Rightarrow h = h'$ by left cancellation, so f is injective. Secondly, let $h' \in gH$. Then $h' = gh$ for some $h \in H$ and $f(h) = h'$, so f is surjective and a bijection. It follows that $|H| = |gH|$, as was to be shown. □

Lemma 9: The relation $\sim$ defined by $a \sim b \Leftrightarrow a^{-1}b \in H$ is an equivalence relation.

Proof: Reflexivity and symmetry are immediate. For transitivity, let $a \sim b$ and $b \sim c$. Then $a^{-1}b, b^{-1}c \in H$, so $a^{-1}c \in H$ and we are done. □

Lemma 10: Let G be a group and H be a subgroup of G. Then the left cosets of H partition G.

Proof: Note that $aH = bH \Leftrightarrow ah = bh' \Leftrightarrow a^{-1}bh' = h \Leftrightarrow a^{-1}b \in H$ for some $h, h' \in H$. Since $a \sim b \Leftrightarrow a^{-1}b \in H$ is an equivalence relation and the equivalence classes are the left cosets of H, these automatically partition G. □

Theorem 11 (Lagrange's theorem): Let G be a finite group and H be a subgroup of G. Then $|G| = |G/H||H|$.

Proof: By the previous lemmas, each left coset has the same number of elements $|H|$ and every $g \in G$ is included in a unique left coset gH. In other words, G is partitioned by $|G/H|$ left cosets, each contributing an equal number of elements $|H|$. The theorem follows. □

Note 12: Each of the previous theorems have analagous versions for right cosets, the proofs of which use identical reasoning. Stating these theorems and writing out their proofs are left as an exercise to the reader.

Corollary 13: Let G be a group and H be a subgroup of G. Then right and left cosets of H have the same number of elements.

Proof: Since H is a left and a right coset we immediately have $|gH| = |H| = |Hg'|$ for all $g, g' \in G$. □

Corollary 14: Let G be a group and H be a subgroup of G. Then the number of left cosets of H in G and the number of right cosets of H in G are equal.

Proof: By Lagrange's theorem and its right coset counterpart, we have $|H||H\backslash G| = |G| = |G/H||H|$. We immediately obtain $|H\backslash G| = |G/H|$, as was to be shown. □

Now that we have developed a reasonable body of theory, let us look at our first important family of groups, namely the cyclic groups.

Problems

Problem 1 (Matrix groups): Show that:

i) The group $GL(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid \det(A) \neq 0\}$ of invertible $n \times n$ matrices is a subgroup of $M_n(\mathbb{R})$. This group is called the *general linear group of order n*.

ii) The group $O(n) = \{A \in M_n(\mathbb{R}) \mid AA^T = I\}$ of $n \times n$ orthogonal matrices is a subgroup of $GL(N, \mathbb{R})$. This group is called the *orthogonal group of order n*.

iii) The group $SO(n) = \{A \in M_n(\mathbb{R}) \mid AA^T = I \wedge \det(A) = 1\}$ is a subgroup of $O(n)$. This group is called the *special orthogonal group of roder n*.

iv) The group $U(n) = \{A \in M_n(\mathbb{C}) \mid AA^* = I\}$ of unitary matrices is a subgroup of $GL(n, \mathbb{C})$. This is called the *unitary group of order* n .

v) The group $SU(n) = \{A \in M_n(\mathbb{C}) \mid AA^* = I \land \det(A) = 1\}$ is a subgroup of $U(n)$. This is called the *special unitary group of order* n .

Problem 2: Show that if H, K are subgroups of G , then $H \cup K$ is a subgroup of G if and only if $H \subseteq K$ or $K \subseteq H$.

- Definition of a Subgroup
  - Subgroup Inherits Identity
  - Intersection of Subgroups is a Subgroup
  - Definition of a Coset
    - a Subgroup and its Cosets have Equal Orders
    - a Group is Partitioned by Cosets of Its Subgroup
  - Lagrange's Theorem
  - Definition of a Cyclic Subgroup
    - Order of a Cyclic Subgroup Help!
    - Euler's Totient Theorem
  - Definition of a Normal Subgroup

Cyclic groups

Definition

- A cyclic group generated by g is
  - $\langle g \rangle = \{g^n \mid n \in \mathbb{Z}\}$
  - where $g^n = \begin{cases} \underbrace{g * g \cdots * g}_{n}, & n \in \mathbb{Z}, n \geq 0 \\ \underbrace{g^{-1} * g^{-1} \cdots * g^{-1}}_{-n}, & n \in \mathbb{Z}, n < 0 \end{cases}$
  - Induction shows: $g^{m+n} = g^m * g^n$ and $g^{mn} = [g^m]^n$

a Cyclic Group of Order n is Isomorphic to the Integers Moduluo n with Addition

Theorem

Let C$_m$ be a cyclic group of order *m* generated by *g* with $*$

Let $(\mathbb{Z}/m, +)$ be the group of integers modulo m with addition

C$_m$ is isomorphic to $(\mathbb{Z}/m, +)$

Lemma

Let *n* be the minimal positive integer such that *g*^{n} = *e*

$$g^i = g^j \leftrightarrow i = j \bmod n$$

Proof of Lemma

Let *i* < *j*. Let *i* - *j* = *sn* + *r* where 0 ≤ *r* < *n* and s,r,n are all integers.

1. $g^i = g^j$
2. $e = g^{i-j} = g^{sn+r} = [g^n]^s * g^r = [e]^s * g^r = g^r$ as *i* - *j* = *sn* + *r*, and $g^n = e$
3. $g^r = e$
4. $r = 0$ as n is the minimal positive integer such that $g^n = e$ and 0 ≤ *r* < *n*
5. $i - j = sn$ 0. and 7.
6. $i = j \bmod n$

Proof

0. Define $$\begin{aligned} f\colon C_m &\to \mathbb{Z}/m \\ g^i &\mapsto i \bmod m \end{aligned}$$

Lemma shows *f* is well defined (only has one output for each input).

f is homomorphism:

$$f(g^i) + f(g^j) = i + j \bmod m = f(g^{i+j}) = f(g^i * g^j)$$

f is injective by lemma

f is surjective as both $\mathbb{Z}/m$ and C_m has *m* elements and m is injective

Cyclic groups

In the previous section about subgroups we saw that if G is a group with $g \in G$, then the set of powers of g, $\langle g \rangle = \{g^n \mid n \in \mathbb{Z}\}$ constituted a subgroup of G , called the *cyclic subgroup generated by* g . In this, section, we will generalize this concept, and in the process, obtain an important family of groups which is very rich in structure.

Definition 1: Let G be a group with an element $g \in G$ such that $\langle g \rangle = G$. Then G is called a *cyclic* group, and g is called a *generator* of G . Alternatively, g is said to *generate* G . If there exists an integer n such that $g^n = e$, and n is the smallest positive such integer, G is denoted C_n, the cyclic group of order n . If no such integer exists, G is denoted C_∞ , the infinite cyclic group.

The infinite cyclic group can also be denoted $F_{\{g\}}$, the *free group* with one generator. This is foreshadowing for a future section and can be ignored for now.

Theorem 2: Any cyclic group is abelian.

Proof: Let G be a cyclic group with generator g. Then if $h, k \in G$, then $h = g^n$ and $k = g^m$ for some $n, m \in \mathbb{Z}$. To show commutativity, observe that $hk = g^n g^m = g^{n+m} = g^{m+n} = g^m g^n = kh$ and we are done. □

Theorem 3: Any subgroup of a cyclic group is cyclic.

Proof: Let G be a cyclic group with generator g, and let $H \leq G$. Since $G = \langle g \rangle$, in paticular every element of H equals g^n for some $n \in \mathbb{Z}$. We claim that if a the lowest positive integer such that $g^a \in H$, then $H = \langle g^a \rangle$. To see this, let $g^n \in H$. Then $n = qa + r$ and $0 \leq r < a$ for unique $q, r \in \mathbb{Z}$. Since H is a subgroup and $g^a \in H$, we must have $g^n (g^a)^{-q} = g^{qa+r} g^{-qa} = g^r \in H$. Now, assume that $r > 0$. Then $g^r \in H$ contradicts our assumption that a is the least positive integer such that $g^a \in H$. Therefore, $r = 0$. Consequently, $g^n \in H$ only if $n = qa$, and $H = \langle g^a \rangle$ and is cyclic, as was to be shown. □

As the alert reader will have noticed, the preceeding proof invoked the notion of *division with remainder* which should be familiar from number theory. Our treatment of cyclic groups will have close ties with notions from number theory. This is no coincidence, as the next few statements will show. Indeed, an alternative title for this section could have been "Modular arithmetic and integer ideals". The notion of an *ideal* may not yet be familiar to the reader, who is asked to wait patiently until the chapter about rings.

Theorem 4: Let $\mathbb{Z}_n = \{0, 1, 2, ..., n-1\}$ with addition defined modulo n. That is $a + b \equiv r \mod n$, where $a + b = qn + r$. We denote this operation by $a +_n b = r$. Then $(\mathbb{Z}_n, +_n)$is a cyclic group.

Proof: We must first show that $(\mathbb{Z}_n, +_n)$is a group, then find a generator. We verify the group axioms. Associativity is inherited from the integers. The element $0 \in Z_n$is an identity element with respect to $+_n$. An inverse of $a \in Z_n$is an element b such that $a +_n b = 0$. Thus $b + a|n$. Then, $b \equiv n - a \equiv -a \mod n$, and so $b = n - a = -a$, and $(\mathbb{Z}_n, +_n)$is a group. Now, since $a = a \cdot 1 = \underbrace{1 + ... + 1}_{a \text{ terms}}$, $1$generates $\mathbb{Z}_n$ and so $(\mathbb{Z}_n, +_n)$is cyclic. □

Unless we explicitly state otherwise, by $\mathbb{Z}_n$we will always refer to the cyclic group $(\mathbb{Z}_n, +_n)$. Since the argument for the generator of $\mathbb{Z}_n$can be made valid for any integer a, this shows that also $\mathbb{Z}$ is cyclic with the generator 1 .

Theorem 5: An element $a \in \mathbb{Z}_n$is a generator if and only if $\gcd(a, n) = 1$.

Proof: We will need the following theorem from number theory: If m, n are integers, then there exists integers r, s such that $rm + sn = 1$, if and only if $\gcd(m, n) = 1$. We will not prove this here. A proof can be found in the number theory section.

For the right implication, assume that $\langle a \rangle = \mathbb{Z}_n$. Then for all $b \in \mathbb{Z}_n$, $b \equiv pa \mod n$ for some integer p. In paticular, there exists an integer s such that $sa \equiv 1 \mod n$. This implies that there exists another integer r such that $sa - rn = 1$. By the above-mentioned theorem from number theory, we then have $\gcd(a, n) = 1$. For the left implication, assume $\gcd(a, n) = 1$. Then there exists integers s, rsuch that $sa - rn = 1 \;\Rightarrow\; sa \equiv 1 \mod n$, implying that $sa = 1$in $\mathbb{Z}_n$. Since $1$generates $\mathbb{Z}_n$, it must be true that $a$is also a generator, proving the theorem. □

We can generalize Theorem 5 a bit by looking at the orders of the elements in cyclic groups.

Theorem 6: Let $H = \langle g \rangle \leq \mathbb{Z}_n$. Then, $|H| = |a| = \dfrac{n}{\gcd(a, n)}$.

Proof: Recall that the order of $a \in \mathbb{Z}_n$ is defined as the lowest positive integer r such that $ra = 0$in $\mathbb{Z}_n$. Since $\langle g \rangle$ is cyclic, there exists an integer $s$such that $ra = sn$is minimal and positive. This is the definition of the least common multiple; $\operatorname{lcm}(a, n) = ra = |a|a$. Recall from number theory that $\operatorname{lcm}(a, n) \cdot \gcd(a, n) = an$. Thus, $|a| = \dfrac{\operatorname{lcm}(a, n)}{a} = \dfrac{an}{a \cdot \gcd(a, n)} = \dfrac{n}{\gcd(a, n)}$, as was to be proven. □

Theorem 7: Every subgroup of $\mathbb{Z}$ is of the form $n\mathbb{Z} = \{na \mid a \in \mathbb{Z}\}$.

Proof: The fact that any subgroup of $\mathbb{Z}$ is cyclic follows from Theorem 3. Therefore, let $n$generate $H \leq G$. Then we see immediately that $H = n\mathbb{Z}$. □

Theorem 8: Let $a, b \in \mathbb{Z}$ be fixed, and let $H = \{ra + sb | r, s \in \mathbb{Z}\}$. Then H is a subgroup of $\mathbb{Z}$ generated by $\gcd(a, b)$.

Proof: We msut first show that $H$is a subgroup. This is immediate since $(ra + sb) - (r'a + s'b) = (r - r')a + (s - s')b \in H$. From the proof of Theorem 3, we see that any subgroup of $\mathbb{Z}$ is generated by its lowest positive element. It is a theorem of number theory that the lowest positive integer d such that $d = ra + sb$ for fixed integers a, b and $r, s \in \mathbb{Z}$ equals the greatest common divisor of a and b or $\gcd(a, b)$. Thus $\gcd(a, b)$generates H . □

Theorem 9: Let $n\mathbb{Z}$ and $m\mathbb{Z}$ be subgroups of $\mathbb{Z}$. Then $n\mathbb{Z} \cap m\mathbb{Z}$ is the subgroup generated by $\operatorname{lcm}(n, m)$.

Proof: The fact that $n\mathbb{Z} \cap m\mathbb{Z}$ is a subgroup is obvious since $n\mathbb{Z}$ and $m\mathbb{Z}$ are subgroups. To find a generator of $n\mathbb{Z} \cap m\mathbb{Z}$, we must find its lowest positive element. That is, the lowest positive integer p such that p is both a multiple of n and of m . This is the definition of the least common multiple of n and m , or $\operatorname{lcm}(n, m)$, and the result follows. □

It should be obvious by now that C_n and $\mathbb{Z}_n$, and C_∞ and $\mathbb{Z}$ are the *same groups*. This will be made precise in a later section but can be visualized by denoting any generator of C_n or C_∞ by 1.

We will have more to say about cyclic groups later, when we have more tools at our disposal.

Permutation groups

Symmetric groups

Theorem 1: Let A be any set. Then, the set S_A of bijections from A to itself, $f\ :, A \longrightarrow A$, form a group under composition of functions.

Proof: We have to verify the group axioms. Associativity is fulfilled since composition of functions is always associative: $(f \circ g) \circ h(x) = f \circ g(h(x)) = f(g(h(x))) = f(g \circ h(x)) = f \circ (g \circ h)(x)$ where the composition is defined. The identity element is the identity function given by $\mathrm{id}_A(a) = a$ for all $a \in A$. Finally, the inverse of a function f is the function f^{-1} taking $f(a)$ to a for all $a \in A$. This function exists and is unique since f is a bijection. Thus S_A is a group, as stated. □

S_A is called *the symmetric group on* A . When $A = \{1, 2, ..., n\}$, $n \in \mathbb{N}$, we write its symmetric group as S_n, and we call this group the *symmetric group on* n *letters*. It is also called the group of *permutations* on n letters. As we will see shortly, this is an appropriate name.

Instead of e , we will use a different symbol, namely ι , for the identity function in S_n .

When $\sigma \in S_n$, we can specify σ by specifying where it sends each element. There are many ways to encode this information mathematically. One obvious way is to indentify σ as the unique $n \times n$ matrix with value 1 in the entries $(i, \sigma(i))$ and 0 elsewhere. Composition of functions then corresponds to multiplication of matrices. Indeed, the matrix corresponding to ρ has value 1 in the entries $(i, \rho(i))$, which is the same as $(\sigma(j), \rho(\sigma(j)))$, so the product has value 1 in the entries $(j, \rho(\sigma(j)))$. This notation may seem cumbersome. Luckily, there exists a more convenient notation, which we will make use of.

We can represent any $\sigma \in S_n$ by a $2 \times n$ matrix $\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$. We obviously lose the correspondence between function composition and matrix multiplication, but we gain a more readable notation. For the time being, we will use this.

Remark 2: Let $\sigma, \rho \in S_n$. Then the product $\sigma\rho \equiv \sigma \circ \rho$ is the function obtained by *first* acting with ρ, and *then* by σ . That is, $\sigma\rho(x) = \sigma(\rho(x))$. This point is important to keep in mind when computing products in S_n. Some textbooks try to remedy the frequent confusion by writing functions like $(x)\rho\sigma$, that is, writing arguments on the *left* of functions. We will *not* do this, as it is not standard. The reader should use the next example and theorem to get a feeling for products in S_n.

Example 3: We will show the multiplication table for S_3. We introduce the special notation for S_3: $\iota = \rho_0$ $\rho_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, $\rho_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$, $\mu_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$, $\mu_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ and $\mu_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$. The multiplication table for S_3 is then

$\circ$	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_2	ρ_1
μ_2	μ_2	μ_3	μ_1	ρ_1	ρ_0	ρ_2
μ_3	μ_3	μ_1	μ_2	ρ_2	ρ_1	ρ_0

Theorem 4: S_n has order $n!$.

Proof: This follows from a counting argument. We can specify a unique element in S_n by specifying where each $i \in \{1, 2, ..., n\}$ is sent. Also, any permutation can be specified this way. Let $\sigma \in S_n$. In choosing $\sigma(1)$ we are completely free and have n choices. Then, when choosing $\sigma(2)$ we must choose from $\{1, ..., n\} - \{\sigma(1)\}$, giving a total of $n - 1$ choices. Continuing in this fashion, we see that for $\sigma(i)$ we must choose from $\{1, ..., n\} - \{\sigma(1), ..., \sigma(i-1)\}$, giving a total of $n + 1 - i$ choices. The total number of ways in which we can specify an element, and thus the number of elements in S_n is then $|S_n| = \prod_{i=1}^{n}(n+1-i) = n(n-1)... \cdot 2 \cdot 1 = n!$, as was to be shown. □

Theorem 5: S_n is non-abelian for all $n \geq 3$.

Proof: Let $\sigma = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ 2 & 1 & 3 & \ldots & n \end{pmatrix}$ be the function only interchanging 1 and 2, and $\rho = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ 1 & 3 & 2 & \ldots & n \end{pmatrix}$ be the function only interchanging 2 and 3. Then $\sigma\rho = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ 2 & 3 & 1 & \ldots & n \end{pmatrix}$ and $\rho\sigma = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ 3 & 1 & 2 & \ldots & n \end{pmatrix}$. Since $\sigma\rho \neq \rho\sigma$, S_n is not abelian. □

Definition 6: Let $\sigma \in S_n$ such that $\sigma^n = \iota$ for some $n \in \mathbb{Z}$. Then σ is called an n*-cycle*, where n is the smallest positive such integer. Let σ^* be the set of integers a such that $\sigma(a) \neq a$. Two cycles σ, ρ are called *disjoint* if $\sigma^* \cap \rho^* = \emptyset$. Also, a 2-cycle is called a *transposition*.

Theorem 7: Let $\sigma, \rho \in S_n$. If $\sigma^* \cap \rho^* = \emptyset$, then $\sigma\rho = \rho\sigma$.

Proof: For any integer $1 \leq a \leq n$ such that $a \in \sigma^*$ but $a \notin \rho^*$ we have $\sigma\rho(a) = \sigma(\iota(a)) = \sigma(a) = \iota(\sigma(a)) = \rho(\sigma(a)) = \rho\sigma(a)$. A similar argument holds for $a \in \rho^*$ but $a \notin \sigma^*$. If $a \notin \sigma^* \cup \rho^*$, we must have $\sigma\rho(a) = \sigma(a) = a = \rho(a) = \rho\sigma(a)$. Since $\sigma^* \cap \rho^* = \emptyset$, we have now exhausted every $a \in \{1, ..., n\}$, and we are done. □

Theorem 8: Any permutation can be represented as a composition of disjoint cycles.

Proof: Let $\sigma \in S_n$. Choose an element $a \in \sigma^*$ and compute $\sigma(a), \sigma^2(a), ..., \sigma^k(a) = a$. Since S_n is finite of order $n!$, we know that k exists and $k \leq n!$. We have now found a k-cycle σ_1 including a. Since $\sigma_1^* = \{a, \sigma(a), ..., \sigma^{k-1}(a)\}$, this cycle may be factored out σ, obtaining $\sigma = \sigma_1\sigma'$. Repeat this process, which terminates since σ^* is finite, and we have contructed a composition of disjoint cycles which equal σ. □

Now that we have shown that all permuations are just compositions of disjoint cycles, we can introduce the ultimate shorthand notation for permutations. For an n-cycle σ, we can show its action by choosing any element $a \in \sigma^*$ and writing $\sigma = \left(a \; \sigma(a) \; \sigma^2(a) \; ... \; \sigma^{n-1}(a)\right)$.

Theorem 9: Any n-cycle can be represented as a composition of transpositions.

Proof: Let $\sigma = \left(a \; \sigma(a) \; \sigma^2(a) \; ... \; \sigma^{n-1}(a)\right)$. Then, $\sigma = \left(a \; \sigma^2(a) \; ... \; \sigma^{n-1}(a)\right)\left(a \; \sigma(a)\right)$ (check this!), omitting the composition sign $\circ$. Interate this process to obtain $\sigma = \left(a \; \sigma^{n-1}(a)\right) ... \left(a \; \sigma^2(a)\right)\left(a \; \sigma(a)\right)$. □

Note 10: This way of representing σ as a product of transpositions is *not* unique. However, as we will see now, the "parity" of such a representation is well defined.

Definition 11: The *parity* of a permutation is *even* if it can be expressed as a product of an even number of transpositions. Otherwise, it is *odd*. We define the function $\operatorname{sgn}(\sigma) = 1$ if σ is even and $\operatorname{sgn}(\sigma) = -1$ if σ is odd.

Lemma 12: The indentity ι has even parity.

Proof: Observe first that $\iota \neq (a\ b)$ for $a \neq b$. Thus the minimum number of transpositions neccesary to represent ι is 2: $\iota = (a\ b)(a\ b)$. Now, assume that for any representation using less than k transpositions must be even. Thus, let $\iota = (a_1\ b_1)...(a_k\ b_k)$. Now, since in paticular $\iota(a_1) = a_1$, we must have $a_1 \in (a_i\ b_i)^*$ for some $1 < i \leq k$. Since disjoint transpositions commute, and $(a_r\ a_s)(a_i\ a_r) = (a_i\ a_s)(a_r\ a_s)$ where $a_i \neq a_r \neq a_s$, it is always possible to configure the transpositions such that the first two transpositions are either $(a_1\ b_1)(a_1\ b_1) = \iota$, reducing the number of transposition by two, or $(a_1\ b_1)(a_1\ b_2) = (a_1\ b_2)(b_1\ b_2)$. In this case we have reduced the number of transpositions involving a_1 by 1. We restart the same process as above. with the new representation. Since only a finite number of transpositions move a_1, we will eventually be able to cancel two permutations and be left with $k-2$ transpositions in the product. Then, by the induction hypothesis, $k-2$ must be even and so k is even as well, proving the lemma. □

Theorem 13: The parity of a permutation, and thus the sgn function, is well-defined.

Proof: Let $\sigma \in S_n$ and write σ as a product of transposition in two different ways: $\sigma = \tau_1...\tau_k = \tau_1'...\tau_{k'}'$. Then, since ι has even parity by Lemma 11 and $\iota = \sigma\sigma^{-1} = \tau_1...\tau_k\tau_{k'}'...\tau_1'$. Thus, $k + k' \equiv 0 \bmod 2$, and $k \equiv k' \bmod 2$, so σ has a uniquely defined parity, and consequentially sgn is well-defined. □

Theorem 14: Let $\sigma, \rho \in S_n$. Then, $\operatorname{sgn}(\sigma\rho) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\rho)$.

Proof: Decompose σ and ρ into transpositions: $\sigma = \mu_1...\mu_k$, $\rho = \nu_1...\nu_l$. Then $\sigma\rho = \mu_1...\mu_k\nu_1...\nu_l$ has parity given by $k+l$. If both are even or odd, $k+l$ is even and indeed $\operatorname{sgn}(\sigma)\operatorname{sgn}(\rho) = 1 \cdot 1 = 1 = \operatorname{sgn}(\sigma\rho)$. If one is odd and one is even, $k+l$ is odd and again $\operatorname{sgn}(\sigma)\operatorname{sgn}(\rho) = (-1) \cdot 1 = -1 = \operatorname{sgn}(\sigma\rho)$, proving the theorem. □

Lemma 15: The number of even permutations in S_n equals the number of odd permutations.

Proof: Let σ be any even permutation and τ a transposition. Then $\tau\sigma$ has odd parity by Theorem 14. Let E be the set of even permutations and O the set of odd permutations. Then the function $f : E \longrightarrow O$ given by $f(\sigma) = \tau\sigma$ for any $\sigma \in E$ and a fixed transposition τ, is a bijection. (Indeed, it is a transposition in S_{S_n}!) Thus E and O have the same number of elements, as stated. □

Definition 16: Let the set of all even permutations in S_n be denoted by A_n. A_n is called the *alternating group* on n letters.

Theorem 17: A_n *is* a group, and is a subgroup of S_n of order $\dfrac{n!}{2}$.

Proof: We first show that A_n is a group under composition. Then it is automatically a subgroup of S_n. That A_n is closed under composition follows from Theorem 14 and associativity is inherited from S_n. Also, the identity permutation is even, so $\iota \in A_n$. Thus A_n is a group and a subgroup of S_n. Since the number of even and odd permutations are equal by Lemma 14, we then have that $|A_n| = \dfrac{|S_n|}{2} = \dfrac{n!}{2}$, proving the theorem. □

Theorem 18: Let $n \geq 3$. Then A_n is generated by the 3-cycles in S_n.

Proof: We must show that any even permutation can be decomposed into 3-cycles. It is sufficient to show that this is the case for pairs of transpositions. Let $a, b, c, d \in S_n$ be distinct. Then, by some casework,

i) $(a\ b)(a\ b) = (a\ b\ c)^3$,

ii) $(a\ b)(b\ c) = (c\ a\ b)$, and

iii) $(a\ b)(c\ d) = (a\ d\ c)(a\ b\ c)$,

proving the theorem. □

In a previous section we proved Lagrange's Theorem: The order of any subgroup divides the order of the parent group. However, the converse statement, that a group has a subgroup for every divisor of its order, is false! The smallest group providing a counterexample is the alternating group A_4, which has order 12 but no subgroup of order 6. It has subgroups of orders 3 and 4, corresponding respectively to the cyclic group of order 3 and the Klein 4-group. However, if we add any other element to the subgroup corresponding to C_3, it generates the whole group A_4. We leave it to the reader to show this.

Dihedral Groups

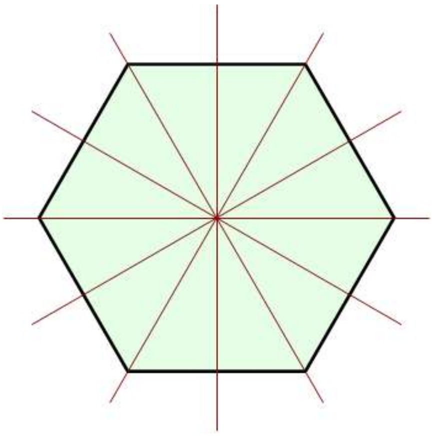
The dihedral groups are the symmetry groups of regular polygons. As such, they are subgroups of the symmetric groups. In general, a regular n-gon has n rotational symmetries and n reflection symmetries. The dihedral groups capture these by consisting of the associated rotations and reflections.

Definition 19: The *dihedral group of order* $2n$, denoted D_{2n}, is the group of rotations and reflections of a regular n-gon.

Theorem 20: The order of D_{2n} is precisely $2n$.

Proof: Let ρ be a rotation that generates a subgroup of order n in D_{2n}. Obviously, $\langle\rho\rangle$ then captures all the pure rotations of a regular n-gon. Now let μ be any rotation in The rest of the elements can then be found by composing each element in $\langle\rho\rangle$ with μ. We get a list of elements $D_{2n} = \{\iota, \rho, ..., \rho^{n-1}, \mu, \mu\rho, ..., \mu\rho^{n-1}\}$. Thus, the order of D_{2n} is $2n$, justifying its notation and proving the theorem. ▯

Remark 21: From this proof we can also see that $\{\rho, \mu\}$ is a generating set for G, and all elements can be obtained by writing arbitrary products of ρ and μ and simplifying the expression according to the rules $\rho^n = \iota$, $\mu^2 = \iota$ and $\rho\mu = \mu\rho^{-1}$. Indeed, as can be seen from the figure, a rotation composed with a reflection is new reflection.



The lines represent the reflection symmetries of a regular hexagon

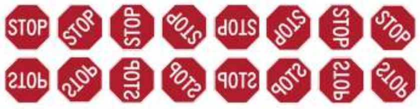


Illustration of the elements of the dihedral group D_{16} as rotations and reflections of a stop sign.

Homomorphism

We are finally making our way into the meat of the theory. In this section we will study *structure-preserving maps* between groups. This study will open new doors and provide us with a multitude of new theorems.

Up until now we have studied groups at the "element level". Since we are now about to take a step back and study groups at the "homomorphism level", readers should expect a sudden increase in abstraction starting from this section. We will try to ease the reader into this increase by keeping one foot at the "element level" throughout this section.

From here on out the notation e_G will denote the identity element in the group G unless otherwise specified.

Group homomorphisms

Definition 1: Let $(G,*)$ and $(H,\cdot)$ be groups. A *homomorphism* from G to H is a function $\phi : G \to H$ such that for all $g_1, g_2 \in G$,

$$\phi(g_1 * g_2) = \phi(g_1) \cdot \phi(g_2).$$

Thus, a homomorphism *preserves the group structure*. We have included the multiplication symbols here to make explicit that multiplication on the left hand side occurs in G, and multiplication on the right hand side occurs in H .

Already we see that this section is different from the previous ones. Up until now we have, excluding subgroups, only dealt with one group at a time. No more! Let us start by deriving some elementary and immediate consequences of the definition.

Theorem 2: Let G, H be groups and $\phi : G \to H$ a homomorphism. Then $\phi(e_G) = e_H$. In other words, the identity is mapped to the identity.

Proof: Let $g \in G$. Then, $\phi(g) = \phi(e_G g) = \phi(e_G)\phi(g)$, implying that $\phi(e_G)$ is the identity in H, proving the theorem. □

Theorem 3: Let G, H be groups and $\phi : G \to H$ a homomorphism. Then for any $g \in G$, $\phi(g^{-1}) = \phi(g)^{-1}$. In other words, inverses are mapped to inverses.

Proof: Let $g \in G$. Then $e_H = \phi(e_G) = \phi(gg^{-1}) = \phi(g)\phi(g^{-1})$ implying that $\phi(g^{-1}) = \phi(g)^{-1}$, as was to be shown. □

Theorem 4: Let G, G' be groups, $\phi : G \to G'$ a homomorphism and let H be a subgroup of G. Then $\phi(H) = \{\phi(h) \mid h \in H\}$ is a subgroup of G'.

Proof: Let $h_1, h_2 \in H$. Then $\phi(h_1), \phi(h_2) \in \phi(H)$ and $\phi(h_1)\phi(h_2)^{-1} = \phi(h_1)\phi(h_2^{-1}) = \phi(h_1 h_2^{-1})$. Since $h_1 h_2^{-1} \in H$, $\phi(h_1)\phi(h_2)^{-1} = \phi(h_1 h_2^{-1}) \in \phi(H)$, and so $\phi(H)$ is a subgroup of G'. □

Theorem 5: Let G, G' be groups, $\phi : G \to G'$ a homomorphism and let H' be a subgroup of G'. Then $\phi^{-1}(H') = \{g \in G \mid \phi(g) \in H'\}$ is a subgroup of G.

Proof: Let $g_1, g_2 \in \phi^{-1}(H')$. Then $\phi(g_1), \phi(g_2) \in H'$, and since H' is a subgroup, $\phi(g_1)\phi(g_2)^{-1} = \phi(g_1)\phi(g_2^{-1}) = \phi(g_1 g_2^{-1}) \in H'$. But then, $g_1 g_2^{-1} \in \phi^{-1}(H')$, and so $\phi^{-1}(H')$ is a subgroup of G. □

From Theorem 4 and Theorem 5 we see that homomorphisms *preserve subgroups*. Thus we can expect to learn a lot about the subgroup structure of a group G by finding suitable homomorphisms into G.

In particular, every homomorphism $\phi : G \to G'$ has associated with it two important subgroups.

Definition 6: A homomoprhism is called an *isomorphism* if it is bijective and its inverse is a homomorphism. Two groups are called *isomorphic* if there exists an isomorphism between them, and we write $G \approx H$ to denote " G is isomorphic to H ".

Theorem 7: A bijective homomorphism is an isomorphism.

Proof: Let G, H be groups and let $\phi\ :\ G \longrightarrow H$ be a bijective homomorphism. We must show that the inverse $\phi^{-1}$is a homomorphism. Let $h_1, h_2 \in H$. then there exist unique $g_1, g_2 \in G$ such that $\phi(g_1) = h_1$and $\phi(g_2) = h_2$. Then we have $\phi(g_1 g_2) = h_1 h_2$since ϕ is a homomorphism. Now apply $\phi^{-1}$to all equations. We obtain $\phi^{-1}(h_1) = g_1$, $\phi^{-1}(h_2) = g_2$and $\phi^{-1}(h_1 h_2) = g_1 g_2 = \phi^{-1}(h_1)\phi^{-1}(h_2)$, so $\phi^{-1}$is a homomorphism and thus ϕ is an isomorphism. □

Definition 8: Let G, G' be groups. A homomorphism that maps every element in G to $e' \in G'$is called a *trvial homomorphism* (or *zero homomorphism*), and is denoted by $0_{GG'}\ :\ G \longrightarrow G'$

Definition 9: Let H be a subgroup of a group G . Then the homomorphism $i\ :\ H \longrightarrow G$ given by $i(h) = h$ is called the *inclusion* of H into G . Let G'be a group isomorphic to a subgroup H of a group G . Then the isomorphism $\phi\ :\ G' \longrightarrow H$ induces an injective homomorphism $i'\ :\ G' \longrightarrow G$ given by $i'(g') = \phi(g')$, called an *imbedding* of G'into G . Obviously, $i' = i \circ \phi$.

Definition 10: Let G, G'be groups and $\phi\ :\ G \longrightarrow G'$a homomorphism. Then we define the following subgroups:

i) $\ker\ \phi = \{g \in G \mid \phi(g) = e'\} \le G$, called the *kernel* of ϕ , and

ii) $\operatorname{im} \phi = \{g' \in G' \mid (\exists g \in G)\phi(g) = g'\} \le G'$, called the *image* of ϕ .

Theorem 11: The composition of homomorphisms is a homomorphism.

Proof: Let G, H, K be groups and $\phi : G \longrightarrow H$ and $\psi\ :\ H \longrightarrow K$ homomorphisms. Then $\psi \circ \phi\ :\ G \longrightarrow K$ is a *function*. We must show it is a homomorphism. Let $g, h \in G$. Then $\psi \circ \phi(gh) = \psi(\phi(gh)) = \psi(\phi(g)\phi(h)) = \psi(\phi(g))\psi(\phi(h)) = \psi \circ \phi(g)\psi \circ \phi(h)$, so $\psi \circ \phi$ is indeed a homomorphisms. □

Theorem 12: Composition of homomorphisms is associative.

Proof: This is evident since homomorphisms are *functions*, and composition of functions is associative. □

Corollary 13: The composition of isomorphisms is an isomorphism.

Proof: This is evident from Theorem 11 and since the composition of bijections is a bijection. □

Theorem 14: Let G, H be groups and $\phi\ :\ G \longrightarrow H$ a homomorphism. Then ϕ is injective if and only if $\ker\ \phi = \{e\} \subseteq G$.

Proof: Assume $\ker\ \phi = \{e\}$ and $g_1, g_2 \in G$. Then $\phi(g_1) = \phi(g_2)\ \Leftrightarrow \phi(g_1)\phi(g_2)^{-1} = \phi(g_1 g_2^{-1}) = e_H$, implying that $g_1 g_2^{-1} \in \ker\ \phi$. But by assumption then $g_1 g_2^{-1} = e\ \Leftrightarrow g_1 = g_2$, so ϕ is injective. Assume now that $\ker\ \phi \ne \{e\}$ and $g \in G$. Then there exists another element $k \in \ker\ \phi$ such that $k \neq e$. But then $\phi(g) = \phi(kg)$. Since both g and $kg \neq g$ map to $\phi(g) = \phi(kg)$, ϕ is not injective, proving the theorem. □

Corollary 15: Inclusions are injective.

Proof: The result is immediate. Since $i(h) = h$ for all $h \in H \le G$, we have $i^{-1}(e) = \{e\}$. □

The kernel can be seen to satisfy a universal property. The following theorem explains this, but it is unusually abstract for an elementary treatment of groups, and the reader should not worry if he/she cannot understand it immediately.

Theorem 16: Let G, G' be groups and $\phi : G \to G'$ a group homomorphism. Also let H be a group and $\psi : H \to G$ a homomorphism such that $\phi \circ \psi = 0_{HG'}$. Also let $i : \ker \phi \to G$ is the inclusion of $\ker \phi$ into G. Then there exists a unique homomorphism $\bar{\psi} : H \to \ker \phi$ such that $\psi = i \circ \bar{\psi}$.
Proof: Since $\phi \circ \psi = 0_{HG'}$, by definition we must have $\psi(H) \leq \ker \phi$, so $\bar{\psi}$ exists. The commutativity $\psi = i \circ \bar{\psi}$ then forces $\bar{\psi}(h) = \psi(h) \in \ker \phi$, so $\bar{\psi}$ is unique. □

$$\begin{tikzcd}
H \arrow[r, "\psi"] \arrow[rr, bend left, "0_{HG'}"] \arrow[dr, dotted, "\exists!\bar{\psi}"'] & G \arrow[r, "\phi"] & G' \\
& \ker \phi \arrow[u, hook, "i"] \arrow[ur, "0"'] &
\end{tikzcd}$$

Commutative diagram showing the universal property of kernels.

Definition 17: A *commutative diagram* is a pictorial presentation of a network of functions. Commutativity means that when several routes of function composition from one object lead to the same destination, the two compositions are equal as functions. As an example, the commutative diagram on the right describes the situation in Theorem 16. In the commutative diagrams (or diagrams for short, we will not show diagrams which no not commute) shown in this chapter on groups, all functions are implicitly assumed to be group homomorphisms. Monomorphisms in diagrams are often emphasized by hooked arrows. In addition, epimorphims are often emphasized by double headed arrows. That an inclusion is a monomorphism will be proven shortly.

Remark 18: From the commutative diagram on the right, the kernel can be *defined* completely without reference to elements. Indeed, Theorem 16 would become the definition, and our Definition 10 i) would become a *theorem*. We will not entertain this line of thought in this book, but the advanced reader is welcome to work it out for him./herself.

Automorphism Groups

In this subsection we will take a look at the homomorphisms from a group to itself.

Definition 19: A homomorphism from a group G to itself is called an *endomorphism of* G. En endomorphism which is also an isomorphism is called an *automorphism*. The set of all endomorphisms of G is denoted $\text{End}(G)$, while the set of all automorphisms of G is denoted $\text{Aut}(G)$.

Theorem 20: $\text{End}(G)$ is a monoid under composition of homomorphisms. Also, $\text{Aut}(G)$ is a submonoid which is also a group.

Proof: We only have to confirm that $\text{End}(G)$ is closed and has an identity, which we know is true. For $\text{Aut}(G)$, the identity homomorphism $\text{id}_G : G \to G$ is an isomorphism and the composition of isomorphisms is an isomorphism. Thus $\text{Aut}(G)$ is a submonoid. To show it is a group, note that the inverse of an automorphism is an automorphism, so $\text{Aut}(G)$ is indeed a group. □

Groups with Operators

An endomorphism of a group can be thought of as a *unary operator* on that group. This motivates the following definition:

Definition 21: Let G be a group and $\Omega \subset \text{End}(G)$. Then the pair (G, Ω) is called a *group with operators*. Ω is called the *operator domain* and its elements are called the *homotheties* of G. For any $\omega \in \Omega$, we introduce the shorthand $\omega(g) = g^\omega$ for all $g \in G$. Thus the fact that the homotheties of G are endomorphisms can be expressed thus: for all $a, b \in G$ and $\omega \in \Omega$, $(ab)^\omega = a^\omega b^\omega$.

Example 22: For any group G, the pair $(G, \emptyset)$ is trivially a group with operators.

Lemma 23: Let (G, Ω) be a group with operators. Then Ω can be extended to a submonoid Ω' of $\text{End}(G)$ such that the structure of (G, Ω') is identical to (G, Ω).

Proof: Let Ω' include the identity endomorphism and let Ω be a generating set. Then Ω' is closed under compositions and is a monoid. Since any element of Ω' is expressible as a (possibly empty) composition of elements in Ω, the structures are identical. ▯

In the following, we assume that the operator domain is always a monoid. If it is not, we can extend it to one by Lemma 23.

Definition 24: Let (G, Ω) and (H, Ω) be groups with operators with the same operator domain. Then a homomorphism $\phi : (G, \Omega) \to (H, \Omega)$ is a group homomorphism $\phi : G \to H$ such that for all $\omega \in \Omega$ and $g \in G$, we have $\phi(g^\omega) = \phi(g)^\omega$.

Definition 25: Let (G, Ω) be a group with operators and H a subgroup of G. Then H is called a *stable subgroup* (or a Ω-invariant subgroup) if for all $h \in H$ and $\omega \in \Omega$, $h^\omega \in H$. We say that H *respects the homotheties* of G. In this case (H, Ω) is a *sub-group with operators*.

Example 26: Let V be a vector space over the field F. If we by V_+ denote the underlying abelian group under addition, then $V = (V_+, F)$ is a group with operators, where for any $k \in F$ and $v \in V_+$, we define $v^k = kv$. Then the stable subgroups are precisely the linear subspaces of V (show this).

Problems

Problem 1: Show that there is no nontrivial homomorphism from $\mathbb{Z}_5$ to S_3.

Normal subgroups and Quotient groups

In the preliminary chapter we discussed equivalence classes on sets. If the reader has not yet mastered this notion, he/she is advised to do so before starting this section.

Normal Subgroups

Recall the definition of kernel in the previous section. We will exhibit an interesting feature it possesses. Namely, let ak, $k \in \ker \phi \leq G$ be in the coset $a \ker \phi$. Then there exists a $k' \in \ker \phi$ such that $k'a = ak$ for all $a \in G$. This is easy to see because a coset of the kernel includes all elements in G that are mapped to a paticular element. The kernel inspires us to look for what are called *normal subgroups*.

Definition 1: A subgroup $N \leq G$ is called *normal* if $gNg^{-1} = N$ for all $g \in G$. We may sometimes write $N \trianglelefteq G$ to emphasize that N is normal in G.

Theorem 2: A subgroup $N \leq G$ is normal if and only if $gN = Ng$ for all $g \in G$.

Proof: By the definition, a subgroup is normal if and only if $gNg^{-1} = N$ since conjugation is a bijection. The theorem follows by multiplying on the right by g. ▯

We stated that the kernel is a normal subgroup in the introduction, so we had better well prove it!

Theorem 3: Let $\phi : G \to G'$ be any homomorphism. Then $\ker \phi \leq G$ is normal.

Proof: Let $k \in \ker \phi$ and $g \in G$. Then $\phi(gkg^{-1}) = \phi(g)e'\phi(g)^{-1} = e'$, so $gkg^{-1} \in \ker \phi$, proving the theorem. ▯

Theorem 4: Let G, G' be groups and $\phi : G \to G'$ a group homomorphism. Then if H is a normal subgroup of $\operatorname{im} \phi$, then $\phi^{-1}(H)$ is normal in G.

Proof: Let $g \in G$ and $h \in \phi^{-1}(H)$. Then $\phi(ghg^{-1}) = \phi(g)\phi(h)\phi(g)^{-1} \in H$ since H is normal in $\operatorname{im} \phi$, and so $ghg^{-1} \in \phi^{-1}(H)$, proving the theorem. ▯

Theorem 5: Let G, G' be groups and $\phi : G \to G'$ a group homomorphism. Then if H is a normal subgroup of G, $\phi(H)$ is normal in $\operatorname{im} \phi$.

Proof: Let $g' \in \operatorname{im}\phi$ And $h \in H$. Then if $g \in G$ such that $\phi(g) = g'$, we have $g'\phi(h)g'^{-1} = \phi(g)\phi(h)\phi(g)^{-1} = \phi(ghg^{-1}) = \phi(h') \in \phi(H)$ for some $h' \in H$ since H is normal. Thus $g'\phi(H)g'^{-1} = \phi(H)$ for all $g' \in \operatorname{im}\phi$ and so H is normal in $\operatorname{im}\phi$. □

Corollary 6: Let G, G' be groups and $\phi : G \to G'$ a surjective group homomorphism. Then if H is a normal subgroup of G, $\phi(H)$ is normal in G'.

Proof: Replace $\operatorname{im}\phi$ with G' in the proof of Theorem 5. □

Remark 7: If H is a normal subgroup of G and K is a normal subgroup of H, it does **not** necessarily imply that K is a normal subgroup of G. The reader is invited to display a counterexample of this.

Theorem 8: Let G be a group and H, K be subgroups. Then

i) If H is normal, then $HK = KH$ is a subgroup of G.

ii) If both H and K are normal, then $HK = KH$ is a normal subgroup of G.

iii) If H and K are normal, then $H \cap K$ is a normal subgroup of G.

Proof: i) Let H be normal. First, since for each $k \in K$, there exists $h, h' \in H$ such that $kh = h'k$, so $KH = HK$. To show HK is a subgroup, let $hk, h'k' \in HK$. Then $h'k'(hk)^{-1} = h'k'k^{-1}h^{-1} = h'h''k'k^{-1} \in HK$ for some $h'' \in H$ since H is normal, and so HK is a subgroup.

ii) Let $g \in G$ and $hk \in HK$. Then since both H and K are normal, there exists $h' \in H$, $k' \in K$ such that $ghkg^{-1} = ghg^{-1}k' = gg^{-1}h'k' = h'k' \in HK$. It follows that $gHKg^{-1} = HK$ and so HK is normal.

iii) Let $g \in G$ and $h \in H \cap K$. Then $ghg^{-1} \in H$ since H is normal, and similarly $ghg^{-1} \in K$. Thus $ghg^{-1} \in H \cap K$ and it follows that $H \cap K$ is normal. □

Examples of Normal Subgroups

In the following, let G be any group. Then G has associated with it the following normal subgroups.

i) The *center* of G, denoted $Z(G)$, is the subgroup of elements which commute with all others. $Z(G) = \{z \in G \mid (\forall g \in G) zg = gz\}$. That $Z(G)$ is a normal subgroup is easy to verify and is left to the reader.

ii) The *commutator subgroup* of G, denoted $G^{(1)}$ or $[G, G]$, is the subgroup generated by the subset $\{[g, h] \mid g, h \in G\}$ where $[g, h] = g^{-1}h^{-1}gh$ for all $g, h \in G$. For $s \in G$, we introduce the shorthand $sgs^{-1} = g^s$. Then we have $s[g, h]s^{-1} = [g^s, h^s]$, such that for any product of commutators $[g_1, h_1][g_2, h_2]...[g_n, h_n]$ where all elements are in G, we have $s[g_1, h_1][g_2, h_2]...[g_n, h_n]s^{-1} = [g_1^s, h_1^s][g_2^s, h_2^s]...[g_n^s, h_n^s]$, and so $G^{(1)}$ is normal.

Remark 9: We can iterate the commutator subgroup construction and define $G^{(0)} = G$ and $G^{(n)} = [G^{(n-1)}, G^{(n-1)}]$ for all $n \in \mathbb{N}$. We will not use the commutator subgroup in future results in this book, so for us it is merely a curiosity.

Equivalence Relations on Groups

Why are normal subgroups important? In the preliminary chapter we discussed equialence relations and the associated set of equivalence classes. If G is a group and $\sim$ is an equivalence relation, when does $G/\sim$ admit a group structure? Of course we have to specify the multiplication on $G/\sim$. We will do so now.

Definition 10: Let G be a group and $\sim$ is an equivalence relation on G , we define multiplication on the equivalence classes in $G/\sim$ such that for all $a, b \in G$,

$$[a][b] = [ab]$$

This is indeed the only natural way to do it. Take the two equivalence classes, choose representatives, compute their product and take its equivalence class. The alert reader will have only one thing on his/her mind: is this well defined? For a general equivalence relation, the answer is no. The reader is invited to come up with an example. What is more interesting is *when* is it well defined? By the definition above, we obviously need the projection map $\pi : G \to G/\sim$ defined by $a \mapsto [a]$ to be a homomorphism. We can in fact condense the requirements down to two, both having to do with cancellation laws.

Theorem 11: Let G be a group and $\sim$ an equivalence relation on G . Then $G/\sim$ is a group under the natural multiplication if and only if for all $a, b, g \in G$

$$a \sim b \Leftrightarrow ag \sim bg \wedge ga \sim gb .$$

Proof: Assume $G/\sim$ is a group. Since $a \sim b \Leftrightarrow [a] = [b]$, the property follows from the cancellation laws in G . Assume now that the property holds. Then its multiplication rule is well defined, and must verify that $G/\sim$ is a group. Let $a, b, c \in G$, then associativity is inherited from G :

$$([a][b])[c] = [ab][c] = [(ab)c] = [a(bc)] = [a][bc] = [a]([b][c]).$$

The identity in $G/\sim$ is the equivalence class of $e \in G$, $[e]$:

$$[e][a] = [ea] = [a] = [ae] = [a][e].$$

Finally, the inverse of $[a]$is $[a^{-1}]$:

$$[g][g^{-1}] = [gg^{-1}] = [e] = [g^{-1}g] = [g^{-1}][g].$$

So $G/\sim$ really defines a group structure, proving the theorem. □

We will call an equivalence relation $\sim$ *compatible* with G if $G/\sim$ is a group. Then, $G/\sim$ is called the *quotient group* of G by $\sim$. Also, as an immediate consequence, this makes $\pi : G \to G/\sim$ into a homomorphism, but not just any homomorphism! It satisfies a universal property!

Theorem 12: Let $\sim$ be en equivalence relation compatible with G , and $\phi : G \to H$ a group homomorphism such that $a \sim b \Rightarrow \phi(a) = \phi(b)$. Then there exists a unique homomorphism $\tilde{\phi} : G/\sim \to H$ such that $\phi = \tilde{\phi} \circ \pi$.

Proof: In the preliminary chapter on set theory, we showed the corresponding statement for sets, so we know that $\tilde{\phi}$ exists as a function between sets. We have to show that it is a homomorphism. This follows immediately: since $\tilde{\phi}([a]) = \phi(a)$by commutativity, we have $\tilde{\phi}([a])\tilde{\phi}([b]) = \phi(a)\phi(b) = \phi(ab) = \tilde{\phi}([ab]) = \tilde{\phi}([a][b])$. As stated already, $\tilde{\phi}([a]) = \phi(a)$shows uniqueness, proving the theorem. □

$$\begin{tikzcd} G \arrow[r, "\phi"] \arrow[d, "\pi"'] & H \\ G/\sim \arrow[ur, dotted, "\exists!\tilde{\phi}"'] & \end{tikzcd}$$

Commutative diagram showing the universal property satisfied by the projection homomorphism.

Lemma 13: Let $\sim$ be an equivalence relation on a group G such that $a \sim b \Leftrightarrow ga = gb$. Then $H = [e]$is a subgroup of G and $a \sim b \Leftrightarrow a^{-1}b \in H \Leftrightarrow aH = bH$.

Proof: First off, H is nonempty since $e \in H$. Let $a, b \in H$. Then $b \sim e \Leftrightarrow e \sim b^{-1}$by multiplying on the left by b^{-1}. Then since $e \sim a$ we have $a^{-1} \sim e$ by the same argument. Applying transitivity gives $a^{-1} \sim b^{-1}$. Finally, multiplying on the left by a gives $e \sim ab^{-1}$, giving $ab^{-1} \in H$ and so $H = [e]$is a subgroup.
Assume $a \sim b$ for $a, b \in G$. Then $[a] = [b]$implying $[a^{-1}b] = [e]$. Thus $a^{-1}b \in [e]$. Now assume $a^{-1}b \in [e]$. Then $[a^{-1}b] = [e]$and so $[a] = [b]$and finally $a \sim b$.
Assume $a^{-1}b \in H$. Then since H is a subgroup, we have $a^{-1}bH = H$ and so $aH = bH$. Finally, assume $aH = bH$. Then $a^{-1}bH = H$. Since in paticular $e \in H$, this implies $a^{-1}b \in H$, completing the proof. ▯
The mirror version using right cosets and the equivalence relation $a \sim b \Leftrightarrow ag = bg$and $a \sim b \Leftrightarrow ab^{-1} \in H \Leftrightarrow Ha = Hb$ is completely analogous. Stating the theorem and writing out the proof is left to the reader as an exercise.
We have showed how an equialence relation defines a subgroup of G . In fact the equivalence classes are all cosets of this subgroup. We will now go the other way, and show how a subgroup defines an equivalence relation on G .

Lemma 14: Let H be a subgroup of a group G . Then,

i) $a \sim_L b \Leftrightarrow a^{-1}b \in H \Leftrightarrow aH = bH$ is an equivalence relation such that $a \sim_L b \Leftrightarrow ga \sim_L gb$ for all $g \in G$.

ii) $a \sim_R b \Leftrightarrow ab^{-1} \in H \Leftrightarrow Ha = Hb$ is an equivalence relation such that $a \sim_R b \Leftrightarrow ag \sim_R bg$ for all $g \in G$.

Proof: We will prove i). The proof for ii) is similar and is left as an exercise for the reader.

The fact that $\sim$ is an equivalence relation and that $a^{-1}b \in H \Leftrightarrow aH = bH$ was proven in the section on subgroups. Assume $a \sim_L b$. Then for all $g \in G$, $(ga)^{-1}(gb) = a^{-1}g^{-1}gb = a^{-1}b \in H$ such that $ga \sim_L gb$. Now assume $ga \sim_L gb$, Then $a^{-1}b = a^{-1}g^{-1}gb = (ga)^{-1}(gb) \in H$ such that $a \sim_L b$, completing the proof. ▯
Theorem 15: For every equivalence relation $\sim_L$ on G such that $a \sim_L b \Leftrightarrow ga \sim_L gb$, there exists a unique subgroup H of G such that $G/\sim$ are precisely the left cosets of H .
Proof: This follows from Lemma 13 and Lemma 14.

Again, the mirror statement is completely analogous. Stating the theorem is left to the reader as an exercise.

Quotients with respect to Normal Subgroups

Lemma 16: An equivalence relation $\sim$ on G is compatible with G if and only if $[e] = H$ is a normal subgroup of G .
Proof: Assume $\sim$ is compatible, $g \in G$ and $a \in H$. Then $a \sim e$, and compatibility gives us $gag^{-1} \sim gg^{-1} = e$, and so $gag^{-1} \in H$. Since a is arbitrary, we obtain $gHg^{-1} = H$ for all $g \in G$ and so H is normal. Assume now that H is normal. Then $aH = bH \Leftrightarrow a^{-1}b \in H$, $Ha = Hb \Leftrightarrow ab^{-1} \in H$ and $aH = Ha$ for all $a, b \in G$. Using this, we obtain $a \sim b \Leftrightarrow ab^{-1} = ab^{-1}bg(bg)^{-1} \sim e \Leftrightarrow ab^{-1}bg = ag \sim bg$ and similarily for the right hand case, so $\sim$ is compatible with G . ▯
Definition 17: When an equivalence relation is given by specifying a normal subgroup H , the quotient group with respect to this equivalence relation is denoted G/H . We then refer to G/H as the quotient of G with respect to H , or G modulo H . Note that this complies with previous definitions of this notation.
Multiplication in G/H is given as before as $(aH)(bH) = (ab)H$, with identity H and $(aH)^{-1} = a^{-1}H$ for all $a, b \in G$.
Definition 18: Let H be a normal subgroup of G . Then we define the *projection homomorphism* $\pi : G \to G/H$ by $\pi(a) = aH$ for all $a \in G$.

Theorem 19: A subgroup is normal if and only if it is the kernel of some homomorphism.

Proof: We have already covered the left implication. For the right implication, assume H is normal. Then G/H is a group and we have the projection homomorphism $\pi : G \to G/H$ as defined above. Since for all $h \in H$ we have $\pi(h) = hH = H$, $\ker \pi = H$ and so H is the kernel of a homomorphism. □

Theorem 20: Let G, G' be groups, $\phi : G \to G'$ a homomorphism and H a normal subgroup of G such that $H \subseteq \ker \phi$. Then there exists a unique homomorphism $\tilde{\phi} : G/H \to G'$ such that $\tilde{\phi} \circ \pi = \phi$.

Proof: This follows from Theorem 12 by letting $a \sim b \Leftrightarrow aH = bH$. □

The Isomorphism Theorems

Theorem 21 (First Isomorphism Theorem): Let G, G' be groups and $\phi : G \to G'$ a homomorphism. Then $G/\ker \phi \approx \operatorname{im} \phi$.

Proof: From Theorem 20 we have that there exists a unique homomorphism $\tilde{\phi} : G/\ker \phi \to G'$ such that $\phi = \tilde{\phi} \circ \pi$. We have to show that $\tilde{\phi}$ is an isomorphism when we corestrict to $\operatorname{im} \phi$. This is immediate, since $\phi(a) = \phi(b) \Leftrightarrow a \ker \phi = b \ker \phi$ by Lemma 13, so that $\tilde{\phi}$ is injective, and for any $g' \in \operatorname{im} \phi$ there is a $g \in G$ such that $\phi(g) = \tilde{\phi}(g \ker \phi) = g'$ so that it is surjective and therefore an isomorphism. □

$$\begin{tikzcd}
G \arrow[r, "\phi"] \arrow[d, "\pi"] & \operatorname{im} \phi \\
G/\ker \phi \arrow[ur, "\tilde{\phi}"'] &
\end{tikzcd}$$

Commutative diagram showing the first isomorphism theorem. $\tilde{\phi}$ is an isomorphism.

Lemma 22: Let G be a group, H a subgroup and K a normal subgroup of G. Then $H \cap K$ is a normal subgroup of H.

Proof: Let $h \in H$ and $k \in H \cap K$. Then $hkh^{-1} \in H$ since $h, k \in H$ and H is a subgroup and $hkh^{-1} \in K$ since $k \in K$, $h \in G$ and K is normal in G. Thus $hkh^{-1} \in H \cap K$ and $H \cap K$ is normal in H. □

Theorem 23 (Second Isomorphism Theorem): Let G be a group, H a subgroup and K a normal subgroup of G. Then $HK/K \approx \dfrac{H}{H \cap K}$.

Proof: Define $\phi : H \to HK/K$ by $\phi(h) = hK$ for all $h \in H$. ϕ is surjective since any element in HK can be written as hk with $h \in H$ and $k \in K$, so $\phi(h) = \pi(hk) = hkK = hK$. We also have that $\ker \phi = \{h \in H \mid hK = K\} = \{h \in H \mid h \in K\} = H \cap K$ and so $HK/K \approx \dfrac{H}{H \cap K}$ by the first isomorphism theorem. □

Lemma 24: Let G be a group, and let H, K be normal subgroups of G such that $K \subseteq H \subseteq G$. Then H/K is a normal subgroup of G/K.

Proof: Let $h \in H$ and $g \in G$. Then $ghg^{-1} = h'$ for some $h' \in H$ since H is normal. Thus $(gK)(hK)(gK)^{-1} = (ghg^{-1})K = h'K$, showing that H/K is normal in G/K. □

Theorem 25 (Third Isomorphism Theorem) Let G be a group, and let H, K be normal subgroups of G such that $K \subseteq H \subseteq G$. Then $\dfrac{G/K}{H/K} \approx G/H$.

Proof: Let $\phi : G/K \to G/H$ be given by $\phi(gK) = gH$. This is well defined and surjective since $K \subseteq H$, and is a homomorphism. Its kernel is given by $\ker \phi = \{gK \in G/K \mid gH = H\} = \{gK \in G/K \mid g \in H\} = H/K$, so by the first isomorphism theorem, $\dfrac{G/K}{H/K} \approx G/H$. □

Theorem 26 (The Correspondence Theorem): Let G be a group and K be a normal subgroup. Now let $A = \{H \leq G \mid K \leq H\}$ and $B = \{H' \mid H' \leq G/K\}$. Then $\pi : H \mapsto \pi(H)$ is an order-preserving bijection from A to B.

Proof: We must show injectivity and surjectivity. For injectivity, note that if $K \leq H$, then $\pi^{-1}(\pi(H)) = HK = H$, so if $H_1, H_2 \in A$ such that $\pi(H_1) = \pi(H_2)$, then $H_1 = \pi^{-1}(\pi(H_1)) = \pi^{-1}(\pi(H_2)) = H_2$, proving injectivity. For surjectivity, let $H' \in B$. Then $K \leq \pi^{-1}(H') \leq G$, so that $\pi^{-1}(H') \in A$, and $\pi(\pi^{-1}(H')) = H'$, proving surjectivity. Lastly, since $H_1 \subseteq H_2$ implies that $\pi(H_1) \subseteq \pi(H_2)$, the bijection is order-preserving. □

Note 27: The correspondence Theorem is sometimes called *The Forth Isomorphism Theorem.*

Theorem 28: Let $H \in A$ from Theorem 26. Then H is normal if and only if $\pi(H)$ is normal in G/K, and then $G/H \approx \dfrac{G/K}{\pi(H)}$.

Proof: Since π is surjective, H normal implies $\pi(H)$ normal. Assume that $\pi(H)$ is normal. Then $\pi^{-1}(\pi(H)) = H$ and so H is normal since it is the preimage of a normal subgroup. To show the isomorphism, let $\phi : G \longrightarrow \dfrac{G/K}{\pi(H)}$ be given by a composition of projections: $\phi : \pi_{\pi(H)} \circ \pi_K$. Then $\ker \phi = \{g \in G \mid \phi(g) = \pi(H)\} = \{g \in G \mid \pi(g) \in \pi(H)\} = \pi^{-1}(\pi(H)) = H$, so by the first isomorphism teorem, $G/H \approx \dfrac{G/K}{\pi(H)}$. □

Corollary 29: Let G be a group and H a normal subgroup. Then for any $K' \leq G/H$ there exists a unique subgroup $K \leq G$ such that $H \leq K \leq G$ and $K = K'H$. Also, K is normal in G if and only if K' is normal in G/H.

Proof: From Theorem 26 we have that the projection $K \mapsto \pi(K) = K'$ is a bijection, and since $\pi(g) = gH$ for all $g \in G$, we have $K = K'H$. The second part follows from Theorem 28. □

Simple Groups

Definition 30: A group is called *simple* is it has no non-trivial proper normal subgroups.

Example 31: Every cyclic group $\mathbb{Z}_p$, where p is prime, is simple.

Definition 32: Let G be a group and H a normal subgroup. H is called a *maximal normal subgroup* if for any normal subgroup K of G, we have $K \subseteq H$.

Theorem 33: Let G be a group and H a normal subgroup. Then H is a maximal normal subgroup if and only if the quotient G/H is simple.

Proof: By Theorem 26 and Theorem 28, G/H has a nontrivial normal subgroup if and only if there exists a proper normal subgroup K of G such that $H \leq K \leq G$. That is, H is not maximal if ans only if G/H is not simple. The theorem follows. □

Problems

Problem 1: Recall the unitary and special unitary groups from the section about subgroups. Define the *projective unitary group of order* n as the group $PU(n) = U(n)/Z(U(n))$. Similarily, define the *projective special unitary group of order* n as $PSU(n) = SU(n)/Z(SU(n))$.

i) Show that $Z(SU(n)) = SU(n) \cap Z(U(n) \approx \mathbb{Z}_n$

ii) Using the second isomorphism theorem, show that $PSU(n) \approx PU(n)$.

Products and Free Groups

During the preliminary sections we introduced two important constructions on sets: the direct product and the disjoint union. In this section we will construct the analogous constructions for groups.

Product Groups

Definition 1: Let G and H be groups. Then we can define a group structure on the direct product $G \times H$ of the *sets* G and H as follows. Let $(g_1, h_1), (g_2, h_2) \in G \times H$. Then we define the multiplication componentwise: $(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2)$. This structure is called the *direct product* of G and H .

Remark 2: The product group *is* a group, with identity $(e_G, e_H)$and inverses $(g, h)^{-1} = (g^{-1}, h^{-1})$. The order of $G \times H$ is $|G \times H| = |G||H|$.

Theorem 3: Let G and H be groups. Then we have homomorphisms $\pi_1 : G \times H \to G$ and $\pi_2 : G \times H \to H$ such that $\pi_1(g, h) = g$ and $\pi_2(g, h) = h$ for all $(g, h) \in G \times H$. These are called the *projections* on the first and second factor, respectively.

Proof: The projections are obviously homomorphisms since they are the identity on one factor and the trivial homomorphism on the other. ▯

Corollary 4: Let G and H be groups. Then $\frac{G \times H}{H} \approx G$ and $\frac{G \times H}{G} \approx H$.

Proof: This follows immediately from plying the first isomorphism theorem to Theorem 3 and using that $G \times \{e_H\} \approx G$ and $\{e_G\} \times H \approx H$. ▯

Theorem 5: Let G and H be groups. Then $G \times \{e_H\}$ and $\{e_G\} \times H$ are normal subgroups of $G \times H$.

Proof: We prove the theorem for $G \times \{e_H\}$. The case for $\{e_G\} \times H$ is similar. Let $g, g' \in G$ and $h \in H$. Then $(g, h)(g', e_H)(g, h)^{-1} = (gg'g^{-1}, hh^{-1}) = (gg'g^{-1}, e_H) \in G \times \{e_H\}$. ▯

We stated that this is an analogous construction to the direct product of sets. By that we mean that it satiesfies the same universal property as the direct product. Indeed, to be called a "product", a construction should have to satisfy this universal property.

Theorem 6: Let G and H be groups. Then if K is a group with homomorphisms $\phi_1 : K \to G$ and $\phi_2 : K \to H$, then there exists a unique homomorphism $u : K \to G \times H$ such that $\phi_1 = \pi_1 \circ u$ and $\phi_2 = \pi_2 \circ u$.

$$\begin{tikzcd}
 & K \arrow[dl, "\phi_1"'] \arrow[dr, "\phi_2"] \arrow[d, dotted, "\exists! u"] & \\
G & G \times H \arrow[l, "\pi_1"'] \arrow[r, "\pi_2"] & H
\end{tikzcd}$$

Commutative diagram showing the universal property satisfied by the direct product.

Proof: By the construction of the direct product, $u : K \to G \times H$ is a homomorphism if and only if $\pi_1 \circ u$ and $\pi_2 \circ u$ are homomorphisms. Thus $u : K \to G \times H$ defined by $u((g, h)) = (\phi_1(g), \phi_2(h))$ is one homomorphism satisfying the theorem, proving existence. By the commutativity condition this is the only such homomorphism, proving uniqueness. ▯

Products of Cyclic Groups

Theorem 7: The order of an element $(a, b) \in \mathbb{Z}_m \times \mathbb{Z}_n$ is $|(a, b)| = \text{lcm}(|a|, |b|)$.

Proof: The lowest positive number c such that $(a, b)^c = (ac, bc) = (0, 0)$is the smallest number such that $ac = rm$ and $bc = sn$ for integers r, s. It follows that c divides both $|a|$ and $|b|$ and is the smallest such number. This is the definition of the least common divider. ▯

Theorem 8: $\mathbb{Z}_m \times \mathbb{Z}_n$ is isomorphic to $\mathbb{Z}_{mn}$ if and only if m and n are relatively prime.

Proof: We begin with the left implication. Assume $\mathbb{Z}_m \times \mathbb{Z}_n \approx \mathbb{Z}_{mn}$. Then $\mathbb{Z}_m \times \mathbb{Z}_n$ is cyclic, and so there must exist an element with order mn. By Theorem 7 we there must then exist a generator $(a, b) \neq (0, 0)$in $\mathbb{Z}_m \times \mathbb{Z}_n$ such that $\text{lcm}(|a|, |b|) = mn$. Since each factor of the generator must generate its group, this implies $\text{lcm}(m, n) = mn$, and so $\gcd(m, n) = 1$, meaning that m and n are relatively prime. Now assume that m and n are relatively prime and that we have generators a of $\mathbb{Z}_m$ and b of $\mathbb{Z}_n$. Then since $\gcd(m, n) = 1$, we have $\text{lcm}(m, n) = mn$ and so $|(a, b)| = mn$. this implies that (a, b)generates $\mathbb{Z}_m \times \mathbb{Z}_n$, which must then be isomorphic to a cyclic group of order mn, im particular $\mathbb{Z}_{mn}$. ▯

Theorem 9 (Characterization of finite abelian groups): Let G be an abelian group. Then there exists prime numbers $p_1, \ldots, p_n$ and positive integers $r_1, \ldots, r_n$, unique up to order, such that

$$G \approx \mathbb{Z}_{p_1^{r_1}} \times \ldots \times \mathbb{Z}_{p_n^{r_n}}$$

Proof: A proof of this theorem is currenly beyond our reach. However, we will address it during the chapter on modules. ▯

Subdirect Products and Fibered Products

Definition 10: A *subdirect product* of two groups G and H is a proper subgroup K of $G \times H$ such that the projection homomorphisms are surjective. That is, $\pi_1(K) = G$ and $\pi_2(K) = H$.

Example 11: Let G be a group. Then the diagonal $\Delta = \{(g, g) \mid g \in G\} \subseteq G \times G$ is a subdirect product of G with itself.

Definition 12: Let G, H and Q be groups, and let the homomorphisms $\phi : G \longrightarrow Q$ and $\psi : H \longrightarrow Q$ be epimorphisms. The *fiber product* of G and H *over* Q, denoted $G \times_Q H$, is the subgroup of $G \times H$ given by $G \times_Q H = \{(g, h) \in G \times H \mid \phi(g) = \psi(h)\}$.

In this subsection, we will prove the equivalence between subdirect products and fiber products. Specifically, every subdirect product is a fiber product and vice versa. For this we need *Goursat's lemma.*

Theorem 13 (Goursat's lemma): Let G and G' be groups, and $H \subseteq G \times G'$ a subdirect product of G and G'. Now let $N = \ker\ \pi_2$and $N' = \ker\ \pi_1$. Then N can be identified with a normal subgroup of G, and N' with a normal subgroup of G', and the image of H when projecting on $G/N \times G'/N'$ is the graph of an isomorphism $G/N \approx G'/N'$.

Proof:

Further Reading

More on the automorphism groups of finite abelian groups. Some results require theory of group actions and ring theory, which is developed in a later section.

http://arxiv.org/pdf/math/0605185v1.pdf

Free Groups

In order to properly define the free group, and thereafter the free product, we need some preliminary definitions.

Definition 10: Let A be a set. Then a *word* of elements in A is a finite sequence $a_1a_2...a_n$ of elements of A , where the positive integer n is the *word length.*

Definition 11: Let $x = a_1...a_n$ and $y = a_{n+1}...a_{n+k}$ be two words of elements in A . Define the *concatenation* of the two words as the word $xy = a_1...a_na_{n+1}...a_{n+k}$.

Now, we want to make a group consisting of the words of a given set A , and we want this group to be the most general group of this kind. However, if we are to use the concatenation operation, which is the only obvious operation on two words, we are immediately faced with a problem. Namely, deciding when two words are equal. According to the above, the length of a product is the sum of the lengths of the factors. In other words, the length cannot decrease. Thus, a word of length n multiplied with its inverse has length at least n , while the identity word, which is the empty word, has length 0. The solution is an algorithm to *reduce* words into *irreducible* ones. These terms are defined below.

Definition 12: Let A be any set. Define the set $W(A)$ as the set of words of *powers* of elements of A . That is, if $a_1, ..., a_n \in A$ and $r_1, ..., r_n \in \mathbb{Z}$, then $a_1^{r_1}...a_n^{r_n} \in F(A)$.

Definition 13: Let $x = a_1^{r_1}...a_n^{r_n} \in W(A)$. Then we define a *reduction* of x as follows. Scan the word from the left until the first pair of indices $j, j+1$ such that $a_j = a_{j+1}$ is encountered, if such a pair exists. Then replace $a_j^{r_j}a_{j+1}^{r_{j+1}}$ with $a_j^{r_j+r_{j+1}}$. Thus, the resulting word is $x_{(1)} = a_1^{r_1}...a_{j-1}^{r_{j-1}}a_j^{r_j+r_{j+1}}a_{j+2}^{r_{j+2}}...a_n^{r_n}$. If no such pair exists, then $x = x_{(1)}$ and the word is called *irreducible*.

It should be obvious if $x \in W(A)$ with length n , then $x_{(n)}$ will be irreducible. The details of the proof is left to the reader.

Definition 14: Define the *free group* $F(A)$ on a set A as follows. For each word $x \in W(A)$ of length n , let the reduced word $x_{(n)} \in F(A)$. Thus $F(A) \subseteq W(A)$ is the subset of irreducible words. As for the binary operation on $F(A)$, if $x, y \in F(A)$ have lengths n and m respectively, define $x * y$ as the completely reduced concatenation $(xy)_{(n+m)}$.

Theorem 15: $F(A)$ *is* a group.

Proof:

Example 16: We will concider free groups on 1 and 2 letters. Let $A_1 = \{a\}$ and $A_2 = \{a, b\}$. Then

$F(A_1) = \{a^n \mid n \in \mathbb{Z}\}$ with $a^na^m = a^{n+m}$.

$F(A_2) = \{\prod_{i=1}^{n} a_ib_i \mid a_i \in F(\{a\}), b_i \in F(\{b\})\}$ such that $a_i \neq e$ for any $i > 1$ and $b_i \neq e$ for any $i < n$. Example product: $(a^2b^{-3}a)(a^{-1}ba) = a^2b^{-3}aa^{-1}ba = a^2b^{-3}ba = a^2b^{-2}a$.

Group Presentations

In this subsection we will breifly introduce another method used for defining groups. This is by prescribing a *group presentation*.

Definition 17: Let G be a group and H a subgroup. Then define the *normal closure* of H in G as the intersection of all normal subgroups in G containing H. That is, if N is the normal closure of H , then

$$N = \bigcap_{\substack{K \trianglelefteq G \\ H \subseteq K}} K \quad .$$

Definition 18: Let S be a set and $R \subseteq F(S)$. Let N be the normal closure of R in $F(S)$and define the group $\langle S \mid R \rangle = F(S)/N$. The elements of S are called *generators* and the elements of R are called *relators*. If G is a group such that $G \approx \langle S \mid R \rangle$, then $\langle S \mid R \rangle$ is said to be a *presentation* of G .

The Free Product

Using the previously defined notion of a group presentation, we can now define another type of group product.

Defintion : Let G and G' be groups with presentations $\langle S \mid R \rangle$ and $\langle S' \mid R' \rangle$. Define the *free product* of G and G' , denoted $G * G'$, as the group with the presentation $\langle S \cup S' \mid R \cup R' \rangle$.

Remark : Depending on the context, spesifically if we only deal with abelian groups, we may require the free product of abelian groups to be abelian. In that case, the free product equals the direct product. This is another example of abelian groups being better behaved than nonabelian groups.

Lemma : The free product includes the component groups as subgroups.

Remark : The free product is not a product in the sense discussed previously. It does not satifsy the universal property other products do. Instead, it satisfies the "opposide", or *dual* property, obtained by reversing the direction of all the arrows in the commutative diagram. We usually call a construction satisfying this universal property a *coproduct*.

Problems

Problem 1: Let H and K be groups of relatively prime orders. Show that any subgroup of $H \times K$ is the product of a subgroup of H with a subgroup of K .

Answer

Coming soon. ▯

Group actions on sets

In this section, we will encounter most important application of group theory. This is the notion of an *action* on some object. Over-generalizing slightly, we only care about groups *because* they act on things.

Group Actions

There are two equivalent ways to define a group action. We will present both, then prove their equivalence.

Definition 1a: Let G be a group and X a set. Then a *group action* (or just *action*) of G on X is a binary operation $* : G \times X \to X$ such that for any $g, h \in G$ and any $x \in X$,

i) $e * x = x$

ii) $g * (h * x) = (gh) * x$

Definition 1b: Let G be a group and X a set. Then a *group action* (or just *action*) of G on X is a homomorphism $\sigma : G \to S_X$.

Theorem 2: Definitions 1a and 1b are equivalent.

Proof: We must show that for each binary operationas in Definition 1a there is a unique homomorphism as in Definition 1b and vice versa. Given the binary operation, define the permutation $\sigma(g) \in S_X$ given by $\sigma(g)(x) = g * x$. Then $\sigma(g) \circ \sigma(h)(x) = \sigma(g)(\sigma(h)(x)) = g * (h * x) = (gh) * x = \sigma(gh)(x)$ for all $x \in X$ and $g, h \in G$. Thus σ is a homomorphism $G \to S_X$. Now, given a homomorphism σ, define a binary operation $* : G \times X \to X$ by $g * x = \sigma(g)(x)$. Then $e * x = \sigma(e)(x) = \mathrm{id}_X(x) = x$ and $g * (h * x) = \sigma(g) \circ \sigma(h)(x) = \sigma(gh)(x) = (gh) * x$. Thus $*$ is a binary operation satisfying the axioms in Definition 1a, and we are done. □

Example 3: Let G be a group. Then G acts on itself by left multiplication, that is, $g * h = gh$ for any $g, h \in G$.

Example 4: Let G be a group with H a any subgroup. Then G acts on the left cosets G/H by left multiplication, that is, $g * (g'H) = (gg')H$ for any $g, h \in G$.

Definition 5: A set which is acted on by a group G is called a G *-set*.

Definition 6: Let G be a group acting on a set X. Then the action of G is said to be

i) *faithfull* if e is the only element in G such that $g * x = x$ for *all* $x \in X$, and

ii) *free* if e is the only element in G such that $g * x = x$ for *any* element in X.

Remark 7: A group acting on itself by left multiplication is a free action. Thus is it also faithfull.

Definition 8: Let X be a G -set. The action is called *transitive* if for any $x, y \in X$, there exists a $g \in G$ sich that $y = g * x$.

Remark 9: We see that the actions in Example 3 and Example 4 are transitive.

Definition 10: Let X be a G -set and $x \in X$. Define the *orbit of* x to be the set $G * x = \{g * x \mid g \in G\}$.

Definition 11: Let X be a G -set and $x, y \in X$. Define the equivalence relation $x \sim y \Leftrightarrow (\exists g \in G) y = g * x \Leftrightarrow y \in G * x$. (Check that this *is* an equivalence relation!) Then $X/\sim$ is called the *orbit space* of X with respect to G, and we write $X/\sim \equiv X/G$.

Remark 12: Since the action is obviously transitive on each orbit, to "understand" group actions we only have to understand *transitive* actions.

Definition 13: Let X be a G -set. The *stabilizer* of $x \in X$ is the set $G_x = \{g \in G \mid g * x = x\}$.

Lemma 14: Let X be a G -set and let $x \in X$. Then G_x is a subgroup of G.

Proof: Since $e * x = x$, $e \in G_x$. Let $a, b \in G_x$. Then $(ab) * x = a * (b * x) = a * x = x$, so $ab \in G_x$. Finally, $x = e * x = (b^{-1}b) * x = b^{-1} * (b * x) = b^{-1} * x$, so $b^{-1} \in G_x$. Thus G_x is a subgroup of G. □

Lemma 15: Let X be a G-set, $a \in X$, $g \in G$ and $b = g * a \in X$. Then $G_b = gG_ag^{-1}$.

Proof: Let $h \in G_a$. Then $(ghg^{-1}) * b = (ghg^{-1}g) * a = (gh) * a = g * (h * a) = g * a = b$, showing $gG_ag^{-1} \subseteq G_b$. To show $G_b \subseteq gG_ag^{-1}$, note that $a = g^{-1} * b$ and follow the same arugment, obtaining $g^{-1}G_bg \subseteq G_a$. □

A G-set is an algebraic structure, and as in any situation where we have several instances of a structure, we cannot resist the temptation to introduce maps between them.

Definition 16: Let X and X' be G-sets with binary operations $*$ and $*'$ respectively. Then a function $f : X \longrightarrow X'$ is called an *equivariant* function if $f(g * x) = g *' f(x)$ for all $x \in X$. f is called an isomorphism if f is bijective.

Lemma 17: Compositions of equivariant functions are equivariant.

Proof: Trivial. □

Theorem 18: Every transitive action of G on a set X is isomorphic to left multiplication on G/H, where $H = G_x$ for any $x \in X$.

Proof: By Lemma 15, the stabilizers of any two elements in an orbit are conjugate and so isomorphic. Thus the chosen element *is* arbitrary. Let $f : G/H \longrightarrow X$ by $f(gH) = g * x$ for all $g \in G$. Then if $g_1H = g_2H$, we have $g_1g_2^{-1} \in H$, so $(g_1g_2^{-1}) * x = x$ and $g_1 * x = g_2 * x$ and so f is well-defined. Now let $h : X \longrightarrow G/H$ be defined by $h(g * x) = gH$. If $g_1 * x = g_2 * x$, then $(g_1g_2^{-1}) * x = x$, so $g_1g_2^{-1} \in H$ and $g_1H = g_2H$ showing that h is well-defined. Since f and h are obviously inverses of each other, this shows that f is a bijection. To prove equivariance, observe that $f(g_1gH) = (g_1g) * x = g_1 * (g * x) = g_1 * f(gH)$. □

Corollary 19 (Orbit stabilizer theorem): Let X be a G-set and $G * x$ the orbit of $x \in X$. Then $|G * x||G_x| = |G|$.

Proof: Indeed, from Theorem 18 and Lagrange's theorem we obtain $|G * x||G_x| = |G|$. □

Definition 20: The orbit of $x \in X$ where X is a G-set is said to be *trivial* if $G * x = \{x\}$. Let Z denote the set of elements of X whose orbits are trivial. Equivalently, $Z = \{x \in X \mid (\forall g \in G) g * x = x\}$.

Lemma 21: Let S be a finite set and G be a group acting on S. Next, let A be a set containing exactly one element from each nontrivial orbit of the action. Then, $|S| = |Z| + \sum_{a \in A} |G/G_a|$.

Proof: Since the orbits partition S, the cardinality of S is the sum of cardinalities of the orbits. All the trivial orbits are contained in Z, so $|S| = |Z| + \sum_{a \in A} |G * a|$. By Corollary 19, $|G * a| = |G/G_a|$, proving the lemma. □

Definition 22: A p-*group* is a group whose order is a finite power of a prime integer p.

Corollary 23: Let G be a p-group acting on a set S. Then $|S| \equiv |Z| \bmod p$.

Proof: Since G is a p-group, $|G * a|$ divides p for each $a \in A$ with A defined as in Lemma 21. Thus $\sum_{a \in A} |G * a| \equiv 0 \bmod p$. □

Group Representations

Linear group actions on vector spaces are especially interesting. These have a special name and comprise a subfield of group theory on their own, called *group representation theory*. We will only touch slightly upon it here.

Definition 24: Let G be a group and Vbe a vector space over a field F. Then a *representation* of G on $V$is a map $\Phi : G \times V \to V$such that

i) $\Phi(g) : V \to V$given by $\Psi(g)(v) = \Psi(g, v)$, $v \in V$, is linear in $v$over F.

ii) $\Phi(e, v) = v$

iii) $\Phi\left(g_1, \Phi(g_2, v)\right) = \Phi(g_1 g_2, v)$for all $g_1, g_2 \in G$, $v \in V$.

V is called the *representation space* and the dimension of V, if it is finite, is called the *dimension* or *degree* of the representation.

Remark 25: Equivalently, a representation of Gon $V$is a homomorphism $\phi : G \to GL(V, F)$. A representation can be given by listing $V$and ϕ , (V, ϕ).

As a representation is a special kind of group action, all the concepts we have introducd for actions apply for representations.

Definition 26: A representation of a group Gon a vector space $V$is called *faithful* or *effective* if $\phi : G \to GL(V, F)$is injective.

Rings

Rings

This section builds upon and expands the theory covered in the previous chapter on groups. The reader is strongly advised to master the material presented in the sections up to and including Products and Free Groups before continuing.

Motivation

The standard motivation for the study of rings is as a generalization of the set of integers $\mathbb{Z}$ with addition and multiplication, in order to study integer-like structures in a more general and less restrictive setting. However, we will also present the following motivation for the study of rings, based on the theory of Abelian groups.

Let G and H be Abelian groups. Then the set $\mathrm{Hom}_{\mathrm{Ab}}(G,H)$(Please don't pay much attention to the subscript for now.) of group homomorphisms $\phi : G \longrightarrow H$ naturally forms an abelian group in the following way. If $\phi, \psi \in \mathrm{Hom}_{\mathrm{Ab}}(G,H)$, define $(\phi+\psi)(g) = \phi(g)+\psi(g)$ for all $g \in G$. It should be obvious where each addition is taking place. In particular, we can consider the set $\mathrm{End}_{\mathrm{Ab}}(G) = \mathrm{Hom}_{\mathrm{Ab}}(G,G)$ of endomorphisms of G. That is, the set of homomorphisms from G to itself. This set is obviously a group from the above discussion, but it is also closed under composition. By endowing the set $\mathrm{End}_{\mathrm{Ab}}(G)$ with the operations of addition, $+$, and composition, $\circ$, we note that it has the following properties:

i) It is an Abelian group under addition.

ii) It is a monoid under multiplication.

iii) Addition distributes over composition.

Indeed, for the third property, note that if $\phi, \psi, \xi \in \mathrm{End}_{\mathrm{Ab}}(G)$ and $g \in G$, then $\phi \circ (\psi + \xi)(g) = (\phi\circ\psi + \phi\circ\xi)(g)$ and $(\phi+\psi)\circ\xi(g) = (\phi\circ\xi + \psi\circ\xi)(g)$. The following material is a generalization of this situation.

Introduction to Rings

Definition 1: A *ring* $(R,+,\cdot)$ is a set R with two binary operations $+$ and $\cdot$ that satisfies the following properties:
For all $a,b,c \in R$,

i) $(R,+)$ is an abelian group.

ii) $(R,\cdot)$ is a monoid.

The definition of ring homomorphism does not include the existence of 1.

iii) $\cdot$ is distributive over $+$:

1) $a\cdot(b+c) = (a\cdot b)+(a\cdot c)$

2) $(a+b)\cdot c = (a\cdot c)+(b\cdot c)$

We will denote the additive identity in a ring by 0_R or 0 if the ring is understood. Similarily, we denote the multiplicative identity by 1_R or 1 when the ring is understood. We'll often use juxtaposition in place of $\cdot$, i.e., ab for $a\cdot b$.

Remark 2: Some authors do not require their rings to have a multiplicative identity element. We will call a ring without an idenitity a *rng*. *Pseudo-rings* is another term used for rings without unity. Authors who do not require a multiplicative identity usually call a ring a *ring with unity*. Unless otherwise stated, we will assume that $0 \neq 1$in our rings. A major part of noncommutaive ring theory was developed without assuming every ring has an identity element.

Example 3: The reader is already familiar with several examples of rings. For instance $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ with the usual addition and multiplication operations. We have a familiy of finite rings given by the sets $\mathbb{Z}_n$for integer $n \geq 2$with addition and multiplication defined modulo n . Finally we have an example of a *rng* given by the sets $n\mathbb{Z}$ for integer $n \geq 2$with the usual addition and multiplication. The reader is invited to confirm the ring axioms for these examples.

Let us now prove some very basic properties about rings. This is analogous to what we did for groups when we first introduced them.

Theorem 4: Let R be a ring, and let $a, b, c \in R$. Then the following are true:

1. If $a + b = a + c$, then $b = c$.
2. The equation $a + x = b$ has a unique solution.
3. $-(-a) = a$
4. $0a = 0$
5. $(-a)b = -(ab)$
6. $(-a)(-b) = ab$

Proof: (1), (2), and (3) all strictly concern addition, and are all previous results from $(R, +)$being a group. The other three parts all concern both addition and multiplication (since 0 and - are additive concepts), so as a proof strategy we expect to use the distributive law in some way to link the two operations. For (4), observe that $0a + 0a = (0 + 0)a = 0a = 0a + 0$. But then by (1), 0a=0. For (5), Note that $(-a)b + ab = (-a + a)b = 0b = 0$. For (6) note that $(-a)(-b) + -(ab) = (-a)(-b) + (-a)b = -a(-b + b) = -a0 = 0$. □

Remark 5: Take another look at the examples in Example 3. Notice that for all those rings, multiplication is a commutative opration. However, the axioms say nothing about this. Thus we should expect to find counter-examples to this.

Definition 6: A ring is called *commutative* if multiplication is commutative.

Example 7: An example of a non-commutative ring is the set $M_n(\mathbb{R})$of $n \times n$ square matrices with real coefficients under standard addition and multiplication of matrices, where $n \geq 2$is an integer. The reader can easily check this for $n = 2$and conclude that it holds for all other n (why?).

Theorem 8: A ring has a unique multiplicative identity.

Proof: During our brief discussion of monoids earlier, we showed that in any monoid the identity is unique. Since a ring sans addition is a monoid, this applies here. □

Example 9: The singleton set $\{*\}$ with addition and multiplication defined by $* + * = *$ and $* \cdot * = *$is a ring, called the *trivial ring* or the *zero ring*. Note that in the trivial ring, $0 = 1$. The reader is invited to show that $0 = 1$in a ring if and only if it is the trivial ring.

If the reader has tried to construct some of the rings $= \mathbb{Z}_n$, he/she may have realised that certain non-zero elements have product zero. We formalize this concept as follows.

Definition 10: Let R be a ring and $a \in R \setminus \{0\}$. $a$is called a *left(resp.right)-zero-divisor* if there exists a $b \in R$ such that $ab = 0$ $(ba = 0)$.

Lemma 11: Let R be a ring with $a \in R$. Define the function $\rho_a\ :\ R \longrightarrow R$ given by $\rho_a(r) = ar$ for all $r \in R$. Then $\rho_a$is injective if and only if $a$is not a left-zero-divisor.

Proof: Assume a is not a left-zero-divisor, and assume we have $ab = \rho_a(b) = \rho_a(c) = ac$ for some $b, c \in R$. This implies $ab - ac = a(b - c) = 0$, giving $b = c$ since a is not a left-zero-divisor, so ρ_a is injective. Conversely, assume a is a left-zero-divisor. Then there exists a $b \in R$ such that $b \neq 0$and $\rho_a(b) = ab = 0 = a0 = \rho_a(0)$, so ρ_a is not injective. ▯

Remark 12: Thus, multiplication by a is left-cancellative if and only if a is not a zero-divisor. The reader is invited to state and prove the equivalent lemma for right-zero-divisors.

Example 13: $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ are all examples of commutative rings without zero divisors. These rings motivate the next definition.

Definition 14: Let R be a commutative ring without zero divisors. Then R is called an *integral domain.*

Just like Definition 14, the majority of special types of rings will be motivated by properties of $\mathbb{Z}$.

Example 15:

1. The set $\mathrm{Hom}_{\mathrm{Set}}(\mathbb{R}, \mathbb{R})$of functions on $\mathbb{R}$ with pointwise addition and multiplication is a ring.
2. More generally, if R is a ring, the set $\mathrm{Hom}_{\mathrm{Set}}(R, R)$ of functions from R to itself is also a ring.
3. The set $\mathrm{Hom}_{\mathrm{Set}}(R, R)$ with function composition for multiplication is **not** a ring since the statement $f \circ (g + h) = f \circ g + f \circ h$ is not true in general.
4. The set of integrable functions on the real numbers, L^1, is a rng under pointwise addition and multiplication given by *convolution*: $(f * g)(t) = \int_{\mathbb{R}} f(\tau)g(t - \tau)d\tau$. This rng is important to the study of linear systems and differential equations. If the reader has enough calculus under his/her belt, he/she reader is invited to show that it does not have an identity, and that it is commutative.
5. The set of Gaussian integers $\mathbb{Z}[i] = \{a + bi | a, b \in \mathbb{Z}\}$ with standard addition and multiplication is a ring.

Definition 16: Let R be a ring. An element $a \in R$ is a *unit* and is *invertible* if there is an element $b \in R$ such that $ab = ba = 1$. The set of all units is denoted by R^*.

Exercise 17: Prove that R^* is a group under multiplication.

Exercise 18:: Show that a zero-divisor is not a unit.

Theorem 19: (Cancellation Law for Integral Domains): Let R be an integral domain, and let $a, b, c \in R$ be nonzero. Then $ab = ac$ if and only if $b = c$.

Proof: Evidently $ab = ac$ if $b = c$. To see the other direction, we rearrange the equality as $ab - ac = 0$. But then $a(b - c) = 0$. Since a is nonzero, and R contains no zero divisors, it must be the case that $b - c = 0$, which is to say that $b = c$.

Definition 20: A ring R is a *division ring* or *skew field* if all non-zero elements are units, i.e. if it forms a group under multiplication with its nonzero elements.

Definition 21: A field is a commutative division ring. Alternatively, a field F is a ring where $(F - 0, \cdot)$ is an abelian group under multiplication. As another alternative, a field is an integral domain where all non-zero elements are invertible.

As stated before, integral domains are easy to work with because they are so close to being fields. In fact, the next theorem shows just how close the two are:

Theorem 22: Let R be a finite integral domain. Then R is a field.

Proof: Let $a \in R$ be nonzero and let $S = \{ab | b \in R\}$. Clearly S is a subset of R . From the cancellation law, we can see that $|S| = |R|$ (since if two elements ab and ac are equal, then $b = c$). But then $S = R$. So then there must be some b such that $ab = 1$. So a is a unit.

Of course proving that a set with two operations satisfy all of the ring axioms can be tedious. So, just as we did for groups, we note that if we're considering a subset of something that's already a ring, then our job is easier.

Definition 23: A *subring* S of a ring R is a subset of R that is also a ring (under the same two operations as for R) and $1_S = 1_R$. We denote " S is a subring of R " by $S \leq R$. Note many mathematicians do not require rings or subrings to have an identity.

Theorem 24: Let $S \neq \emptyset$ be a subset of a ring R . Then $S \leq R$ if and only if for all $a, b \in S$,

1. $a - b \in S$,
2. $ab \in S$,
3. $1 \in S$.

Example 25:

1. $\mathbb{Z} \leq \mathbb{Q} \leq \mathbb{R} \leq \mathbb{C}$.
2. The trivial ring $\{0\}$ is a subring of every ring.
3. The set of Gaussian integers $\mathbb{Z}\left[i\right]$is a subring of the complex numbers $\mathbb{C}$.

Ring Homomorphisms

Just as with groups, we can study *homomorphisms* to understand the similarities between different rings.

Homomorphisms

Definition

Let *R* and *S* be two rings. Then a function $f : R \longrightarrow S$ is called a **ring homomorphism** or simply **homomorphism** if for every $r_1, r_2 \in R$, the following properties hold:

$$f(r_1 r_2) = f(r_1) f(r_2),$$
$$f(r_1 + r_2) = f(r_1) + f(r_2).$$

In other words, *f* is a ring homomorphism if it preserves additive and multiplicative structure.

Furthermore, if *R* and *S* are rings with unity and $f(1_R) = 1_S$, then *f* is called a **unital ring homomorphism**.

Examples

1. Let $f : \mathbb{Z} \longrightarrow M_2(\mathbb{Z})$be the function mapping $a \mapsto \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$. Then one can easily check that f is a homomorphism, but not a unital ring homomorphism.
2. If we define $g : a \mapsto \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$, then we can see that g is a unital homomorphism.
3. The *zero homomorphism* is the homomorphism which maps ever element to the zero element of its codomain.

Theorem: Let R and S be integral domains, and let $f : R \longrightarrow S$ be a nonzero homomorphism. Then f is unital.

Proof: $1_S f(1_R) = f(1_R) = f(1_R^2) = f(1_R) f(1_R)$. But then by cancellation, $f(1_R) = 1_S$.
In fact, we could have weakened our requirement for R a small amount (How?).

Theorem: Let R, S be rings and $\varphi : R \longrightarrow S$ a homomorphism. Let R' be a subring of R and S' a subring of S. Then $\varphi(R')$is a subring of S and $\varphi^{-1}(S')$ is a subring of R . That is, the kernel and image of a homomorphism are subrings.
Proof: Proof omitted.

Theorem: Let R, S be rings and $\varphi : R \longrightarrow S$ be a homomorphism. Then φ is injective if and only if $\ker \varphi = 0$.

Proof: Consider φ as a group homomorphism of the additive group of R .

Theorem: Let F, E be fields, and $\varphi : F \longrightarrow E$ be a nonzero homomorphism. Then φ is injective, and $\varphi(x)^{-1} = \varphi(x^{-1})$.

Proof: We know $\varphi(1) = 1$ since fields are integral domains. Let $x \in F$ be nonzero. Then $\varphi(x^{-1})\varphi(x) = \varphi(x^{-1}x) = \varphi(1) = 1$. So $\varphi(x)^{-1} = \varphi(x^{-1})$. So $\varphi(x) \neq 0$ (recall you were asked to prove units are nonzero as an exercise). So $\ker \varphi = 0$.

Isomorphisms

Definition

Let R, S be rings. An **isomorphism** between R and S is an invertible homomorphism. If an isomorphism exists, R and S are said to be *isomorphic*, denoted $R \cong S$. Just as with groups, an isomorphism tells us that two objects are *algebraically the same*.

Examples

1. The function g defined above is an isomorphism between $\mathbb{Z}$ and the set of integer scalar matrices of size 2, $S = \{\lambda I_2 | \lambda \in \mathbb{Z}\}$.
2. Similarly, the function $\varphi : \mathbb{C} \longrightarrow M_2(\mathbb{R})$ mapping $z \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ where $z = a + bi$ is an isomorphism.

   This is called the *matrix representation* of a complex number.
3. The *Fourier transform* $\mathcal{F} : L^1 \longrightarrow L^1$ defined by $\mathcal{F}(f) = \int_{\mathbb{R}} f(t)e^{-i\omega t}dt$ is an isomorphism mapping

   integrable functions with pointwise multiplication to integrable functions with convolution multiplication.

Excercise: An isomorphism from a ring to itself is called an **automorphism**. Prove that the following functions are automorphisms:

1. $f : \mathbb{C} \rightarrow \mathbb{C}, f(a + bi) = a - bi$
2. Define the set $\mathbb{Q}(\sqrt{2}) = \left\{a + b\sqrt{2} | a, b \in \mathbb{Q}\right\}$, and let $g : \mathbb{Q}(\sqrt{2}) \rightarrow \mathbb{Q}(\sqrt{2}), g(a + b\sqrt{2}) = a - b\sqrt{2}$

Ideals

Motivation

In ../Rings/ we saw that the set of even integers $2\mathbb{Z} \subset \mathbb{Z}$ was a subring of the integers.

We can also see very easily that the integers $\mathbb{Z} \subset \mathbb{Q}$ are a subring of the rational numbers under the usual operations of addition and multiplication.

The even integers, when taken as a subring of the integers have a property that the integers when taken as a subring of the rationals do not. The even integers taken as a subring of the rationals also lack this property.

The property is that the even integers, taken as a subring of the integers, *absorb* multiplication. Let's call the even integers $I = 2\mathbb{Z}$ for ease of notation.

Consider the following: For all $i \in I$, we can see by the definition of I that $i = 2k$ for some $k \in \mathbb{Z}$.

For all $a \in \mathbb{Z}$ see that $ai = a(2k) = 2ak \in I$.

In English, regardless of which even integer is chosen, multiplying it by any integer will give us a different *even* integer.

Definition of an Ideal

Definition: Given a ring R , a subset $I \subseteq R$ is said to be a *left ideal of* R if it absorbs multiplication from the left; that is, if $\forall i \in I, \forall r \in R, ri \in I$.

Definition: Given a ring R , a subset $I \subseteq R$ is said to be a *right ideal of* R if it absorbs multiplication from the right; that is, if $\forall i \in I, \forall r \in R, ir \in I$.

Definition: We define an *ideal* I to be something that is both a left ideal and a right ideal. We also require that $(I, +)$ is a subgroup of $(R, +)$.

We write $I \triangleleft R$ as shorthand for this.

To verify that a subset of a ring is an ideal, it is only necessary to check that it is closed under subtraction and that it absorbs multiplication; this is because of the subgroup criterion from Abstract_Algebra/Group_Theory/Subgroup.

Definition: An ideal $I \triangleleft R$ is *proper* if $I \subsetneq R$.

Definition: An ideal $I \triangleleft R$ is *trivial* if I={0}.

Lemma: An ideal I is proper if and only if $1 \notin I$.

Proof: If $1 \in I$ then $r = r1 \in I$ so $I = R$.

The converse is obvious.

Theorem: In a division ring, the only proper ideal is trivial.

Proof: Suppose we have an ideal in a division with a nonzero element a. Take any element b in our division ring. Then a^{-1}b is in the division ring as well, and aa^{-1}b = b is in the ideal. Therefore, it is not a proper ideal.

Definition: Let S be a nonempty subset of a ring R. Then the ideal generated S is defined to be the smallest ideal in R containing S, which would be the intersection of all such ideals. We can characterize this ideal by the collection of all finite sums

$$\left\{\sum_{i=1}^{n} r_{i1} s_i r_{i2} | r_{i1}, r_{i2} \in R, s_i \in S\right\}$$

And one can easily verify that this is an ideal, and that all ideals containing S must contain this ideal. If it is commutative, then one can simply characterize it as

$\left\{\sum_{i=1}^{n} r_i 1 s_i | r_i \in R, s_i \in S\right\}$

The ideal generated by a single element a is called a **principal ideal**. If the ring is commutative, it consists of all elements of the ring of the form ra where r is any element in the ring.

Example: Let $R = \mathbb{Z}$ be the ring of integers. The principal ideal (n) is the subset of $\mathbb{Z}$ consisting of positive and negative multiples of n. For example (2) is the subset of even integers. Then one can view the factor ring $\mathbb{Z}/(n)$ simply as the set $\{0, 1, \ldots, n-1\}$ under addition and multiplication modulo n.

Operations on Ideals

Given a collection of ideals we can generate other ideals. For instance it is easy to check that the intersection of any family of ideals is again an ideal. We write this simply as $\cap_{j \in J} I_j$.

Given any set $S \subset R$ we can construct the smallest ideal of R containing S which we denote by $\langle S \rangle$. It is determined by $\langle S \rangle = \cap_{S \subset I \triangleleft R} I$, though often we can be more explicit than this.

If $I_{j \in J}$ is a collection of ideals we can determine the *sum*, written $\sum_{j \in J} I_j$, as the smallest ideal containing all the ideals I_j. One can check explicitly that its elements are *finite* sums of the form $\sum_{j \in J} x_j$.

Finally if I, J are two ideals in R one can determine the ideal-theoretic product as the smallest ideal containing the set-theoretic product $\{ij \mid i \in I, j \in J\}$. Note that the ideal-theoretic product is in general strictly larger than the set-theoretic product, and that it simply consists of finite sums of the form $\sum i_r j_r$ where $i_r \in I$ $j_r \in J$

Example: Let (m) and (n) the principal ideals in $\mathbb{Z}$ just given. Then one can check explicitly that $(m) \cap (n) = (r)$, where r is the lcm of m and n. Moreover $(m)(n) = (mn)$, and $(m) + (n) = (s)$ where s is the hcf of m and n. Observe that $(m)(n) = (m) \cap (n)$ if and only if s = mn if and only if m and n are co-prime if and only if $(m) + (n) = (1)$.

Homomorphisms and Ideals

Rings, like groups, have factor objects that are kernels of homomorphisms. Let $f : R \longrightarrow S$ be a ring homomorphism. Let us determine the structure of the kernel of *f* which is defined to be all elements which map to the identity.

If *a* and *b* are in the kernel of *f*, i.e. $f(a) = f(b) = 0$, and *r* is any element of *R*, then

$$f(a-b) = f(a) - f(b) = 0,$$
$$f(ar) = f(a)f(r) = 0,$$
$$f(ra) = f(r)f(a) = 0.$$

Therefore $\ker(f)$ is an ideal of *R*.

Also note that the homomorphism will be a monomorphism i. e. it is injective or one-to-one when the kernel consists only of the identity element.

We also have the following

Theorem: If the only proper ideal of R is the trivial ideal {0}, then if f is a homomorphism from R to S, and it does not map all elements of R to the identity in S, then it is injective.

Proof: The kernel of the homomorphism must be an ideal, and since the only ideals are R and the trivial ideal, one of these two must be the kernel. However, since not all elements of R map to the identity of S, R is not the kernel, so the trivial ideal must be.

Since this condition is satisfied for all division rings, it is true for all division rings.

The construction of factor rings in the next section will prove that there exists a homomorphism with I as its kernel for any ideal I.

Factor Rings

Definition: Given a ring R and an ideal $I \subseteq R$, the ring of cosets of R , r+I where r is within R, is written R/I , where each coset is defined to be the set {r+ili is an element of I}, and by Lagrange's theorem, it partitions R. This set of cosets, called the *factor ring* (or *quotient ring*) of R modulo I is a ring with addition defined the same way as one would define it for groups (since the ring is a group under addition), and with multiplication defined as follows:

- $(a+I)(b+I) = ab+I$.

To show that this is independent of the choice of a and b (or, the operations are well-defined), suppose that a' and b' are elements of the same respective coset. Then a'=a+j and b'=b+k for some element j,k within I. Then a'b'=ab+ak+jb+jk and since ak, jb, and jk are elements of I, a'b' and ab must belong to the same coset, so ab+I=a'b'+I. Obviously the cosets form a group under addition because of what was proved earlier about factor groups, and furthermore the factor ring forms an abelian group under addition because the ring forms an abelian group under addition. Since the product rs+I is analogous to the multiplication in the ring, it obviously has all the properties of a ring. Furthermore, if the ring is commutative, then the factor ring will also be commutative.

Observe that there is a canonical ring homomorphism $\pi : R \to R/I$ determined by $\pi(r) = r+I$, called the projection map. We record some properties of this homomorphism in the next section of the isomorphism theorems.

Ring Isomorphism Theorems

We have already proved the isomorphism theorems for groups. Now we can use analogous arguments to prove the isomorphism theorems for rings, substitution the notion of "normal subgroups" with ideals.

Factor Theorem

Let I be an ideal of a ring R, Let $\pi(x) = x+I$ be the usual homomorphism from R to R/I. Now let f be a homomorphism from R to S. Observe that if $\tilde{\phi} : R/I \to S$ is a ring homomorphism, then the composition $\phi = \tilde{\phi} \circ \pi : R \to R/I \to S$ is a ring homomorphism such that $I \subset \ker \phi$ (because $x \in I \implies \phi(x) = \tilde{\phi} \circ \pi(x) = \tilde{\phi}(0_{R/I}) = 0_S$). This characterizes all such morphisms in the following sense

Factor Theorem: Let $\phi : R \to S$ be a ring homomorphism such that $I \subset \ker \phi$. Then there is a unique homomorphism $\tilde{\phi} : R/I \to S$ such that $\phi = \tilde{\phi} \circ \pi$. Furthermore, $\tilde{\phi}$ is an epimorphism if and only if ϕ is an epimorphism, $\tilde{\phi}$ is a monomorphism if and only if its kernel is I.

Proof We prove it the same way we did for groups. Define $\tilde{\phi}(x+I)$ to be $\phi(x)$. To see that this is well-defined, let a+I=b+I, and so that a-b is an element of I, so that $\phi(a-b) = 0_S$ so that $\phi(a) = \phi(b)$. Now ϕ is a homomorphism, implying that $\tilde{\phi}$ is also. The proofs of the additional statements can be carried over from the proofs of the additional statements of the factor theorem from groups.

First Isomorphism Theorem

Let R be a ring, and let f be a homomorphism from R to S with kernel K. Then the image of f is isomorphic to R/K.

Proof

Using the factor theorem, we can find a homomorphism from R/K to S, and since the kernel is the same as the ideal used in forming the quotient group, and since the f is an epimorphism over its image, this homomorphism is an isomorphism.

Second Isomorphism Theorem

Let R be a ring, let I be an ideal, and let S be a subring.

1. S+I, the set of all s+i with s within S and i within I, is a subring of R.
2. I is an ideal of S+I.
3. The intersection of S and I is an ideal of S.
4. (S+I)/I is isomorphic to $S/(S \cap I)$.

Proof

1. It can be verified that it contains 1, and is closed under multiplication.
2. Of course, since I is an ideal of R, then it must be an ideal under any subring.
3. From a similar argument for groups, it can only contain elements of I, but restricted to S, so it must be an ideal of S.
4. Let $f : R \to R/I$ be a function restricted to the domain S, and define $f(x) = x + I$. It is obvious that its kernel is $S \cap I$ and that its image is (S+I)/I.

Third Isomorphism Theorem

Let I be an ideal of a ring R, and let J be an ideal of the same ring R that contains I. J/I is an ideal of R/I, and R/J is isomorphic to (R/I)/(J/I).

Proof Define the function $f : R/I \to R/J$ to be $f(a + I) = a + J$ which is well-defined because since I is an ideal that is within J. This is also obviously a homomorphism. Its kernel is all elements that map onto J, and is thus all a+I such that a is within J, or J/I. Moreover, its image is R/J, and so we can use the first isomorphism theorem to prove the result.

Correspondence Theorem

Let I be an ideal of a ring R. Define the function π to map the set of rings and ideals containing I to the set of rings and ideals of R/I, where $\pi(X)$ = the set of all cosets x+I where x is an element of X. This function is one-to-one, and the image of rings or ideals containing I are rings or ideals within R/I.

Proof Define the function f from rings or ideals containing I to the rings or ideals of R/I, by f(A)=A/I. We have already proved the correspondence for addition because rings form an abelian group under addition. Thus, we need only to check for multiplication. Suppose S is a subring of R containing I. S/I is obviously closed under addition and subtraction. For multiplication, suppose that x and y are elements of S. Then (x+I)(y+I)=xy+I which is also an element of S/I, proving that it is closed under multiplication. The identity 1 is within S, and we have it that 1+I is also within S/I. Thus, S/I is a subring of R/I. Now suppose that S/I is a subring of R/I. Then it is also obvious that S is closed under addition and subtraction and multiplication, proving that S is a subring of R. Now suppose that J is an ideal of R containing I. Then by the third isomorphism theorem, J/I is an ideal of R/I. Now suppose that J/I is an ideal of R/I. Let r be any element of R, and let j be any element of J. Then since J/I is an ideal of R/I, (r+I)(j+I)=rj+I must be an element of J/I. This indicates that rj must be an element of J, proving that J is an ideal of R.

Chinese Remainder Theorem

Definitions

Definition: Two elements a, b are said to be *congruent* in an ideal I if and only if they belong to the same coset in R/I. This is true when a-b is within I. Write $a = b\mathrm{mod}I$ to mean that a is congruent to b modulo I .

Lemma: Given an ideal $I \subseteq R$, a subset of a ring R , the congruence class $[r]$ modulo I of an element $r \in R$ is $[0]$ if and only if $r \in I$. To see this, simply note that $a = b\mathrm{mod}I$ means $a - b \in I$; plugging in gives $r - 0 \in I$.

Definition: Two natural numbers are relatively prime when ax+by=1 for integers x and y. We do the same for rings - two ideals I and J are relatively prime when ai+bj=1 for ring elements a, b, and for an element i within I and an element j within J. In other words, two ideals are relatively prime if their sum contains the identity element i. e. if I+J is the whole ring R.

We will now prove the

Chinese Remainder Theorem

Let R be a ring, and let I_n be n pairwise (i. e. when considering any two pairs) relatively prime ideals.

1. Let a be a number from 1 to n. There exists an element r within R that is within all ideals I_n such that $n \neq a$, and such that $r = 1\mathrm{mod}I_a$
2. Let $r_1, r_2, ..., r_n$ be elements of R. Then there exists an element r within R such that $r = r_i\mathrm{mod}I_i$ for all i=1,2,3,...,n.
3. Let I be the intersection of the ideals. Another element of R, s satisfies $s = r_i\mathrm{mod}I_i$ for all i=1,2,3,...,n if and only if $r = s\mathrm{mod}I$.
4. R/I is isomorphic to the product ring $\Pi_{i=1}^{n} R/I_i$

Proof

1. Since I_1 and I_i (i>1) are relatively prime, there will exist an elements $b_i \in I_1$ and $c_i \in I_i$ (i>1) such that $b_i + c_i = 1$. This implies that $\Pi_{i=2}^{n}(b_i + c_i) = 1$. Now we expand this product on the left side. All terms of the product other than $c_2c_3c_4...c_n$ belong to I_1 while $c_2c_3c_4...c_n$ itself belongs the set S of all finite sums of products $x_2x_3x_4...x_n$ with $x_i \in I_i$. Thus, it can be written in the form b+a=1, where b is an element of I_1 , and where a is an element of S. Then $a \equiv 1 \pmod{I_1}$ and $a \equiv 0 \pmod{I_i}$ for i>1.

Prime and Maximal Ideals

There are two important classes of ideals in a ring - *Prime* and *Maximal.*

Definition: An ideal $I \triangleleft R$ is *prime* if it satisfies:

for any ideals A and B such that AB is a subset of I , implies A is in I or B is in I.

Definition: An ideal $I \triangleleft R$ is *maximal* if it is proper (i.e. $I \subsetneq R$ and it satisfies:

$I \subset J \triangleleft R \implies I = J$ or $J = R$

That is, there are no proper ideals between I and R .

The following Lemma is important for many results, and it makes essential use of Zorn's Lemma (or equivalently the Axiom of Choice)

Lemma: Every non-invertible element of a ring is contained in some maximal ideal

Proof: Suppose x is the non-invertible element. Then the first observation is that (x) is a proper ideal, for if $(x) = R$, then in particular $1 \in (x)$ so $1 = ax$ contradiction the assumption. Let S be the set of proper ideals in R containing x ordered by inclusion. The first observation implies that S is non-empty, so to apply Zorn's

Lemma we need only show that every increasing set of ideals contains an upper-bound. Suppose $\{I_j\}_j$ is such an increasing set, then the least upper bound is $\sum_{j\in j} I_j$ as this is the smallest ideal containing each ideal. If one checks that the union $\cup_{j\in J} I_j$ is an ideal, then this must be $\sum_{j\in J} I_j$. To show it's proper, we need only show $1 \notin \cup_{j\in J} I_j$ for all j. But this f because each I_j is proper.

Therefore by Zorn's Lemma there is a maximal element J of S. It is clearly maximal for if J' were any ideal satisfying $J \subset J' \subsetneq R$ then J' would be an element of S, and by maximality of J we would have $J' \subset J$ whence $J = J'$.

Properties of rings may be naturally restated in terms of the ideal structure. For instance

Proposition: A commutative ring R is an Integral Domain if and only if (0) is a prime ideal.

Proof: This follows simply because $x = 0 \iff x \in (0)$.

This explains why an Integral Domain is also referred to as a *Prime Ring*. Similarly, we may give a necessary and sufficient condition for a ring to be a field :

Proposition: A commutative ring R is a *Field* if and only if (0) is a maximal ideal (that is there are no proper ideals)

Proof: We only need to show that every element $0 \neq x \in R$ is invertible. Suppose not then by Lemma ... x is contained in some (proper) maximal ideal, a contradiction.

Corollary: An ideal $I \lhd R$ is *maximal* if and only if R/I is a field.

Proof: By the previous Proposition we know R/I is a field if and only if its only proper ideal is (0). By the correspondence theorem (...) this happens if and only if there are no proper ideals containing I.

Corollary: The kernel of a homomorphism f from R to S is a maximal ideal when S is a field. The proof of this follows from the first isomorphism theorem because S is isomorphic to R/ker f.

It's also clear that

Lemma: An ideal $I \lhd R$ is *prime* if and only if R/I is an integral domain.

Proof: Write $\bar{x}$ for the element of R/I corresponding to the equivalence class $[x]$. Clearly every element of R/I can be written in this form.

$\Rightarrow)$ $\bar{x}\bar{y} = 0 \iff xy \in I \iff x \in I$ or $y \in I \iff \bar{x} = 0$ or $\bar{y} = 0$where the second equivalence follows directly because I is prime.

$\Leftarrow)$This follows in exactly the same way.

Corollary: The kernel of a homomorphism f from R to S is a prime ideal when S is an integral domain. The proof of this follows from the first isomorphism theorem because S is isomorphic to R/ker f.

Lemma: A maximal ideal is also prime.

Proof: Suppose $I \lhd R$ is a maximal ideal, and $xy \in I$. Suppose further that $x \notin I$. Then the ideal $I + (x)$is an ideal containing $I$and x, so is strictly larger than I. By maximality $I + (x) = R \ni 1$. So $1 = i + rx \implies y = iy + rxy \in I$.

Alternatively, we can use the above two results, and the fact that all fields are integral domains to prove this.

Glossary

Please see the extensive Wikipedia:Glossary of ring theory.

Integral domains

Integral Domains

Motivation: The concept of divisibility is central to the study of ring theory. Integral domains are a useful tool for studying the conditions under which concepts like divisibility and unique factorization are well-behaved. In fact, they are very important for polynomial rings as well.

The integral domain was already defined before on the page on rings. We provide the definition again for reference.

Definition An *integral domain* is a commutative ring R with $1_R \neq 0_R$ such that for all $a, b \in R$, the statement $ab = 0$implies either $a = 0$or $b = 0$.
An equivalent definition is as follows:

Definition Given a ring R , a *zero-divisor* is an element $a \in R$such that $\exists x \in R, x \neq 0$such that $a * x = 0_R$.

Definition An *integral domain* is a commutative ring R with $1_R \neq 0_R$ and with no non-zero zero-divisors.

Remark An integral domain has a useful cancellation property: Let R be an integral domain and let $a, b, c \in R$ with $a \neq 0$. Then $ab = ac$implies $b = c$. For this reason an integral domain is sometimes called a cancellation ring.

Examples:

1. The set $\mathbb{Z}$ of integers under addition and multiplication is an integral domain. However, it is not a field since the element $2 \in Z$ has no multiplicative inverse.
2. The set trivial ring {0} is not an integral domain since it does not satisfy $0 \neq 1$.
3. The set $\mathbb{Z}_6$of congruence classes of the integers modulo 6 is **not** an integral domain because $[2] * [3] = [0]$in $\mathbb{Z}_6$.

Theorem: Any field $F$is an integral domain.

Proof: Suppose that $F$is a field and let $a \in F, a \neq 0$. If $ax = 0$for some x in F, then multiply by $a^{-1}$to see that $ax = 0 \Rightarrow a^{-1}(ax) = a^{-1}0 \Rightarrow 1x = 0 \Rightarrow x = 0$. Fcannot, therefore, contain any zero divisors. Thus, $F$is an integral domain. □

Definition If R is a ring, then the set of polynomials in powers of x with coefficients from R is also a ring, called the polynomial ring of R and written $R[x]$. Each such polynomial is a finite sum of terms, each term being of the form $rx^n$where $r \in R$ and x^n represents the n -th power of x . The leading term of a polynomial is defined as that term of the polynomial which contains the highest power of x in the polynomial.

Remark A polynomial equals 0if and only if each of its coefficients equals 0.

Theorem: Let R be an integral domain and let $R[x]$be the ring of polynomials in powers of x whose coefficients are elements of R . Then $R[x]$is an integral domain if and only if R is.

Proof If commutative ring R is not an integral domain, it contains two non-zero elements $a$and b such that $ab = 0$. Then the polynomials ax and bx are non-zero elements of $R[x]$and $axbx = abxx = 0xx = 0$. Thus if R is not an integral domain, neither is $R[x]$.

Now let R be an integral domain and let A and Bbe polynomials in $R[x]$. If the polynomials are both non-zero, then each one has a non-zero leading term, call them ax^m and bx^n. That these are the leading terms of polynomials A and B means that the leading term of the product AB of these polynomials is abx^{m+n}. Since

R is an integral domain and $a, b \in R$, $ab \neq 0$. This means, by the Remark above, that the product AB is not zero either. This
an integral domain.

Unique Factorization Domains, Principal Ideal Domains, and Euclidean Domains

Unique Factorization Domains, Prime Ideal Domains, and Euclidean Domains are ideas that work only on integral domains.

Some definitions

- Two ring elements a and b are **associates** if a=ub for some unit u.
- A nonzero nonunit a is **irreducible** if it cannot be expressed as a=bc where b and c are both nonunits.
- a divides b if b=ar for some r within R. When this happens, we write a|b.
- A nonzero nonunit is **prime** when a|bc implies that a|b or a|c.

Theorem: If a is prime, then a is irreducible.

Let a be prime, and let a=bc, so that either a|b or a|c. Without loss of generality, assume that a|b, so that b=ad for some element d. Then you can factor a=bc into a=adc, implying that cd=1, or that c is a unit.

Now that we have proven that all prime elements are irreducible, is the converse true? The answer to that is no, for we can easily obtain counterexamples to it. However, we will prove a sufficient and necessary condition for all irreducible elements to be prime.

Unique Factorization Domains

Definition: Let R be an integral domain. If the following two conditions hold:

1. If a is nonzero, then $a=up_1p_2...p_n$ where u is a unit, and p_i are irreducible.
2. Let $a=uq_1q_2...q_m$ be another factorization of irreducibles. Then n = m and after a suitable re-ordering, each p_i and q_i are associates.

Then we call (the integral domain) R a **unique factorization domain (UFD)**.

The converse to the above theorem holds true in a UFD.

Theorem: In a UFD, all irreducibles are prime.

Proof

Let a|bc, where a is irreducible. Then ad=bc for some element d. Taking the factorization, a = $ud_1d_2...d_l$ = $vb_1b_2...b_mwc_1c_2...c_n$ = bc where u, v, and w are units. Because it is a UFD, a must be an associate of some b_i or c_i, implying that a|b or a|c.

The following theorem provides a sufficient and necessary condition for an integral domain R to be an integral domain.

Theorem:

1. Let R be a UFD. R satisfies the following *ascending chain condition* on *principal ideals*: let $a_1, a_2, a_3, \ldots$ be a sequence of elements of R such that the principal ideals $(a_1), (a_2), (a_3), \ldots$ satisfy the condition that $(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq, \ldots$. Then there exists an N such that for all n>N, all the (a_n) are the same.
2. If an integral domain R satisfies the ascending chain condition, then every nonzero element can be factored into irreducible elements, meaning that it satisfies the first condition for being a UFD.
3. If, in addition to satisfying the ascending chain condition, all irreducible elements are prime, then the integral domain is a UFD.

Proof

1. Consider a sequence $a_1, a_2, a_3, \dots$ of elements of R such that $(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq, \dots$. Then obviously $a_{n+1}|a_n$ for all natural numbers n, since $a_n \in (a_{n+1})$. Then due to unique factorization, all the factors of a_{n+1} are associates of the factors a_n, counting multiplicity of factors. Therefore, the number of non-unit factors is a decreasing sequence on the whole numbers. However, a_1 has finitely many factors, so there is an N such that for all n>N, all the factors are a_n associates, meaning that all the (a_n) are also the same.
2. Clearly any nonzero irreducible a_1 can be factored into irreducibles, which is itself. Otherwise, let $a_1 = a_2 b_2$ be a product of nonunits. If this is not a product of irreducibles, then suppose that one of them is not irreducible, say a_2. Then obviously $a_2|a_1$ so the principal ideals satisfy the relations $(a_1) \subseteq (a_2)$. We can factor a_2 in the same way, to obtain $a_2 = a_3 b_3$ as a product of nonunits. Thus, if a_1 cannot be factored into irreducibles, we can get an increasing chain $(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq \dots$ of principal ideals, meaning that it does not satisfy the ascending chain condition.
3. Let $a = sp_1 p_2 p_3 \dots p_n = rq_1 q_2 q_3 \dots q_n$ where r and s are units and each p_i and q_i are irreducible, and thus prime. Since p_1 divides a, it divides one of the factors, and after suitably re-arranging the second factorization, p_1 can divide q_1. However, q_1 is irreducible, so they must be associates, and thus can be factored out and replaced by a unit. We can continue this process until there are no factors left, at which point we conclude that all the factors are associates.

Principal Ideal Domains

Definition: a **principal ideal domain** (PID) is an integral domain such that every ideal can be generated by a single element (i. e. every ideal is a principal ideal).

Theorem: All PIDs are UFDs.

Proof:

Suppose we have an ascending chain of principal ideals $(a_1) \subseteq (a_2) \subseteq \dots$ and let I be the union $I = \bigcup_{i=1}^{\infty} (a_i)$.

Obviously I is an ideal, and is a principal ideal because it is in a PID. Therefore, it is generated by a single element, $I = (a)$. Since $a \in I$, $a \in (a_N)$ for some N. Then if $i \geq N$, then we have $(a) = (a_N)$, so it satisfies the ascending chain condition of principal ideals.

Let an element a be irreducible. If $1 \in (a)$, then a would be a unit, so (a) must be a proper ideal. If there is no maximal proper ideal containing (a), then the ascending chain condition would not be satisfied, so we can conclude that there is a maximal ideal proper ideal I containing (a) (Note: This does not require the Zorn's lemma or axiom of choice, since we did not use the theorem on maximal ideals). This ideal must be a principal ideal (b), but since $a \in (b)$, b|a, and since a is irreducible, b must either be a unit or an associate of a. Since (b) is a proper ideal, b must not be a unit, so it must be an associate of a. Therefore, (a)=(b), so (a) is maximal. However, all maximal ideals are clearly prime, so (a) is a prime ideal, which implies that a is prime.

Theorem: A UFD is a PID if and only if every nontrivial prime ideal is maximal.

Proof:

Suppose R is a PID, so that consequently, it is a UFD. Let (a) be an ideal of R, which in turn must be contained in a maximal proper ideal (b) due to the ascending chain condition (Note: again, this does not make use of Zorn's lemma). Since $a \in (b)$, b|a. Since a is irreducible, b must either be a unit or an associate of a. However, since (b) is a proper ideal, it must not be a unit, so it must be an associate of a. Therefore, (a)=(b), so (a) is maximal. Conversely,

Euclidean Domains

Definition: An integral domain R is a **Euclidean domain** (ED) if there is a function f from the nonzero elements of R to the whole numbers such that for any element $a \in R$ and any nonzero element b, that a=bq+r for some $q, r \in R$ and such that f(r)<f(b) or such that r=0.

Note: In an ED, the Euclidean algorithm to find the greatest common divisor is applicable.

Theorem: All EDs are PIDs.

Proof:
Suppose we have an ideal of R. If it contains only 0, then it is principal. Otherwise, it contains elements other than 0. Then f(I), the image of I under f, is a nonempty set of nonnegative integers. Choose the minimum x of this set, and consider an element b within I which mapped to this x. Let a be another element of I, and there exists $q, r \in R$ such that a=qb+r and such that either f(r)<f(b) or r=0. Since both a and b belong to I, r must also belong to I since r=a-qb. However, f(b) is the minimum, so it must be less than or equal to f(r). Thus, r must be 0, so a=qb, proving that b is the generator of the principal ideal (b).

Fraction Fields

We know from experience that we arrive at the idea of fractions by merely considering the idea of the quotient of two integers. The motivation behind this is simply to arrive at a multiplicative inverse for every non-zero element. Thus, we can consider an integral domain R and construct its *field of fractions*. However, we can also try to make this work for any commutative ring, even if it has zero divisors other than 0. There is a slight alteration required because we cannot define $\frac{a}{b}\frac{c}{d}$ when bd=0. Thus, we must place restrictions in case b and d are zero divisors in the case of multiplication. In this case, it is called the *localization of a ring*.

Definitions

A *multiplicative* subset of a commutative ring R is a subset that does not contain 0, does contain 1, and is closed under multiplication. Some examples of multiplicative sets are the set of nonzero elements of an integral domain, the set of elements of a commutative ring that are not zero divisors, and R\P where P is a prime ideal of the commutative ring R.

Let S be a multiplicative subset. We will consider the Cartesian product R×S. Define the equivalence relation on this product: (a,b)~(c,d) whenever there exists an s such that s(ad-bc)=0.

If it is an integral domain, then (a,b) could be regarded as a/b. Now to check that this is an equivalence relation, it is obvious that it is reflexive and symmetric. To prove that it is transitive, let (a,b)~(c,d) and let (c,d)~(e,f). Then there are elements s and t within S such that s(ad-bc)=0 and such that t(cf-de)=0. This implies that stfad-stfbc=0 and that sbtcf-sbtde=0. Adding the two, we get stfad-sbtde=0, or std(af-be)=0, implying that (a,b)~(e,f).

We can thus use these equivalence classes to define the fraction: $\frac{a}{b}$ is the equivalence class containing (a,b).

Now we set this to be a ring. First, we define addition to be $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$ and multiplication to be $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$. The additive identity is $\frac{0}{1}$and the additive inverse is $\frac{-a}{b}$. The multiplicative identity is simply $\frac{1}{1}$.
Now we prove below that it is indeed a ring:

Theorem

The set of fractions with addition and multiplication as defined is a commutative ring, and if R is an integral domain, then the fractions are also. And if additionally S is R\{0}, then the set of fractions is a field.

Proof

First, we note that

1. $(\frac{a}{b}+\frac{c}{d})+\frac{e}{f} = \frac{ad+bc}{bd}+\frac{e}{f} = \frac{(ad+bc)f+bde}{bdf} = \frac{adf+b(cf+de)}{bdf} = \frac{a}{b}+\frac{cf+de}{df} = \frac{a}{b}+(\frac{c}{d}+\frac{e}{f})$
2. $\frac{a}{b}+\frac{0}{1} = \frac{a\cdot 1+b\cdot 0}{b\cdot 1} = \frac{a}{b}$
3. $\forall c \in S : 1_R(0\cdot 1 - 0\cdot c) = 0 \Rightarrow \frac{0}{1} = \frac{0}{c}(*)$ and therefore $\frac{a}{b}+\frac{-a}{b} = \frac{ab+b(-a)}{bb} = \frac{0}{bb} = \frac{0}{1}$

, from which follows that $(R \times S, +)$is a group.
It is abelian because of the definition of the sum in S and R is commutative.

Furthermore, $(R \times S, \cdot)$is a monoid because

1. $\frac{1}{1}\cdot\frac{a}{b} = \frac{1a}{1b} = \frac{a}{b}$
2. and $(\frac{a}{b}\cdot\frac{c}{d})\cdot\frac{e}{f} = \frac{ace}{bdf} = \frac{a}{b}\cdot(\frac{c}{d}\cdot\frac{e}{f})$, where two (not difficult) intermediate steps are left to the reader.

And, also the distributive laws hold, because

$$\frac{a}{b}\left(\frac{c}{d}+\frac{e}{f}\right) = \left(\frac{c}{d}+\frac{e}{f}\right)\frac{a}{b} = \frac{acf+ade}{bdf}$$ and
$$\frac{acf+ade}{bdf} = \frac{acbf+adbe}{bdfb} = \frac{ac}{bd}+\frac{ae}{bf}$$

, which shows that we have indeed found a ring.

The ring is commutative because of the definition of the product in S R is commutative.

Let now R be an integral domain, and let $\frac{a}{b}, \frac{c}{d} \in S$. Then, because of $(*)$and since $bd \in S$, $\frac{a}{b}\cdot\frac{c}{d} = 0 \Leftrightarrow \frac{ac}{bd} = 0 \Leftrightarrow \frac{ac}{bd} = \frac{0}{bd} \Leftrightarrow \exists s \in S : s(ac\cdot bd - 0\cdot bd) = sacbd = 0$. But since R was assumed to be an integral domain, and $s, b, d \neq 0$since $s, b, d \in S$, the last statement is exactly equivalent to $\exists s \in S : a = 0 \vee c = 0$, which is in turn equivalent to $\exists s \in S : sab = s(ab - 0c) = 0 \vee \exists s \in S : scd = s(cd - 0d) = 0$, and this is equivalent to $\frac{a}{b} = 0 \vee \frac{c}{d} = 0$, which shows that the fraction set is an integral domain if R is one.
Let's assume now that S = R \ {0}, and let $\frac{a}{b} \neq 0 \Leftrightarrow a \neq 0$, where the last equivalence is due to (*) and $\frac{a}{b} = 0 \Leftrightarrow \exists s \in S : s(a1 - b0) = as1 = as = 0 \Leftrightarrow a = 0$, where the last equivalence is in turn due to the fact that R is an integral domain and S does not contain zero. Then $ab \neq 0$due to the fact that R is an integral domain, and thereofore $ab \in S$ since S = R \ {0}, and $\frac{a}{b}\cdot\frac{b}{a} = \frac{ab}{ab}$. But since $\forall s \in S : s(ab1 - ba1) = 0$, we have $\frac{ab}{ab} = 1$and $\frac{a}{b}\cdot\frac{b}{a} = 1$and therefore, by noting that we have assumed R to be commutative, we have that every element of R\{0} is invertible.
From this follows that the set of fractions is indeed a field, because we have already checked all field axioms, QED.

Polynomial Rings

Although there is a theory of non-commutative polynomial rings, it presents some difficulties and will not be treated on this page. Thus, we will work only with commutative rings for their polynomial rings.

The degree of a polynomial $a_0 + a_1X + ... + a_nX^n$ is defined to be n. If R is a field, and f and g are polynomials of R[X], then we can divide f by g to get f=gq+r. However, we can also do this for any arbitrary ring if the leading coefficient of g is 1.

Fields

Fields

We will first define a field.

Definition. A *field* is a non empty set F with two binary operations $+$ and $\cdot$ such that $(F, +, \cdot)$ has commutative unitary ring structure and satisfy the following property:
$\forall x \in F - \{0\} \exists y \in F : x \cdot y = 1$
This means that every element in F except for 0 has a multiplicative inverse.

Essentially, a field is a commutative division ring.

Examples:

1. $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ (rational, real and complex numbers) with standard $+$ and $\cdot$ operations have field structure. These are examples with infinite cardinality.
2. $\mathbb{Z}_p$, the integers modulo p where p is a prime, and $+$ and $\cdot$ are mod p; is a family of finite fields.

Fields and Homomorphisms

Definition (embedding)

An **embedding** is a ring homomorphism $f : F \longrightarrow G$ from a field F to a field G. Since the kernel of a homomorphism is an ideal, a field's only ideals are 0 and the field itself, and $f(1_F) = 1_G$, we must have the kernel equal to 0, so that f is injective and F is isometric to its image under f. Thus, the embedding deserves its name.

Field Extensions

Definition (Field Extension and Degree of Extension)

- Let F and G be fields. If $F \subseteq G$ and there is an embedding from F into G, then G is a **field extension** of F.
- Let G be an extension of F. Consider G as a vector space over the field F. The dimension of this vector space is the **degree** of the extension, $[G : F]$. If the degree is finite, then G is a **finite extension** of F, and G is of degree $n = [G : F]$ over F.

Examples (of field extensions)

- The real numbers $\mathbb{R}$ can be extended into the complex numbers $\mathbb{C}$.
- Similarly, one can add the imaginary number i to the field of rational numbers to form the field of Gaussian integers.

Theorem (Existence of Unique embedding from the integers into a field)

Let F be a field, then there exists a unique homomorphism $\alpha : \mathbb{Z} \to F$.

Proof: Define α such that $\alpha(1) = 1_F, \alpha(2) = 1_f + 1_F$etc. This provides the relevant homomorphism.
Note: The Kernel of α is an ideal of $\mathbb{Z}$. Hence, it is generated by some integer m . Suppose $m = ab$ for some $a, b \in \mathbb{Z}$ then $0 = \alpha(m) = \alpha(a)\alpha(b)$and, since $F$is a field and so also an integral domain, $\alpha(a) = 0$or $\alpha(b) = 0$. This cannot be the case since the kernel is generated by m and hence m must be prime or equal 0.

Definition (Characteristic of Field)

The **characteristic** of a field can be defined to be the generator of the kernel of the homomorphism, as described in the note above.

Algebraic Extensions

Definition (Algebraic Elements and Algebraic Extension)

- Let K be an extension of Fthen $\lambda \in K$ is **algebraic** over Fif there exists a non-zero polynomial $f(x) \in F[x]$such that $f(\lambda) = 0.$
- K is an **algebraic extension** of Fif K is an extension of F, such that every element of K is algebraic over F.

Definition (Minimal Polynomial)

If x is algebraic over Fthen the set of polynomials in $F[x]$which have x as a root is an ideal of $F[x]$. This is a principle ideal domain and so the ideal is generated by a unique monic non-zero polynomial, $m(x)$. We define the $m(x)$to be the minimal polynomial.

Splitting Fields

Definition (Splitting Field)

Let Fbe a field, $f(x) \in F[x]$and $a_1, a_2, ..., a_n$are roots of F. Then a smallest Field Extension of Fwhich contains $a_1, ..., a_n$is called a splitting field of $f(x)$over F.

Finite Fields

Theorem (Order of any finite field)

Let F be a finite field, then $|F| = p^n$for some prime p and $n \in \mathbb{N}$.
proof: The field of integers mod $p$is a subfield of $F$where $p$is the characteristic of F. Hence we can view Fas a vector space over $\mathbb{Z}_p$. Further this must be a finite dimensional vector space because $F$is finite. Hence any $x \in F$can be expressed as a linear combination of $n$members of $F$with scalers in $\mathbb{Z}_p$and any such linear combination is a member of F. Hence $|F| = p^n$.

Theorem (every member of F is a root of $x^p - x$)

let F be a field such that $|F| = p^n$, then every member is a root of the polynomial $x^p - x$.

proof: Consider $F^* = F/0$ as a the multiplicative group. Then by la grange's theorem $\forall x \in F^*, x^{p^n-1} = 1$. So multiplying by x gives $x^{p^n} = x$, which is true for all $x \in F$, including 0.

Theorem (roots of $x^p - x$ are distinct)

Let $x^p - x$ be a polynomial in a splitting field E over $\mathbb{Z}_p$ then the roots $a_1, ...a_n$ are distinct.

Splitting Fields and Algebraic Closues

Splitting Fields

Let F be a field and p(x) be a nonconstant polynomial in F(x). We already know that we can find a field extension of F that contains a root of p(x). However, we would like to know whether an extension E of F containing all of the roots of p(x) exists. In other words, can we find a field extension of F such that p(x) factors into a product of linear polynomials? What is the "smallest" extension containing all the roots of p(x)?

Let F be a field and $p(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n$ be a nonconstant polynomial in F[x]. An extension field E of F is a **splitting field** of p(x) if there exist elements $\alpha_1, \cdots, \alpha_n$ in E such that $E = F(\alpha_1, \cdots, \alpha_n)$ and $p(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n)$ in E[x].

A polynomial $p(x) \in F[x]$ *splits* in E if it is the product of linear factors in E[x].

Example 1: Let $p(x) = x^4 + 2x^2 - 8$ be in $\mathbb{Q}[x]$. Then p(x) has irreducible factors $x^2 - 2$ and $x^2 + 4$. Therefore, the field $\mathbb{Q}(\sqrt{2}, i)$ is a splitting field for p(x).

Example 2: Let $p(x) = x^3 - 3$ be in $\mathbb{Q}[x]$. Then p(x) has a root in the field $\mathbb{Q}(\sqrt[3]{3})$. However, this field is not a splitting field for p(x) since the complex cube roots of 3, $\dfrac{-\sqrt[3]{3} \pm (\sqrt[6]{3})^5 i}{2}$ are not in $\mathbb{Q}(\sqrt[3]{3})$.

Splitting Fields **Theorem** *Let p(x) ∈ F(x) be a nonconstant polynomial. Then there exists a splitting field E for p(x).*

Proof. We will use mathematical induction on the degree of p(x). If $degp(x) = 1$, then p(x) is a linear polynomial and $E = F$. Assume that the theorem is true for all polynomials of degree k with $1 \leq k < n$ and let $degp(x) = n$. We can assume that p(x) is irreducible; otherwise, by our induction hypothesis, we are done. There exists a field K such that p(x) has a zero α_1 in K. Hence, $p(x) = (x - \alpha_1)q(x)$, where $q(x) \in K(x)$. Since $degq(x) = n - 1$, there exists a splitting field $E \supset K$ of q(x) that contains the zeros $\alpha_2, \cdots, \alpha_n$ of p(x) by our induction hypothesis. Consequently

$$E = K(\alpha_2, \cdots, \alpha_n) = F(\alpha_1, \cdots, \alpha_n)$$

is a splitting field of p(x).

The question of uniqueness now arises for splitting fields. This question is answered in the affirmative. Given two splitting fields K and L of a polynomial $p(x) \in F(x)$, there exists a field isomorhpism $\phi : K \longrightarrow L$ that preserves F. In order to prove this result, we must first prove a lemma.

Lemma Theorem *Let $\phi : E \longrightarrow F$ be an isomorphism of fields. Let K be an extension field of E and $\alpha \in K$ be algebraic over E with minimal polynomial p(x). Suppose that L is an extension field of F such that β is root of the polynomial in F[x] obtained from p(x) under the image of ϕ . Then ϕ extends to a unique isomorhpism $\psi : E(\alpha) \longrightarrow F(\beta)$ such that $\psi(\alpha) = \beta$ and ψ agrees with ϕ on E.*

Lemma Proof. If p(x) has degree n, then we can write any element in $E(\alpha)$ as a linear combination of $1, \alpha, \cdots, \alpha^{n-1}$. Therefore, the isomorphism that we are seeking must be

$\phi(a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1}) = \psi(a_0) + \psi(a_1)\beta + \cdots + \psi(a_{n-1})\beta^{n-1}$,
where

$a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1}$
is an element in $E(\alpha)$. The fact that ϕ is an isomorphism could be checked by direct computation; however, it is easier to observe that ϕ is a composition of maps that we already know to be isomorphisms.
We can extend ψ to be an isomorphism from E[x] to F[x], which we will also denote by ψ , by letting

$\psi(a_0 + a_1x + \cdots + a_nx^n) = \psi(a_1)x + \cdots + \psi(a_n)x^n$.
This extension agrees with the original isomorphism $\psi : E \longrightarrow F$, since constant polynomials get mapped to constant polynomials. By assumption, $\psi(p(x)) = q(x)$; hence, ψ maps $\langle p(x)\rangle$ onto $\langle q(x)\rangle$. Consequently, we have an isomorphism $\overline{\psi} : E[x]/\langle p(x)\rangle \rightarrow F[x]/\langle q(x)\rangle$. We have isomorphisms $\sigma : E[x]/\langle p(x)\rangle \longrightarrow F(\alpha)$and $\tau : F[x]/\langle q(x)\rangle \longrightarrow F(\beta)$, defined by evaluation at $\alpha$and β, respectively. Therefore, $\psi = \tau^{-1}\overline{\phi}\sigma$ is the required isomorphism.
Now write $p(x) = (x - \alpha)f(x)$and $q(x) = (x - \beta)g(x)$, where the degrees of f(x) and g(x) are less than the degrees of p(x) and q(x), respectively. The field extension K is a splitting field for f(x) over E(α), and L is a splitting field for g(x) over F(β). By our induction hypotheses there exists an isomorphism $\psi : K \longrightarrow L$ such that ψ agrees with $p\bar{h}i$ on E(α). Hence, there exists an isomorphism $\psi : K \longrightarrow L$ such that ψ agrees with psi on E.
Corollary *Let p(x) be a polynomial in F[x]. Then there exists a splitting field K of p(x) that is unique up to isomorphism.*

Algebraic Closures

Given a field *F*, the question arises as to whether or not we can find a field *E* such that every polynomial *p(x)* has a root in *E*. This leads us to the following theorem.

Theorem 21.11 *Let E be an extension field of F. The set of elements in E that are algebraic over F form a field.*

Proof. Let $\alpha, \beta \in E$be algebraic over *F*. Then $F(\alpha, \beta)$is a finite extension of *F*. Since every element of $F(\alpha, \beta)$is algebraic over $F, \alpha \pm \beta, \alpha/\beta$, and $\alpha/\beta\ (\beta \neq 0)$are all algebraic over *F*. Consequently, the set of elements in *E* that are algebraic over *F* forms a field.
Corollary 21.12 *The set of all algebraic numbers forms a field; that is, the set of all complex numbers that are algebraic over* $\mathbb{Q}$ *makes up a field.*

Let *E* be a field extension of a field *F*. We define the ***algebraic closure*** of a field *F* in *E* to be the field consisting of all elements in *E* that are algebraic over *F*. A field *F* is ***algebraically closed*** if every nonconstant polynomial in *F[x]* has a root in *F*.

Theorem 21.13 *A field F is algebraically closed if and only if every nonconstant polynomial in F[x] factors into linear factors over F[x].*

Proof. Let *F* be an algebraically closed field. If $p(x) \in F[x]$is a nonconstant polynomial, then *p(x)* has a zero in *F*, say α. Therefore, $x - \alpha$ must be a factor of *p(x)* and so $p(x) = (x - \alpha)q_1(x)$, where deg $q_1(x) = \deg p(x) - 1$. Continue this process with $q_1(x)$ to find a factorization

$$p(x) = (x - \alpha)(x - \beta)q_2(x),$$

where deg $q_2(x) = \deg p(x) - 2$. The process must eventually stop since the degree of *p(x)* is finite.
Conversely, suppose that every nonconstant polynomial *p(x)* in *F[x]* factors into linear factors. Let $ax - b$ be such a factor. Then $p(b/a) = 0$. Consequently, *F* is algebraically closed.

Corollary 21.14 *An algebraically closed field F has no proper algebraic extension E.*

Proof. Let *E* be an algebraic extension of *F*; then $F \subset E$. For $\alpha \in E$, the minimal polynomial of α is $x - \alpha$. Therefore, $\alpha \in F$and $F = E$.

Theorem 21.15 *Every field F has a unique algebraic closure.*

It is a nontrivial fact that every field has a unique algebraic closure. The proof is not extremely difficult, but requires some rather sophisticated set theory. We refer the reader to [3], [4], or [8] for a proof of this result.

We now state the Fundamental Theorem of Algebra, first proven by Gauss at the age of 22 in his doctoral thesis. This theorem states that every polynomial with coefficients in the complex numbers has a root in the complex numbers. The proof of this theorem will be given in Abstract Algebra/Galois Theory.

Theorem 21.16 (Fundamental Theorem of Algebra) *The field of complex numbers is algebraically closed.*

Vector Spaces

Vector Spaces

Definition

Definition (Vector Space)

> Let *F* be a field. A set *V* with two binary operations: + (addition) and $\times$ (scalar multiplication), is called a **Vector Space** if it has the following properties:

1. $(V, +)$ forms an abelian group
2. $v(a+b) = va + vb$ for $v \in V$ and $a, b \in F$
3. $a(v+u) = av + au$ for $v, u \in V$ and $a \in F$
4. $(ab)v = a(bv)$
5. $1_F v = v$

The scalar multiplication is formerly defined by $F \times V \xrightarrow{\phi} V$, where $\phi((f, v)) = fv \in V$.

Elements in *F* are called scalars, while elements in *V* are called vectors.

Some Properties of Vector Spaces

1. $0_F v = 0_V = a0_V$
2. $(-1_F)v = -v$
3. $av = 0 \iff a = 0$ or $v = 0$

> Proofs:

1. $0_F v = 0_F v + 0_V = (0_F + 0_F)v = 0_F v + 0_F v \Rightarrow 0_V = 0_F v. Also, a0_V = a0_V + 0_V = a(0_V + 0_V) = a0_V v + a0_V \Rightarrow a0_V = 0_V v$
2. We want to show that $v + (-1_F)v = 0$, but
   $v + (-1_F)v = 1_F v + (-1_F)v = (1_F + (-1_F))v = 0_F v = 0_V$
3. Suppose $av = 0$ such that $a \neq 0$, then $a^{-1}(av) = a^{-1}0 = 0 \Rightarrow 1_F v = v = 0$

Modules

Modules

Motivation

Let *G* be an abelian group under addition. We can define a sort of multiplication on *G* by elements of $\mathbb{Z}$ by writing $ng = \underbrace{g+g+\cdots+g}_{n}$ for $n \in \mathbb{Z}^{+}$ and $g \in G$. We can extend this to the case where *n* is negative by writing $(-n)g = \underbrace{-g-g-\cdots-g}_{n}$. We would, however, like to be able to define a sort of multiplication of a group by an arbitrary ring.

Definition

Definition 1 (Module)

Let *R* be a ring and *M* an abelian group. We call *M* a **left *R*-module** if there is a function $R \times M \to M\ :\ (r,m) \mapsto rm$, called a *scalar multiplication*, satisfying

1. $(r+s)m = rm + sm$,
2. $r(m+n) =$, and
3. $(rs)m = r(sm)$

for all $r, s \in R,\ m, n \in G$.

We call *R* the *ring of scalars* of *M*.

Note: We can also define a right *R*-module analogously by using a function $M \times R \to M\ :\ (m,r) \mapsto mr$. In particular the third property then reads:

$m(rs) = (mr)s$

Note that the two notions coincide if *R* is a commutative ring, and in this case we can simply say that *M* is an *R*-module.

Definition 2: Given any ring *R*, we can define it's *opposite ring*, R^{op}, having the same elements and addition operation as *R*, but opposite multiplication. Their multiplication rules are related by $r \cdot s = sr$. In contrast to group theory, there is no reason in general for a ring to be isomorphic to its opposite ring.

The obervant reader will have noticed that the scalar multiplication in a left *R*-module *M* is simply a ring homomorphism $\phi\ :\ R \to \text{End}(M)$ such that $rm = \phi(r)(m)$ for all $r \in R$, $m \in M$. We leave it as an excercise to verify that the scalar multiplication in a right *R*-module is a ring homomorphism $\phi'\ :\ R^{\text{op}} \to \text{End}(M)$. Thus a right *R*-module is simply a left R^{op}-module. As a consequence of this, all the results we will formulate for left *R*-modules are automatically true for right *R*-modules as well. There are no assumptions that the module is unital, namely that 1m = m for all m in M.

Examples of Modules

1. Any ring *R* is trivially an *R*-module over itself. More interestingly, any left ideal *I* of *R* is also a left *R*-module with the obvious scalar multiplication. In addition, if *I* is a two-sided ideal of *R*, then the quotient ring R/I is an *R*-module with the induced scalar multiplication $r(s+I) = rs + I$, $(s+I)r = sr + I$.

2. If *R* is a ring, then the set $M_{n,m}(R)$ of $n \times m$ matrices with entries in *R* is an *R*-module under componentwise addition and scalar multiplication. More generally, for any set *X*, the set R^X of functionf from *X* to *R*, with or without finite support, is an *R*-module in an obvious way.

3. The *k*-modules over a field *k* are simply the *k*-vector spaces.

4. As was shown in the introduction of this chapter, any abelian group is a $\mathbb{Z}$ -module in a natural way. ("Natural" here has a rigorous mathematical meaning which will be explained later.

5. Let *S* be a subring of a ring *R*. Then *R* is an *S*-module in a natural way. We can extend this as follows. Let S,R be rings and $\phi : S \longrightarrow R$ a ring homomorphism. Then *R* is an *S*-module with scalar multiplication $sr = \phi(s)r$ and $rs = r\phi(s)$ for all $s \in S$, $r \in R$.

6. Any matrix ring of a ring *R* is a *R*-module under componentwise scalar multiplication.

7. Any polynomial ring of a ring *R* is an *R*-module with the obvious scalar multiplication.

8. If *S* is a subring of a ring *R*, then any left *R*-module is also a left *S*-module with the restricted scalar multiplication. We will treat this more generally later.

Submodules

Definition 3: (Submodule)

Given a left R -module M a submodule of M is a subset $N \subseteq M$ satisfying

1. N is a subgroup of M, and
2. for all $r \in R$ and all $n \in N$ we have $rn \in N$.

The second condition above states that submodules are closed under left multiplication by elements of R ; it is implicit that they inherit their multiplication from their containing module; $R \times N \longrightarrow N$ must be the restriction of $R \times M \longrightarrow M$.

Example 4: Any module *M* is a submodule of itself, called the improper submodule, and the zero submodule consisting only of the additive identity of *M*, called the trivial submodule.

Example 5: A left ideal *I* is a submodule of *R* viewed as an *S*-module, where *S* is any (not neccesarily proper) subring of *R*.

Lemma 6: Let *M* be a left *R*-module. Then the following are equivalent.

i) *N* is a submodule of *M*

ii) If $r_i \in R$ and $n_i \in N$ for all $i \in I = \{1, ..., k\}$, $k \in \mathbb{N}$, then $\sum_{i \in I} r_i n_i \in N$.

iii) If $r_1, r_2 \in R$ and $n_1, n_2 \in N$, then $r_1 n_1 + r_2 n_2 \in N$.

Proof: i) => iii): $r_1 n_1$and $r_2 n_2$are in N by the second property, then $r_1 n_1 + r_2 n_2 \in N$ by the first property of Definition 3.

iii) => ii): Follows by induction on k.

ii) => i): Let $k = 2$, $r_1 = 1$, $r_2 = -1$, then for arbitrary $n_1, n_2 \in N$ be have $n_1 - n_2 \in N$, proving N is a subgroup. Now let $k = 1$, then for arbitrary $r_1, n_1 \in N$, $r_1 n_1 \in N$, proving property 2 in Definition 3. ▯

The lemma gives an alternative characterisation of submodules, and those sets closed onder linear combinations of elements.

Analogously to the case of vector spaces, we have ways of creating new subspaces from old ones. The rest of this subsection will be concerned with this.

Lemma 7: Let *M* be a left *R*-module, and let *N* and *L* be submodules of *M*. Then $N \cap L$ is a submodule in *M*, and it is the largest submodule contained in both *N* in *L*.

Proof: Let $a, b \in N \cap L$ and $r, s \in R$. Then $ra + sb \in N$ and $ra + sb \in L$ since *N* and *L* are submodules, so $ra + sb \in N \cap L$ and $N \cap L$ is a submodule of *M*. Now, assume that *S* is a submodule of *M* contained in *N* and *L*. Then any $a \in S$ must be in both *N* and *L* and therefore in $N \cap L$ such that $S \subseteq N \cap L$, proving the lemma. □

Now, as the reader should expect at this point, given submodules *N* and *L* of *M*, the union $N \cup L$ is in general not a submodule. In fact, we have the following lemma:

Lemma 8: Let *M* be a left *R*-module and let *N* and *L* be submodules. Then $N \cup L$ is a submodule if and only if $L \subseteq N$ or $N \subseteq L$.

Proof: The left implication is obvious. For the right implication, assume $N \cup L$ is a submodule of *M*. Then if $n \in N$ and $l \in L$, then $n + l \in N \cup L$, which implies that $n + l \in M$ or $n + l \in L$. Assume without loss of generality that $n + l \in N$. Then, since *N* is a submodule, we must have $(n + l) - n = l \in N$, proving $L \subseteq N$. □

Definition 9: Let *M* be a left *R*-module, and let N_i be submodules for $1 \le i \le k \in \mathbb{N}$. Then define their *sum*,

$$\sum_{i=1}^{k} N_i = \left\{ \sum_{j=1}^{k} n_j \ : \ n_j \in N_j \right\}.$$

Definition 9 has a straightforward extension to sums over arbitrary index sets. This definition is left for the reader to state. We will only need the finite case in this chapter.

Lemma 10: Let *M* be a left *R*-module and let *N* and *L* be submodules. Then $N + L$ is a submodule of *M*, and it is the smallest submodule containing both *N* and *L*.

Proof: It is straightforward to see that $N + L$ is a submodule. To see that it is the smallest submodule containing both *N* and *L*, let *S* be a submodule containing both *N* and *L*. Then for any $n \in N \subseteq S$ and $l \in L \subseteq S$, we must have $n + l \in S$. But this the the same as saying that $N + L \subseteq S$, proving the lemma. □

With Lemma 7 and Lemma 10 established, we can state the main result of this subsection.

Definition 11: Let *M* be a left *R*-module. Then let $\mathcal{S}(M)$ be the set of submodules ordered by set inclusion.

Lemma 12: Let *M* be a left *R*-module. Then $\mathcal{S}(M)$ forms a lattice, the join of $N, L \in \mathcal{S}(M)$ being given by $N \vee L = N + L$ and their meet by $N \wedge L = N \cap L$.

Proof: Most of the work is already done. All that remains is to check assosiativity, the absorption axioms and the idempotency axioms. The associativity is trivially satisfied, $A \cap (B \cap C) = (A \cap B) \cap C$ and $(A + B) + C = A + (B + C)$ for all $A, B, C \in \mathcal{S}(M)$. As for absorption, We have to check $A + (B \cap A) = A$ and $A \cap (A + B) = A$ for all $A, B \in \mathcal{S}(M)$, but this is also trivially true. Lastly, we obviously have $A + A = A$ and $A \cap A = A$ for all $\mathcal{S}(M)$, so we are done. □

Corollary 13: Let *M* be a left *R*-module. Then $\mathcal{S}(M)$ is a modular lattice.

Note: Recall that $\mathcal{S}(M)$ is modular if and only if whenever $A, B, C \in \mathcal{S}(M)$ such that $A \subseteq C$, we have $A + (B \cap C) = (A + B) \cap C$.

Proof: Let $a \in A$, $b \in B$, $c \in C$ such that $a + b = c$. Since $A \subseteq C$, we have $a = c'$ for some $c' \in C$, such that $b \in C$. Thus $b \in B \cap C$ and $a + b \in A + (B \cap C)$. On the other hand, we have $a + b \in A + B$ and $a + b = c \in C$, so $a + b \in (A + B) \cap C$. □

Definition 14: Let *M* be a left *R*-module. A submodule *N* is called *maximal* if whenever *L* is a submodule satisfying $N \subseteq L \subseteq M$, then $L = N$ or $L = M$.

Theorem 15: Every submodule of a left *R*-module is contained in a maximal submodule.

Proof: Let *N* be a submodule, and let $S = \{L \in \mathcal{S}(M) \mid N \subseteq L \subset M\}$. Then *S* is a poset under set inclusion. Let $\{U_1, U_2, ...\}$ be a chain in *S*, and note that $U = U_1 + U_2 + ...$ is a submodule containing each U_i , such that *U* is an upper bound for the chain. Then, since each chain in *S* has an upper bound, by Zorn's Lemma *S* has a maximal element, *P*, say. *P* is obviously an ideal containing *N*. By the definition of *S*, *P* is also a maximal submodule of *M*, proving the theorem. ▯

Generating Modules

Given a subset A of a left R -module M, we define the left submodule generated by A to be the smallest submodule (w.r.t. set containment) of M that contains A . It is denoted by RA for a reason which will become clear in a moment.

The existence of such a submodule comes from the fact that an intersection of R -modules is again an R -module: Consider the set S of all submodules of M containing A . Since M contains A , we see that S is non-empty. The intersection of the modules in S clearly contains A and is a submodule of M. Further, any submodule of M containing A also contains the intersection. Thus $RA = \cap S$.

Assuming that M is unitary, the elements of RA have a simple description;

$$RA = \left\{ \sum_{i=1}^{n} r_i a_i \mid n \in \mathbb{N}, r_i \in R, a_i \in A \right\}.$$

That is, every element of RA can be written as a finite left linear combination of elements of A . This equality can be justified by double inclusion: First, any submodule containing A must contain all left R -linear combinations of elements of A since modules are closed under addition and left multiplication by elements of R . Thus, $RA \supseteq \{\sum_{i=1}^{n} r_i a_i \mid n \in \mathbb{N}, r_i \in R, a_i \in A\}$. Secondly, the set of all such linear combinations forms a submodule of M containing A (use $n = 1$ and $r_1 = 1_R$) and hence it contains RA .

Generating Submodules by Ideals

Consider any ring R , left ideal $I \subseteq R$, and left R -module M. One can think of I as a subring of R (non-unitary when $I \neq R$) and hence M is an I -module using the regular multiplication by elements of R .

If we consider the set $IM = \{\sum_{i=1}^{n} r_i m_i \mid n \in \mathbb{N}, r_i \in I, m_i \in M\}$ we obtain a submodule of M. This follows from our discussion of generated submodules. However, since I is not unitary, it is not necessary that $IM = M$.

Thus, we may consider the quotient module M/IM . Clearly this is an R -module but it is also an R/I module under the obvious action.

Proposition

> Given an R -module M and ideal I of R, the R module M/IM is an R/I -module with multiplication $(r+I)(m+IM) = rm + IM$.

proof.

To show that this is well defined, we observe that if $r + I = s + I$ then $r - s \in I$ and hence

$$(rm + IM) - (sm + IM) = rm - sm + IM = (r - s)m + IM = 0 + IM$$

since $(r - s)m \in IM$. Thus,

$$(r + I)(m + IM) = rm + IM = sm + IM = (s + I)(m + IM)$$

which proves that the action of R/I on M/IM is well defined. It follows now that M/IM is an R/I -module simply because it is an R -module.

Quotient Modules

Recall that any subgroup N of an abelian group M allows one to construct an equivalence relation; for $m, m' \in M$,

$$m \sim m' \iff m - m' \in N \cdot$$

Cosets of N, equivalence classes under the relation above, can then be endowed with a group structure, derived from the original group, and is given the name M/N. The sum of two cosets $m + N = [m]$ and $m' + N = [m']$ is simply $(m + m') + N = [m + m']$.

Lemma 16 Let *M* be a left *R*-module and *N* be a submodule. Then *M/N*, defined above, is a left *R*-module.

Proof: M/N is obviously an abelian group, so we just have to check that it has a well-defined *R*-action. Let $r \in R$ and $m \in M$. Then we define $r(m + N) = rm + N$. The distributivity and associativity properties of the action are inherited from *M*, so we just need well-definedness. Let $m, m' \in M$ with $m - m' \in N$. Then $r(m - m' + N) = r(m - m') + N = N$ since *N* is a submodule, and we are done. ▯

Module Homomorphisms

Like all algebraic structures, we can define maps between modules that preserve their algebraic operations.

Definition (Module Homomorphism)

An R -module homomorphism $\phi : M \to N$ is a function from M to N satisfying

1. $\phi(m + m') = \phi(m) + \phi(m')$ (it is a group homomorphism), and
2. $\phi(rm) = r\phi(m)$.

When a map between two algebraic structures satisfies these two properties then it called an R -linear map.

Definition (Kernel, Image)

Given a module homomorphism $\phi : M \to N$ the kernel of ϕ is the set

$$\ker \phi = \{m \in M \mid \phi(m) = 0\}$$

and the image of ϕ is the set

$$\phi(M) = \{n \in N \mid \exists m \in M, \phi(m) = n\}.$$

The kernel of ϕ is the set of elements in the domain that are sent to zero by ϕ. In fact, the kernel of any module homomorphism is a submodule of M. It is clearly a subgroup, from group theory, and it is also closed under multiplication by elements of R : $\phi(rm) = r\phi(m) = r(0) = 0$ for $m \in \ker \phi$.

Similarly, one can show that the image of ϕ is a submodule of N.

Hypercomplex numbers

Hypercomplex numbers are numbers that use the square root of -1 to create more than 1 extra dimension.

The most basic Hypercomplex number is the one used most often in vector mathematics, the Quaternion, which consists of 4 dimensions. Higher dimensions are diagrammed by adding more roots to negative 1 in a predefined relationship.

Quaternions

A Quaternion consists of four dimensions, one real and the other 3 imaginary. The imaginary dimensions are represented as *i*, *j* and *k*. Each imaginary dimension is a square root of -1 and thus it is not on the normal number line. In practice, the *i*, *j* and *k* are all orthogonal to each other and to the real numbers. As such, they only intersect at the origin (0,0*i*, 0*j*, 0*k*).

The basic form of a quaternion is:

- $q = a + bi + cj + dk$

where a, b, c and d are real number coefficients.

For a quaternion the relationship between *i*, *j* and *k* is defined in this simple rule:

- $i^2 = j^2 = k^2 = i \times j \times k = -1$

From this follows:

- $i \times j = k$, $j \times i = -k$
- $j \times k = i$, $k \times j = -i$
- $k \times i = j$, $i \times k = -j$

As you may have noticed, multiplication is not commutative in hyperdimensional mathematics.

They can also be represented as a 1 by 4 matrix in the form

real	*i*	*j*	*k*
1	1	1	1

...

...

The quaternion is a 4 dimensional number, but it can be used to diagram three dimensional vectors and can be used to turn them without the use of calculus.

see also: Wikipedia's Article on Quaternion [1]

Octonion

8-dimensional. See: Wikipedia's Article on Octonion [2]

Sedenions

16-dimensional. See: Wikipedia's Article on Sedenion [3]

References

[1] http://en.wikipedia.org/wiki/Quaternion
[2] http://en.wikipedia.org/wiki/Octonion
[3] http://en.wikipedia.org/wiki/Sedenion

Matroids

A matroid is an algebraic construct that is related to the notion of independence.

Matroids are an abstraction of several combinatorial objects, among them graphs and matrices. The word matroid was coined by Whitney in 1935 in his landmark paper "On the abstract properties of linear dependence". In defining a matroid Whitney tried to capture the fundamental properties of dependence that are common to graphs and matrices. Almost simultaneously, Birkhoff showed that a matroid can be interpreted as a geometric lattice. Maclane showed that matroids have a geometric representation in terms of points, lines, planes, dimension 3 spaces etc. Often the term combinatorial geometry is used instead of simple matroids. However, combinatorial geometry has another meaning in mathematical literature. Rank 3 combinatorial geometries are frequently called linear spaces. Matroids are a unifying concept in which some problems in graph theory, design theory, coding theory, and combinatorial optimization become simpler to understand.

Algebras

Algebras

In this section we will talk about structures with three operations. These are called algebras. We will start by defining an algebra over a field, which is a vector space with a bilinear vector product. After giving some examples, we will then move to a discussion of quivers and their path algebras.

Algebras over a Field

Definition 1: Let Fa field, and let A be an F-vector space on which we define the vector product $\cdot\ :\ A \times A \longrightarrow A$. Then A is called an *algebra over* Fprovided that $(A, +, \cdot)$ is a ring, where $+$is the vector space addition, and if for all $a, b, c \in A$ and $\alpha \in F$,

1. $a(bc) = (ab)c$,
2. $a(b+c) = ab + ac$and $(a+b)c = ac + bc$,
3. $\alpha(ab) = (\alpha a)b = a(\alpha b)$.

The *dimension* of an algebra is the dimension of A as a vector space.

Remark 2: The appropriate definition of a *subalgebra* is clear from Definition 1. We leave its formal statement to the reader.

Definition 2: If $(A, +, \cdot)$is a commutative ring, A is called a *commutative algebra.* If it is a division ring, A is called a *division algebra.* We reserve the terms *real* and *complex algebra* for algebras over $\mathbb{R}$ and $\mathbb{C}$, respectively. The reader is invited to check that the following examples really are examples of algebras.

Example 3: Let Fbe a field. The vector space F^nforms a commutative F-algebra under componentwise multiplication.

Example 4: The quaternions $\mathbb{H}$ is a 4-dimensional real algebra. We leave it to the reader to show that it is *not* a 2-dimensional complex algebra.

Example 5: Given a field F, the vector space of polynomials $F[x]$is a commutative F-algebra in a natural way.

Example 6: Let Fbe a field. Then any matrix ring over F, for example $\begin{pmatrix} F & 0 \\ F & F \end{pmatrix}$, gives rise to an F -algebra in a natural way.

Quivers and Path Algebras

Naively, a quiver can be understood as a directed graph where we allow loops and parallell edges. Formally, we have the following.

Definition 7: A *quiver* is a collection of four pieces of data, $Q = (Q_0, Q_1, s, t)$,

1. $Q_0$is the set of *vertices* of the quiver,
2. $Q_1$is the set of *edges*, and
3. $s, t\ :\ Q_1 \longrightarrow Q_0$are functions associating with each edge a *source vertex* and a *target vertex*, respectively.

We will always assume that $Q_0$is nonempty and that $Q_0$and $Q_1$are finite sets.

Example 8: The following are the simplest examples of quivers:

1. The quiver with one point and no edges, represented by 1.

2. The quiver with n point and no edges, $1 \quad 2 \quad ... \quad n$.
3. The linear quiver with n points, $1 \xrightarrow{a_1} 2 \xrightarrow{a_2} ... \xrightarrow{a_{n-1}} n$.
4. The simplest quiver with a nontrivial loop, $1 \underset{a}{\overset{b}{\leftrightarrows}} 2$.

Definition 9: Let Q be a quiver. A *path* in Q is a sequence of edges $a = a_m a_{m-1} ... a_1$ where $s(a_i) = t(a_{i-1})$ for all $i = 2, ..., m$. We extend the domains of s and t and define $s(a) \equiv s(a_0)$ and $t(a) \equiv t(a_m)$. We define the *length* of the path to be the number of edges it contains and write $l(a) = m$. With each vertex i of a quiver we associate the *trivial path* e_i with $s(e_i) = t(e_i) = i$ and $l(e_i) = 0$. A nontrivial path a with $s(a) = t(a) = i$ is called an *oriented loop* at i.

The reason quivers are interesting for us is that they provide a concrete way of constructing a certain family of algebras, called *path algebras*.

Definition 10: Let Q be a quiver and F a field. Let FQ denote the free vector space generated by all the paths of Q. On this vector space, we define a vector porduct in the obvious way: if $u = u_m ... u_1$ and $v = v_n ... v_1$ are paths with $s(v) = t(u)$, define their product vu by concatenation: $vu = v_n ... v_1 u_m ... u_1$. If $s(v) \neq t(u)$, define their product to be $vu = 0$. This product turns FQ into an F-algebra, called the *path algebra of* Q.

Lemma 11: Let Q be a quiver and F field. If Q contains a path of length $|Q_0|$, then FQ is infinite dimensional.

Proof: By a counting argument such a path must contain an oriented loop, a, say. Evidently $\{a^n\}_{n \in \mathbb{N}}$ is a linearly independent set, such that FQ is infinite dimensional.

Lemma 12: Let Q be a quiver and F a field. Then FQ is infinite dimensional if and only if Q contains an oriented loop.

Proof: Let a be an oriented loop in Q. Then FQ is infinite dimensional by the above argument. Conversely, assume Q has no loops. Then the vertices of the quiver can be ordered such that edges always go from a lower to a higher vertex, and since the length of any given path is bounded above by $|Q_0| - 1$, there dimension of FQ is bounded above by $\dim FQ \leq |Q_0|^2 - |Q_0| < \infty$.

Lemma 13: Let Q be a quiver and F a field. Then the trivial edges e_i form an orthogonal idempotent set.

Proof: This is immediate from the definitions: $e_i e_j = 0$ if $i \neq j$ and $e_i^2 = e_i$.

Corollary 14: The element $\sum_{i \in Q_0} e_i$ is the identity element in FQ.

Proof: It sufficed to show this on the generators of FQ. Let a be a path in Q with $s(a) = j$ and $t(a) = k$.

Then $\left(\sum_{i \in Q_0} e_i \right) a = \sum_{i \in Q_0} e_i a = e_j a = a$. Similarily, $a \left(\sum_{i \in Q_0} e_i \right) = a$.

To be covered:

- General R-algebras

Boolean algebra

Boolean Algebra

Boolean algebra is a deductive mathematical system closed over the values zero and one (false and true). A *binary operator* " o " defined over this set of values accepts a pair of ator accepts two boolean inputs and produces a single boolean output (the logical AND of the two inputs).

For any given algebra system, there are some initial assumptions, or *postulates* that the system follows. You can deduce additional rules, theorems, and other properties of the system from this basic set of postulates:

- *Closure*. The boolean system of *closed* with respect to a binary operator if for every pair of boolean values. It produces a boolean result. For example, logical AND is closed in the boolean system because it accepts only boolean operands and produces only boolean results.
- *Commutativity*. A binary operator " o " is said to be commutative if $A^oB = B^oA$ for all possible boolean values A and B.
- *Associativity*. A binary operator " o " is said to be associative if $$(A^oB)^oC = A^o(B^oC)$$ for all boolean values A, B, and C.
- *Distribution*. Two binary operators " o " and "%" are distributive if $$A^o(B\%C) = (A^oB)\%(A^oC)$$ for all boolean values A, B, and C.
- *Identity*. A boolean value I is said to be the *identity element* with respect to some binary operator " o " if $A^oI = A$
- *Inverse*. A boolean value I is said to be the *inverse element* with respect to some binary operator " o " if $A^oI = B$ and $B/neqA$ (i.e., B is the opposite value of A in a boolean system).

For our purposes, we will base boolean algebra on the following set of operators and values:

The two possible values in the boolean system are zero and one. Often we will call these values false and true (respectively).

The symbol " $\cdot$ " represents the logical AND operation; e.q., $A \cdot B$ is the result of logically ANDing the boolean values A and B. When using single letter variable names, this text will drop the " $\cdot$ " symbol; Therefore, AB also represents the logical AND of the variables A and B (we will also call this the *product<\i> of A and B).*

The symbol "+" represents the logical OR operation; e.g., A + B is the result of logical ORing the boolean values A and B. (We will also call this the sum of A and B.)

Logical *complement, negation,* or *not*, is a unary operator. This text will use the (') symbol to denote logical negation. For example, A ' denotes the logical NOT of A.

If several different operators appear in a single boolean expression, the result of the expression depends on the precedence of the operators. We'll use the following precedences (from highest to lowest) for the boolean operators: parenthesis, logical NOT, logical AND, then logical OR. The logical AND and OR operators are *left associative*. If two operators with the same precedence are adjacent, you must evaluate them from left to right. The logical NOT operation is right associative, although it would produce the same result using left or right associativity since it is a unary operator.

We will also use the following set of postulates:

P1 Boolean algebra is closed under the AND, OR, and NOT operations.

P2 The identity element with respect to $\cdot$ is one and + is zero. There is no identity element with respect to logical NOT.

P3 The $\cdot$ and + operators are commutative.

P4 $\cdot$ and + are distributive with respect to one another. That is, $A \cdot (B + C) = (A \cdot B) + (A \cdot C)$ and $A + (B \cdot C) = (A + B) \cdot (A + C)$.

P5 For every value A there exists a value A' such that $A \cdot A' = 0$ and $A + A' = 1$. This value is the logical complement (or NOT) of A.

P6 $\cdot$ and + are both associative. That is, $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ and $(A + B) + C = A + (B + C)$.

You can prove all other theorems in boolean algebra using these postulates. This text will not go into the formal proofs of these theorems, however, it is a good idea to familiarize yourself with some important theorems in boolean algebra. A sampling include:

Th1: A + A = A

Th2: $A \cdot A = A$

Th3: A + 0 = A

Th4: $A \cdot 1 = A$

Th5: $A \cdot 0 = 0$

Th6: A + 1 = 1

Th7: $(A + B)' = A' \cdot B'$

Th8: $(A \cdot B)' = A' + B'$

Th9: $A + A \cdot B = A$

Th10: $A \cdot (A + B) = A$

Th11: $A + A'B = A + B$

Th12: $A' \cdot (A + B') = A'B'$

Th13: $AB + AB' = A$

Th14: $(A' + B') \cdot (A' + B) = A'$

Th15: $A + A' = 1$

Th16: $A \cdot A' = 0$

Theorems seven and eight above are known as *DeMorgan's Theorems* after the mathematician who discovered them.

The theorems above appear in pairs. Each pair (e.g. Th1 & Th 2, Th3 & Th4, etc.) form a *dual*. An important principle in the boolean algebra system is that of *duality*. Any valid expression you can create using the postulates and theorems of boolean algebra remains valid if you interchange the operators and constants appearing in the expression. Specifically, if you exchange the $\cdot$ and + operators and swap the 0 and 1 values in an expression. you will wind up with an expression that obeys all the rules of boolean algebra. *This does not mean the dual expression computes the same values,* it only means that both expressions are legal in the boolean algebra system. Therefore, this is an easy way to generate a second theorem for any fact you prove in the boolean algebra system.

Although, we will not be proving any theorems for the sake of boolean algebra in this text, we will use these theorems to show that two boolean equations are identical. This is an important operation when attempting to produce *canonical representations* of a boolean expression or when simplifying a boolean expression.

Boolean Functions and Truth Tables

A boolean expression is a sequence of zeros, ones, and *literals* separated by boolean operators. A literal is a primes (negated) or unprimed variable name. For our purposes, all variable names will be a single alphabetic character. A boolean function is a specific boolean expression; we will generally give boolean functions the name "F" with a possible subscript. For example, consider the following boolean:

$$F_0 = AB + C$$

This function computes the logical AND of A and B and then logically ORs this result with C. If A=1, B=O, and C=1, then F_0 returns the value one $(1 \cdot 0 + 1 = 1)$.

Another way to represent a boolean function is a via a *truth table*. The previous chapter used truth tables to represent the AND and OR functions. Those truth tables took the forms:

Table 6: AND Truth Table

AND	0	1
0	0	0
1	0	1

Table 7: OR Truth Table

OR	0	1
0	0	1
1	1	1

For binary operators and twp input variables, this form of a truth table is very natural and convenient. However, reconsider the boolean function F_0 above. That function has *three* input variables, not two. Therefore, one cannot use the truth table format given above. Fortunately, it is still very easy to construct truth tables for three or more variables, The following example shows one way to do this for functions of three or four variables:

Clifford Algebras

In mathematics, **Clifford algebras** are a type of associative algebra. They can be thought of as one of the possible generalizations of the complex numbers and quaternions. The theory of Clifford algebras is intimately connected with the theory of quadratic forms and orthogonal transformations. Clifford algebras have important applications in a variety of fields including geometry and theoretical physics. They are named for the English geometer William Clifford.

Some familiarity with the basics of multilinear algebra will be useful in reading this section.

Introduction and basic properties

Specifically, a Clifford algebra is a unital associative algebra which contains and is generated by a vector space *V* equipped with a quadratic form *Q*. The Clifford algebra $C\ell(V,Q)$ is the "freest" algebra generated by *V* subject to the condition[1]

$$v^2 = Q(v) \quad \text{for all } v \in V.$$

If the characteristic of the ground field *K* is not 2, then one can rewrite this fundamental identity in the form

$$uv + vu = \langle u, v \rangle \text{ for all } u, v \in V,$$

where $<u, v> = Q(u + v) - Q(u) - Q(v)$ is the symmetric bilinear form associated to Q. This idea of "freest" or "most general" algebra subject to this identity can be formally expressed through the notion of a universal property (see below).

Clifford algebras are closely related to exterior algebras. In fact, if $Q = 0$ then the Clifford algebra $C\ell(V,Q)$ is just the exterior algebra $\Lambda(V)$. For nonzero *Q* there exists a canonical *linear* isomorphism between $\Lambda(V)$ and $C\ell(V,Q)$ whenever the ground field *K* does not have characteristic two. That is, they are naturally isomorphic as vector spaces, but with different multiplications (in the case of characteristic two, they are still isomorphic as vector spaces, just not naturally). Clifford multiplication is strictly richer than the exterior product since it makes use of the extra information provided by *Q*. More precisely, they may be thought of as quantizations of the exterior algebra, in the same way that the Weyl algebra is a quantization of the symmetric algebra.

Quadratic forms and Clifford algebras in characteristic 2 form an exceptional case. In particular, if char $K = 2$ it is not true that a quadratic form is determined by its symmetric bilinear form, or that every quadratic form admits an orthogonal basis. Many of the statements in this article include the condition that the characteristic is not 2, and are false if this condition is removed.

Universal property and construction

Let *V* be a vector space over a field *K*, and let $Q : V \to K$ be a quadratic form on *V*. In most cases of interest the field *K* is either **R** or **C** (which have characteristic 0) or a finite field.

A Clifford algebra $C\ell(V,Q)$ is a unital associative algebra over *K* together with a linear map $i : V \to C\ell(V,Q)$ defined by the following universal property: Given any associative algebra *A* over *K* and any linear map $j : V \to A$ such that

$$j(v)^2 = Q(v)1 \text{ for all } v \in V$$

(where 1 denotes the multiplicative identity of *A*), there is a unique algebra homomorphism $f : C\ell(V,Q) \to A$ such that the following diagram commutes (i.e. such that $f \circ i = j$):

$$\begin{tikzcd}
V \arrow[r, "i"] \arrow[dr, "j"'] & C\ell(V,Q) \arrow[d, dashed, "f"] \\
 & A
\end{tikzcd}$$

Working with a symmetric bilinear form <·,·> instead of Q (in characteristic not 2), the requirement on j is

$j(v)j(w) + j(w)j(v) = $ <v, w> *for all* v, w V.

A Clifford algebra as described above always exists and can be constructed as follows: start with the most general algebra that contains V, namely the tensor algebra $T(V)$, and then enforce the fundamental identity by taking a suitable quotient. In our case we want to take the two-sided ideal I_Q in $T(V)$ generated by all elements of the form

$v \otimes v - Q(v)1$ for all $v \in V$

and define $C\ell(V,Q)$ as the quotient

$C\ell(V,Q) = \mathrm{T}(V)/I_{Q}.$

It is then straightforward to show that $C\ell(V,Q)$ contains V and satisfies the above universal property, so that $C\ell$ is unique up to isomorphism; thus one speaks of "the" Clifford algebra $C\ell(V, Q)$. It also follows from this construction that i is injective. One usually drops the i and considers V as a linear subspace of $C\ell(V,Q)$.

The universal characterization of the Clifford algebra shows that the construction of $C\ell(V,Q)$ is *functorial* in nature. Namely, $C\ell$ can be considered as a functor from the category of vector spaces with quadratic forms (whose morphisms are linear maps preserving the quadratic form) to the category of associative algebras. The universal property guarantees that linear maps between vector spaces (preserving the quadratic form) extend uniquely to algebra homomorphisms between the associated Clifford algebras.

Basis and dimension

If the dimension of V is n and $\{e_1,\ldots,e_n\}$ is a basis of V, then the set

$$\{e_{i_1}e_{i_2}\cdots e_{i_k} \mid 1 \le i_1 < i_2 < \cdots < i_k \le n \text{ and } 0 \le k \le n\}$$

is a basis for $C\ell(V,Q)$. The empty product ($k = 0$) is defined as the multiplicative identity element. For each value of k there are n choose k basis elements, so the total dimension of the Clifford algebra is

$$\dim C\ell(V,Q) = \sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

Since V comes equipped with a quadratic form, there is a set of privileged bases for V: the orthogonal ones. An orthogonal basis in one such that

$$\langle e_i, e_j \rangle = 0 \qquad i \neq j.$$

where <·,·> is the symmetric bilinear form associated to Q. The fundamental Clifford identity implies that for an orthogonal basis

$$e_i e_j = -e_j e_i \qquad i \neq j.$$

This makes manipulation of orthogonal basis vectors quite simple. Given a product $e_{i_1}e_{i_2}\cdots e_{i_k}$ of *distinct* orthogonal basis vectors, one can put them into standard order by including an overall sign corresponding to the number of flips needed to correctly order them (i.e. the signature of the ordering permutation).

If the characteristic is not 2 then an orthogonal basis for V exists, and one can easily extend the quadratic form on V to a quadratic form on all of $C\ell(V,Q)$ by requiring that distinct elements $e_{i_1}e_{i_2}\cdots e_{i_k}$ are orthogonal to one another whenever the $\{e_i\}$'s are orthogonal. Additionally, one sets

$$Q(e_{i_1}e_{i_2}\cdots e_{i_k}) = Q(e_{i_1})Q(e_{i_2})\cdots Q(e_{i_k}).$$

The quadratic form on a scalar is just $Q(\lambda) = \lambda^2$. Thus, orthogonal bases for V extend to orthogonal bases for $C\ell(V,Q)$. The quadratic form defined in this way is actually independent of the orthogonal basis chosen (a basis-independent formulation will be given later).

Examples: Real and complex Clifford algebras

The most important Clifford algebras are those over real and complex vector spaces equipped with nondegenerate quadratic forms.

Every nondegenerate quadratic form on a finite-dimensional real vector space is equivalent to the standard diagonal form:

$$Q(v) = v_1^2 + \cdots + v_p^2 - v_{p+1}^2 - \cdots - v_{p+q}^2$$

where $n = p + q$ is the dimension of the vector space. The pair of integers (p, q) is called the signature of the quadratic form. The real vector space with this quadratic form is often denoted $\mathbf{R}^{p,q}$. The Clifford algebra on $\mathbf{R}^{p,q}$ is denoted $C\ell_{p,q}(\mathbf{R})$. The symbol $C\ell_n(\mathbf{R})$ means either $C\ell_{n,0}(\mathbf{R})$ or $C\ell_{0,n}(\mathbf{R})$ depending on whether the author prefers positive definite or negative definite spaces.

A standard orthonormal basis $\{e_i\}$ for $\mathbf{R}^{p,q}$ consists of $n = p + q$ mutually orthogonal vectors, p of which have norm +1 and q of which have norm −1. The algebra $C\ell_{p,q}(\mathbf{R})$ will therefore have p vectors which square to +1 and q vectors which square to −1.

Note that $C\ell_{0,0}(\mathbf{R})$ is naturally isomorphic to $\mathbf{R}$ since there are no nonzero vectors. $C\ell_{0,1}(\mathbf{R})$ is a two-dimensional algebra generated by a single vector e_1 which squares to −1, and therefore is isomorphic to $\mathbf{C}$, the field of complex numbers. The algebra $C\ell_{0,2}(\mathbf{R})$ is a four-dimensional algebra spanned by $\{1, e_1, e_2, e_1e_2\}$. The latter three elements square to −1 and all anticommute, and so the algebra is isomorphic to the quaternions $\mathbf{H}$. The next algebra in the sequence is $C\ell_{0,3}(\mathbf{R})$ is an 8-dimensional algebra isomorphic to the direct sum $\mathbf{H} \oplus \mathbf{H}$ called Clifford biquaternions.

One can also study Clifford algebras on complex vector spaces. Every nondegenerate quadratic form on a complex vector space is equivalent to the standard diagonal form

$$Q(z) = z_1^2 + z_2^2 + \cdots + z_n^2$$

where $n = \dim V$, so there is essentially only one Clifford algebra in each dimension. We will denote the Clifford algebra on $\mathbf{C}^n$ with the standard quadratic form by $C\ell_n(\mathbf{C})$. One can show that the algebra $C\ell_n(\mathbf{C})$ may be obtained as the complexification of the algebra $C\ell_{p,q}(\mathbf{R})$ where $n = p + q$:

$$C\ell_n(\mathbb{C}) \cong C\ell_{p,q}(\mathbb{R}) \otimes \mathbb{C} \cong C\ell(\mathbb{C}^{p+q}, Q \otimes \mathbb{C}).$$

Here Q is the real quadratic form of signature (p,q). Note that the complexification does not depend on the signature. The first few cases are not hard to compute. One finds that

$$C\ell_0(\mathbf{C}) = \mathbf{C}$$

$$C\ell_1(\mathbf{C}) = \mathbf{C} \oplus \mathbf{C}$$

$$C\ell_2(\mathbf{C}) = M_2(\mathbf{C})$$

where $M_2(\mathbf{C})$ denotes the algebra of 2×2 matrices over $\mathbf{C}$.

It turns out that every one of the algebras $C\ell_{p,q}(\mathbf{R})$ and $C\ell_n(\mathbf{C})$ is isomorphic to a matrix algebra over $\mathbf{R}$, $\mathbf{C}$, or $\mathbf{H}$ or to a direct sum of two such algebras. For a complete classification of these algebras see classification of Clifford algebras.

Properties

Relation to the exterior algebra

Given a vector space V one can construct the exterior algebra $\Lambda(V)$, whose definition is independent of any quadratic form on V. It turns out that if F does not have characteristic 2 then there is a natural isomorphism between $\Lambda(V)$ and $C\ell(V,Q)$ considered as vector spaces (and there exists an isomorphism in characteristic two, which may not be natural). This is an algebra isomorphism if and only if $Q = 0$. One can thus consider the Clifford algebra $C\ell(V,Q)$ as an enrichment (or more precisely, a quantization, cf. the Introduction) of the exterior algebra on V with a multiplication that depends on Q (one can still define the exterior product independent of Q).

The easiest way to establish the isomorphism is to choose an *orthogonal* basis $\{e_i\}$ for V and extend it to an orthogonal basis for $C\ell(V,Q)$ as described above. The map $C\ell(V,Q) \to \Lambda(V)$ is determined by

$$e_{i_1}e_{i_2}\cdots e_{i_k} \mapsto e_{i_1}\wedge e_{i_2}\wedge\cdots\wedge e_{i_k}.$$

Note that this only works if the basis $\{e_i\}$ is orthogonal. One can show that this map is independent of the choice of orthogonal basis and so gives a natural isomorphism.

If the characteristic of K is 0, one can also establish the isomorphism by antisymmetrizing. Define functions $f_k : V \times \ldots \times V \to C\ell(V,Q)$ by

$$f_k(v_1,\cdots,v_k) = \frac{1}{k!}\sum_{\sigma\in S_k} \operatorname{sgn}(\sigma)\, v_{\sigma(1)}\cdots v_{\sigma(k)}$$

where the sum is taken over the symmetric group on k elements. Since f_k is alternating it induces a unique linear map $\Lambda^k(V) \to C\ell(V,Q)$. The direct sum of these maps gives a linear map between $\Lambda(V)$ and $C\ell(V,Q)$. This map can be shown to be a linear isomorphism, and it is natural.

A more sophisticated way to view the relationship is to construct a filtration on $C\ell(V,Q)$. Recall that the tensor algebra $T(V)$ has a natural filtration: $F^0 \subset F^1 \subset F^2 \subset \ldots$ where F^k contains sums of tensors with rank $\leq k$. Projecting this down to the Clifford algebra gives a filtration on $C\ell(V,Q)$. The associated graded algebra

$$\bigoplus_k F^k/F^{k-1}$$

is naturally isomorphic to the exterior algebra $\Lambda(V)$. Since the associated graded algebra of a filtered algebra is always isomorphic to the filtered algebra as filtered vector spaces (by choosing complements of F^k in F^{k+1} for all k), this provides an isomorphism (although not a natural one) in any characteristic, even two.

Grading

The linear map on V defined by $v \mapsto -v$ preserves the quadratic form Q and so by the universal property of Clifford algebras extends to an algebra automorphism

$$\alpha : C\ell(V,Q) \to C\ell(V,Q).$$

Since α is an involution (i.e. it squares to the identity) one can decompose $C\ell(V,Q)$ into positive and negative eigenspaces

$$C\ell(V,Q) = C\ell^0(V,Q) \oplus C\ell^1(V,Q)$$

where $C\ell^i(V,Q) = \{x \in C\ell(V,Q) \mid \alpha(x) = (-1)^i x\}$. Since α is an automorphism it follows that

$$C\ell^i(V,Q)C\ell^j(V,Q) = C\ell^{i+j}(V,Q)$$

where the superscripts are read modulo 2. This means that $C\ell(V,Q)$ is a $\mathbf{Z}_2$-graded algebra (also known as a superalgebra). Note that $C\ell^0(V,Q)$ forms a subalgebra of $C\ell(V,Q)$, called the **even subalgebra**. The piece $C\ell^1(V,Q)$ is called the **odd part** of $C\ell(V,Q)$ (it is not a subalgebra). This $\mathbf{Z}_2$-grading plays an important role in the analysis and application of Clifford algebras. The automorphism α is called the **main involution** or **grade involution**.

Remark. In characteristic not 2 the algebra $C\ell(V,Q)$ inherits a **Z**-grading from the canonical isomorphism with the exterior algebra $\Lambda(V)$. It is important to note, however, that this is a *vector space grading only*. That is, Clifford multiplication does not respect the **Z**-grading only the $\mathbf{Z}_2$-grading. Happily, the gradings are related in the natural way: $\mathbf{Z}_2 = \mathbf{Z}/2\mathbf{Z}$. The *degree* of a Clifford number usually refers to the degree in the **Z**-grading. Elements which are pure in the $\mathbf{Z}_2$-grading are simply said to be even or odd.

If the characteristic of F is not 2 then the even subalgebra $C\ell^0(V,Q)$ of a Clifford algebra is itself a Clifford algebra. If V is the orthogonal direct sum of a vector a of norm $Q(a)$ and a subspace U, then $C\ell^0(V,Q)$ is isomorphic to $C\ell(U,-Q(a)Q)$, where $-Q(a)Q$ is the form Q restricted to U and multiplied by $-Q(a)$. In particular over the reals this implies that

$C\ell^0_{p,q}(\mathbb{R}) \cong C\ell_{p,q-1}(\mathbb{R})$ for $q > 0$, and

$C\ell^0_{p,q}(\mathbb{R}) \cong C\ell_{q,p-1}(\mathbb{R})$ for $p > 0$.

In the negative-definite case this gives an inclusion $C\ell_{0,n-1}(\mathbf{R}) \subset C\ell_{0,n}(\mathbf{R})$ which extends the sequence

$\mathbf{R} \subset \mathbf{C} \subset \mathbf{H} \subset \mathbf{H}\oplus\mathbf{H} \subset \ldots$

Likewise, in the complex case, one can show that the even subalgebra of $C\ell_n(\mathbf{C})$ is isomorphic to $C\ell_{n-1}(\mathbf{C})$.

Antiautomorphisms

In addition to the automorphism α, there are two antiautomorphisms which play an important role in the analysis of Clifford algebras. Recall that the tensor algebra $T(V)$ comes with an antiautomorphism that reverses the order in all products:

$$v_1 \otimes v_2 \otimes \cdots \otimes v_k \mapsto v_k \otimes \cdots \otimes v_2 \otimes v_1.$$

Since the ideal I_Q is invariant under this reversal, this operation descends to an antiautomorphism of $C\ell(V,Q)$ called the **transpose** or **reversal** operation, denoted by x^t. The transpose is an antiautomorphism: $(xy)^t = y^t x^t$. The transpose operation makes no use of the $\mathbf{Z}_2$-grading so we define a second antiautomorphism by composing α and the transpose. We call this operation **Clifford conjugation** denoted $\bar{x}$

$$\bar{x} = \alpha(x^t) = \alpha(x)^t.$$

Of the two antiautomorphisms, the transpose is the more fundamental.[3]

Note that all of these operations are involutions. One can show that they act as ±1 on elements which are pure in the **Z**-grading. In fact, all three operations depend only on the degree modulo 4. That is, if x is pure with degree k then

$$\alpha(x) = \pm x \qquad x^t = \pm x \qquad \bar{x} = \pm x$$

where the signs are given by the following table:

***k* mod 4**	**0**	**1**	**2**	**3**	
$\alpha(x)$	+	−	+	−	$(-1)^k$
x^t	+	+	−	−	$(-1)^{k(k-1)/2}$
$\bar{x}$	+	−	−	+	$(-1)^{k(k+1)/2}$

The Clifford scalar product

When the characteristic is not 2 the quadratic form *Q* on *V* can be extended to a quadratic form on all of *Cℓ*(*V*,*Q*) as explained earlier (which we also denoted by *Q*). A basis independent definition is

$$Q(x) = \langle x^t x \rangle$$

where <*a*> denotes the scalar part of *a* (the grade 0 part in the **Z**-grading). One can show that

$$Q(v_1 v_2 \cdots v_k) = Q(v_1)Q(v_2)\cdots Q(v_k)$$

where the v_i are elements of *V* — this identity is *not* true for arbitrary elements of *Cℓ*(*V*,*Q*).

The associated symmetric bilinear form on *Cℓ*(*V*,*Q*) is given by

$$\langle x, y \rangle = \langle x^t y \rangle.$$

One can check that this reduces to the original bilinear form when restricted to *V*. The bilinear form on all of *Cℓ*(*V*,*Q*) is nondegenerate if and only it is nondegenerate on *V*.

It is not hard to verify that the transpose is the adjoint of left/right Clifford multiplication with respect to this inner product. That is,

$$\langle ax, y \rangle = \langle x, a^t y \rangle, \text{and}$$
$$\langle xa, y \rangle = \langle x, y a^t \rangle.$$

Structure of Clifford algebras

In this section we assume that the vector space *V* is finite dimensional and that the bilinear form of *Q* is non-singular. A central simple algebra over *K* is a matrix algebra over a (finite dimensional) division algebra with center *K*. For example, the central simple algebras over the reals are matrix algebras over either the reals or the quaternions.

- If *V* has even dimension then *Cℓ*(*V*,*Q*) is a central simple algebra over *K*.
- If *V* has even dimension then $C\ell^0$(*V*,*Q*) is a central simple algebra over a quadratic extension of *K* or a sum of two isomorphic central simple algebras over *K*.
- If *V* has odd dimension then *Cℓ*(*V*,*Q*) is a central simple algebra over a quadratic extension of *K* or a sum of two isomorphic central simple algebras over *K*.
- If *V* has odd dimension then $C\ell^0$(*V*,*Q*) is a central simple algebra over *K*.

The structure of Clifford algebras can be worked out explicitly using the following result. Suppose that *U* has even dimension and a non-singular bilinear form with discriminant *d*, and suppose that *V* is another vector space with a quadratic form. The Clifford algebra of *U*+*V* is isomorphic to the tensor product of the Clifford algebras of *U* and $(-1)^{\dim(U)/2}dV$, which is the space *V* with its quadratic form multiplied by $(-1)^{\dim(U)/2}d$. Over the reals, this implies in particular that

$$Cl_{p+2,q}(\mathbb{R}) = M_2(\mathbb{R}) \otimes Cl_{q,p}(\mathbb{R})$$
$$Cl_{p+1,q+1}(\mathbb{R}) = M_2(\mathbb{R}) \otimes Cl_{p,q}(\mathbb{R})$$
$$Cl_{p,q+2}(\mathbb{R}) = \mathbb{H} \otimes Cl_{q,p}(\mathbb{R})$$

These formulas can be used to find the structure of all real Clifford algebras; see the classification of Clifford algebras.

The Clifford group Γ

In this section we assume that V is finite dimensional and the bilinear form of Q is non-singular.

The Clifford group Γ is defined to be the set of invertible elements x of the Clifford algebra such that

$$xv\alpha(x)^{-1} \in V$$

for all v in V. This formula also defines an action of the Clifford group on the vector space V that preserves the norm Q, and so gives a homomorphism from the Clifford group to the orthogonal group. The Clifford group contains all elements r of V of nonzero norm, and these act on V by the corresponding reflections that take v to $v - \langle v,r\rangle r/Q(r)$ (In characteristic 2 these are called orthogonal transvections rather than reflections.)

Many authors define the Clifford group slightly differently, by replacing the action $xv\alpha(x)^{-1}$ by xvx^{-1}. This produces the same Clifford group, but the action of the Clifford group on V is changed slightly: the action of the odd elements Γ^1 of the Clifford group is multiplied by an extra factor of −1. This action used here has several minor advantages: it is consistent with the usual superalgebra sign conventions, elements of V correspond to reflections, and in odd dimensions the map from the Clifford group to the orthogonal group is onto, and the kernel is no larger than K^*. Using the action $\alpha(x)vx^{-1}$ instead of $xv\alpha(x)^{-1}$ makes no difference: it produces the same Clifford group with the same action on V.

The Clifford group Γ is the disjoint union of two subsets Γ^0 and Γ^1, where Γ^i is the subset of elements of degree i. The subset Γ^0 is a subgroup of index 2 in Γ.

If V is finite dimensional with nondegenerate bilinear form then the Clifford group maps onto the orthogonal group of V and the kernel consists of the nonzero elements of the field K. This leads to exact sequences

$$1 \to K^* \to \Gamma \to O_V(K) \to 1,$$
$$1 \to K^* \to \Gamma^0 \to SO_V(K) \to 1.$$

In arbitrary characteristic, the **spinor norm** Q is defined on the Clifford group by

$$Q(x) = x^t x$$

It is a homomorphism from the Clifford group to the group K^* of non-zero elements of K. It coincides with the quadratic form Q of V when V is identified with a subspace of the Clifford algebra. Several authors define the spinor norm slightly differently, so that it differs from the one here by a factor of −1, 2, or −2 on Γ^1. The difference is not very important.

The nonzero elements of K have spinor norm in the group K^{*2} of squares of nonzero elements of the field K. So when V is finite dimensional and non-singular we get an induced map from the orthogonal group of V to the group K^*/K^{*2}, also called the spinor norm. The spinor norm of the reflection of a vector r has image $Q(r)$ in K^*/K^{*2}, and this property uniquely defines it on the orthogonal group. This gives exact sequences:

$$1 \to \{\pm 1\} \to Pin_V(K) \to O_V(K) \to K^*/K^{*2},$$
$$1 \to \{\pm 1\} \to Spin_V(K) \to SO_V(K) \to K^*/K^{*2}.$$

Note that in characteristic 2 the group {±1} has just one element.

Spin and Pin groups

In this section we assume that V is finite dimensional and its bilinear form is non-singular. (If K has characteristic 2 this implies that the dimension of V is even.)

The **Pin group** $Pin_V(K)$ is the subgroup of the Clifford group Γ of elements of spinor norm 1, and similarly the **Spin group** $Spin_V(K)$ is the subgroup of elements of Dickson invariant 0 in $Pin_V(K)$. When the characteristic is not 2, these are the elements of determinant 1. The Spin group usually has index 2 in the Pin group.

Recall from the previous section that there is a homomorphism from the Clifford group onto the orthogonal group. We define the special orthogonal group to be the image of Γ^0. If K does not have characteristic 2 this is just the group of elements of the orthogonal group of determinant 1. If *K* does have characteristic 2, then all elements of the orthogonal group have determinant 1, and the special orthogonal group is the set of elements of Dickson invariant 0.

There is a homomorphism from the Pin group to the orthogonal group. The image consists of the elements of spinor norm 1 ∈ K^*/K^{*2}. The kernel consists of the elements +1 and −1, and has order 2 unless *K* has characteristic 2. Similarly there is a homomorphism from the Spin group to the special orthogonal group of *V*.

In the common case when *V* is a positive or negative definite space over the reals, the spin group maps onto the special orthogonal group, and is simply connected when *V* has dimension at least 3. **Warning:** This is not true in general: if *V* is $R^{p,q}$ for *p* and *q* both at least 2 then the spin group is not simply connected and does not map onto the special orthogonal group. In this case the algebraic group $Spin_{p,q}$ is simply connected as an algebraic group, even though its group of real valued points $Spin_{p,q}(R)$ is not simply connected. This is a rather subtle point, which completely confused the authors of at least one standard book about spin groups.

Spinors

Suppose that *p+q=2n* is even. Then the Clifford algebra $C\ell_{p,q}$(**C**) is a matrix algebra, and so has a complex representation of dimension 2^n. By restricting to the group $Pin_{p,q}$(**R**) we get a complex representation of the Pin group of the same dimension, called the **spinor representation**. If we restrict this to the spin group $Spin_{p,q}$(**R**) then it splits as the sum of two **half spin representations** (or **Weyl representations**) of dimension 2^{n-1}.

If *p+q=2n+1* is odd then the Clifford algebra $C\ell_{p,q}$(**C**) is a sum of two matrix algebras, each of which has a representation of dimension 2^n, and these are also both representations of the Pin group $Pin_{p,q}$(**R**). On restriction to the spin group $Spin_{p,q}$(**R**) these become isomorphic, so the spin group has a complex **spinor representation** of dimension 2^n.

More generally, spinor groups and pin groups over any field have similar representations whose exact structure depends on the structure of the corresponding Clifford algebras: whenever a Clifford algebra has a factor that is a matrix algebra over some division algebra, we get a corresponding representation of the pin and spin groups over that division algebra. For examples over the reals see the article on spinors.

Applications

Differential geometry

One of the principal applications of the exterior algebra is in differential geometry where it is used to define the bundle of differential forms on a smooth manifold. In the case of a (pseudo-)Riemannian manifold, the tangent spaces come equipped with a natural quadratic form induced by the metric. Thus, one can define a Clifford bundle in analogy with the exterior bundle. This has a number of important applications in Riemannian geometry.

Physics

Clifford algebras have numerous important applications in physics. Physicists usually consider a Clifford algebra to be an algebra spanned by matrices $\gamma_1,\ldots,\gamma_n$ called Dirac matrices which have the property that

$$\gamma_i\gamma_j + \gamma_j\gamma_i = 2\eta_{ij}$$

where η is the matrix of a quadratic form of signature (*p*,*q*) — typically (1,3) when working in Minkowski space. These are exactly the defining relations for the Clifford algebra $Cl_{1,3}(C)$ (up to an unimportant factor of 2), which by the classification of Clifford algebras is isomorphic to the algebra of 4 by 4 complex matrices.

The Dirac matrices were first written down by Paul Dirac when he was trying to write a relativistic first-order wave equation for the electron, and give an explicit isomorphism from the Clifford algebra to the algebra of complex matrices. The result was used to define the Dirac equation. The entire Clifford algebra shows up in quantum field theory in the form of Dirac field bilinears.

Footnotes

1. Mathematicians who work with real Clifford algebras and prefer positive definite quadratic forms (especially those working in index theory) sometimes use a different choice of sign in the fundamental Clifford identity. That is, they take $v^2 = -Q(v)$. One must replace Q with $-Q$ in going from one convention to the other.
2. The opposite is true when uses the alternate (−) sign convention for Clifford algebras: it is the conjugate which is more important. In general, the meanings of conjugation and transpose are interchanged when passing from one sign convention to the other. For example, in the convention used here the inverse of a vector is given by $v^{-1} = v^t/Q(v)$ while in the (−) convention it is given by $v^{-1} = \bar{v}/Q(v)$.

References

- Carnahan, S. *Borcherds Seminar Notes, Uncut.* Week 5, "Spinors and Clifford Algebras".
- Lawson and Michelsohn, *Spin Geometry*, Princeton University Press. 1989. ISBN 0-691-08542-0. An advanced textbook on Clifford algebras and their applications to differential geometry.
- Lounesto, P., *Clifford Algebras and Spinors*, Cambridge University Press. 2001. ISBN 0-521-00551-5.
- Porteous, I., *Clifford Algebras and the Classical Groups*, Cambridge University Press. 1995. ISBN 0-521-55177-3.

External links

- Planetmath entry on Clifford algebras [1]
- A history of Clifford algebras [2] (unverified)

References

[1] http://planetmath.org/encyclopedia/CliffordAlgebra2.html

[2] http://web.archive.org/20010211185832/members.fortunecity.com/jonhays/clifhistory.htm

Further Abstract Algebra

Quaternions

The set of **Quaternions** is an algebraic structure first studied by the Irish mathematician William Rowan Hamilton, in the hopes of constructing a generalization for complex numbers. When first discovered, quaternions generated a lot of excitement among mathematicians and physicists alike, for it was hoped that quaternions would provide a "unified theory" of mechanics and electromagnetism. Although these hopes proved to be unfounded, quaternions are still considered interesting as well as useful mathematical entities.

Definition

A **Quaternion** is an ordered 4-tuple $q = (a, b, c, d)$, where $a, b, c, d \in \mathbb{R}$. A quaternion is often denoted as $q = a + bi + cj + dk$ (Observe the analogy with complex numbers). The set of all quaternions is denoted by $\mathbb{H}$.
It is straightforward to define component-wise addition and scalar multiplication on $\mathbb{H}$, making it a real vector space.

The rule for multiplication was a product of Hamilton's ingenuity. He discovered what are known as the **Bridge-stone Equations**:

$$i^2 = j^2 = k^2 = ijk = -1$$

From the above equations alone, it is possible to derive rules for the pairwise multiplication of i , j , and k :

$ij = k, jk = i, ki = j$ (positive cyclic permutations)

$ji = -k, kj = -i, ik = -j$ (negative cyclic permutations).

Using these, it is easy to define a general rule for multiplication of quaternions. Because quaternion multiplication is not commutative, $\mathbb{H}$ is not a field. However, every nonzero quaternion has a multiplicative inverse (see below), so the quaternions are an example of a non-commutative division ring. It is important to note that the non-commutative nature of quaternion multiplication makes it impossible to define the quotient p/q of two quaternions p and q unambiguously, as the quantities pq^{-1} and $q^{-1}p$ are generally different.
Like the more familiar complex numbers, the quaternions have a conjugation, often denoted by a superscript star: q^* . The conjugate of the quaternion $q = a + bi + cj + dk$ is $q^* = a - bi - cj - dk$. As is the case for the complex numbers, the product qq^* is always a positive real number equal to the sum of the squares of the quaternion's components. Using this fact, it is fairly easy to show that the multiplicative inverse of a general quaternion q is given by

$$q^{-1} = \frac{q^*}{qq^*}$$

where division is defined since qq^* is a scalar. Note that, unlike in the complex case, the conjugate q^* of a quaternion q can be written as a polynomial in q :

$$q^* = -\frac{1}{2}(q + iqi + jqj + kqk).$$

The quaternions are isomorphic to the Clifford algebra $C\ell_2(\mathbf{R})$ and the even subalgebra of $C\ell_3(\mathbf{R})$.

Pauli Spin Matrices

Quaternions are closely related to the Pauli spin matrices of Quantum Mechanics. The Pauli matrices are often denoted as

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

(Where i is the well known quantity $\sqrt{-1}$ of complex numbers)

The 2×2 identity matrix is sometimes taken as σ_0. It can be shown that S , the real linear span of the matrices σ_0, $i\sigma_1$, $i\sigma_2$and $i\sigma_3$, is isomorphic to the set of all quaternions, $\mathbb{H}$. For example, take the matrix product below:

$$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

Or, equivalently:

$$i\sigma_3 i\sigma_2 = i\sigma_1$$

All three of these matrices square to the negative of the identity matrix. If we take $1 = \sigma_0$, $i = i\sigma_3$, $j = i\sigma_2$, and $k = i\sigma_1$, it is easy to see that the span of the these four matrices is "the same as" (that is, isomorphic to) the set of quaternions $\mathbb{H}$.

Exercise

1. Using the Bridge-stone equations, explicitly state the rule of multiplication for general quaternions, that is, given $q_1 = a_1 + b_1 i + c_1 j + d_1 k$ and $q_2 = a_2 + b_2 i + c_2 j + d_2 k$, give the components of their product $q = q_1 q_2$

Reference

- E.T. Bell, *Men of Mathematics*, Simon & Schuster, Inc.
- The Wikipedia article on Pauli Spin Matrices

Category theory

Category theory is the study of *categories*, which are collections of objects and *morphisms* (or arrows), or from one object to another. It generalizes many common notions in Algebra, such as different kinds of products, the notion of kernel, etc. See Category Theory for additional information.

Definitions & Notations

Definition 1: A *category* $\mathcal{C}$ consists of

A set $O(\mathcal{C})$ of *objects*.

For any $X, Y \in \mathrm{Obj}(\mathcal{C})$, a set $\mathrm{Hom}(X, Y)$ of *morphisms* from X to Y.

These obey the following axioms:

There is a notion of *composition*. If $X, Y, Z \in O(\mathcal{C})$, $f \in \mathrm{Hom}(X, Y)$ and $g \in \mathrm{Hom}(Y, Z)$, then f and g are called a composable pair. Their composition is a morphism $g \circ f \in \mathrm{Hom}(X, Z)$.

Composition is associative. $f \circ (g \circ h) = (f \circ g) \circ h$ whenever the composition is defined.

For any object X, there is an identity morphism $\mathrm{id}_X \in \mathrm{Hom}(X, X)$ such that if Y, Z are objects, $f \in \mathrm{Hom}(X, Y)$ and $g \in \mathrm{Hom}(Z, X)$, then $\mathrm{id}_X \circ f = f$ and $g \circ \mathrm{id}_X = g$.

Definition 2: A morphism f has associated with it two functions dom and cod called *domain* and *codomain* respectively, such that $f \in \mathrm{Hom}(X, Y)$ if and only if $\mathrm{dom}\, f = X$ and $\mathrm{cod}\, f = Y$. Thus two morphisms f, g are composable if and only if $\mathrm{cod}\, f = \mathrm{dom}\, g$.

Remark 3: Unless confusion is possible, we will usually not specify which Hom-set a given morphism belongs to. Also, unless several categories are in play, we will usually not write $X \in O(\mathcal{C})$, but just " X is an object". We may write $X \xrightarrow{f} Y$ to implicitly indicate the Hom-set f belongs to. We may also omit the composition symbol, writing simply gf for $g \circ f$.

Basic Properties

Lemma 4: Let X be an object of a category. The the identity morphism for X is unique.

Proof: Assume i and j are identity morphisms for X . Then $i = i \circ j = j$.

Example 5: We present some of the simplest categories:

i) $\mathbf{0}$ is the *empty category*, with no objects and no morphisms.

ii) $\mathbf{1}$ is the category containging only a single object and its identity morphism. This is the *trivial category*.

iii) $\mathbf{2}$ is the category with two objects, X and Y, their identity morphisms, and a single morphism $f \in \mathrm{Hom}(X, Y)$.

iv) We can also have a category like $\mathbf{2}$, but where we have two morphisms $f, g \in \mathrm{Hom}(X, Y)$ with $f \neq g$. Then f and g are called *parallel morphisms*.

v) $\mathbf{3}$ is the category with three objects X, Y, Z . We have $f \in \mathrm{Hom}(X, Y)$, $g \in \mathrm{Hom}(Y, Z)$ and $h = gf \in \mathrm{Hom}(X, Z)$.

Some examples of categories

- $\mathbf{Set}$: the category whose objects are sets and whose morphisms are maps between sets.
- $\mathbf{FinSet}$: the category whose objects are finite sets and whose morphisms are maps between finite sets.
- The category whose objects are open subsets of $\mathbb{R}^n$ and whose morphisms are continuous (differentiable, smooth) maps between them.
- The category whose objects are smooth (differentiable, topological) manifolds and whose morphisms are smooth (differentiable, continuous) maps.
- Let k be a field. Then we can define $k-\mathbf{Vect}$: the category whose objects are vector spaces over k and whose morphisms are linear maps between vector spaces over k .
- $\mathbf{Group}$: the category whose objects are groups and whose morphisms are homomorphisms between groups.

In all the examples given thus far, the objects have been sets with the morphisms given by set maps between them. This is not always the case. There are some categories where this is not possible, and others where the category doesn't naturally appear in this way. For example:

- Let $\mathcal{G}$ be any category. Then its opposite category $\mathcal{G}^{op}$ is a category with the same objects, and all the arrows reversed. More formally, a morphism in $\mathcal{G}^{op}$ from an object X to Y is a morphism from Y to X in $\mathcal{G}$.
- Let M be any monoid. Then we can define a category with a single object, with morphisms from that object to itself given by elements of M with composition given by multiplication in M .
- Let G be any group. Then we can define a category with a single object, with morphisms from that object to itself given by elements of G with composition given by multiplication in G .
- Let $\mathcal{C}$ be any small category, and let $\mathcal{D}$ be any category. Then we can define a category $\mathcal{D}^{\mathcal{C}}$ whose objects are functors from $\mathcal{C}$ to $\mathcal{D}$ and whose morphisms are natural transformations between the functors from $\mathcal{C}$ to $\mathcal{D}$.
- $\mathbf{Cat}$: the category whose objects are small categories and whose morphisms are functors between small categories.

Lattice theory

A *lattice* is a *poset* such that each pair of elements has a unique *least upper bound* and a unique *greatest lower bound*.

Article Sources and Contributors

Sets *Source*: http://en.wikibooks.org/w/index.php?oldid=2482158 *Contributors*: Adrignola, Barsamin, Chrislmorin, Espen180, Old Father Time, Panic2k4, Vorziblix, 5 anonymous edits

Equivalence relations and congruence classes *Source*: http://en.wikibooks.org/w/index.php?oldid=2579524 *Contributors*: Barsamin, Derbeth, Espen180, Jguk, MikeBorkowski, Panic2k4, Sigma 7, Thenub314, Vorziblix, 19 anonymous edits

Functions *Source*: http://en.wikibooks.org/w/index.php?oldid=2581070 *Contributors*: Adrignola, Barsamin, Bomazi, Espen180, Mathmensch, Panic2k4, TheKing44, 5 anonymous edits

Binary Operations *Source*: http://en.wikibooks.org/w/index.php?oldid=2542963 *Contributors*: Barsamin, Espen180, Panic2k4, Reekdeb Bhunia, TheKing44, 4 anonymous edits

Linear Algebra *Source*: http://en.wikibooks.org/w/index.php?oldid=2469433 *Contributors*: Debutante, Espen180, Richardhp, Swift, 1 anonymous edits

Number Theory *Source*: http://en.wikibooks.org/w/index.php?oldid=2506986 *Contributors*: Greenbreen, Panic2k4, TheKing44, Vorziblix, 1 anonymous edits

Group *Source*: http://en.wikibooks.org/w/index.php?oldid=2586412 *Contributors*: Arydye001, Chrislmorin, Espen180, Josef Sniatecki, Mathmensch, Mjf19, TheKing44, 6 anonymous edits

Subgroup *Source*: http://en.wikibooks.org/w/index.php?oldid=2543646 *Contributors*: Arydye001, Espen180, 2 anonymous edits

Cyclic groups *Source*: http://en.wikibooks.org/w/index.php?oldid=2580783 *Contributors*: Espen180, Mathmensch, QuiteUnusual, Sigma 7

Permutation groups *Source*: http://en.wikibooks.org/w/index.php?oldid=2580796 *Contributors*: Espen180, Mathmensch, 2 anonymous edits

Homomorphism *Source*: http://en.wikibooks.org/w/index.php?oldid=2581036 *Contributors*: Abramsky, Arydye001, Espen180, Mathmensch, Thenub314, 4 anonymous edits

Normal subgroups and Quotient groups *Source*: http://en.wikibooks.org/w/index.php?oldid=2581066 *Contributors*: Espen180, Mathmensch, 4 anonymous edits

Products and Free Groups *Source*: http://en.wikibooks.org/w/index.php?oldid=2453066 *Contributors*: Alexjbest, Espen180

Group actions on sets *Source*: http://en.wikibooks.org/w/index.php?oldid=2586673 *Contributors*: Espen180, Mathmensch

Rings *Source*: http://en.wikibooks.org/w/index.php?oldid=2582884 *Contributors*: Avicennasis, Badanedwa, Barnaby dawson, Dysprosia, Espen180, Hyad, Jguk, Mathmensch, Ndriscoll, Panic2k4, Remag12, Rschwieb, Sigma 7, The Scarlet Letter, Thenub314, 26 anonymous edits

Ring Homomorphisms *Source*: http://en.wikibooks.org/w/index.php?oldid=2582902 *Contributors*: Az1568, ComplexZeta, Jguk, Mathmensch, Ndriscoll, Panic2k4, The Scarlet Letter, Thenub314, 1 anonymous edits

Ideals *Source*: http://en.wikibooks.org/w/index.php?oldid=2584780 *Contributors*: Ashsong, Avicennasis, Drufino, Jguk, Mathmensch, Panic2k4, Remag12, RrsunN, The Scarlet Letter, Thebhgg, Thenub314, Xodarap00, 41 anonymous edits

Integral domains *Source*: http://en.wikibooks.org/w/index.php?oldid=2584751 *Contributors*: Alsocal, Ashsong, Jguk, Mathmensch, Panic2k4, The Scarlet Letter, Thenub314, 17 anonymous edits

Fraction Fields *Source*: http://en.wikibooks.org/w/index.php?oldid=2587449 *Contributors*: Mathmensch, Thenub314, Xxagile, 8 anonymous edits

Polynomial Rings *Source*: http://en.wikibooks.org/w/index.php?oldid=1598961 *Contributors*: Thenub314, Xxagile, 1 anonymous edits

Fields *Source*: http://en.wikibooks.org/w/index.php?oldid=2305095 *Contributors*: Crl0x7c2, HiYouAll11235, Jacinto Paramo, Jguk, Panic2k4, The Scarlet Letter, Thenub314, Tom.kemp90, Uffejensen, 10 anonymous edits

Splitting Fields and Algebraic Closues *Source*: http://en.wikibooks.org/w/index.php?oldid=2323914 *Contributors*: Adrignola, Jkagarise01, 3 anonymous edits

Vector Spaces *Source*: http://en.wikibooks.org/w/index.php?oldid=2023759 *Contributors*: Adakkak, Thenub314, Xxagile, 2 anonymous edits

Modules *Source*: http://en.wikibooks.org/w/index.php?oldid=2497864 *Contributors*: Ahhoefel, Avicennasis, ComplexZeta, Espen180, Jguk, Panic2k4, Remag12, Samsita, Thenub314, Wikimi-dhiann, 5 anonymous edits

Hypercomplex numbers *Source*: http://en.wikibooks.org/w/index.php?oldid=2280946 *Contributors*: Emperorbma, Jguk, Thenub314, 4 anonymous edits

Matroids *Source*: http://en.wikibooks.org/w/index.php?oldid=1599000 *Contributors*: Jguk, Sigma 7, Thenub314, 4 anonymous edits

Algebras *Source*: http://en.wikibooks.org/w/index.php?oldid=2400869 *Contributors*: Espen180, Ezrakilty, Jguk, Thenub314, 2 anonymous edits

Boolean algebra *Source*: http://en.wikibooks.org/w/index.php?oldid=2106730 *Contributors*: QuiteUnusual, ShadowsAssociate

Clifford Algebras *Source*: http://en.wikibooks.org/w/index.php?oldid=2580387 *Contributors*: Adrignola, ChenJunZhou, CommonsDelinker, Fishpi, Gro-Tsen, Panic2k4, Rotlink, Sigma 7, Thenub314

Quaternions *Source*: http://en.wikibooks.org/w/index.php?oldid=2307762 *Contributors*: Erigena, Nejssor, Panic2k4, Rcamp004, SPat, Thenub314, Xxagile, 4 anonymous edits

Category theory *Source*: http://en.wikibooks.org/w/index.php?oldid=2421129 *Contributors*: Adrignola, Dysprosia, Elroch, Espen180, Hoogli, J Bytheway, Jguk, Ndave, Sigma 7, Thenub314, 21 anonymous edits

Lattice theory *Source*: http://en.wikibooks.org/w/index.php?oldid=2305100 *Contributors*: Alsocal, Jguk, Panic2k4, Sigma 7, Thenub314, 4 anonymous edits

Image Sources, Licenses and Contributors

File:Hexagon Reflections.png *Source*: http://en.wikibooks.org/w/index.php?title=File:Hexagon_Reflections.png *License*: Public Domain *Contributors*: Grafite, Incnis Mrsi, Mate2code, Nbarth

File:Dihedral8.png *Source*: http://en.wikibooks.org/w/index.php?title=File:Dihedral8.png *License*: Public Domain *Contributors*: Jim.belk

File:Diagramkernel.png *Source*: http://en.wikibooks.org/w/index.php?title=File:Diagramkernel.png *License*: Creative Commons Zero *Contributors*: Espen180

File:Diagramquotient.png *Source*: http://en.wikibooks.org/w/index.php?title=File:Diagramquotient.png *License*: Creative Commons Attribution-Sharealike 3.0 *Contributors*: Espen180

File:FirstIsoThm.png *Source*: http://en.wikibooks.org/w/index.php?title=File:FirstIsoThm.png *License*: Creative Commons Attribution-Sharealike 3.0 *Contributors*: Espen180

File:ProdGrp.png *Source*: http://en.wikibooks.org/w/index.php?title=File:ProdGrp.png *License*: Creative Commons Attribution-Sharealike 3.0 *Contributors*: Espen180

Image:CliffordAlgebra-01.png *Source*: http://en.wikibooks.org/w/index.php?title=File:CliffordAlgebra-01.png *License*: Public Domain *Contributors*: Original uploader was Fropuff at en.wikipedia

License