Binary Operations, Monoids, and Groups

W. E. Aitken

September 2022 Edition*

This is one in a planned series of documents that survey the foundations of algebra. The documents of this series aim to cover the same sort of topics one would see in an introductory abstract algebra class, but to explore these topics perhaps a bit more abstractly and in more depth than one would necessarily want in a first course. So this series should be thought of as *foundational* but not introductory. As such the target audience is a reader who has already had some exposure to abstract algebra, but wishes to explore or review the foundations of the subject.

This document covers ideas related to the concept of *binary operations*. This includes examples, various properties (commutative, associative) that binary operations can have, the ideas of identity and inverse, and so on. This leads to the definitions of monoids and groups. It also covers basic laws of exponents, the general associativity laws, and the general commutative law. The ideas related to binary operations are first step to algebra, and so this document can be thought of as the "ground floor" for modern algebra built on a solid (but elementary) set-theoretical base. As such it is not intended to cover group theory in any degree beyond an introduction to a few basic notions, and it does not cover ring theory at all.

1 Background and Notation

I assume familiarity with some basic set theory including the idea of a Cartesian product $A \times B$ of two sets A and B, basic ideas related to functions, the notion of a family, and the notion of a total ordering of a set. I will assume as known the basic number systems (see, for example, in my Number Systems text). These constitute the logical prerequisites, but, as mentioned above, some previous exposure to groups or rings, say, is also useful. Otherwise, some of this document might seem a bit abstract and unmotivated. I use matrices in some examples as well, and the notion of continuous and differentiable functions come up in a few exercises, but these concepts are not central to the development of the theory.

In particular, as part of the set-theoretic background and notion of function, I use the concept of domain and codomain of a function, and use the notation $A \to B$ to indicate that the function f has domain A and codomain B (we sometimes say

that $A \to B$ is the *type* of f). When it comes time to define exponentiation, I will take as established the concept and basic properties of iteration of functions of the form $f\colon S \to S$. Such properties are established in my Number Systems text.

I will use $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ for the set of natural numbers, the set of integers, the set of rational numbers, the set of real numbers, and the set of complex numbers respectively. We have the inclusion $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R} \subsetneq \mathbb{C}$. There are two definitions of $\mathbb{N}$ that are fairly common. I use the version where $\mathbb{N}$ includes 0 as an element:

$$\mathbb{N} = \{0, 1, 2, \ldots\}.$$

The other popular version of the natural numbers excludes 0 and starts at 1; I will write this set as $\mathbb{Z}_+$. The set $\mathbb{Z}$ differs from $\mathbb{N}$ by including negative integers.

2 Binary Operations

Definition 1. Let S be a set. A *binary operation* on S is just a function $S \times S \to S$.

Example 1. Let $S = \mathbb{R}$. Multiplication $\times\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a binary operation since it takes as input two real numbers (thought of as an ordered pair) and outputs a real number. Addition and subtraction also give binary operations on $\mathbb{R}$, but division does not.

Let $S = \mathbb{N}$. Multiplication and addition are binary operations on $\mathbb{N}$. (Subtraction and division are not). There are many, many other examples. For example, the function $f\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined by the rule $f(x, y) = 17$ is a binary operation. This operation ignores the inputs x and y and always outputs the constant value of 17. Such constant operations are not as interesting as the usual operations in arithmetic, but they do satisfy the official definition.

Example 2. Let $X_n = \{1, \ldots, n\}$. Let $S = \mathcal{F}(X_n)$ be the set of functions $X_n \to X_n$. Thus S has n^n elements. Then composition is a binary operation. To see this, suppose $g, f \in S$. In other words, f, g are both functions $X_n \to X_n$. Then the composition $g \circ f$ is also a function $X_n \to X_n$. In other words, if $g, f \in S$ then also $g \circ f \in S$. So composition $\circ$ is a binary operation on $S = \mathcal{F}(X_n)$.

Let S_n be the set of bijective functions $X_n \to X_n$. We also call such functions *permutations*. If $g, f \in S_n$ then $g \circ f$ is also a bijection, so is in S_n. Thus composition is a binary operation

$$S_n \times S_n \to S_n.$$

Note that S_n has $n!$ elements.

Example 3. Let $M_{m,n}(\mathbb{R})$ be the set of real matrices with m rows and n columns. For example, two elements of $M_{2,3}(\mathbb{R})$ are

$$\begin{pmatrix} 1 & 0 & \sqrt{2} \\ -1 & \frac{1}{2} & 11 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1 & e & \pi \\ 2 & 2e & 2\pi \end{pmatrix}.$$

Then matrix addition $+$ is a binary operation on $M_{m,n}(\mathbb{R})$. When $m = n$ we write $M_n(\mathbb{R})$ instead of $M_{m,n}(\mathbb{R})$. Matrix multiplication is a binary operation on $M_n(\mathbb{R})$.

Remark. There are two possible notations we can use with binary operations: (i) ordinary functional notation or (ii) infix notation. Infix notation is the standard notation for binary operations, but it is sometimes illuminating to use functional notation from time to time.

First we describe *ordinary functional notation.* If $f\colon S \times S \to S$ is a binary operation, and if $a, b \in S$, then in ordinary functional notation we write

$$f(a, b)$$

for the result or output of the function. The element $f(a, b) \in S$ is often called the *value.* (Ordinary functional notion is sometimes called "prefix" notation, and sometimes we write $f\,a\,b$ as a variant for $f(a, b)$ when using this notation.)

For *infix* notation we put the function name in between the two inputs. Typically the function name is not a letter, but a symbol such as $*$, $+$, or $\times$. Suppose, for example, that $*\colon S \times S \to S$ is a binary operation, and that $a, b \in S$. Then instead of writing the value as $*(a, b)$ or $*\,a\,b$, in infix notation we write

$$a * b$$

where we put $*$ between the two inputs. In infix notation we require parenthesis in what is otherwise an ambiguous expression. For example, $a * b * c$ could be interpreted as $(a*b)*c$ or $a*(b*c)$. If we have established grouping conventions we can leave the parentheses off and let the reader mentally supply them. For instance the infix expression $x + y \cdot z$ in various numbers systems (or rings in general) is interpreted as $x + (y \cdot z)$, not as $(x + y) \cdot z$.

We use infix notation in the next definition:

Definition 2. Let $*\colon S \times S \to S$ be a binary operation. We say that this operation is *commutative* if

$$a * b = b * a$$

for all $a, b \in S$.

Remark. To show that $*\colon S \times S \to S$ is *not* commutative, you just need to find a specific counterexample. For example, you probably suspect that subtraction

$$-\colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$$

is not commutative. To prove it you need to give a specific counterexample. We choose 1 and 0 as our inputs in $\mathbb{Z}$. This is a counterexample since $1 - 0 \neq 0 - 1$. (Note that choosing 1 for both does not give a counterexample, since $1 - 1 = 1 - 1$. So not every choice of $a, b \in S$ will give you a counterexample).

Remark. If we use ordinary functional notation instead of infix notation for a binary operation $f\colon S \times S \to S$, the definition of commutative can be written as requiring

$$f(a, b) = f(b, a)$$

for all $a, b \in S$. An example of such a commutative binary operation (with ordinary functional notation) would be the function defined by the equation

$$f(x, y) = \sin(x^2 y^2)$$

where $S = \mathbb{R}$. On the other hand the function $g(x, y) = \sin(xy^2)$ does not look commutative because of the asymmetry between x and y in the formula. But to prove it is not commutative you need a specific counterexample. (In this case it is not difficult to find such a counterexample).

Remark. Suppose S_n is the set of permutations $X_n \to X_n$. Then the binary operation of composition $\psi : S_n \times S_n \to S_n$ applied to (g, f) can be written in two ways: as $\psi(g, f)$ or as $g \circ f$. The first is ordinary functional notation The second is infix notation using the symbol $\circ$ to denote the (infix version) of the operation.

2.1 Closure

Sometimes people describe a binary operation as "closed" to indicate that if $a, b \in S$ then $a * b \in S$. For us this term is redundant for a binary operation since $*$ is required to be a function $S \times S \to S$, so if $a, b \in S$ then $a * b$ must always be in the codomain S.

But the concept is useful if we are in a slightly different situation. Suppose you have a function mapping ordered pairs in S to values in some set T. In other words, suppose that we have a function $*\colon S \times S \to T$. Suppose also that S is a subset of T. We say this function is *closed* on S if $a * b \in S$ for all $a, b \in S$. In other words, for any two inputs in S the value "lands in" S and never at a value which is in T but outside S. When we are in this situation, we can restricted the codomain to S and then we would have a true binary operation.[1]

Informally it is common for people to says that "binary operations must be closed", but as mentioned above closure build into the definition of binary operation, and not an extra condition that needs to be imposed on binary operations. However, when you define a binary operation $S \times S \to S$, you do need to check that the values are always in S, and you may call this "checking for closure" if you would like. This is part of making sure the definition is well-defined, and often this check is left (tacitely) to the reader.

The concept of closed becomes more important when we work with submonoids and subgroups. Suppose $*\colon S \times S \to S$ is a binary operation on S and that A is a subset of S. We say that $*$ *is closed on* A if $a * b \in A$ for all $a, b \in A$. If that holds we get a binary operation $A \times A \to A$ by restricting the domain and codomain of the original operation.

Example 4. Subtraction $\mathbb{N} \times \mathbb{N} \to \mathbb{Z}$ is not closed since the image of the function contains negative integers which are outside $\mathbb{N}$. So subtraction cannot be thought of as a binary operation on $\mathbb{N}$. Subtraction is a binary operation on $\mathbb{Z}$, though.

[1] In mathematics, restriction of codomain is often done tacitly, or without much fanfare, since the restricted operation is in some sense "the same" and is represented by the same collection of ordered pairs. But the "type" of the operation is changed since the codomain is changed. More often authors will be more explicit about restriction of domain, but even this more standard type of restriction is sometimes done tacitly, even though the "type" of the function has changed.

3 Associativity

Suppose that $*\colon S \times S \to S$ is a binary operation. An expression such as

$$a * b * c$$

that involves more than two elements is ambiguous since a binary operation combines only two elements at a time. To avoid ambiguity, you need parentheses to show what two elements are being combined. For example $a*b*c$ can be interpreted as $(a * b) * c$, which really means $d * c$ where $d = a * b$. On the other hand, $a * b * c$ can be interpreted as $a * (b * c)$, which really means $a * d$ where $d = b * c$. It turns out that many important operations are associative, and the two interpretations of $a * b * c$ give equal values.

Remark. For general operations, that may or may not be associative, a common convention is to group from left to right. So $a * b * c * d$ would be interpreted as

$$((a * b) * c) * d.$$

For example, subtraction is not associative (for $\mathbb{Z}$), so what does

$$3 - 4 - (-10) - 5$$

mean? According to the left-to-right convention, it is equal to 4. The grouping convention is very important for non-associative operations such as subtraction. For associative operations it doesn't really make a difference since the value is the same.

Definition 3. Suppose $*\colon S \times S \to S$ is a binary operation. Then $*$ is said to be *associative* if $(a * b) * c = a * (b * c)$ for all $a, b, c \in S$.

Remark. If we use ordinary functional notation, and if $f\colon S \times S \to S$ is the binary operation, then associativity means

$$f(f(a, b), c) = f(a, f(b, c))$$

for all $a, b, c \in S$.

4 Operation Tables

Suppose $*\colon S \times S \to T$ is given where S is a finite set. We can describe the operation fully by giving a square table giving all the values of $*$. This involves (1) ordering the set S, (2) making a row and column for each element of S in the order from step 1, (3) for each $a, b \in S$ writing the value $a * b$ in the a-row and b-column.

If all elements of the table are in S then we conclude that $*$ is closed and so (by choosing $T = S$), the table can be regarded as describing a binary operation

$$S \times S \to S.$$

Exercise 1. Given an operation table for addition modulo 4 for the set

$$S = \{0, 1, 2, 3\}.$$

Exercise 2. Given an operation table for multiplication modulo 7 for the set

$$S = \{1, 2, 3\}.$$

is this a binary operation (is it closed)?

Exercise 3. Given an operation table for multiplication modulo 7 for the set

$$S = \{1, 2, 4\}.$$

is this a binary operation (is it closed)?

Exercise 4. What has to be true about the operation tables for a binary operation to be commutative?

Remark. There is no quick criteria that I know of to decide if an operation is associative by just looking at the operation table.

Exercise 5. Suppose $S = \{T, F\}$ be a set with two distinct elements called "true" and "false". This set is important in logic. Define the "and" the "or" and the "exclusive or" binary operations. Find the tables for these operators.

Exercise 6. Give a table for the composition operation on S_3. (Use cycle notation if you are familiar with this).

5 Identity Elements

Definition 4. Let $*\colon S \times S \to S$ be a binary operation. An *identity element* for $*$ is an element $e \in S$ such that

$$e * a = a \qquad \text{and} \qquad a * e = a$$

for all $a \in S$.

Exercise 7. Give examples of identity elements. Not all binary operations have identity elements. Can you give examples of this as well?

Exercise 8. What are the identity elements (if any) for the logical operations of Exercise 5.

Theorem 1. *Let $*\colon S \times S \to S$ be a binary operation. If there is an identity for $*$, then it is unique.*

Proof. Suppose that $e_1, e_2 \in S$ are both identities and consider $e_1 * e_2$. Since e_1 is an identity, we have

$$e_1 * e_2 = e_2.$$

Since e_2 is an identity, we have

$$e_1 * e_2 = e_1.$$

Thus

$$e_1 = e_2.$$

□

Remark. From now on if we know that an element is an identity for a given binary operation, then we can call this element *the* identity because of the above uniqueness theorem.

Identities are unique, but left or right identities are not necessarily unique. We begin with the definition of left and right identities.

Definition 5. Let $*\colon S \times S \to S$ be a binary operation. A *left identity element* for $*$ is an element $e \in S$ such that

$$e * a = a$$

for all $a \in S$.

Definition 6. Let $*\colon S \times S \to S$ be a binary operation. A *right identity element* for $*$ is an element $e \in S$ such that

$$a * e = a$$

for all $a \in S$.

Remark. Observe that, by definition, every identity element is a left identity and a right identity. In fact our definitions make it clear that an element is an identity if and only if it both a left-identity and a right identity.

Exercise 9. Suppose that $*\colon S \times S \to S$ is given by a operation table. Suppose that $e_1 \in S$. What has to be true about the e_1 row in order for e_1 to be a left identity? Suppose $e_2 \in S$. What has to be true about the e_2 column for e_2 to be a right identity?

What has to be true about rows and columns for e to be a (two-sided) identity?

Exercise 10. Define a binary operation with an operation table so that the operation has exactly three distinct left-identities. This shows that left identities are not unique. Does your operation have any right identities? Is it commutative?

Exercise 11. Suppose that $*\colon S \times S \to S$ is commutative. Show that every left identity is automatically an identity (so left identities are unique in this case since identities are unique).

As seen above, it is possible in a noncommutative binary operation to have more than one left identity. In this case there will be no right identities. In fact, the following theorem gives uniqueness whenever there exists *both* left and right identities (even if at first we do not know if the right and left identities are the same element).

Theorem 2. *Let $*\colon S \times S \to S$ be a binary operation. Suppose $e_1 \in S$ is a left identity and $e_2 \in S$ is a right identity. Then $e_1 = e_2$. In particular, e_1 and e_2 are the unique identity for $*$.*

Proof. Use the same type of argument as used for Theorem 1. (In fact, we could have started with the above theorem, and then made Theorem 1 a corollary). □

6 Monoids

Definition 7. A *monoid* M is a set with a given choice of operation $*\colon M \times M \to M$ such that (1) $*$ is associative, and (2) $*$ has an identity element in M.

Definition 8. A *commutative monoid* M is a monoid whose binary operation is commutative.

Example 5. We have many examples of monoids:

- The natural numbers $\mathbb{N}$ is a commutative monoid under addition. In this case the unique identity is 0. (In this document, we accept 0 as a natural number.)
- The natural numbers $\mathbb{N}$ is a commutative monoid under multiplication. In this case the unique identity is 1.
- The set of functions $\mathcal{F}(X_n)$ is a monoid under composition. What is the identity?
- The set of permutations S_n is a monoid under composition. What is the identity?
- The integers $\mathbb{Z}$ form a commutative monoid under addition. What is the identity?
- The integers $\mathbb{Z}$ form a commutative monoid under multiplication. What is the identity?
- The set $\mathbb{Q}$ is a commutative monoid under addition. What is the identity?
- The set $\mathbb{Q}$ is a commutative monoid under multiplication. What is the identity?
- The set $\mathbb{R}$ is a commutative monoid under addition. What is the identity?
- The set $\mathbb{R}$ is a commutative monoid under multiplication. What is the identity?
- The set $M_{n,m}(\mathbb{R})$ of commutative matrices is a commutative monoid under addition. What is the identity?
- The set $M_n(\mathbb{R})$ of matrices is a monoid under multiplication. What is the identity?
- The set $\mathbb{C}$ is a commutative monoid under addition. What is the identity?
- The set $\mathbb{C}$ is a commutative monoid under multiplication. What is the identity?
- The set $\{T, F\}$ is a commutative monoid under the "and" operation. What is the identity?
- The set $\{T, F\}$ is a commutative monoid under the "or" operation. What is the identity?

- The set $\{T, F\}$ is a commutative monoid under the "exclusive or" operation. What is the identity?
- The set $\mathbb{Z}_n$ of integers modulo n is a commutative monoid under addition. What is the identity?
- The set $\mathbb{Z}_n$ of integers modulo n is a commutative monoid under multiplication. What is the identity?

Remark. Two common (infix) notations for monoids are *multiplicative notation* (using juxtaposition or some sort of product sign for the binary operation) and *additive notation* (using $+$ for the binary operation). In multiplicative notation 1 is a common notation for the unique identity. In additive notation 0 is the common notation for the identity element. Sometimes, for example in logic, we might want to avoid either of these standard conventions.

Remark. You may also see the term "semigroup". A *semigroup* S is a set with a given choice of operation $*\colon S \times S \to S$ such that $*$ is associative. Note that every monoid is a semigroup, but it is easy to find examples of semigroups that are not monoids.

Remark. A monoid has two pieces of data: (i) the underlying set M and (ii) the binary operation $*$ on M. We sometimes write this package (or "structure") as an ordered pair $\langle M, * \rangle$. If you change either component, you have a different monoid. For example $\langle \mathbb{Z}, + \rangle$ is distinct from $\langle \mathbb{N}, + \rangle$. Likewise, $\langle \mathbb{Z}, + \rangle$ is distinct from $\langle \mathbb{Z}, \cdot \rangle$.

If the operation on M is clear from context, we sometimes denote the monoid $\langle M, * \rangle$ simply as M. In other words, it is common to use the name of the underlying set as also the name for the monoid as well.

7 Inverses

Warning: we cannot talk about inverses until we first have an identity element. So in this section we will always assume we have an identity element; often we will work in a monoid where we also have associativity.

Definition 9. Let $*\colon M \times M \to M$ be a binary operation with an identity element $e \in M$. Let $a \in M$. We say that a *is invertible* if there is an element $b \in M$ such that

$$a * b = b * a = e.$$

Theorem 3. *Let M be a monoid with identity element $e \in M$. If $a \in M$ is invertible then there is a unique element $b \in M$ such that*

$$a * b = b * a = e.$$

Proof. Suppose there are two such elements $b_1, b_2 \in M$. Consider the resulting elements $b_1 * (a * b_2)$ and $(b_1 * a) * b_2$ in M. We have

$$b_1 * (a * b_2) = b_1 * e = b_1.$$

and

$$(b_1 * a) * b_2 = e * b_2 = b_2.$$

The associativity law now implies that $b_1 = b_2$. □

Definition 10. Let M be a monoid with identity element $e \in M$. Suppose $a \in M$ is invertible. Then the *inverse of* a is defined to be the unique element $b \in M$ such that

$$a * b = b * a = e.$$

In additive notation (where the operation is written $+$), we write $-a$ for the inverse of a. If we are not using additive notation, we usually write a^{-1} for the inverse of a.

Exercise 12. Suppose that $*\colon M \times M \to M$ is given by a operation table. What do you have to check about the table to verify that $b \in M$ is the inverse of $a \in M$?

Exercise 13. When talking about inverses, we usually want S to be a monoid (so that inverses are unique). But we can define the notion of *an* inverse, even if S is not a monoid: all we need is that S has an identity. However, inverses are not necessarily unique when S is not a monoid. Give an example of non-uniqueness of inverses. Hint: write an operation table for a set $S = \{e, a, b, c\}$ where e is an identity, but where b and c are are both inverses of a.

Exercise 14. There is a notion of left inverse and right inverse. In this case you do not always have uniqueness even in a monoid. Give the definition of left inverse. Give the definition of right inverse.

Exercise 15. Assume that $a \in M$ where M is a monoid. Show that if a has a left inverse, and that a has a right inverse, then the left inverse is equal to the right inverse and a is invertible.

Here are some basic laws about inverses. (The proofs are straightforward):

Theorem 4. *Suppose $e \in M$ is the identity element of a monoid. Then e is invertible, and its inverse is itself:*

$$e^{-1} = e.$$

Theorem 5. *Suppose $c \in M$ is an invertible element of a monoid. Then c^{-1} is invertible, and*

$$(c^{-1})^{-1} = c.$$

Corollary 6. *Suppose $c \in M$ is an invertible element of a monoid with $c^{-1} = d$. Then d is invertible with $d^{-1} = c$.*

Theorem 7. *Suppose $c, d \in M$ be invertible elements of a monoid. Then $c * d$ is invertible, and*

$$(c * d)^{-1} = d^{-1} * c^{-1}.$$

Definition 11. We sometimes call invertible element in a monoid *units*.

We can rephrase Theorems 4, 5, and 7 in terms of units:

Theorem 8. *Suppose M is a monoid. Then the identity element is a unit, the inverse of a unit is a unit, and the product of units is a unit (where here we are using multiplicative notion for convenience).*

8 Cancellation laws

Theorem 9 (Left-cancellation Law)**.** *Suppose M is a monoid and $c \in M$ is invertible. Suppose that*

$$c * x = c * y$$

with $x, y \in M$. Then

$$x = y.$$

Proof. Since $c * x = c * y$ we get the equation

$$c^{-1} * (c * x) = c^{-1} * (c * y).$$

By the associative law, we can rewrite this equation as

$$(c^{-1} * c) * x = (c^{-1} * c) * y.$$

By the definition of inverses, we can write this equation as

$$e * x = e * y.$$

where $e \in M$ is the identity. By the definition of identity, we get

$$x = y.$$

□

Similarly we have the following:

Theorem 10 (Right-cancellation Law)**.** *Suppose M is a monoid, and $c \in M$ is invertible. Suppose that*

$$x * c = y * c$$

with $x, y \in M$. Then

$$x = y.$$

Remark. For the left-cancellation law, it is enough that c is left-invertible. For the right-cancellation law, it is enough that c is right-invertible.

Example 6. In the monoid $\langle \mathbb{R}, + \rangle$, every element is invertible (it turns out to be a "group"). So the cancellation law for addition holds for all $c \in \mathbb{R}$:

$$c + x = c + y \implies x = y.$$

However, in the monoid $\langle \mathbb{R}, \times \rangle$, only the nonzero elements are invertible. So the cancellation law

$$cx = cy \implies x = y$$

only applies for $c \neq 0$.

9 Groups

Definition 12. A *group* is a monoid such that every element is invertible.

Definition 13. An *Abelian group* is a group whose operation is commutative. In other words, it is a commutative monoid such that every element has an inverse. We use the term "Abelian" in honor of the mathematician Abel (1802–1829).[2]

Example 7. Some examples:

- The set of natural numbers $\mathbb{N}$ is not a group under addition nor is it a group under multiplication.
- The monoid of integers $\mathbb{Z}$ under addition is an Abelian group, but the corresponding monoid $\mathbb{Z}$ under multiplication is not a group (since 2, for example, has no inverse).
- The set of permutations S_n is a group under composition. The inverse of a permutation $\alpha \in S_n$ is the inverse function α^{-1}. This group is non-Abelian if $n > 2$. (In contrast, the monoid of functions $\mathcal{F}(X_n)$ is not a group if $n > 1$).
- The set $M_{n,m}(\mathbb{R})$ of matrices is a group under addition. Describe inverses. Is it Abelian?
- The set $M_n(\mathbb{R})$ of matrices is not a group under multiplication. However, the subset $GL_n(\mathbb{R})$ of matrices with determinate not equal to zero is a group. The inverse of A is the matrix A^{-1}. If $n > 1$ then $GL_n(\mathbb{R})$ is a non-Abelian group.
- The set $\mathbb{R}$ is a group under addition, but not under multiplication. However, if we remove 0 the set that remains $\mathbb{R}^\times = \mathbb{R} - \{0\}$ is a group.
- The set $\mathbb{Q}$ is a group under addition, but not under multiplication. What element (or elements) of $\mathbb{Q}$ under multiplication do not have inverses? What if you remove just 0 from $\mathbb{Q}$?
- The set $\mathbb{C}$ is a group under addition, but not under multiplication. If you remove 0 from $\mathbb{C}$ you do get a group under multiplication.
- The set $\{T, F\}$ is a group under the "exclusive or" operation. Observe that the inverse of any element is itself. (Are there any other logical operators on $\{T, F\}$ for which we get a group?)
- The set $\mathbb{Z}_n$ of integers modulo n is a group under addition, but not under multiplication. Is it Abelian? Is there a connection between $\mathbb{Z}_2$ under addition and $\{T, F\}$ under exclusive or?
- Let p be a prime. Then the set $\mathbb{Z}_p^\times$ of integers modulo p not congruent to zero turns out to be an Abelian group under multiplication with $p - 1$ elements.

For groups we have the cancellation law:

Theorem 11. *Let $c \in G$ where G is a group. If $c * x = c * y$ with $x, y \in G$, then $x = y$. Similarly, if $x * c = y * c$ with $x, y \in G$, then $x = y$.*

[2] I capitalize "Abelian" to honor Abel, but many authors write "abelian" with a lower-case 'a'.

Proof. This follows from cancellation laws for monoids. □

Theorem 12. *If G is a group with identity element e then*

$$e^{-1} = e$$

For all $c \in G$

$$(c^{-1})^{-1} = c.$$

For all $c, d \in G$

$$(c * d)^{-1} = d^{-1} * c^{-1}.$$

If G is an Abelian group then for all $c, d \in G$

$$(c * d)^{-1} = c^{-1} * d^{-1}.$$

Proof. This follows from the inverse laws for monoids. □

Recall that if the operation of a monoid M is written additively using $+$, then it is customary to write the identity element as 0 and the inverse of $a \in M$ as $-a$, if it exists. We can rephrase various results using additive notation. For example, the above theorem can be stated as follows:

Corollary 13. *If G is a group with additive notation then*

$$-0 = 0$$

For all $c \in G$

$$-(-c) = c.$$

For all $c, d \in G$

$$-(c + d) = (-d) + (-c)$$

If G is an Abelian group then for all $c, d \in G$

$$-(c + d) = (-c) + (-d).$$

Remark. When we use additive notation we are usually working in an Abelian group or at least a commutative monoid.

Definition 14. Let G be a group with additive notation. Given $a, b \in G$ we define the subtraction operation as follows:

$$a - b \stackrel{\text{def}}{=} a + (-b).$$

Corollary 14. *If G is an Abelian group with additive notation then for all $c, d \in G$*

$$-(c + d) = (-c) - d.$$

Exercise 16. Suppose M is monoid such that every element has a left-inverse. Show that M is a group.

9.1 The One-Side Inverse Shortcut

According to the definition of inverse we must show both $a * b = e$ and $b * a = e$ in order to conclude that b is the inverse of a. If we are working in a group then this is overkill, and there is a short-cut: we only need to show one of the two equations. (Of course if $*$ is commutative one equation is enough; but the point here is that in a group G one equation is enough even if G is non-Abelian.)

Theorem 15. *Let G be a group with identity element e. If*

$$a * b = e$$

where $a, b \in G$ then

$$b * a = e$$

and so

$$b = a^{-1} \quad \textit{and} \quad a = b^{-1}.$$

Proof. Note that

$$a * b = e = a * a^{-1}.$$

By the cancellation law, we get

$$b = a^{-1}.$$

In particular,

$$b * a = a^{-1} * a = e$$

and so

$$a = b^{-1}.$$

□

10 Translation Functions

Think of $\mathbb{R}^2$ as the collection of vectors in the plane under the operation of vector addition. If you fix a vector (c_1, c_2) then the function

$$(x, y) \mapsto (c_1, c_2) + (x, y) = (c_1 + x, c_2 + y)$$

corresponds to the geometric idea of translation: each point in $\mathbb{R}^2$ maps to its translation by the vector (c_1, c_2). This operation is used to move points, and so subsets of $\mathbb{R}^2$, in a way that preserves lengths and angles and so on. In particular, this map is an example of what we call an "isometry".

We can generalize this translation map to any monoid $\langle M, * \rangle$. To do so fix an element $c \in M$ and define a function $T_c \colon M \to M$ by the rule

$$T_c(x) \stackrel{\text{def}}{=} c * x.$$

We call this *left-translation* by c. We define right-translation in a similar manner. Of course if the monoid is commutative, as in the case of $\mathbb{R}^2$ under addition, left and right translations by $c \in M$ agree. If unspecified, a *translation* will be understood as left-translation.

The following two results are straightforward:

Lemma 16. *Let e be the identity in the monoid M. Then the translation function T_e is the identity function $M \to M$.*

Lemma 17. *Let T_a and T_b be translation functions $M \to M$ associated to elements $a, b \in M$ where $\langle M, * \rangle$ is a monoid. Then the composition $T_a \circ T_b \colon M \to M$ is also a translation function. In fact*

$$T_a \circ T_b = T_{a*b}.$$

Exercise 17. Show that had we defined T_a and T_b as right-translations, we would get $T_a \circ T_b = T_{b*a}$ instead.

We can use the above two Lemmas to prove the following:

Lemma 18. *Suppose that $c \in M$ is an invertible element of a monoid M. Then the translation functions T_c and $T_{c^{-1}}$ are inverse functions to each other. In other words the compositions $T_c \circ T_{c^{-1}}$ and $T_{c^{-1}} \circ T_c$ are both the identity function. We can express this fact nicely with the identity*

$$T_c^{-1} = T_{c^{-1}}.$$

Corollary 19. *Suppose that $c \in M$ is an invertible element of a monoid M. Then the translation functions T_c defined by $T_c(x) = c * x$ is a bijection.*

We can apply this result to operation tables of finite monoids M:

Theorem 20. *Suppose $\langle M, * \rangle$ is a finite monoid with invertible element c. Suppose M is ordered once and for all, and consider the associated operation table for $*$. Then the row associated to c has every element of M appearing exactly once. Similarly the column associated to c has every element of M appearing exactly once.*

Proof. Suppose $a_1, a_2, \ldots, a_n$ are the elements of M listed in order. Then the row associated to c consists of the following finite sequence of elements:

$$c * a_1, c * a_2 \ldots, c * a_n.$$

Think of this sequence as

$$T_c(a_1), T_c(a_2), \ldots, T_c(a_n).$$

Every element appears at most once because T_c is injective. Every element appears at least once because T_c is surjective. The argument for the c-column is similar. □

11 Exponentiation in a Monoid

Throughout this section, let M be a monoid with operation $*$ and identity element e. For most of what follows we will employ multiplicative notation, allowing us to write $a * b$ simply as ab. Our goal is to introduce powers a^u where $n \in \mathbb{N}$ and, in fact, when a is invertible for all $n \in \mathbb{Z}$. (We translate all the main results to additive notation later on where we write na).

Let $a \in M$ be an element of the monoid. Informally, if $n \in \mathbb{N}$ then we define a^n to be

$$a * a * \cdots * a$$

where a is repeated n times. If $a \in M$ is invertible, we define a^{-n} for negative integers $-n < 0$ by defining a^{-n} either as $(a^{-1})^n$ or $(a^n)^{-1}$. If one follows this path, then one can derive the usual laws of exponentiation, even for negative powers (if a is invertible). However, proving them often requires a bunch of different cases which depends on the sign (positive or negative) of each of the exponents involved in the given law.

We take another approach to exponentiation using iteration of translation functions. This approach treats positive and negative powers in a more unified manner. Our plan is to iterate the translation function $x \mapsto ax$ starting with $x = e$ yielding the sequence

$$e \quad a * e \quad a * a * x \quad a * a * a * e \quad \ldots.$$

The nice thing about this approach is that we can use negative iteration when the translation function is invertible. In what follows I will take the concept of iteration, and some of the basic results for iterated functions, as part of the set-theoretical background for abstract algebra.[3] Here are three propositions summarizing what we need:

Proposition 21. *Suppose $f\colon S \to S$ is a function from a set S to itself. Then for any nonnegative integer $n \in \mathbb{N}$ we have the nth iteration f^n of f. Properties of iteration for such a function $f\colon S \to S$ include the following:*

- *For all $n \in \mathbb{N}$, the iteration f^n is a function $S \to S$.*
- *The zeroth iteration f^0 is the identity map $S \to S$.*
- *The first iteration f^1 is just $f\colon S \to S$ itself.*
- *For $m, n \in \mathbb{N}$, we have $f^m \circ f^n = f^{m+n}$. In particular, for $n \in \mathbb{N}$ we have $f^{n+1} = f^n \circ f = f \circ f^n$.*
- *For $m, n \in \mathbb{N}$, we have $(f^m)^n = f^{mn}$.*

Proposition 22. *Suppose $f\colon S \to S$ is a bijection from a set S to itself. Then we can extend the definition of interation f^u to all integers $u \in \mathbb{Z}$. Properties of iteration for such a bijection $f\colon S \to S$ include the following:*

- *For all $u \in \mathbb{Z}$, the iteration f^u is also a bijection $S \to S$.*
- *The iteration f^{-1} is, as the notation suggests, the inverse of f.*
- *For $u, v \in \mathbb{Z}$, we have $f^u \circ f^v = f^{u+v}$.*
- *For $u, v \in \mathbb{Z}$, we have $(f^u)^v = f^{uv}$.*
- *If f is the identity function, then f^u is the identity function for all $u \in \mathbb{Z}$.*

[3] See my number systems textbook for a formal account of iteration and a rigorous development of the properties of iteration.

Proposition 23. *Suppose $f, g\colon S \to S$ are commuting functions. In other words, suppose $f \circ g = g \circ f$. Then*

$$(f \circ g)^n = f^n \circ g^n$$

for all $n \in \mathbb{N}$. If in addition f and g are both bijections $S \to S$, then

$$(f \circ g)^u = f^u \circ g^u$$

for all $u \in \mathbb{Z}$.

With this we are ready for the formal definition of exponentiation. Recall that from Corollary 19 that the translation function T_a is a bijection if a is invertible.

Definition 15 (Exponentiation in a multiplicative monoid)**.** Let M be a monoid that employs multiplicative notation. If $n \in \mathbb{N}$ then

$$a^n \stackrel{\text{def}}{=} T_a^n(e).$$

where $e \in M$ is the identity element, and where T_a is the translation map $x \mapsto a*x$.

If a is invertible in M, then we extend the above definition for all $u \in \mathbb{Z}$:

$$a^u \stackrel{\text{def}}{=} T_a^u(e)$$

(This is well-defined since T_a is a bijection in this case).

Theorem 24. *Suppose that $a \in M$ where M is a multiplicative monoid. Then*

$$a^0 = e$$

where $e \in M$ is the identity element, and

$$a^1 = a.$$

If a is invertible in M then a^{-1} is the inverse of a in M.

Proof. Observe that $a^0 = T_a^0(e) = e$ since T_a^0 is the identity map (Proposition 21). Next, by Proposition 21 and the above definition,

$$a^1 = T_a^1(e) = T_a(e) = ae = a.$$

Finally, suppose a is invertible, and let b be the inverse of a in M. By Lemma 18, the translation T_b is the inverse of T_a, so $T_a^{-1} = T_b$ by Proposition 22 (where T_a^{-1} here mean the $-1 \in \mathbb{Z}$ iteration of T_a). Thus

$$a^{-1} = T_a^{-1}(e) = T_b(e) = be = b$$

where a^{-1} means the -1 power of a (not, a priori, the inverse of a which we denote simply as b). □

Remark. The last part of the above shows that both meanings of a^{-1} agree: as an inverse and as a negative power.

To prove additional properties of exponentiation, we will need a few lemmas about translations functions. The first states that the iteration of translation is a translation.

Lemma 25. *Consider the iteration T_a^u of the translation T_a where $a \in M$ and where $u \in \mathbb{N}$ or more generally, if a is invertible, where $u \in \mathbb{Z}$. Then there is an element $c \in M$ such that $T_a^n = T_c$.*

Proof. Fix $a \in M$. First we show that the claim holds when $u \in \mathbb{N}$ by induction. The base case holds since $T_a^0 = T_e$ where e is the idenitity. Suppose that $T_a^n = T_c$ for a particular $n \in \mathbb{N}$ and $c \in M$. Then

$$T_a^{n+1} = T_a^n \circ T_a = T_c \circ T_a = T_{ca}. \qquad \text{(Prop 21 and Lemma 17)}$$

Thus the claim holds for $n+1$ with ca as the associated element of M. By induction we accept the claim for $u \geq 0$.

We have established the claim for for all $a \in M$ and all $u \geq 0$, and wish to establish the claim for negative u. So let $a \in M$ be invertible with inverse $b \in M$, and suppose $u = -n$ for some $n \in \mathbb{N}$. We have $T_b = T_a^{-1}$ by Lemma 18. So, using the identity $(f^u)^v = f^{uv}$ from Proposition 22,

$$T_a^u = T_a^{-n} = \left(T_a^{-1}\right)^n = T_b^n.$$

We have established that $T_b^n = T_c$ for some $c \in M$. Thus $T_a^u = T_c$ as desired. □

We can actually be more specific about what c is in the previous lemma.

Lemma 26. *Suppose that $a \in M$. Then*

$$T_a^u = T_{a^u}$$

for all $u \in \mathbb{N}$. If a is invertible, then this holds more generally for all $u \in \mathbb{Z}$.

Proof. By Lemma 25, $T_a^u = T_c$ for some $c \in M$. Apply T_a^u and T_c to the identity element $e \in M$:

$$a^u \overset{\text{def}}{=} T_a^u(e) = T_c(e) = ce = c.$$

□

Now we are ready to prove some important identities.

Theorem 27. *Let $a \in M$ where M is a multiplicative monoid. If $m, n \in \mathbb{N}$ then*

$$a^{m+n} = a^m a^n.$$

More generally, if a is invertible in M and if $u, v \in \mathbb{Z}$ then

$$a^{u+v} = a^u a^v.$$

Proof. Observe that

$$\begin{aligned} T_{a^{m+n}} &= T_a^{m+n} \quad \text{(Lemma 26)} \\ &= T_a^m \circ T_a^n \quad \text{(Proposition 21)} \\ &= T_{a^m} \circ T_{a^n} \quad \text{(Lemma 26)} \\ &= T_{a^m a^n} \quad \text{(Lemma 17)} \end{aligned}$$

From $T_{a^{m+n}} = T_{a^m a^n}$ we get

$$a^{m+n} = T_{a^{m+n}}(e) = T_{a^m a^n}(e) = a^m a^n.$$

The above proof is for $m, n \in \mathbb{N}$. To generalize to $u, v \in \mathbb{Z}$, replace m with u, replace n with v, and replace the reference to Proposition 21 with Proposition 22 in the above argument. □

We are not assuming that M is a commutative monoid. However, we have the following:

Corollary 28. *Let $a \in M$. Then a power of a commutes with any power of a.*

Proof. This follows since addition is commutative in $\mathbb{Z}$. □

Theorem 29. *Let $a \in M$ where M is a multiplicative monoid. If $m, n \in \mathbb{N}$ then*

$$(a^m)^n = a^{mn}.$$

If $a \in M$ is invertible and if $u, v \in \mathbb{Z}$ then

$$(a^u)^v = a^{uv}.$$

Proof. We have $T_{a^m} = T_a^m$ (Lemma 26), so

$$(T_{a^m})^n = (T_a^m)^n = T_a^{mn}$$

where the last equality is based on a property of iterations (Proposition 21). This, together with Definition 15, yields

$$(a^m)^n = (T_{a^m})^n(e) = T_a^{mn}(e) = a^{mn}.$$

To generalize to $u, v \in \mathbb{Z}$, when a is invertible, replace m with u, n with v, and the reference to Proposition 21 with Proposition 22 in the above argument. □

Theorem 30. *Let $a \in M$ where M is a monoid. If a is invertible, then so is a^u for all $u \in \mathbb{Z}$.*

Proof. Observe that a^{-u} is the inverse of a^u by previous exponentiation laws. □

In a commutative monoid we have that $(ab)^n = a^n b^n$. We will show this, but first we need a simple lemma.

Lemma 31. *Let $a, b \in M$. Suppose $ab = ba$ (which holds, for example, when M is a commutative monoid). Then T_a and T_b commute as functions: $T_a \circ T_b = T_b \circ T_a$.*

Proof. This follows from Lemma 17 since $T_{ab} = T_{ba}$ in this case. □

Theorem 32. *Let $a, b \in M$ in a multiplicative monoid. If $ab = ba$ (which is true, for example, if M is a commutative monoid or an Abelian group) then for all $n \in \mathbb{N}$*

$$(ab)^n = a^n b^n.$$

If in addition a and b are invertible, then so is ab, and for all $u \in \mathbb{Z}$

$$(ab)^u = a^u b^u.$$

Proof. By Lemma 31, T_a and T_b commute. So

$$\begin{aligned} \left(T_{ab}\right)^n &= \left(T_a \circ T_b\right)^n \qquad \text{(Lemma 17)} \\ &= T_a^n \circ T_b^n \qquad \text{(Prop. 23)} \\ &= T_{a^n} \circ T_{b^n} \qquad \text{(Lemma 26)} \\ &= T_{a^n b^n} \qquad \text{(Lemma 17)} \end{aligned}$$

This, together with Definition 15, yields

$$(ab)^n \stackrel{\text{def}}{=} \left(T_{ab}\right)^n(e) = T_{a^n b^n}(e) = a^n b^n e = a^n b^n.$$

If a and b are invertible, then ab is invertible since it has inverse $b^{-1}a^{-1}$. The above argument for $n \in \mathbb{N}$ generalizes to $u \in \mathbb{Z}$ in this case. □

Theorem 33. *Let $e \in M$ be the identity element in a multiplicative monoid. Then $e^u = e$ for all $u \in \mathbb{Z}$.*

Proof. We know that T_e is the identity map by Lemma 16. So, by Proposition 22, T_e^u is also the identity map. Thus

$$e^u \stackrel{\text{def}}{=} T_e^u(e) = e.$$

□

11.1 Exponentiation in Additive Notation

When M is a monoid or group written in additive notation, we use a different notation for exponentiation. The results are the same, but are just expressed in a different notation.

When M is a monoid with its binary operation written as $+$ then the identity element is generally written as 0. For $n \in \mathbb{N}$ and $a \in M$ we write na for the additive version of exponentiation. Similarly if a is invertible and $u \in \mathbb{Z}$, we write ua for exponentiation. All the results of the previous section hold for M, but one just has to translate to additive notation. For the convenience of the reader, we summarize the main results when the monoid is a group.

Theorem 34 (Exponentiation in an additive group)**.** *Suppose A is a group written in additive notation. Suppose $a \in A$. Then we have the following*

- $0\,a = 0$. *(Here the first* 0 *is in* $\mathbb{Z}$ *and the second is in* G*).*
- $1\,a = a$
- $(-1)a = -a$
- $(u+v)a = ua + vb$ *for all* $u, v \in \mathbb{Z}$
- $(uv)a = u(va)$ *for all* $u, v \in \mathbb{Z}$
- ua *commutes with* va *for all* $u, v \in \mathbb{Z}$.

If $a, b \in G$ *commute,* $a + b = b + a$*, then*

- $u(a+b) = ua + ub$ *for all* $u \in \mathbb{Z}$

Finally,

- $u\,0 = 0$ *for all* $u \in \mathbb{Z}$ *(here* 0 *is in* G*)*

Remark. We use the normal elementary algebra notational conventions for supressing parenthesis. For example, the expression $ua + vb$ really means $(ua) + (vb)$ (and not $u(a + (vb))$ say).

Remark. Most groups that use additive notation are Abelian, so the law

$$u(a+b) = ua + ub$$

holds for all $u \in \mathbb{Z}$ and $a, b \in G$.

12 The Induced Operation on Closed Subsets

Definition 16. Let $*\colon S \times S \to S$ be a binary operation. If A is a subset of S then we say that A is *closed* under $*$ if the following holds:

$$\forall a, b \in A, \quad a * b \in A.$$

Example 8. Consider $\mathbb{N}$ under addition, and let A be the subset of even numbers greater than 11. Then A is closed under $+$. Let B be the set of odd natural numbers. Then B is not closed under $+$ since $7 + 3$ is not in B.

Definition 17. Let $*\colon S \times S \to S$ be a binary operation. Let $A \subseteq S$ be a subset that is close under $*$. Then we define a binary operation on $*_A\colon A \times A \to A$ using the following rule: if $a, b \in A$ then $a *_A b$ is defined to be $a * b$. This operation is called the *induced operation* on A.

We warn the reader that the term *induced operation* can refer to different things depending on context. The above is what is meant when A is a subset of S. There are also induced operation corresponding to sets of functions, or sets of cosets. The common idea is that an induced binary operation closely matches, in some sense, a previously given binary operation, but often has a different, albeit related, domain or codomain.

The induced operation of Definition 17 is often thought of as the operation obtained by restricting the operation $*$ from S to the subset A. The assumption that A is closed plays a key role here. It is always legal to restrict the domain of a function. So we can always restrict $*\colon S \times S \to S$ to $A \times A$ to obtain a function $A \times A \to S$. It is not always legal, however, to restrict a codomain. You can only restrict to a smaller codomain if the this proposed codomain contains the image. So we can only restrict from $A \times A \to S$ to $A \times A \to A$ if the image of the restriction $A \times A \to S$ is a subset of A. This is exactly the requirement that A is closed under $*$: when we say "A is closed" we just mean that the image of $*$ restricted to $A \times A$ is contained in A.

By definition we have the law

$$a *_A b = a * b$$

for all $a, b \in A$. So $*$ and $*_A$ look like the same operation; the only real difference is the domain and codomain. The bottom line is that, as long as $a, b \in A$, the expressions $a * b$ and $a *_A b$ are completely interchangeable.

The operation $*_A\colon A \times A \to A$ is sometime called the *restriction* of $*$ to A. We also say that the operation $S \times S \to S$ an *extension* of the operation $A \times A \to A$ to S. It is common practice to use the same symbol for both operations $S \times S \to S$ and $A \times A \to A$. When we do so, we rely on context to tell us which is meant by the symbol in any given usage. (When applied to elements in A, both notions give the same result, so one can be be safely ambiguous. For example, if $A = \mathbb{N}$ and $S = \mathbb{Z}$ then $7 + 3$ is 10 whether or not you are thinking of $+\colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ or the induced operation $+\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$.)

Lemma 35. *Let $*\colon S \times S \to S$ be a binary operation. Let $A \subseteq S$ be a subset that is close under $*$.*

If $*\colon S \times S \to S$ is associative, then the induced operation $A \times A \to A$ is also associative.

If $*\colon S \times S \to S$ is commutative, then the induced operation $A \times A \to A$ is also commutative.

Proof. We will prove the claim related to the associative law. The proof of the commutative claim is similar. So let $a, b, c \in A$. We have

$$(a *_A b) *_A c \stackrel{\text{def}}{=} (a * b) * c$$

by definition of $*_A$. Similarly, we have

$$a *_A (b *_A c) \stackrel{\text{def}}{=} a * (b * c).$$

Since $a, b, c \in A$, and since A is a subset of S, we have $a, b, c \in S$. Since $*$ is associative on S by assumption, we have

$$(a * b) * c = a * (b * c).$$

Putting these equations together we get the desired equation:

$$(a *_A b) *_A c = a *_A (b *_A c).$$

□

From this we get the following:

Theorem 36. *Suppose M is a monoid with operation $*$ and identity element e. If A is a subset closed under $*$ and if A contains e then A is itself a monoid under the induced operation. The identity of A with the induced operation is the identity of M.*

13 Submonoids

Informally a submonoid is a subset A of a monoid M that is itself a monoid under the induced operation. This will only make sense if the subset A is closed under the operation of M, so this will always be required. One question is whether we want the identity of A to be the same element as the identity of M. Based on the answer to this question we have two notions of submonoid. Here we will pursue the notion where M and A have the same identity element: this will be a second requirement.[4] As we saw in Theorem 36, these two requirements suffice. So we define submonoid as follows:

Definition 18. Suppose $\langle M, * \rangle$ is a monoid with identity element e. A *submonoid* is a subset $A \subseteq M$ such that (1) A is closed under $*$ and (2) $e \in A$. As we saw above, every submonoid is itself a monoid under the induced operation with identity e.

Example 9. The set of even integers forms a submonoid of $\langle \mathbb{Z}, + \rangle$. The set of odd integers fails to be a submonoid. If A is the set of even numbers greater than 11 then A is a closed subset of $\mathbb{Z}$ under addition, but is not a submonoid since it does not contain 0. (And in fact, this A has no identity element).

We restate Theorem 36:

Theorem 37. *Suppose A is a submonoid of a monoid M. Then A is itself a monoid under the induced operation.*

The following is clear:

Theorem 38. *Every submonoid of a commutative monoid is itself a commutative monoid (under the induced operation).*

Exercise 18. Consider the monoid $\langle \mathbb{Z}_{10}, + \rangle$. Let A be the submonoid $\{0, 2, 4, 6, 8\}$. Make an operation table for A using the induced operation $+_A$.

Theorem 39. *Suppose M is a monoid and that B is a submonoid of M. Let A be a subset of B. Then A is a submonoid of M if and only if A is a submonoid of B.*

Proof. Let e be the identity element of M. Since B is a submonoid, e is the identity of B. Let $*$ be the binary operation on M. Then, as we have seen, the induced operation $*_B$ is the binary operation on the submonoid B.

[4]The set $M = \mathbb{N} \times \mathbb{N}$ under multiplication is a monoid under componentwise products with identity $(1, 1)$. The subset $A = \mathbb{N} \times \{0\}$ is a monoid under the induced operation with identity element $(1, 0)$. However, by our official definition, A is not a submonoid of M since the identity of A is different from that of M.

Suppose that A is a submonoid of M. Then $e \in A$. But e is the the identity of B. Suppose $a, b \in A$. Then $a *_B b = a * b$ since $*_B$ is the induced operation. Since A is a submonoid of M, the element $a * b$ must be in A. Thus $a *_B b$ is in A. We conclude that A is closed in B. We have established both conditions needed for A to be a submonoid of B.

Suppose that A is a submonoid of B. Then the identity e of B is in A. So the identity of M is in A since e is also the identity of M. Suppose $a, b \in A$. Then $a *_B b = a * b$ since $*_B$ is the induced operation. Since A is a submonoid of B, the element $a *_B b$ must be in A. Thus $a * b$ is in A. We conclude that A is closed in M. We have established both conditions needed for A to be a submonoid of M. □

The following lemmas give compatibility between notions in a monoid and in a submonoid:

Lemma 40. *Suppose A is a submonoid of M. If $a \in A$ has an inverse b in the monoid A, then b is the inverse of a in M as well.*

Similarly, if a has inverse b in M, and if $a, b \in A$, then a has inverse b in A.

Proof. Let e be the identity of M, which is also the identity of A since A is a submonoid.

Suppose $a \in A$ has inverse b in A. By definition of inverse $a *_A b = b *_A a = e$. But $a *_A b = a * b$ and $b *_A a = b * a$ since $a, b \in A$. Thus $a * b = b * a = e$ in M.

Now suppose a has inverse b in M, so $a * b = b * a = e$. Suppose $a, b \in A$. Then $a *_A b = a * b$ and $b *_A a = b * a$. Thus $a *_A b = b *_A a = e$. □

Lemma 41 (Multiplicative Version)**.** *Suppose A is a submonoid of M. Let $a \in A$. Then a^n as defined in M is the same as a^n as defined in A where here $n \in \mathbb{N}$. In particular A is closed under powers: $a^n \in A$ for all $a \in A$ and $n \in \mathbb{N}$ where a^n is as defined in M.*

If in addition a is invertible in A then the above extends to all $n \in \mathbb{Z}$.

Proof. This follows by induction for $n \in \mathbb{N}$. The more general statement for a invertible follows from the properties of exponents. □

Exercise 19. Suppose that M is a multiplicative monoid. Let $a \in M$. Show that $A = \{a^n \mid n \in \mathbb{N}\}$ is a submonoid of M. Show that this A is the smallest submonoid containing a in the sense that any submonoid of M containing a must contain A as a submonoid.

Show that if $a \in M$ is invertible, then $A' = \{a^u \mid u \in \mathbb{Z}\}$ is also a submonoid, and that in A' is an Abelian group. Show that it is the smallest submonoid containing a that is a group.

14 The Unit Group

In this section $\langle M, * \rangle$ will be a monoid with identity e. There is a submonoid of M that is of particular interest:

Definition 19. Let $\langle M, * \rangle$ be a monoid. Then M^{inv} is defined to be the set of elements of M that are invertible. Elements of M^{inv} are called *units*. As we will see, M^{inv} forms a group under the induced operation, and we call M^{inv} the *unit group* of M.

The following is just a restatement of the definition of group:

Theorem 42. *A monoid M is a group if and only if $M^{inv} = M$.*

Example 10. (Availing ourselves of some linear algebra). Let $M = M_n(\mathbb{R})$ be the set of n-by-n matrices under matrix multiplication. This is a monoid. The unit group M^{inv} consists of invertible matrices is a group. It is an important group in mathematics called *the general linear group*, and is often denoted as $GL_n(\mathbb{R})$. As we learn in linear algebra, $GL_n(\mathbb{R})$ is the set of n-by-n matrices with determinant not equal to zero.

Example 11. (Assuming familarity with rings) Let $M = R$ be a ring. If we ignore addition, and use multiplication as the operation, we get a monoid. In this case R^{inv} consists of all elements in R with multiplicative inverses. We sometimes write this subset as $R^\times$, and call this the *group of units of R*. Note that if F is a field, then $F^\times$ is $F - \{0\}$.

Exercise 20. Consider $M = \mathbb{N}$, $M = \mathbb{Z}$, and $M = \mathbb{Q}$ under multiplication. What is M^{inv} in each of these cases? Note that we get a group under multiplication in each case.

Now do the same exercise for addition: identify M^{inv} in each of these cases.

As mentioned above we will determine that M^{inv} is a group. First we convince ourselves that M^{inv} is a submonoid, and hence M^{inv} is itself a monoid.

Lemma 43. *Let $\langle M, * \rangle$ be a monoid. Then M^{inv} is a submonoid.*

Proof. The follows from Theorem 8. □

Theorem 44. *Let $\langle M, * \rangle$ be a monoid. Then M^{inv} is a submonoid of M that is in fact a group under the induced operation. Furthermore, M^{inv} is the largest such group in the following sense: if G is a submonoid of M that is a group under the induced operation, then $G \subseteq M^{inv}$.*

Proof. Since M^{inv} is a submonoid, it is in fact a monoid under the induced operation. Does every element of M^{inv} have an inverse in M^{inv}? By definition, every element of M^{inv} has an inverse in M, but Theorem 8 tells us more: the inverse is guaranteed to be in M^{inv} itself.

Suppose G is a submonoid of M that is a group under the induced operation. Then every element of G is a unit in M (see Lemma 40). Thus $G \subseteq M^{inv}$. □

Exercise 21. If $\langle M, * \rangle$ be a commutative monoid, then can you conclude that M^{inv} is an Abelian group under the induced operation?

Exercise 22. Find the group of invertible elements in $\langle \mathbb{Z}_8, \times \rangle$. Make an operation table for this group.

15 Subgroups

Suppose $\langle G, * \rangle$ is a group with identity element e. Some subsets $A \subseteq G$ will be groups under the induced operation, but others will not. If we want A to be a group under the induced operation, we will certainly need A to be closed under $*$. If we want A and G to share the same identity element (as we did with monoids), we would want A to contain e, and then we would need A to be "closed under inverses" as well in the sense that $a^{-1} \in A$ for all $a \in A$ where here a^{-1} is the inverse of a in G. The following shows that these properties are indeed necessary and sufficient:

Lemma 45. *Suppose $\langle G, * \rangle$ is a group with identity element e, and that $A \subseteq G$ is a subset closed under $*$. Then A is a group under the induced operation if and only if* (1) *$e \in A$, in other words A is a submonoid, and* (2) *if $a \in A$ then $a^{-1} \in A$ where a^{-1} is the inverse of a in G.*

Proof. Suppose A is a group under the induced operation. Then A must contain an identity element e'. Note that

$$e' * e' = e' = e' * e.$$

So $e' = e$ by the cancellation law in G. So A is a submonoid of G. By assumption each $a \in A$ has an inverse in A. This inverse is a^{-1} (the inverse of a in G) by Lemma 40. Thus (1) and (2) both hold.

Conversely suppose (1) and (2) hold. By (1) A is a submonoid of G, so is a monoid under the induced operation. By (2) and Lemma 40 every element of A is invertible. Thus A is a group. □

The above lemma motivates the following definitions:

Definition 20. Let $\langle G, * \rangle$ be a group. We say that a subset $A \subseteq G$ is *closed under inverses* if the inverse of a is in A for all $a \in A$.

Definition 21. Suppose $\langle G, * \rangle$ is a group. A *subgroup* is a subset $A \subseteq G$ such that (1) A is closed under $*$, (2) the identity element of G is in A and (3) A is closed under inverses.

By Lemma 45 together with Lemma 40 and Lemma 41 we have the following:

Theorem 46. *If A is a subgroup of G, then A is itself a group under the induced (restricted) operation. The identity of A is the identity in G. The inverse of an element $a \in A$ in A is equal to its inverse in G. For each $a \in A$ and $n \in \mathbb{Z}$ the nth power of a in A is equal to the nth power of a in G.*

Exercise 23. Show that if A is a subgroup of an Abelian group, then A is itself an Abelian group under the induced operation. (Hint: Lemma 35).

Remark. The idea of a subgroup makes it easy to construct examples of groups without proving all the equations that must hold in a group: the associative law equation, the identity law equations, and the inverse law equations. Now that we have the above theorem, we get these equations for free when dealing with a subgroup of a known group.

If A is not a subgroup of a known group, and you want to show it is a group, you have to do it the hard way and prove all the required equations (including associativity, identity equations, and inverse equations).

There are some short-cuts to checking that A is a subgroup. The first allows you to replace checking that the identity is in A with checking that A nonempty.

Proposition 47. *Suppose $\langle G, * \rangle$ is a group. Then a subset $A \subseteq G$ is a subgroup if and only if* (1) *A is nonempty,* (2) *A is closed under $*$,* (3) *A is closed under inverses.*

Proof. First suppose that A is a subgroup. Then since the identity e is in A we know that it is (1) nonempty. Conditions (2) and (3) also hold by the definition of subgroup.

Conversely suppose (1), (2), and (3) hold. Since A is nonempty, it contains an element $a \in A$. By (3) we have $a^{-1} \in A$. So by (2) we have $a * a^{-1} \in A$. In other words $e \in A$. This, together with (2) and (3), means that A is a subgroup by the definition of subgroup. □

The following gives you a way to check two conditions instead of three (we use multiplicative notation for convenience, but it can be translated into additive notation):

Proposition 48. *Suppose $\langle G, * \rangle$ is a group. Then a subset $A \subseteq G$ is a subgroup if and only if* (1) *A is nonempty and* (2) *$a * b^{-1} \in A$ for all $a, b \in A$.*

Exercise 24. Prove the above proposition. Hint: assuming (1) and (2), first show that $e \in A$, then show A is closed under inverses. Finally show A is closed under $*$.

For finite groups the situation is easier and there is a nice shortcut. Actually the enclosing group G does not even need to be finite, just the subset A.

Proposition 49. *Suppose $\langle G, * \rangle$ is a group, and that A is a finite subset of G. Then A is a subgroup of G if and only if* (1) *A contains the identity, and* (2) *A is closed under $*$.*

Proof. One direction is clear, so we focus on the other direction. Suppose $a \in A$. Consider the translation function $T_a : A \to A$ given by the rule

$$T_a(x) = a * x.$$

This function is well-defined since A is closed under $*$. This function is injective by the cancellation law. Since A is finite, this function must be surjective as well by properties of finite sets. Since $e \in A$, there must be an element b mapping to e. Since $T_a(b) = e$ we conclude that $a * b = e$. Thus b is the inverse of a (Theorem 15).

Thus A is closed under inverses. This together with (1) and (2) gives the result. □

Corollary 50. *Every finite submonoid of a group is a subgroup.*

Exercise 25. Suppose $\langle G, * \rangle$ is a group, and that A is a finite subset of G. Suppose that (1) A is nonempty and (2) A is closed under $*$. Show that the identity must be in A. Conclude that A must be a subgroup.

In other words, (1) in the above Proposition can be replaced by "(1) A is nonempty". Hint: if $a \in A$, then all its powers are in A. Two powers with distinct exponents are equal so the identity e is a power of a.

Exercise 26. Suppose M is a monoid with identity e, then show that M and $\{e\}$ are submonoids of M. Suppose G is a group with identity e. Then show that G and $\{e\}$ are subgroups of G.

16 Products and Sums of Finite Sequences

We will use the term *finite ordered sequence* in a set S for any family $(a_i)_{i \in I}$ of elements of S such that I is a finite, totally ordered set. We define an equivalence relation for such finite ordered sequences: suppose $(a_i)_{i \in I}$ and $(b_J)_{i \in J}$ are two finite ordered sequences in S then we say that these finite ordered sequences are *order equivalent* if (i) the cardinality of I and J are equal, and (ii) the unique order-preserving bijection $\gamma \colon I \to J$ has the property that $b_{\gamma(i)} = a_i$ for all $i \in I$. The reader can verify the following:

Lemma 51. *Order equivalence is an equivalence relation among finite ordered sequences in a set S.*

The definition of order equivalence makes use of the following fact from set theory (that can be proved by induction):

Lemma 52. *If I and J are finite totally ordered sets of the same size, then there is a unique order-preserving bijection $I \to J$.*

We say that a finite ordered sequence is *nonempty* if I is nonempty. We now define the product of finite ordered sequences:

Definition 22. Fix a set S and a binary operation $* \colon S \times S \to S$. Then we define the *product* $\prod (a_i)_{i \in I}$ of any nonempty finite ordered sequence $(a_i)_{i \in I}$ in S recursively as follows. If I has one element, call it α, then $\prod (a_i)_{i \in I}$ is defined to be just a_α. If I has more than one element then let ω be the last element of I, and let I' be $I - \{\omega\}$. Then we define the product by the recursive equation

$$\prod (a_i)_{i \in I} \stackrel{\text{def}}{=} \left(\prod (a_i)_{i \in I'} \right) * a_\omega$$

where $(a_i)_{i \in I'}$ denotes the restriction of the given family to I'.

We will sometimes follow the common convention and write $\prod (a_i)_{i \in I}$ as

$$\prod_{i \in I} a_i.$$

If $I = [m, n]_{\mathbb{Z}}$ is the interval of integers from m to n (inclusive) then we can employ the notation

$$\prod_{i=m}^{n} a_i$$

for the product $\prod(a_i)_{i\in I}$. If there is no danger of ambiguity, we can use abbreviations for products such as

$$\prod a_i \qquad \text{or} \qquad \prod_i a_i.$$

If the operation is written + then we usually use additive notation where we replace $\prod$ with $\sum$ and use the term *sum* instead of *product*.

Definition 23. If M is a monoid or a group, then we extend the above definition to include the empty sequence: we define the product of the empty finite ordered sequence to be the identity element of M.

Theorem 53. *Let S be a set with binary operation $*\colon S \times S \to S$. If $(a_i)_{i\in I}$ and $(b_J)_{i\in J}$ are order equivalent nonempty finite ordered sequences, then*

$$\prod(a_i)_{i\in I} = \prod(b_j)_{i\in J}.$$

(If $S = M$ is a monoid, then this extends to the empty sequence.)

Proof. A straightforward induction argument sufficies. □

17 The General Associativity Laws

Now we consider various generalizations of the associativity law. These are all based on the following result:

Theorem 54 (General Associative Law: First Form)**.** *Let S be a set with an associative binary operation $*\colon S\times S \to S$. Suppose $(a_i)_{i\in I}$ is a nonempty finite ordered sequence in S. Suppose also that I is the disjoint union of two nonempty subsets I_1 and I_2 and that every element of I_2 is an upper bound of I_1. Then*

$$\prod_{i\in I} a_i \;= \prod_{i\in I_1} a_i * \prod_{i\in I_2} a_i.$$

Proof. If $I_2 = \{\omega\}$ is a singleton then ω must be the maximum in I, and the result now follows from the recursive definition of the product. We proceed by induction and assume I_2 has more than one element. Let ω be the maximum element of I_2 and let $I_2' = I_2 - \{\omega\}$. Note also that ω is the maximum of all of I. Observe that, by the associative law and the induction hypothesis (on the size of I_2)

$$\left(\prod_{i\in I_1\cup I_2'} a_i\right) * a_\omega = \left(\prod_{i\in I_1} a_i * \prod_{i\in I_2'} a_i\right) * a_\omega = \prod_{i\in I_1} a_i * \left(\left(\prod_{i\in I_2'} a_i\right) * a_\omega\right)$$

By definition of product, the left-hand side is just $\prod(a_i)_{i\in I}$ and the right-hand side simplifies as desired:

$$\prod_{i\in I_1} a_i * \left(\left(\prod_{i\in I_2'} a_i\right) * a_\omega\right) = \prod_{i\in I_1} a_i * \prod_{i\in I_2} a_i.$$

□

For monoids we can extend this law to possibly empty sequences:

Corollary 55 (General Associative Law: First Form for Monoids)**.** *Let $\langle M, * \rangle$ be a monoid. Suppose $(a_i)_{i \in I}$ is a finite ordered sequence in M. Suppose also that I is the disjoint union of two subsets I_1 and I_2 where every element of I_2 is an upper bound of I_1. Then*

$$\prod_{i \in I} a_i = \prod_{i \in I_1} a_i * \prod_{i \in I_2} a_i.$$

Proof. If I_1 and I_2 are nonempty then use the theorem. If I_1 or I_2 are empty, replace the corresponding product with the identity element, and the result is clear. □

We can freely remove or add terms that are the identity element:

Theorem 56. *Let $\langle M, * \rangle$ be a monoid with identity element. Suppose $(a_i)_{i \in I}$ is a finite ordered sequence in S. Suppose I' is a subset of I such that $a_i = e$ if $i \in I - I'$, in other words, the ordered subsequence $(a_i)_{i \in I'}$ contains all the nontrivial terms of $(a_i)_{i \in I}$. Then*

$$\prod_{i \in I} a_i = \prod_{i \in I'} a_i.$$

In particular, if $a_i = e$ for all $i \in I$ then $\prod_{i \in I} a_i$ is equal to e, the product of the empty subsequence.

Proof. This can be proved by induction on the size of I. We distinguish two cases: where I' contains the maximum element of I, and where I' does not contain this maximum. □

We can extend this law to more than two terms:

Theorem 57 (General Associative Law: Second Form)**.** *Let S be a set with an associative binary operation $*\colon S \times S \to S$. Suppose $(a_i)_{i \in I}$ is a nonempty finite ordered family in S. Suppose also that I is the disjoint union of $k \geq 1$ nonempty subsets $I_1, \ldots, I_k$ such that every element of a given I_t is an upper-bound for all I_s with $s < t$. Then*

$$\prod_{i \in I} a_i = \prod_{j=1}^{k} \prod_{i \in I_j} a_i.$$

Proof. If $k = 1$ then the result is clear by the definition of singleton products. We proceed by induction to $k \geq 2$. Suppose $I' = I_1 \cup \cdots \cup I_{k-1}$. Then by Theorem 54, and the induction hypothesis,

$$\prod_{i \in I} a_i = \prod_{i \in I'} a_i * \prod_{i \in I_k} a_i = \left(\prod_{j=1}^{k-1} \prod_{i \in I_j} a_i \right) * \prod_{i \in I_k} a_i.$$

The right-hand side simplifies to the desired express (using the definition of product). □

For monoids we can drop the requirement that each I_j is nonempty. We can even allow $k = 0$:

Corollary 58 (General Associative Law: Second Form for Monoids)**.** *Let $\langle M, * \rangle$ be a monoid with identity element e. Suppose $(a_i)_{i \in I}$ is a nonempty finite ordered family in M. Suppose also that I as the disjoint union of $k \geq 0$ subsets $I_1, \ldots, I_k$ such that every element of a given I_t is an upper-bound for all I_s with $s < t$. Let J be the set of nonnegative integers up to k. Then*

$$\prod_{i \in I} a_i = \prod_{j \in J} \prod_{i \in I_j} a_i.$$

Proof. If I is empty, then we get a product $\prod_{j \in J} e$ of identity elements which is the identity, and the result follows. If I is nonempty then we can remove from J all j such that I_j is empty (Theorem 56), so we reduce to the case where each I_j is nonempty. We then can use the above theorem to establish the result. □

Next we consider the idea that we can rearrange the parenthesis for associative operators without affecting the result. Suppose $*$ is associative, then we regard expressions such as $((a_1 * a_2) * a_3) * a_4$ and $(a_1 * a_2) * (a_3 * a_4)$ as representing different calculations of the same result (by associativity). We formalize the notion of "calculation" as follows:[5]

Definition 24. Let S be a set with a binary operation $*\colon S \times S \to S$. We define the concept of a "calculation" of a nonempty finite sequence with terms in S recursively as follows: If $I = \{i_0\}$ is a singleton set, the only calculation of the sequence $(a_i)_{i \in I}$ is a_{i_0}. If I has size $n \geq 2$ then the calculations of $(a_i)_{i \in I}$ are the elements of S of the form $b_1 * b_2$ where I_1 and I_2 are nonempty subsets partitioning I such that every element of I_2 is an upper bound of I_1, where b_1 is a calculation of $(a_i)_{i \in I_1}$, and where b_2 is a calculation of $(a_i)_{i \in I_2}$.

Example 12. Consider the sequence $4, -3, 2, 1, 3$ with terms in $\mathbb{Z}$. Consider subtraction as our binary operation. Then

$$\Big(4 - \big((-3) - (2 - 1)\big)\Big) - 3 = 5$$

and

$$(4 - (-3)) - ((2 - 1) - 3)) = 9$$

are two calculations. The results are not equal, which illustrates that subtraction is not associative.

Theorem 59 (General Associative Law: Third Form)**.** *Let S be a set with an associative binary operation $*\colon S \times S \to S$. Suppose $(a_i)_{i \in I}$ is a nonempty finite ordered family in S. Then all calculations of $(a_i)_{i \in I}$ using $*$ are equal to the product $\prod(a_i)_{i \in I}$. In particular, all calculations of $(a_i)_{i \in I}$ are equal.*

Proof. If $I = \{i_0\}$ then the only calculation of $(a_i)_{i \in I}$ is a_{i_0}, which is equal to $\prod(a_i)_{i \in I}$ by definition. So we can assume the size of I is greater than one, and proceed by induction. Suppose b is a calculation of $(a_i)_{i \in I}$, so $b = b_1 * b_2$

[5] With a bit more work we could formalize "calculation" in terms of the binary treelike structure used to describe how to calculate the result. Then the theorem would say that any two such trees associated to a given sequence $a_1, a_2, \ldots, a_n$.

where $b_1 = \prod(a_i)_{i\in I_1}$, where $b_2 = \prod(a_i)_{i\in I_2}$, where I_1 and I_2 partition I, and where every element of I_2 is an upper bound of I_1. By the induction hypothesis and by Theorem 54

$$b = b_1 * b_2 = \prod_{i\in I_1} a_i * \prod_{i\in I_2} a_i = \prod_{i\in I} a_i.$$

□

Finally we mention that our definition of a^n is equivalent to the other popular definition:[6]

Theorem 60. *Let $a \in M$ where M is a multiplicative monoid. If n is a positive integer then*

$$a^n = \prod_{i=1}^{n} a.$$

Proof. We can prove this by induction on n. The base case reduces to the equation $a^1 = a$ established above. The induction step can be proved as follows:

$$a^{n+1} = a^n a^1 = \left(\prod_{i=1}^{n} a\right) a = \prod_{i=1}^{n+1} a.$$

□

Exercise 27. Translate the results of this section to additive notation.

18 General Commutativity Laws

In this section we will develop general commutativity laws for multiplicative monoids. We use monoids since the statements and proofs are a bit more elegant than in a more general setting. But we can certainly generalize these results to "commutative semigroups" where we do not require an identity element (you just need to be careful about empty sequences). Also it is straightforward to convert the results to additive notation.

First we show we can move any term to the end:

Lemma 61. *Let $\langle M, *\rangle$ be a commutative monoid. If $(a_i)_{i\in I}$ is a finite ordered sequence of elements in M and if $i_0 \in I$ then*

$$\prod_{i\in I} a_i = \left(\prod_{i\in I'} a_i\right) * a_{i_0}$$

where $I' = I - \{i_0\}$ with the restricted ordering.

[6] We used iteration to define powers, but other authors such as Bourbaki in *Algebra* use the equation in Theorem 60 instead, at least for positive n.

Proof. Let I_1 be the set of elements of I strictly smaller than i_0 and let I_2 be the set of elements of I strictly larger than i_0. From Corollary 55 and the recursive definition of products we have

$$\prod_{i\in I} a_i = \prod_{i\in I_1\cup\{i_0\}} a_i * \prod_{i\in I_2} a_i = \left(\left(\prod_{i\in I_1} a_i\right) * a_{i_0}\right) * \prod_{i\in I_2} a_i.$$

Using associativity and commutivity we have

$$\left(\left(\prod_{i\in I_1} a_i\right) * a_{i_0}\right) * \prod_{i\in I_2} a_i = \prod_{i\in I_1} a_i * \left(a_{i_0} * \prod_{i\in I_2} a_i\right) = \prod_{i\in I_1} a_i * \left(\left(\prod_{i\in I_2} a_i\right) * a_{i_0}\right).$$

Finally, by associativity and Corollary 55 we have

$$\prod_{i\in I_1} a_i * \left(\left(\prod_{i\in I_2} a_i\right) * a_{i_0}\right) = \left(\prod_{i\in I_1} a_i * \prod_{i\in I_2} a_i\right) * a_{i_0} = \left(\prod_{i\in I'} a_i\right) * a_{i_0}.$$

□

From this lemma we can prove the first commutative law which states that the order of the index set does not affect the product:

Theorem 62 (General Commutative Law: First Form)**.** *Let $\langle M, *\rangle$ be a commutative monoid. Let I and J be two ordered finite sets with the same underlying set which we call K. Let $(a_k)_{k\in K}$ be a family of elements of M indexed by K, which we we can think of as a finite ordered sequence either as $(a_i)_{i\in I}$ or as $(a_j)_{j\in J}$. Note these sequences have the same terms, but perhaps given in a different order. Then*

$$\prod_{i\in I} a_i = \prod_{j\in J} a_j.$$

Proof. The proof is by induction on the size of K. If K is empty, then both sides of the equation in question are the identity element by definition of the empty product. So assume K is not empty, and fix an element $k \in K$.

Let $I' = I - \{k\}$ and $J' = J - \{k\}$ where these sets have the restricted order from I and J respectively. Note that I' and J' have the same underlying set. By the above lemma and the induction hypothesis we have

$$\prod_{i\in I} a_i = \left(\prod_{i\in I'} a_i\right) * a_k = \left(\prod_{j\in J'} a_j\right) * a_k = \prod_{j\in J} a_j.$$

□

Definition 25. Let $\langle M, *\rangle$ be a commutative monoid and let $(a_k)_{k\in K}$ be a family of elements of M indexed by a finite set K. Then we define $\prod(a_k)_{k\in K}$ to be the value resulting from any chosen total ordering of K. By the above theorem, all orderings give the same value, so this new type of product is well-defined, and agrees with the old definition (Definition 22) for any ordering of K.

Recall that in Section 16 we considered an equivalence relation for finite ordered sequences called *order equivalence*, and showed that order equivalent sequences have the same product. Recall that two finite ordered sequence $(a_i)_{i\in I}$ and $(b_J)_{i\in J}$ with values in a given set are *order-equivalent* if (i) the cardinality of I and J are equal, and (ii) the unique order-preserving bijection $\gamma\colon I \to J$ has the property that $b_{\gamma(i)} = a_i$ for all $i \in I$. We now define another equivalence relation where we drop the requirement that γ be order-preserving: we declare two finite families $(a_i)_{i\in I}$ and $(b_J)_{i\in J}$ with values in a given set to be *multiset-equivalent* if there is a bijection $\gamma\colon I \to J$ with the property that $b_{\gamma(i)} = a_i$ for all $i \in I$. (Note: we can define a multiset as an equivalence class under this relation. However, we will not need multisets, per se, in this document.)

Theorem 63 (General Commutative Law: Second Form)**.** *Let $(a_i)_{i\in I}$ and $(b_j)_{j\in J}$ be two finite families of elements of M where $\langle M, *\rangle$ is a commutative monoid. If $(a_i)_{i\in I}$ and $(b_j)_{j\in J}$ are multiset-equivalent then*

$$\prod_{i\in I} a_i = \prod_{j\in J} a_j.$$

Proof. Let γ be a bijection $\gamma\colon I \to J$ with the property that $b_{\gamma(i)} = a_i$ for all $i \in I$. Fix any ordering on I. If J comes with an order, then ignore it. Instead, impose the unique order on J such that γ is an order preserving map. Under this choice of order we have that $(a_i)_{i\in I}$ and $(b_j)_{j\in J}$ are order-equivalent sequences and so, using these orderings,

$$\prod_{i\in I} a_i = \prod_{j\in J} b_j.$$

Note that this equality continuous to hold when we drop the orderings on I and J (Theorem 62 and Definition 25).

□

19 Induced Operation on Functions

Above in Section 12 we used the term "induced operation" for the restriction of a binary operation to a closed subset. We will switch gears a bit here and define another kind of "induced operations", namely induced operations on functions. The motivation is observation that if we have a binary operation on a set M then we get a natural binary operation on the set of functions $S \to M$ where S is any given domain:

Definition 26. Suppose S and M are sets. Then let $\mathcal{F}(S, M)$ be the set of functions $S \to M$.

Definition 27. Let M be a set with a binary operation $*\colon M \times M \to M$ and let S be be a fixed set to serve as a domain. Then we define a binary operation $*_{\mathcal{F}}$ on the set of functions $\mathcal{F}(S, M)$ as follows: given a pair of functions $f, g \in \mathcal{F}(S, M)$ define $f *_{\mathcal{F}} g$ by the rule

$$(f *_{\mathcal{F}} g)(x) \stackrel{\text{def}}{=} f(x) * g(x).$$

We call $*_{\mathcal{F}}$ the *induced operation on functions*. We will often write $*_{\mathcal{F}}$ as just $*$ and use context to distinguish this operation from that on M.

Remark. Observe that $*_{\mathcal{F}}$ (or just $*$) gives a binary operation

$$\mathcal{F}(S, M) \times \mathcal{F}(S, M) \to \mathcal{F}(S, M).$$

The following lemmas are proved in a straightforward manner.

Lemma 64. *Suppose that $*\colon M \times M \to M$ is an associative operation. Then the induced operation on $\mathcal{F}(S, M)$ is also associative.*

Lemma 65. *Suppose that $*\colon M \times M \to M$ is a commutative operation. Then the induced operation on $\mathcal{F}(S, M)$ is also commutative.*

Lemma 66. *Suppose that $*\colon M \times M \to M$ is a binary operation with identity element $e \in M$. Let $\mathbf{e}\colon S \to M$ be the constant function defined by the equation $\mathbf{e}(x) = e$ for all $x \in S$. Then $\mathbf{e}$ is an identify element for the induced operation on $\mathcal{F}(S, M)$.*

Theorem 67. *Suppose M is a monoid and that S is a set. Then $\mathcal{F}(S, M)$ is a monoid under the induced operation. The identity element is the constant function $S \to M$ whose values are given by the identity element of M. If M is a commutative monoid then so is $\mathcal{F}(S, M)$.*

Theorem 68. *Suppose $\langle G, * \rangle$ is a group and that S is a set. Then $\mathcal{F}(S, G)$ is a group under the induced operation. The identity element is the constant function $S \to G$ whose values are given by the identity element of G.*

Proof. By Theorem 67, $\mathcal{F}(S, G)$ is a monoid. To show it is a group we need to show every element $f \in \mathcal{F}(S, G)$ has an inverse. Let $g\colon S \to G$ be defined by the rule $x \mapsto f(x)^{-1}$ (for convenience we adopt multiplicative notation for G, and use e for the identity in G). Then for any $x \in S$ we have

$$(f * g)(x) = f(x) * g(x) = f(x) * f(x)^{-1} = e.$$

In other words, $f * g$ is the constant function $\mathbf{e}$ whose values are e. Similarly, we can show $g * f = \mathbf{e}$. Thus g is the inverse of f in $\mathcal{F}(S, G)$. □

Exercise 28. Suppose M monoid and that S is a set. Describe the elements in the group $\mathcal{F}(S, M)^{inv}$ of invertible elements.

Exercise 29. (Taking as given the notions of continuity and differentiability) Consider $\langle \mathbb{R}, + \rangle$. Let $G = \mathcal{F}(\mathbb{R}, \mathbb{R})$. As we have seen, G is a group under the induced addition $+$. Let H_1 be the subset of continuous functions. Show that H_1 is a subgroup of G. Let H_2 be the subset of differentiable functions. Show that H_2 is a subgroup of H_1 and G.

Exercise 30. (Taking as given the notions of continuity and differentiability) Consider $\langle \mathbb{R}, \cdot \rangle$. Let S be the open interval $(0, 1)$, and let $M = \mathcal{F}(S, \mathbb{R})$. As we have seen, M is a monoid under the induced multiplication. A typical element of M is the function defined by $f(x) = 1/x$. Let H_1 be the subset of continuous functions. Show that H_1 is a submonoid of M. Let H_2 be the subset of differentiable functions. Show that H_2 is a submonoid of H_1 and G.

19.1 Cartesian Powers of a Monoid

We now consider the special case of sets of functions $\mathcal{F}(S, M)$ where $S = \{1, \ldots, n\}$ where n is a positive integer. Then an element f of $\mathcal{F}(S, M)$ can be completely described simply by specifying n values:

$$f(1), f(2), \ldots, f(n).$$

In other words, elements of $\mathcal{F}(S, M)$ can be thought of as finite sequences $(a_1, \ldots, a_n)$ with $a_i \in M$. Here we identify $(a_1, \ldots, a_n)$ with the function $i \mapsto a_i$ which is an element of $\mathcal{F}(S, M)$ where $S = \{1, \ldots, n\}$.

Definition 28. Let M be a set, and let n be a positive integer. Then M^n is defined to be $\mathcal{F}(\{1, \ldots, n\}, M)$. As discussed above, we think of M^n as the set of finite sequences of the form $(a_1, \ldots, a_n)$ where each $a_i \in M$. We call M^n the nth *Cartesian power of* M.

If M has a binary operation $*\colon M \times M \to M$, then we get an *induced operation* $*$ on M^n using Definition 27. The induced operation can be written as

$$(a_1, \ldots, a_n) * (b_1, \ldots, b_n) = (a_1 * b_1, \ldots, a_n * b_n)$$

where the operation $*$ in the left-hand side is the induced operation, and the operation in the right hand side is the operation in M.

The following is just a special case of previous results for $\mathcal{F}(G, S)$:

Theorem 69. *Let n be a positive integer. If M is a monoid, then so is M^n (under the induced operation). If M is a commutative monoid, then so is M^n. If G is a group then so is G^n. If G is an Abelian group, then so is G^n.*

Exercise 31. (Taken as given the vector spaces $\mathbb{R}^2$ and $\mathbb{R}^3$). Consider $\mathbb{R}$ as a group under addition. Show that the induced operation on $\mathbb{R}^2$ and $\mathbb{R}^3$ are the usual vector addition. Conclude that $\mathbb{R}^2$ and $\mathbb{R}^3$ are Abelian groups under vector addition.

Exercise 32. Suppose H is a subgroup of G. By extending codomains, you can identify every element of $\mathcal{F}(S, H)$ with an element of $\mathcal{F}(S, G)$, and so can identify $\mathcal{F}(S, H)$ with a subset of $\mathcal{F}(S, G)$. Show that when you do this, the operation on $\mathcal{F}(S, H)$ is the restriction of the operation on $\mathcal{F}(S, G)$, and that $\mathcal{F}(S, H)$ can be regarded as a subgroup of $\mathcal{F}(S, G)$. Note: as a special cases we see H^n can be regarded as a subgroup of G^n for all positive n.

(Note: the analogous property holds for monoids and submonoids, and this can be shown along the way.)

Remark. For any set G it is common to define G^0 to be a set with one element. When G is a group or monoid, it is common to regard G^0 as a group using the unique binary operation on this one-point set G^0. (This is consistent with thinking of G^0 as the set $\mathcal{F}(\varnothing, G)$).

20 Composition of Functions

Above we considered one sort of operation on functions which gives a monoid, assuming the codomain is a monoid. Now we describe another way to form monoids

whose elements are functions. Let X be a given set and consider $\mathcal{F}(X, X)$, the set of functions from X to itself. In this case we write $\mathcal{F}(X, X)$ simply as $\mathcal{F}(X)$. Note that composition $\circ$ gives a binary operation

$$\circ \colon \mathcal{F}(X) \times \mathcal{F}(X) \to \mathcal{F}(X)$$

and that if X is finite of size n then $\mathcal{F}(X)$ has n^n elements. The following theorems are straightforward consequences of basic set-theoretical facts:

Theorem 70. *Let X be a set. Then $\mathcal{F}(X)$ is a monoid under composition. The identity element is the identity function $X \to X$.*

Theorem 71. *Let X be a set. Then a function $f \in \mathcal{F}(X)$ is invertible in the monoid $\mathcal{F}(X)$ if and only if it is invertible as a function $X \to X$. The inverse of an invertible f in the monoid $\mathcal{F}(X)$ is the inverse function $f^{-1} \colon X \to X$.*

Definition 29. Let X be a set. The group of units of $\mathcal{F}(X)$ is called the *symmetric group* on X, which we can write as $\mathcal{S}(X)$ or $\mathcal{S}_X$, and its elements are called *permutations* of X.

Remark. Note that if X is finite with n elements then the symmetric group on X has $n!$ elements. If $X = \{1, \ldots, n\}$ then the symmetric group is sometimes just written as S_n (we sometimes write it as $\mathcal{S}_n$)

Remark. Observe that if A is a group or monoid then $\mathcal{F}(A, A)$ has two standard operations: composition and the induced operation. We have to be clear of what is meant in any given situation, especially since we sometimes use multiplicative notation fg for both operations. Similarly, f^{-1} can have two possible meanings depending on which monoid, and hence which binary operation, is under consideration.

Exercise 33. Consider $\langle \mathbb{R}, \cdot \rangle$, and let $M = \mathcal{F}(\mathbb{R}, \mathbb{R})$. Let $f, g \in M$ be defined by the equations $f(x) = x + 1$ and $g(x) = x^2$. What is $fg, gf, f \circ g$, and $g \circ f$?